An Examination of the Effect of CAPSI as a Learning System in Developing Knowledge and Critical Thinking in Two Blended Learning Courses

by

Louis Svenningsen

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

Doctor of Philosophy

Department of Psychology

University of Manitoba

Winnipeg, Manitoba, Canada

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION

An Examination of the Effect of CAPSI as a Learning System in Developing Knowledge and Critical Thinking in Two Blended Learning Courses

By

Louis Svenningsen

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree

Of

Doctor of Philosophy

Acknowledgements

I would like to thank my advisor, Dr. Joe Pear, for his wealth of knowledge, infinite patience, and un-ending support. He allowed me the opportunity to develop skills and confidence by listening to and accepting my ideas, and then providing thoughtful and abundant feedback. I am humbled by the never-ending trust he has had in me.

It is hard to support someone in pursuit of a goal without really knowing what is involved in that pursuit. Therefore, I would like to thank the patience and support that has been provided by my family. Thank you to my wife Joy, who, for many years, received continuous statements about completion dates, and then watched those dates slip by. I would also like to thank my parents and children, who unconditionally support me in my pursuits.

The research would not have occurred without the support of Dr. Christine Blais. I am indepted to her for allowing the research to be part of the course, Introduction to University, and for her feedback on this thesis.

I am very lucky to have had Dr. Garry Martin and Dr. Dieter Schonwetter on my committee. This thesis would not have been completed without their feedback, patience, guidance, and optimism. I hope I have proven worthy of their support.

Finally, I would like to thank Michael Obrien-Moran. He was supportive of this research from the beginning, and was the main cog in allowing it to succeed in the Introduction to University courses. At times, it was my belief that he was more optimistic of the research than I was. He took on many extra tasks to insure that this research could be done.

Table of Contents

Table of Contents 1

Abstract 7

An Examination Of The Effects of CAPSI As A Learning System in Developing Knowledge and Critical Thinking in Two Blended Learning Courses 9

The Internet 10

Comparisons of Web-based Courses with Standard Lecture Courses 11

Blended Learning 13

Effectiveness of Blended Learning 15

Research Involving Blended learning 17

A Behavioural Approach to Education 19

A Behavioural Approach to Education: PSI 21

Effectiveness of PSI 23

Criticisms of PSI 24

Computer-Aided Personalized System of Instruction 25

Computer-Aided Personalized System of Instruction in a Blended Learning Design 28

Critical Thinking 28

Teaching Critical Thinking using Bloom's Taxonomy 30

Teaching Critical Thinking with CAPSI Using Bloom's Taxonomy 32

The Present Study 34

EXPERIMENT 1 37

Method 39

Participants 39

Procedure 40

The CAPSI Program 43

The Applied Critical Thinking Measure (ACTM) 44

Data Collection and Reliability Measurement 46

Data Analysis 47

Results 47

Final Exam 47

Critical Thinking Exam Question 49

Questions on Course Content 50

ACTM 51

Discussion 52

EXPERIMENT 2 54

Method 54

Participants 54

Procedure 55

CAPSI Assignment 57

Data Collection and Assessment 59

Inter-rater reliabilities 59

Results 59

Discussion 61

EXPERIMENT 3 62

Method 63

Participants 63
Procedure 64
Quizzes 65
CAPSI 66
Exams 67
Data Collection and Assessment 68
Results 68
Discussion 71
General Discussion 73
Critical Thinking 74
Course Content 75
Decrease in dropouts 76
Limitations 76
ACTM 77
CAPSI 78
Instructor Effects 78
Implications 79
Future Directions 81
ACTM 81
CAPSI 81
References 85
Appendices
A – Final Exam – Introduction to University 103

B - Marking Key Final Exam 2006 – Fall Term 105

C - Applied Critical Thinking Measure 112

D - Scoring Key for the Applied Critical Thinking Measure 119

E - Detailed Description of Procedure for Experiment 1 123

F - Detailed Description Questions Answers For Experiment 2 128

G - Examples of Questions Used At Each Level in Experiment 3 131

H- Consent Form for Experiment 1 132

I- Consent Form for Experiment 2 134

J- Consent Form for Experiment 3 136

List of Tables and Figures

Table 1: Nested design model used in the present study 47

Table 2: Mean scores, standard deviations, and group sizes for final exam scores.... 48

Table 3: Between-subject effects on the final exam 48

Table 4: Pairwise comparisons between instructors on final exam scores 49

Table 5: Participants, means, and standard devuiations on the critical thinking question 49

Table 6: Results of nested design analysis on the critical thinking exam question 50

Table 7: Pair-wise instructor comparisons on the critical thinking exam question 50

Table 8: Participants, means, and standard deviations on the content questions........ 51

Table 9: Results of nested design analysis on the content questions 51

Table 10: Participants, means, and standard deviations for the end of term ACTM . 51

Table 11: Between-subjects effects for the end of term ACTM............................... 52

Table 12: Pair-wise comparisons on the end of term ACTM 52

Table 13: Populations, means, and standard deviations for CAPSI and non-CAPSI sections on the pre-ACTM and ACTM 60

Table 14: Independent samples t-tests for the equality of means between the two lecture sections using a two-tailed test.. 60

Table 15: Number and percentage of students who wrote each of the quizzes 68

Table 16: Number and percentage of students who completed the CAPSI unit tests. 69

Table 17: Mean differences between groups of all students who wrote each exam... 69

Table 18: Mean differences on exam scores between groups of students who completed the course.. 70

Table 19: Repeated measures ANOVA for exams between and within sections 70

Figure 1: Mean scores for the pre-ACTM (Before) and ACTM (After) 61

Figure 2: Mean scores for CAPSI and non-CAPSI sections over the 4 exams 71

Abstract

The growth of the Internet has begun to make an impact on how course information is delivered, leading to a new pedagogical approach called blended learning (Concannon, Flynn & Campbell, 2005; Dzubian et. al. 2004). One pedagogical design that facilitates this changing approach combines Keller's Personalized System of Instruction (Brothen & Wambach, 1999; Keller, 1968) with Web-based technology, resulting in a learning-management method called Computer-Aided Personalized System of Instruction (CAPSI) (Pear & Kinsner, 1988). Three experiments were conducted in order to assess CAPSI in regard to student course knowledge and critical thinking development. These experiments incorporated three different research designs. Experiments 1 and 2 were conducted in a first year *Introduction to University* course at the University of Manitoba. In Experiment 1 two lecture sections received a CAPSI assignment while two received an extra paper assignment. The results indicate that the CAPSI sections performed significantly better on a final exam and a critical thinking question in the final exam. There was also a positive, but non-significant difference between the CAPSI and the non-CAPSI sections on the content questions and on a measure of critical thinking. In Experiment 2, one lecture section of *Introduction to University* received a CAPSI assignment, while another lecture section was assigned a research paper. Students in both lecture sections were assessed at the same level for critical thinking on the ACTM prior to the CAPSI or paper assignment. There was a significant difference in scores between sections in favour of the CAPSI section after the completion of the CAPSI or paper assignment. Experiment 3 involved students in two sections of first year *Introduction*

to Psychology course at the University of Winnipeg. In this experiment, the CAPSI group performed better on four multiple-choice exams administered during the course. These difference, however, were not statistically significant, possibly due to a significantly higher dropout rate in the non-CAPSI section than in the CAPSI section. Over all three experiments, the CAPSI sections consistently outperformed the sections with which they were compared with, indicating that CAPSI is an effective empirically based educational methodology.

An Examination Of The Effects of CAPSI As A Learning System in Developing Knowledge and Critical Thinking in Two Blended Learning Courses

The primary design for promoting learning and critical thinking in higher education has remained relatively unchanged since universities first began: an instructor lecturing to a group of students (Brothen & Wambach, 1998; Levy, 2004). There has also been little change in the lecture method per se, apart from the technological development of new presentation tools, evolving from nothing, to blackboards, to overhead projectors, LCD projectors, video, etc. (Brothen & Wambach, 1998; Dzubian, Hartman,& Moskal, 2004; Levy, 2004; Terenzini, 1999). However, the growth of the Internet, has begun to make an impact on how course information is delivered, leading possibly to a re-thinking of the traditional learning model in higher education that includes blending on-line technology with face-to-face classroom delivery (Dzubian et. al. 2004).

This technological development appears to be driving a change in postsecondary instruction (Chickering & Kytle, 1999; Concannon, Flynn, & Campbell, 2005; Falconer and Littlejohn, 2007; Waschull, 2001). Universities and instructors are starting to include more technology in order to first, provide more flexible designs for knowledge delivery, critical thinking development, and assessment, and second, to allow students the opportunity to pursue an education while being involved with other activities, primarily part time or full time work schedules (Ausburn, 2004; Concannon, Flynn & Campbell, 2005). Another, although perhaps less commonly recognized driving force in the change in higher education is the rise of behaviourism, which puts the focus on measurable behavioural outcomes of the students as opposed to the performance of the instructor (e.g., Kulik, Kulik, & Bangert-Drowns, 1990). The following provides an overview of these changes occurring in higher education – advances in computer technology to allow for more flexibility in course design, and the rise of behaviourism.

The Internet

The Internet has become an integral part of teaching and learning in universities (Ausburn, 2004; Concannon et al., 2005; Daley, Williams, Davis, & Dymock, 2001; Dewhurst, Macleod, & Norris, 2000). In the United States, nearly 3.2 million students enrolled in at least one online course during the fall 2005 term, which is a 39% increase over the 2.3 million from the previous year (Allan & Seaman, 2006). This trend continued into the fall of 2006 with almost 3.5 million students (20 % of all U.S. higher education students) taking at least one online course (Allan & Seaman, 2006); a 9% increase over the number reported the previous year. This growth rate for online enrollments far exceeded the 1.5% growth of the overall higher education student population in the United States. Moreover, this trend is not confined to the United States. This number of online enrolments is expected to grow to an estimated 80 million internationally by 2025 (Hosie, Schibeci, & Backhau, 2005).

The number of online courses and enrolments in those courses will probably continue to increase for four interrelated reasons. First, most university students are comfortable with technology (Borreson & Salaway, 2008; Gustafson, 2004), with 80.5% of university students owning a laptop and 71.1 % of freshmen students owning a laptop (Borreson & Salaway, 2008). Second, many university students believe technology to be an important and necessary part of their course work and university experience (Concannon et al., 2005). Third, in order to attract students, colleges and universities are introducing more technology into their course selections (Ausburn, 2004). Finally, as the costs of establishing and maintaining classroom space increases for universities, and costs such as transportation, food, and parking increases for students (Welker and Berardino, 2006), the virtual classroom has the potential of becoming an increasingly cost effective alternative to both university administrators and students.

Comparisons of Web-based Courses with Standard Lecture Courses

Prior to 2003, several studies compared face-to-face learning with online learning and found no significant differences in learning-outcome measures (Morss, 1999; Quitadamo & Brown, 2001; Skinner, 1990). In 2003, however, a survey found that 57 % of academic leaders rated the learning outcomes in online education as the same or superior to those in face-to-face education, while in 2006 that number grew to 62 % (Allan & Seaman, 2006). In light of these findings, some academic leaders assert that computer-based instruction is at least equal to traditional instructional methods in terms of student achievement in university courses (Ausburn, 2004).

These outcomes, however, may underestimate the true potential of web-based courses (Muellar, Wood, Willoughby, Ross, & Specht, 2008). One reason for this could be due to hasty implementation of technology (Muellar, et. al, 2008; Theall, 1999). A second reason for this may be because many college and university instructors lack a clear understanding of the factors that influence learning in courses with online components (Ma & Runyon, 2004; Quitadamo & Brown, 2001; Salinas, 2008; Wingard, 2004). This is understandable given the number of opinions on how to effectively implement online education (Quitadamo & Brown, 2001), the lack of research in this area (Ausburn, 2004; Budd, 2002; Mehlenbacher , Miller, Covington, & Larson., 2000; Theall 1999), and that much of the research that is done, is based on questionnaires and surveys whose validity and reliability are uncertain (Bernard, Abrami, Lou, Borokhovski, Wade, & Wozney, 2004; Ma & Runyon, 2004; Mehlenbacher et al., 2000; Merisotis & Phipps, 1999; Wingard, 2004).

A third reason to underestimate the potential of web-based courses is the material used for these courses. Often these are prepared by individuals with a narrow conception of how to best use Internet technologies in their courses due to a lack of a theoretical foundation derived from research on courses designed for Internet delivery (Gresh & Mrozowski, 2000; Jerris & Poppie, 2002). In fact, many instructors are unfamiliar with any learning-management system (Salter, Richards, & Carey, 2004; Theall, 1999). Jeris and Poppie (2002) go on to state that faculty members who have developed knowledge of Internet tools have based their course design on what works within a certain electronic format rather than in a context of teaching and learning. All of which results in materials that may be confusing to students, especially when no manual is available to explain the learning strategies (Baker et al., 1997).

A fourth reason to underestimate the Web's potential is that that while the majority of teachers have access to and use computers; they may not be fully prepared to integrate computer technology in their classrooms (Lim & Chai, 2008; Mueller, et. al., 2008). This could be due to lack of experience and lack of time (Eynon, 2008; Power, 2008). Mueller et. al. (2008) note that it takes five or six years for a teacher to gain mastery in integrating technology into the classroom, while Adams (2004) stated that between "30 and 200 hours of development time are needed to produce 1 hour's worth of content" (p. 7).

However, with well-designed implementations of technology incorporating well-designed and sound pedagogy, Web-based instruction may actually be superior to conventional classroom instruction (Olson & Wisher, 2002). This would be consistent with earlier studies on computer-based instruction prior to the rise of the Internet. In their meta-analysis on computer-based instruction, Kulik and Kulik (1987) note that "most of the studies reported that computer-based instruction has positive effects on students" (p. 224). More recent reviews state that students in computer-based instructional classes generally achieve higher examination scores, learn lessons in less time, learn more, and enjoy the course more (Baker Hale, & Gifford, 1997; Grant & Spencer, 2003).

Blended Learning

The demographic on university campuses now includes "full-time part-time" students (Ausburn, 2004; Chickering & Kytle, 1999; Concannon et al., 2005) who are trying to maintain a full course load but are working part-time or even full-time throughout the school year, limiting the time available to attend lectures. The Internet, however, offers a promising alternative for these students by reducing the need to attend three hours of in-class lectures per class every week, while offering the opportunity of being part of a learning community in a face-to-face classroom (Garrison & Kanuka, 2004). This mixture of technology-based online learning activities with in-class instruction maintains flexibility and convenience in a manner that most effectively improves student learning (Akkoyunlu & Yılmaz-Soylu, 2008; Ben-Jacob, 1999; Dzubian, et al, 2004; Garrison & Kanuka, 2007; Lim, Morris, & Kupritz 2006; Osguthorpe & Graham, 2003; Welker & Berardino, 2006).This combination of classroom instruction with Web-based learning is known as blended learning (Carr-Chellman & Duchastel, 2001; Concannon et al., 2005; Merisotis, 2001). It is a method (or group of methods) that combines processes of facilitating student learning with opportunities for student self-instruction (Alonso et al., 2005). Blended learning, therefore, is learning that mixes face-to-face classroom behaviours with internet based activities and self paced learning (Akkoyunlu & Yılmaz-Soylu, 2008; Osguthorpe & Graham, 2003).

Blended learning has been looked at in two ways. On the one hand, it has been seen simply as a way to deliver instruction that offers increased flexibility to both students and instructions. On the other hand, it has been seen as a delivery method in which the online component is a natural extension of traditional classroom learning (Matheos, Daniel, & McCalla, 2006). In either case, Graham Spanier, President of Pennsylvania State University, stated that learning that supports the convergence of online and residential instruction is "the single greatest unrecognized trend in higher education today" (cited in Vaughn, 2007, p. 89). This trend will likely continue to grow due to the competition among universities and colleges for students, combined with the economic climate that forces many students to be in the workplace (Welker & Berardino, 2006). Kim and Bonk (2006) also note this trend in an analysis of survey results from 562 members of the Multimedia Educational Resource for Learning and Online Teaching.

Although it may be particularly advantageous to "full-time part-time" students, its applicability is not limited to them. It is also applicable to both part-time and full-time students who are not working. In fact, with blended learning, the distinction between full-time, part-time, and "full-time part-time" students may diminish in administrative importance. In surveys conducted by Borreson and Salaway (2008), they found 59.3 % of student respondents preferring a moderate amount of information technology in their courses, 25.0 % preferring extensive or exclusive information technology, and 15.8 % preferring no information technology

Effectiveness of Blended Learning

Blended learning allows for consistency in preparation, delivery, organization, and overall course management, leading to a high level of quality (Welker & Berardino, 2006; Wingard, 2004). More generally, it has been claimed that blended learning models are found to be the most effective learning strategy (Skill & Young, 2002). Osguthorpe and Graham (2003) stated that adoption of blended learning leads to improvement in: pedagogy, access to knowledge, social interaction, personal presence, cost effectiveness, and ease of revision, with the overall aim of finding an effective balance between online access and face-to-face interactions. Dzubian, Hartman, and Moskal (2004) found that on average, blended learning courses have a higher percentage of students achieving grades of "C" or better, and lower dropout rates. Their research compared face-to-face courses with fully online and blended courses over three years and seven semesters at the University of Central Florida. Over that time, students receiving an A, B, or C in fully online courses was 91.3%, in face-to-face classes this was 91.6%, while in blended design courses, this was 93.3%.

In addition, Garnham and Kaleta (2002) found that students in blended courses learned more than students in traditional lecture courses, and retained more information than students in comparable completely online courses. Their research involved 17 faculty members at six institutions; the University of Wisconsin, University of Wisconsin-Milwaukee and four University of Wisconsin-College campuses. The instructors represented a wide variety of disciplines, and the courses they converted to blended ranged in size from less than 15 students to over 200. Instructors reported that students in these blended courses wrote better papers, scored higher on exams, and provided more meaningful discussions on course material than students in traditional lectures. Aycock and Kaleta (2002) also reported that instructors found that students in blended courses performed better at learning course material, mastering concepts, and applying what they had learned, than students in traditional lecture courses.

Wingard (2004) surveyed and interviewed 46 experienced faculty members from seven institutions after introducing Web components into the classroom. Almost half reported increased discussion in the classroom, and more than a third reported increased student-student and student-instructor communication inside and outside the classroom. As a result of this increased communication, seven of the 46 faculty members reported feeling "more prepared and more familiar with their students' academic progress" (p. 7). Half reported a decrease in lecturing on the material, with the focus turning instead to providing instruction on how to critically evaluate Web resources. For many of those who reported no change in the amount of lecturing, their lectures became more challenging or involved more extensive coverage of the material, and contained more complex examples, diagrams, and models.

Faculty also noted that adopting to a blended design brought more attention to learning outcomes, allowed for more efficient use of classroom time, provided a greater ability to track assignments and identify students who were not posting work or participating in the course, facilitated quickly updating a course with new information, and provided more opportunities to give individualized feedback accompanied with a permanent record of instructor-student and student-student interchanges (Wingard, 2004).

Research Involving Blended learning

Recent areas of research using a blended design include assessing the effectiveness of blended design from the student's viewpoint, the instructor's and designer's viewpoint, and a general effectiveness viewpoint. Regarding the student's viewpoint, Akkoyunlu and Yılmaz-Soylu, (2008) developed an objective assessment instrument of blended learning based on learners' views on blended learning and its implementation process. So and Brush (2008) examined the relationships of the students' perceived levels of collaborative learning, social presence, and overall satisfaction in a blended learning environment. Regarding instructors and designers, Lim and Chai (2008) studied the impacts of instructional decisions when planning for instruction in a technology-enhanced learning environment. Regarding the general effectiveness of a blended design, Kaczynski, Wood, and Harding (2008) used interviews, observations, and open-ended questionnaires targeting all stakeholders using blended courses in three universities in South Africa, Australia, and the United

States. This was done to assess the extent of blended learning in the university environment in order to compare changes in the learning environment.

Student performance and behaviours in blended learning courses have also recently been studied. For example, Gerber, Grund, and Grote (2008) analyzed tutoring in a corporate finance course via the internet as an element of the learning context and student behaviour and learning performance in a blended learning course. They found that student performance was not related to the quantity of tutors' activities, but to the quantity of students' activities. However, they also stated that the interpersonal messages from the tutors' to the students served as a form of motivational support, which in turn promoted interactions between the students, enhanced their activity in the course, and, as a consequence, also increased their learning performance. Similarly, Hughes (2007) found that blended learning with increased support and visible tutor monitoring can improve module retention by motivating learners to complete coursework on time, without increasing tutor workload.

The use of blended designs has also been researched in a variety of educational environments and contexts. Examples of courses using blended learning include flight education (Robertson 2006), social work (Bellefeuille, 2006; Cooner & Hickman 2008), geography (Balram & Dragićević 2008), public health (Moore, Perlow, Judge, & Koh, 2006; So, 2009), photography (Abrahmov & Ronen, 2008), medical school clerkship curriculum (Szulewski & Davidson, 2008), human anatomy (Pereira, Pleguezuelos, Molina-Tomas, & Masdeu,, 2007), dentistry (Phainis, Stokes, Walsh, & Cannavina, 2007; Phainis, Stokes, Walsh, & Cannavina, 2008), and nursing (Bloomfield, While, & Roberts, 2008).

The research mentioned above is typical of research on instructional innovations, which seems to follow a series of evolutionary phases (Svinicki, 2007). The first phase takes the form of "it works," a phase in which a descriptive series of studies demonstrate proof of a concept. These studies are followed up by other researchers, whose work can be conceptualized as; "it works here too," thus demonstrating generalization and transferability of the relatively new instructional method to new situations (Svinicki, 2007). Eventually the new method, now perhaps no longer innovative, reaches the stage at which researchers try to understand why it works. The present study traverses phases one and three. It is in phase one of an instructional innovation by employing CAPSI as a component of blended learning. It also can be included in phase three as part of the research to more fully understand the reason why blended learning-as a larger concept- works.

A Behavioural Approach to Education

In general, there is a deficiency of research on blended learning designs (Vaughn & Garrison, 2005). However, a review of the literature related to how users learn from multimedia in blended designs identifies three primary learning theories; (a) behaviourist (Pear & Crone-Todd, 1999; Pear, Crone-Todd, Wirth, & Simister, 2001; Pear, Pear, & Novak, 1996; Price, 1999), (b) constructivist (Bellefeuille, 2006; Christensen, 2003; Cooner & Hickman 2008; Dalsgaard & Godsk 2007; Gerber & Grote, 2008), and (c) cognitive (Balram & Dragićević, 2008; Bodiea, Powers, & Fitch-Hauser, 2006; Dettori, Giannetti, & Persico, 2006; Giouvanakis, Thanasisn, Samaras, Tarabanis, & Konstantinos, 2001; Vaugn & Garrison, 2005). The literature however, either does not reveal a generally agreed upon theory that easily applies to all learning environments (Mackey & Ho, 2008), or incorporates more than one pedagogical approach in a learning environment (Adams, 2004). This is because the changes introduced by distance education and blended learning have concentrated on the development of computer tools as complements to traditional teaching methods rather than a focus on the establishment of a learning theory (Pereira, Pleguezuelos, Molina-Ros, Molina-Tomas, & Masdeu, 2007). Therefore, this study will focus on the effectiveness of a behavioural approach to multimedia usage in a blended design.

Behaviourism is a minority position in education and psychology (Cracolice & Roth, 1996); however, it arguably played a major role in the movement away from the traditional approach to higher education (see Pascarella & Terenzini, 1991, 2005). In the 1950s and 1960s, B. F. Skinner (e.g., 1958, 1968) began to advocate strongly for the application of behavioural principles to education through teaching machines and programmed instruction. Other behaviourally orientated psychologists and educators echoed his arguments for designing instructional methods that focuses on the behaviour of the learner.

A behavioural approach to education offers the possibility of combining a well-designed pedagogy for learning with technology in a blended design. The process of learning within this model begins with clearly defined objectives (Svinicki 1999). From this beginning, course material is divided into small steps that build toward the objectives with an increased frequency of correct responses and a decrease in errors (Moallem, 2001). Immediate reinforcement is provided when correct behaviours are demonstrated. This will result in learning, which, according to Alonso (2005), is the gaining of knowledge in order to solve problems more successfully, thereby focusing on "doing as a basis for achieving an effective understanding of the knowledge" (p. 217).

A Behavioural Approach to Education: PSI

One model that follows this process and has achieved clear positive results is the Personalized System of Instruction (PSI), developed by Keller (Brothen & Wambach, 1999; Keller, 1968). It is an application of reinforcement theory involving presentation of material, student performance related to that material, and consequences related to that performance (Sherman, 1992). It is a highly structured student-centred approach (Price, 1999) designed to improve student achievement while replacing the tradition of punishment (failure) with the use of positive consequences (Grant & Spencer, 2003). That is, grades were not given in the original design; instead, the focus was on the eventual attainment of mastering the course material (Cracolice & Roth, 1996).

There is also a stress on the written word in a PSI designed course. Students are provided with a text material and provide written answers to unit questions. The instructor presents material in written form rather than lectures. The instructor also prepares a written study guide, which includes study objectives and questions that focuses students' attention on to the important material to be learned and what students are expected to do (Grant & Spencer, 2003).

The major components of PSI are: (a) student self-pacing corresponding to his or her ability and other demands; (b) clear study objectives; (c) stress on the written word as opposed to lecturing; (d) textual material divided into small study units; (e) demonstration of mastery of one study unit of material before proceeding to the next study unit; (f) use of more advanced students (typically in another course) to administer unit tests and provide feedback; and (g) use of lectures for motivational purposes as opposed to providing new information (Crosbie & Kelly, 1993; Grant & Spencer, 2003; Keller, 1968; Pear & Novak, 1996). In a PSI approach, the instructor takes on the role of facilitator, allowing students to become more interactive with the material. However, many instructors still enjoy the more central role provided by the traditional lecture method and are hesitant to adopt new methods that downplay that role (Saville & Elliott, 2005).

In a PSI-taught course, the instructor and students have specific roles and responsibilities. The responsibilities of the instructor are to select and organize the course material, develop questions for tests and exams, provide the final evaluations, and present lectures. Assisting the instructor are *proctors,* who typically are students in a more advanced course. Proctors assess students' answers on tests, provide feedback on students' answers, and decide whether the answers demonstrate mastery – which will allow the student to proceed to the next unit – or whether the student should restudy the material and try again.

It is the responsibility of the student – and a key to the PSI design – to control the pace of his or her learning. After receiving initial instruction about the course and PSI, students proceed individually through the instructional units at their own pace (Cracolice & Roth, 1996). The student decides when to study the material for a particular unit and when to take a test on that unit. This allows the student the flexibility that is believed by some to be the optimal design for cognitive engagement and learning (Stoney & Oliver, 1999).

Effectiveness of PSI

PSI has been shown to be more effective than the traditional lecture method at developing student learning under a variety conditions and research designs (Cracolice & Roth, 1996; Kulik, Kulik, & Bangert-Downs, 1990; Kulik, Kulik, & Cohen, 1979; Pascarella & Terenzini, 1991, 2005). Students in PSI-taught courses have achieved higher exam scores with less instructional time, professed greater enjoyment toward learning the course material, and, in the case in which PSI was administered through computers, developed an increased positive attitude toward computers (Dewhurst et al., 2000; Kulik et al., 1979). Differences were more pronounced when performance was measured on essay exams as opposed to multiple-choice exams (Kulik et al. 1979).

One reason for this is that in traditional lecture courses, final grades are largely determined by non-repeatable assignments and exams, whereas, in PSI-taught courses, students may rewrite tests on which they did not initially demonstrate mastery (Kulik et al., 1979). The result is that PSI-taught courses generally have negatively skewed grade distributions – i.e., there are typically many more A's and B's than in traditional lecture courses. This is a situation administrators might view as a sign of grade inflation, which is considered undesirable because it appears to devalue grades and hence, in the eyes of some administrators, reflects poor course management (Binder & Watkins, 1990). What may not be given sufficient consideration is that students in PSI-taught courses typically have done a sufficient amount of work and learned a sufficient amount of material to have earned the high marks they receive. In addition, in a PSI-taught course, all students know exactly what is expected of them, which also tends to equalize their performance and thus reduce the variability in the grade distribution (Binder and Watkins, 1990).

Related to the above is a study by Skinner (1990), who compared three types of computer–based instructional conditions involving tutorials, unit tests, a mastery component, and proctors with a traditional teaching method. Overall, he found that low achieving students benefited appreciably more from computer-based instruction than high achievers (Skinner, 1990). Therefore, with PSI, high-achieving students remain high achieving students, whereas lower performing students now achieve higher marks, thus resulting in a skewed distribution of grades (Ironsmith & Eppler, 2007).

PSI is also well suited for research because it is highly systematic and provides a clear record of each student's progress. It allows a researcher to assess its various components, such as the adequacy of materials, optimal frequency for testing, the effectiveness of proctors, appropriate study unit sizes, constraints on self-pacing, as well as the effectiveness of the system as a whole (Brothen, 1996; Brothen & Wambach, 1998; Crosbie & Kelly, 1993; Hambleton, Foster, & Richardson, 1998; Price, 2000; Pear & Crone-Todd, 1999; Lloyd & Lloyd, 1986; Sherman, 1992).

Criticisms of PSI

A number of PSI users have found that many students either procrastinate or proceed so slowly through the material that they do not complete the course on time and drop out (Kulik, Kulik, & Bangert-Drowns, 1990). However, while procrastination and high dropout rates can be argued as a reason for not instituting PSI courses, research shows that there are ways to ameliorate these (e.g., Born et al., 1972; Brothen, 1996; Crosbie & Kelly, 1993; Hambleton, Foster, & Richardson,1998; Price, 1999).

Other criticisms against it are that: (a) PSI runs counter to traditional pedagogical systems in higher education; (b) PSI does not easily fit into the traditional academic calendar; (c) designing PSI courses requires extensive initial preparation; and (d) some instructors who attempted to use PSI misapplied or did not understand the behavioural principles on which it is based (Saville, Zinn, Neef, Norman, & Ferreri, 2006; Schmitt, 1998). However, Gallup and Allen (n.d.) note that "although there are valid complaints about PSI, it has been abundantly clear for many years that there is no pedagogical reason for deprecating PSI or for rejecting it outright" and that "PSI remains a superior method of instruction" (p. 3).

Computer-Aided Personalized System of Instruction

Alonso et al. (2005) pointed out that "computers are the potential saviours of the education system because they can be used to personalise learning" (p.218). Computers can be used to design learning experiences according to personal knowledge and needs, record the progress that is made toward course goals, and provide opportunities to gain quick feedback thereby letting students know which thought process are incorrect and which ones are correct (Alonso et al., 2005). This educational use of computers is instantiated in a learning system called Computer-Aided Personalized System of Instruction (CAPSI), which has evolved from a mainframe to a Web-based environment (Pear & Martin, 2004), combining the assets of computers in learning with the effective learning principles of PSI (Pear & Crone-Todd, 1999, 2002; Pear & Kinsner, 1988; Pear & Novak, 1996; Sherman, 1992).

The major components of CAPSI are: (a) instructor selected textbooks and other textual material; (b) study questions based on the assigned material, (c) short essay unit tests; (d) rapid feedback; (e) use of either teaching assistants, or more advanced students in the course to administer unit tests and provide feedback; and f) retesting when a student does not meet mastery on a unit test. CAPSI also allows students to appeal a restudy – i.e., lack of mastery on a unit test - to the instructor. Mehlenbacher et al. (2000) pointed out that computer-based instruction – such as in CAPSI - facilitates organizing content into explicit instructional goals that are appropriate to a student's level of understanding. CAPSI takes advantage of this by allowing students to select the times at which they are ready to write a unit test, based on their personal judgment that they have mastered the material in the unit. The result of a unit test indicates whether the student is correct in this judgment or whether restudying the unit is necessary. Since students must demonstrate mastery on each study unit before proceeding to the next unit, they do not progress in a course without understanding the previous required material. This helps ensure that the instructional goals are always appropriate to a student's level of understanding of the course material.

CAPSI has been studied to find relationships between various factors and student performance. For example, Springer and Pear (2008) used CAPSI to explore the relationship between objectively rescored final exam grades, peer reviewing, and the rate at which students completed unit tests. Martin, Pear, and Martin (2002) assessed proctor accuracy in marking answers as correct or incorrect. While Crone-Todd, Pear, and Read (2000) assessed study questions in two psychology courses in order to develop operational definitions of the thinking levels required to answer study questions.

Pear and Crone-Todd (1999) also assessed CAPSI and its overall effectiveness in four different psychology courses. On a departmental student evaluation questionnaire administered at the end of the course terms, they found that 37% of the surveyed students reported that the courses were average when compared with other courses, while 53.7% reported that the courses were good or very good compared to other courses. Pear and Novak (1996) conducted a study to evaluate CAPSI in two second-year undergraduate psychology courses. In this study, students in these courses were given a questionnaire at the end of the term about their experience with the courses and an analysis of determinants of performance in the course. They found that most students rated the courses as good as a typical lecture course, and more than one third rated it better. Two issues from this research were included in the current study. First, students in their study expressed dissatisfaction with the lack of lectures. While Pear and Novak provide reasons why lectures are not a vital component of university teaching, they do state that it "may be desirable to build some lectures into the structure of a CAPSI course (p. 122). The second issue is to support their finding that CAPSI enables instructors to develop students' ability to express themselves about a subject in large classes.

Computer-Aided Personalized System of Instruction in a Blended Learning Design

Because CAPSI is Web-based it can be used as a component in blended learning. Carmen (2005) suggested five key ingredients for optimal blended learning: live events, self-paced learning, collaboration, assessment, and performance-support materials. These can all be primary components in a blended learning course that also contains CAPSI as a component. The live events can be included in the form of classroom lectures or discussions, students can take the CAPSI unit tests on their own schedules, collaboration can occur between students and between students and the instructor both in-class and on the Internet, assessment can occur from both in-class quizzes and exams and the CAPSI unit tests, while performance-support materials can be provided within the CAPSI program or from other materials, such as text books.

Critical Thinking

It is important to note that it is more difficult to assess gains in critical thinking than it is to access gains in factual knowledge because there is no universally agreed-upon definition of critical thinking (Botsch & Botsch, 2001; Weis & Guyton-Simmons, 1998; Williams, 1999). However, a variety of definitions exist as to what critical thinking is, especially in its relation to the educational process (Astleitner, 2002; Collucciello, 1997; Ennis 1987; Garrison, 1991; Giancarlo & Facione, 2001; Williams, 1999). For example,

> critical thinking is a higher-order thinking skill, which mainly consists of evaluating arguments. It is a purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanations of the evidential,

> conceptual, methodological, or contextual considerations upon which judgment is based (Astleitner, 2002, p. 53)

Also,

> critical thinking is the intellectually disciplined process of actively and skilfully conceptualizing, applying, analyzing, synthesizing, and/or evaluating information gathered from, or generated by observation, experience, reflection, reasoning, or communication, as a guide to belief and action. In its exemplary form, it is based on universal intellectual values that transcend subject matter divisions: clarity, accuracy, precision, consistency, relevance, sound evidence, good reasons, depth, breadth, and fairness. (Scriven & Paul, 2004 p.1)

In addition, to be a critical thinker is to

> take control of one's thought processes and gain a metacognitive understanding of these processes. (Garrison and Kanuka, 2004 p. 98)

Yet another definition is

> critical thinking is reasonable reflective thinking that is focused on deciding what to believe or do. (Ennis 1987 p. 10)

One final definition, and one that is connected with the use of the assessments of critical thinking for this study, is that critical thinking is characterized as purposeful, self-regulatory judgment (Giancarlo & Facione 2001). This judgment is done in regard to what to believe or do in a given context. In doing this, a person engages in analysis, interpretation, evaluation, and inference, in order to explain the problem (Giancarlo & Facione, 2001).

Teaching Critical Thinking using Bloom's Taxonomy

The technical and pedagogical changes that are occurring in higher-education instruction need to incorporate the development critical thinking skills in students (Aretz, Bolen, & Devereux, 1997; Giancarlo & Facione, 2001; Halpern, 1998; Williams, 1999). However, teaching critical thinking skills is difficult for two reasons: (1) it is an abstract conceptual skill; and (2), as noted above, there is no generally accepted standard for critical thinking. Despite these challenges, there have been a number of systems that have attempted to formalize various aspects of identifying and teaching critical thinking processes. One of the most recognized of these is Bloom's Taxonomy in the Cognitive Domain, which can be used with any learners and in any content area (Chyung & Stepich, 2003) and, as its developers state, was "designed to be a classification of student behaviours which represent the intended outcomes of the educational process" (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956, p. 12). The development of the taxonomy was started by amassing a list of educational objectives and then grouping these objectives into categories organized from the simplest to the most complex (Bloom, et. al., 1956: Manton, Turner, & English). From this, Bloom and his colleagues identified six major thinking levels: (1) knowledge; (2) comprehension; (3) application; (4) analysis; (5) synthesis; and (6) evaluation.

At the knowledge level, students provide answers that may me memorized or closely paraphrased from the assigned material (Crone-Todd, Pear, & Read, 2000). Some key words usually associated with questions at this level are: name, list, and identify (Manton, Turner, & English, 1999). Comprehension requires that students use their own words in answering a question, while still using terminology that is consistent to the course material. Some key words usually associated with questions at this level are: describe, rephrase, and explain (Manton, Turner, & English, 1999). Application requires that students recognize, identify, or apply a concept or principle introduced in the course, or provide an example not found in the course materials in order to answer a question. Some key words usually associated with questions at this level are: relate, employ, and demonstrate (Manton, Turner, & English, 1999). Analysis requires students to break down concepts into their parts, or identify or explain essential components of concepts, principles, or processes. It might also require students to compare and contrast, or explain how an example illustrates a certain concept or principle. Some key words usually associated with questions at this level are: analyze why, categorize, and classify (Manton, Turner, & English, 1999). Synthesis requires students to put together parts to form a whole. It might also require students to generate definitions not identified in the course material, or to explain how to combine principles or concepts to produce something new. Some key words usually associated with questions at this level are: design, create, and construct (Manton, Turner, & English, 1999). Finally, evaluation requires students to present and evaluate reasons for and against a particular position, and to come to a conclusion regarding the validity of that position (Crone-Todd, Pear, & Read, 2000). Some key words usually associated with questions at this level are: evaluate, predict, and argue (Manton, Turner, & English, 1999). These categories are roughly hierarchical. For example, to be able to creatively put together several basic concepts to create a new idea (level 5), one must have a good understanding or comprehension (level 2) of those basic concepts.

Although Bloom's taxonomy is not the only possible way to classify thinking levels, it is widely known and used in education (Chyung & Stepich, 2003), and therefore provides a good starting point for teaching higher-order thinking and studying its development. Bloom's Taxonomy also allows a dynamic instructional process in which the low-level cognitive skills can be learned in combination with higher-level cognitive skills (Chyung & Stepich, 2003). As evidence of this, Pear and his colleagues have demonstrated in a research program that CAPSI is highly adaptable to developing critical or higher-level thinking as based on Bloom's taxonomy (e.g., Crone-Todd & Pear, 2001; Crone-Todd et al., 2000; Pear, 2002; Pear, Crone-Todd, Wirth, & Simister, 2001; Pear & Martin, 2004).

Teaching Critical Thinking with CAPSI Using Bloom's Taxonomy

It has been argued that although PSI; and in turn CAPSI, may be effective for basic knowledge learning, it cannot develop higher-order critical thinking skills (Adams, 2004; Budd, 2002). Hobbs (1987) argued out that students in a PSI-taught course can become book bound "through the use of study questions, objectives, and quiz items that emphasize memorization over critical thinking" (p. 106). Caldwell (1985) raised the concern that study questions may become the focus, and that these study questions could concentrate on low-level content, which in turn could lead to simplistic, low-level learning. Similarly, it has been asserted that learning and being tested on material in small units does not provide for synthesis of information, which is a higher-order or critical thinking skill (Theall, 1999).

However, Grant and Spencer (2003) pointed out that any criticism of PSI in regard to teaching critical thinking is equally applicable to university courses in general. More specifically, Reboy and Semb (1991) make the following four points in answering the criticism that PSI is not suited for teaching critical thinking skills: First, PSI is a delivery system; therefore, content within the PSI delivery system can be focused on developing higher-order cognitive skills by using a systematic questioning technique (Quitadamo & Brown, 2001). Second, PSI has been successfully used in courses requiring critical thinking skills. Third, research has shown that "students enrolled in PSI courses have improved their higher-order cognitive skills" (p. 212). Fourth, PSI allows for "early diagnosis of inadequate reasoning because of frequent student contact and assessment" (p. 213).

Teaching by computer-mediated communications encourages interacting through written forms of communication, which may result in increased reflection, better writing skills, and greater problem-solving performance (Garrison & Kanuka, 2004; King, 1995). The range and quality of interactive dialogue that can be facilitated through CAPSI fits in with this widely accepted means of facilitating critical thinking because CAPSI questions require written answers. The feedback to students' on their answers, both in pointing out correct answers and providing the reasons for assigning a restudy, has features of an intellectual dialogue on specific course information (Carr-Chellman & Duchastel, 2001; King, 1995). The emphasis on written answers in CAPSI, providing feedback while assessing answers, and writing appeals, allows for a process of formulating and challenging ideas, which is rarely accomplished by students on their own.

Another aspect of CAPSI that fits with Bloom is his four elements that determine the quality of instruction; cues, participation, reinforcement, and feedback/correctives (Guskey, 2001). To optimize the use of these four elements, Bloom divided the learning material into smaller units corresponding to chapters in a textbook, then to administer tests at the completion of each of these units to determine how well students learned the material (Guskey, 2001). He also believed that if these tests were accompanied by feedback and corrections, they could serve as valuable learning tools (Guskey, 2001).

All of these, it has been suggested, promote the development of critical thinking skills (Bernard et al., 2004). Moreover, it has been reported that with the increased use of online activities, students spend more time developing and using higher-order thinking skills (Waschull, 2001). Finally, the requirement in CAPSI of students to assess their peer's answers should lead to higher order thinking because it forces students to exercise judgment in complex and uncertain situations (Alonso et al., 2005).

The Present Study

Mortera-Guitierrez (2006) pointed out that blended learning "is so new that there are few academic sources and attempts in the sparse literature to reflect theoretically on its pedagogical and epistemological foundations" (p. 316). He also pointed out that "because of a lack of understanding of the well established theoretical principles, blended learning as an educational and instructional approach makes this urgent need a priority" (p. 316). Others have made similar points (e.g., Ausburn, 2004; Matheos, et.al, 2006; Vaughn and Garrison 2004). Ausburn (2004) stated that there is little in the literature regarding important design features in blended learning classrooms. Matheos et al (2006) stated that: "Despite the promise of blended learning there is limited empirical evidence to support many of the claims" (p. 65), while Vaughn and Garrison (2004) point to a lack of research on the regarding the effectiveness of blended learning.

Still it must be reiterated that research on blended learning is in its infancy, and while the preliminary results are encouraging, they must be interpreted with caution. Skill and Young (2002) stated that understanding how best to integrate online and in-class components will continue to be a considerable challenge for educators. While Ausburn (2004) stated that with the increase in the popularity of blended learning, this "lack of knowledge is problematic" (p. 329). These comments reflect the observation of Svinicki (2007); that blended learning, like other instructional innovations, is in either phases one or two of a developing instructional innovation, highlighting the need for further research.

Even though there has been a growing investment in technology in higher education, there is a lack of evidence so far that it has led to significant educational gains for students (Hosie et al., 2005), or that this inclusion of technology has perhaps been predicated on little more than assumptions about the likely educational effectiveness' of the internet (Selwyn, 2008). From this, it could be argued that there is a need to gain more understanding regarding the realities of internet use in university teaching and learning in order to more successfully integrate the internet into university education (Selwyn, 2008). This may be in part, because many applications of technology, as mentioned earlier, have not yet incorporated the most effective pedagogical principles (Alonso et al, 2005). The present study sought to determine the pedagogical effectiveness of a technological learning-management system – CAPSI – with respect to students learning course material and developing critical thinking skills. This study was designed to help provide an understanding of how to effectively incorporate and develop online technologies to enhance teaching effectiveness in a blended design. To do this, three separate experiments were conducted to assess the use of CAPSI in various course designs, with different students, and in different situations in order to gain an understanding of its possible effectiveness for student learning and critical thinking in university.

In Experiments 1 and 2, the students were in a first year University of Manitoba course called *Introduction to University*. These two experiments incorporated different research designs, were implemented over two different time spans (13 weeks and 5 weeks) using two different sets of materials, in two different years. An additional variable in these experiments was the possible influence of instructors to the learning of course material and critical thinking. In Experiment 1, this possible influence of instructors was assessed by measuring outcomes across four different instructors and 30 teaching assistants. In Experiment 2, this possible influence of instructors was looked at by assessing only one instructor, but in two different course sections, and 16 teaching assistants.

In Experiment 3, a different design was implemented. This experiment again only had one instructor over two sections of the same course. The differences between this experiment and the previous two were that this time the experiment did not include teaching assistants, was conducted over a longer period of time (26 weeks), involved different content, and was conducted at a different university. The course was a first year *Introduction to Psychology* course at the University of Winnipeg.

An additional difference in the three experiments was the focus of the assessments. Developments in both course knowledge and critical thinking were assessed in Experiment 1. Experiments 2 and 3 assessed these separately. Experiment 2 only assessed development of critical thinking skills, while Experiment 3 only assessed development of course knowledge.

EXPERIMENT 1

Introduction to University at the University of Manitoba is a first year university course that introduces students to the skills they will need for other courses in university, thus helping them make the transition from high school to university. A major component of the course consists of instruction, practice, and feedback on writing a proper academic essay. Other aspects of the course were designed for students to develop critical thinking and problem solving skills, to gain knowledge regarding the functions of the university, to prepare for and deliver a class presentation, and to acquire study and learning skills needed for success in university. The course normally requires students to write three essay assignments; however, because of the emphasis on writing answers in CAPSI to indicate course knowledge, the third essay in the CAPSI sections was eliminated as a requirement and the CAPSI assignment was included in its place. Therefore, the hypothesis in Experiment 1 was that completing a CAPSI assignment develops course knowledge and critical thinking skills above that of writing a third essay assignment.

This hypothesis was tested using the final examination (Appendix A) and a critical thinking assessment tool called the Applied Critical Thinking Measurement (ACTM; Appendix C). The final exam consisted of three types of questions exemplifying three levels of Bloom's taxonomy: knowledge questions, an applied question, and a critical thinking (evaluation) question. The final exam was assessed in three ways: the final exam score, the marks associated with the critical thinking question, and the combined marks of the knowledge and applied questions.

The ACTM and the associated scoring sheet were designed and researched by Dr. Robert Renaud of the Education Faculty at the University of Manitoba. This assessment tool had been used in two other courses at the University of Manitoba to assess critical thinking development (Regehr, 2003; Renaud & Mandzuk, 2006). There were five reasons for choosing the ACTM as the critical thinking measure. First, it is a working model that fits with the definition provided earlier by Giancoarlo and Facione (2001); namely, that critical thinking is characterized as purposeful, self-regulatory judgment in regard to what to believe or do in a given context. In doing this, a person engages in analysis, interpretation, evaluation, and inference, in order to explain the problem (Giancarlo & Facione, 2001).

The second reason for choosing this assessment is that it requires written responses. The course places a fundamental focus on to the development of writing skills; therefore, an assessment requiring written responses as opposed to multiple-choice responses correlates with the objectives of the course. Third, the written responses are to various ill-defined or underdetermined problems, which are problems that are never solved once and for all but can be tentatively solved repeatedly as more evidence becomes available (Aretz et al., 1997). Therefore, this measure fits well with this research because another purpose of the course is to help students use critical thinking skills in dealing with the kinds of complex issues and vague problems that they will face later in both their academic and professional (Halpern, 1998). The fourth reason for using this type of assessment is that it requires responses that fit Facione's (1986) suggestion that critical thinking should be thought of as "the ability to properly construct and evaluate arguments" (p. 222). In the ACTM, a student is assessed as having critical thinking skills to the extent that he or she can provide solid support for the arguments presented in a scenario as well as judge whether another argument is well supported (Williams, 1999). The fifth reason for choosing the ACTM is that questions in the CAPSI study units are compatible with the kinds of problems in the ACTM.

Method

Participants

The participants were 364 University of Manitoba students enrolled in *Introduction to University* during the Fall 2006 academic term consisting of 13 weeks. At the University of Manitoba, an academic term is approximately equal to an American semester; therefore "academic term" and "semester" will be used interchangeably in this thesis.

There were no academic requirements for admission to the course other than having been admitted to the University of Manitoba. However, the course was only open to students who had completed less than twelve credit hours.

Procedure

Required Information about the research protocol, a summery of the project, a statement regarding how the project would be introduced to the students, a copy of the proposed student consent form, and a letter of support from the head of was submitted for approval by the ethics prior to the start of the experiment. See Appendix G for a copy of the final consent as presented to the students in the first week of classes

There were four lecture sections in the course: two lecture sections were designated CAPSI sections and two sections were designated non-CAPSI sections. The numbers of students originally registered in the CAPSI sections were 111 and 139, respectively, totalling 250 students. The numbers of students originally registered in the non-CAPSI sections were 137 and 130, respectively, totalling 267 students. The numbers of students in the CAPSI sections who completed the course and wrote the final exam were 108 in one section and 85 in the other, respectively, totalling 193. The numbers of students in the non-CAPSI sections were 70 and 101, respectively, totalling 171.

Consideration was given to Bernard, Abrami, Lou, Borokhovski, Wade, and Wozney's, (2004) concern about lack of random assignment in educational research. While this experiment did not incorporate random assignment, the populations were comparable because students signed up for particular lecture sections and corresponding lab sections without prior knowledge of the different methods to be used in the different lecture sections.

Students in both lecture sections were required to complete an online quiz regarding library usage, a mid-term exam, an oral presentation, and a final exam. Students in two lecture sections of the experimental group (the CAPSI sections) were required to complete two APA-formatted essays (roughly 1000 words each) and a CAPSI assignment, while students in two lecture sections making up the control group (the non-CAPSI sections) were asked to write three APA-formatted essays (roughly 1000 words each) as a part of course requirements. The topics of the essays had to do be chosen from a list of topics, or on approval by the instructor. The third essay of the non-CAPSI group and the CAPSI assignment was worth 15% of the final grade.

Different instructors were responsible for each of the four lecture sections of the course; therefore, the assessment of CAPSI effectiveness included any possible effects that instructors might have on knowledge and critical thinking development. Each instructor delivered the in-class lecture component and supervised eight teaching assistants (TAs) for the eight labs associated with each lecture section. For the two CAPSI sections, one instructor was the author; the other instructor had previous knowledge of the CAPSI program. The instructors for the non-CAPSI sections were experienced university instructors who had been involved with the *Introduction to University* course for many years. All instructors had a minimum of three years experience teaching the course, with the instructors of the non-CAPSI sections having greater experience teaching the course.

The TAs, who were selected by the department that administered the course, were matched to a lab section based only on their availability for a certain section.

They did not have previous knowledge of which sections were associated with the CAPSI assignment and which were the Non-CAPSI sections associated with the extra essay assignment. Their responsibility was to facilitate the lab sections associated with each lecture section. The purpose of the labs was to put theory into practice by providing a vehicle for group exercises, individual presentations, and the development of writing skills. The TAs also marked essay assignments, the final exam, and, in the CAPSI sections, unit tests that were not marked by peer reviewers.

Lectures were 75 minutes long and occurred once a week on either Tuesdays or Thursdays, depending on the lecture section. Labs were 75 minutes long and occurred once a week on the Tuesday or Thursday that was not being used for its associated lecture. For example, if a lecture was scheduled for Tuesday, the lab associated with that lecture was scheduled for the Thursday of that week. Tuesdays and Thursdays were the lecture and lab days, respectively, for one CAPSI and non-CAPSI sections; and Thursdays and Tuesdays were the lab and lecture days, respectively, for the other CAPSI and non-CAPSI sections.

All lecture sections used the same textbook: *Study and Critical Thinking Skills in College* Sixth Edition, by Kathleen McWhorter (2006). They were also given a course outline, which detailed the course requirements. There was, however, one difference in the outlines. Students in the CAPSI sections received information about the CAPSI assignment, while students in the non-CAPSI sections received information about writing a third essay.

All students wrote the same final exam (Appendix A) at the same time in the exam period after lectures had ended. The final exam assessed students on: (a) their knowledge of the course textual material; (b) their ability to apply either problem solving skills or to illustrate their knowledge of Bloom's taxonomy; and (c) their use of Hegel's Dialectic as evidence of critical thinking (see Appendix B for the marking rubric).

The CAPSI Program

The CAPSI program was introduced to the CAPSI sections in a lecture session during the second week of classes. The students in the CAPSI sections were shown and directed to the CAPSI website at http://home.cc.umanitoba.ca/~capsi/ .At the same time, a demonstration of CAPSI was conducted, using a dummy student account to show students how to move through the system in order to take unit tests and to peer review other students unit tests – a major aspect of the CAPSI program. The demonstrations allowed time for questions; however, students could also ask questions in class or through the messaging system within CAPSI at any time during the academic term. Essentially, the CAPSI program provided a Web-based method for students to write short-essay unit tests and peer review unit tests of fellow students.

The CAPSI program required that unit tests be marked by either a TA or by two peer reviewers. When assigning peer reviewers, the CAPSI program selected two peers with the lowest number of peer-reviewing points who passed the unit being marked and who had agreed to be available to mark within 24 hours after the unit test to be marked was submitted. When two peer reviewers assessed the answers, both had to independently agree that the unit test was a pass in order for the program to record it as a pass. There was also a built-in appeal process for arguing the validity of a given answer. The appeal was submitted to the instructor, who then judged the strength of the argument and decided whether or not to accept the appeal.

The CAPSI assignment was designed for students to pass 20 unit tests and to peer review 10 unit tests. Each CAPSI unit corresponded to one chapter from the course textbook. Each unit test was worth .75 % of the final grade, while each peer review was worth .5%. For each unit, students were required to answer 3 randomly generated questions from the CAPSI unit corresponding to the appropriate chapter from the textbook.The questions for each unit test were composed according to the requirements for the six levels of Bloom's taxonomy. The questions were designed in order for students to test their knowledge of the textbook material. Further details on the CAPSI program are provided in Appendix E.

The Applied Critical Thinking Measure (ACTM)

The ACTM was administered to students attending the last lab at the end of the term. There was no specific course requirement that students had to attend the lab or participate in writing the ACTM. Therefore, the number of students writing the ACTM was lower than the number of students writing the exam. As in the lecture sections, time of day, and day of the week was accounted for by taking advantage of the initial lab assignments associated with the lecture sections. CAPSI-assigned students wrote the assessments on either the Tuesday or Thursday of the final week of labs, in either the 8:30 or 10:00 time slots. The Non-CAPSI-assigned students also wrote on either the Tuesday or Thursday of the final week of labs, in either the 8:30 or 10:00 time slots

In this measure, students were asked to respond to three vague situations they were likely to be familiar with (see Appendix C). They were then asked to make a "Yes," "No," or "Not sure" selection based on what they had read and the question that was asked at the end of the scenario. Students were then directed to either write three statements that justified their "Yes" or "No" decision or, if "Not Sure" was selected, to specify three questions regarding what information they would need in order to make a justified decision.

A marking rubric was supplied to assessors in order to rate the responses to the ACTM scenarios (Appendix D). Each written response was rated on a scale from 0-1 for a "Yes" choice, 0-2 for a "No" choice, and 0-3 for a "Not Sure" choice. The "Not Sure" responses were given the higher rating because the vague nature of the scenario opened it to greater critical analysis and questions. Therefore, a student could obtain a critical thinking score from 0 to 9 points on each scenario, achieving a total score from 0-27 for the complete assessment.

At the beginning of the course, students were administered a "Pre-ACTM", which was an assessment that contained problems similar to those they would get on the ACTM at the end of the term. Lab attendance and ACTM participation was similar to the situation regarding the end of term ACTM.

The intention was to compare the differences between the Pre-ACTM and ACTM scores obtained at the end of the term in the CAPSI and Non-CAPSI groups, thereby controlling for the possibility that the groups might have started out at different levels of critical thinking by chance. However, after data collection, and before retrieval of ACTM scores for analysis, the set of ACTM scores for Instructor 1 were misplaced. Therefore, the ACTM scores collected at the end of the term, rather than these differences, were compared.

Data Collection and Reliability Measurement

The final exam consisted of two questions on course content and a critical thinking question. The final exam scores on the two content questions and on the critical thinking question of the two groups were compared, as well as comparisons between instructors, and between instructors nested within groups. The initial scoring was done by the TAs for the students in their respective lab sections using a marking rubric provided by the instructor (Appendix B).

A random sample of 20 final exams was re-scored by an independent grader to assess the reliability of the exam grader. The independent grader used the same marking rubric as the initial scorer. The independent grader was blind with regard to the lecture sections the papers were from. The formula used to assess reliability was the lower score divided by the higher score. If the scores were the same, then the IOR was 1.00 for that assessment. Using this measure a mean reliability score of .85 was obtained.

The scoring of the ACTMs was conducted after the grades for the course were turned in. The ACTMs were scored by the author and two independent markers in order to secure an estimate of the reliability of the ACTM scores. Two reliability scorers were blind as to the lecture sections that the ACTMs were associated with. The total number of ACTMs used for the post ACTM statistical analysis was 132; of these, 45 were rescored, resulting in 34.09% of the assessments being rescored. The formula used for the reliability measure was the lower score divided score by the higher score. If the scores were the same, the reliability measure was 1.00 for that assessment. Using this measure a mean reliability score of .82 was obtained.

Data Analysis

The data was analyzed using an ANOVA Teaching-Method-Nested-Within-Instructors design. This was done to assess the relationship between the CAPSI and non-CAPSI sections as well as assess any instructor effects. In this design, students were nested within instructors, and the instructors were nested within the teaching methods. The instructors were not randomly selected; therefore, this is a fixed design. Table 1 illustrates the layout of the design.

Table 1: Nested design model used in the present study

CAPSI		Non-CAPSI	
Instructor 1	Instructor 2	Instructor 3	Instructor 4
Participant 1,1	Participant 1, 2	Participant 1,3	Participant 1,4
Participant 2,1	Participant 2, 2	Participant 2,3	Participant 2,4
...	...	...	...
Participant n,1	Participant n, 2	Participant n,3	Participant n,4

Four comparisons between the CAPSI and non-CAPSI groups, instructors, and instructors nested-within-designs were made in this experiment and will be discussed in the results sections in this order: (a) scores on the final exam; (b) scores on the critical thinking question of the final exam; (c) scores on the two content questions of the final exam; and (d) scores on the ACTM.

Results

Final Exam

Table 2 shows mean scores and standard deviations related to the overall final exam scores for each instructor and the mean scores for the CAPSI group and the non-CAPSI group. It should be noted that when a lecture section is referred to by instructor, this includes the TAs and labs associated with that lecture section as well as any other variable that differentiated the lecture sections.

As seen in Table 2, the CAPSI group scored higher than the non-CAPSI group on the final exam. The overall average of the CAPSI group was 69.18 and that of the non-CAPSI group was 69.95, resulting in a difference of 3.23 between the groups.

Table 2: Mean scores, standard deviations, and group sizes for final exam scores

	CAPSI			Non-CAPSI		
	Instructor 1	Instructor 2	**Total CAPSI**	Instructor 3	Instructor 4	**Total Non-CAPSI**
N	108	85	**193**	70	101	**171**
Mean	73.16	65.21	**69.18**	62.4821	69.4203	**65.95**
S.D	6.52	6.40	**6.46**	7.5699	5.1279	**6.35**

As can be seen in Table 3, the difference between the two groups on the final exam just reached the .05 level of statistical significance by rounding to two decimal places (p = .053). The effect size as indexed by eta squared would be classified as minimal (Tabachnick & Fidell, 2007). Table 3 also shows that the final exam grade related to at least one instructor is significantly different from the scores of the other instructors F (3, 360) = 7.971, p = .000. Also, the mark associated with at least one instructor is significantly different when instructors nested within classes are considered F (2, 360) = 10.205, p = .000. These effect sizes were moderate (Tabachnick & Fidell, 2007).

Table 3: Between-subject effects on the final exam

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Instructors	5854.952	3	1951.651	7.971	.000	.062
Groups	923.317	1	923.317	3.771	.053	.010
Instructors nested within Groups	4997.471	2	2498.736	10.205	.000	.054
Error	88147.585	360	244.854			

To look more closely at the differences between instructors, a pairwise comparison was done. Table 4 shows the results. Note that Instructor 1 was significantly different from Instructors 2 and 3, and that Instructors 3 and 4 were significantly different from each other.

Table 4: Pairwise comparisons between instructors on final exam scores

Instructor	Instructor	Mean Difference	Std. Error	Sig.
1	2	7.951	2.269	.001
	3	10.675	2.401	.000
	4	3.737	2.166	.512
2	3	2.724	2.526	1.000
	4	-4.214	2.303	.409
3	4	-6.939	2.434	.005

Critical Thinking Exam Question

Because the statistical analysis of Questions 1 and 2 was non-significant, the statistical analysis of Question 3 will be presented first. As seen in Table 5, the overall average of the CAPSI group was higher (M = 28.25) than the non-CAPSI group

(M = 26.74) resulting in a difference of 1.51 between the groups.

Table 5: Participants, means, and standard deviations on the critical thinking question

	CAPSI			Essay		
	Instructor 1	Instructor 2	**Total CAPSI**	Instructor 3	Instructor 4	**Total Non-CAPSI**
N	108	85	**193**	70	101	**171**
Mean	29.12	27.38	**28.25**	23.67	29.80	**26.74**
S.D	6.516	6.399	**6.457**	7.5699	5.1279	**6.3489**

As seen in Table 6, the mean difference between the two groups for the critical thinking question was significant at the .05 level (p = .026). Although this difference was significant, the effect size would be classified as small (Tabachnick & Fidell, 2007). Table 6 also shows that the grade on the critical thinking question associated with at least one instructor is significantly different from the scores of the other instructors

F (3, 360) = 14.843, p = .000. Also, the scores related to at least one instructor is significantly different when instructors are nested within classes F (2, 360) = 20.999, p = .000. These were medium differences (Tabachnick & Fidell, 2007).

Table 6: Results of nested design analysis on the critical thinking exam question

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Instructors	1801.680	3	600.560	14.843	.000	.110
Groups	202.810	1	202.810	5.012	.026	.014
Instructors nested within groups	1699.320	2	849.660	20.999	.000	.104
Error	14565.998	360	40.461			

Table 7 shows that significant differences occurred between Instructors 1 and 3, Instructors 2 and 3, and Instructors 3 and 4. Therefore, Instructor 3 is the one instructor who had a significant difference between all instructors on this question.

Table 7: Pair-wise instructor comparisons on the critical thinking exam question

Instructor	Instructor	Mean Difference	Sig.
1	2	1.749	.353
	3	5.454	.000
	4	-.677	1.000
2	3	3.705	.002
	4	-2.426	.060
3	4	-6.131	.000

Questions on Course Content

As seen in Table 8, the CAPSI group scored higher than the non-CAPSI group on the course content questions. As can be seen in Table 9, however, this difference was not statistically significant at the .05 level F (1, 364) = 1.415. p = .235. Table 9 also indicates that there were no significant differences between instructors in the grades on Questions 1 and 2, F (3, 364) = .903. p = .440.

Table 8: Participants, means, and standard deviations on the content questions

	CAPSI			Content Questions		
	Instructor 1	Instructor 2	**Total CAPSI**	Instructor 3	Instructor 4	**Total Non-CAPSI**
N	108	85	**193**	70	101	**171**
Mean	43.95	41.39	**42.82**	38.76	41.31	**40.26**
S.D	9.1	33.08	**22.94**	13.07	20.76	**17.99**

Table 9: Results of nested design analysis on the content questions

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Instructors	1170.460	3	390.153	.903	.440	.007
Groups	611.489	1	611.489	1.415	.235	.004
Instructors within Groups	578.197	2	289.099	.669	.513	.004
Error	155179.184	364	432.254			

ACTM

As seen in Table 10, the CAPSI group scored higher than the non-CAPSI group on the ACTM. However, as seen in Table 11, the difference between the means of the two groups is not significant at the .05 level. Table 11 also indicates that the ACTM scores from at least one instructor were significantly different at the .05 level from the ACTM scores from the other instructors. Also, the scores related to at least one instructor is significantly different when instructors are nested within classes. In both cases the effect sizes were moderate (Tabachnick & Fidell, 2007). Table 12 shows that the significant difference between instructors on the ACTM was between Instructor 3 and Instructor 4.

Table 10: Participants, means, and standard deviations for the end of term ACTMs

	CAPSI			Essay		
	Instructor 1	Instructor 2	**Total CAPSI**	Instructor 3	Instructor 4	**Total Non-CAPSI**
N	20	53	**73**	31	28	**59**
Mean	9.65	9.02	**9.33**	6.42	10.82	**8.6204**
S.D	5.05	5.212	**5.13**	4.849	7.552	**6.18**

Table 11: Between-subjects effects for the end of term ACTMs

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Instructor	306.109	3	102.036	3.152	.027	.069
Group	14.904	1	14.904	.460	.499	.004
Instructor nested within groups	290.874	2	145.437	4.493	.013	.066
Error	4143.187	128	32.369			

Table 12: Pairwise comparisons scores on the end of term ACTM

Instructor	Instructor	Mean Difference	Sig.
1	2	.631	1.000
	3	3.231	.299
	4	-1.171	1.000
2	3	2.600	.272
	4	-1.803	1.000
3	4	-4.402	.021

Discussion

The major finding of Experiment 1 was that students in the CAPSI group scored significantly higher on the final exam than students in the non-CAPSI group did. This appears to be due primarily to the performance on the critical thinking question, because there was a significant difference between the groups on that question but not on the content questions combined. Although there were non-significant differences between the CAPSI and the non-CAPSI groups on the combined content questions and the ACTM, these differences did favour the CAPSI group.

The outcomes in regard to the critical thinking test question and the exam scores supports the earlier findings of greater exam scores for PSI type learning (Kulik et al, 1990; M. Skinner, 1990). The results also support the findings of Dzubian, Hartman, and Moskal (2004) and Pereira et al. (2007) that students in blended learning courses perform as well or better than students in strictly lecture-based courses.

Although the main differences were in the direction favouring the CAPSI group, the effect sizes were minimal or small, even when the differences were statistically significant. This contrasts with the large effect sizes that have been obtained in the comparisons of PSI with the lecture method (e.g., Kulik et al., 1979, 1990). However, it should be noted that CAPSI was only one component of the instructional method used in the CAPSI group, that lectures were used in both groups, and that the non-CAPSI group had the benefit of instructional methods in addition to straight lectures. In addition, as Terenzini et al. (1995) pointed out, "one possible explanation for the absence of significant effects [in comparing differences between instructional methods in university courses] may be that a semester course experience may be too brief to produce any measurable impact" (p. 24). In other words, the differences observed in this experiment might have been more pronounced in a longer course in which CAPSI could have been used for a longer period of time. Finally, in regard to the ACTM result, there were fewer participants who wrote the ACTM than who wrote the final exam; this could explain the lack of statistical significance for the ACTM, whereas it was obtained for the final exam. Also, conducting the ACTM in the labs could have affected the results. The ACTM's were conducted by the TAs, who, despite instruction as to how to carry out the assessment, could have provided different instructions or responses to possible questions from the students during the time of the in-lab assessments. These possible differences, then, could have changed how a student would have responded to a scenario.

EXPERIMENT 2

The high variability between instructors in Experiment 1 may have tended to mask any differences between the CAPSI and non-CAPSI groups. Therefore, to eliminate variability between instructors, Experiment 2 compared the CAPSI and non-CAPSI treatments in two lecture sections that were both taught by the same instructor. In addition, whereas Experiment 1 examined both the development of knowledge and critical thinking, Experiment 2 focused exclusively on the latter in order to more closely examine the effects of CAPSI on critical thinking development. To this end, the CAPSI questions were only of the higher level (Bloom's level 6); i.e., the CAPSI component used only questions that required students to evaluate a position supported by information from the course textbook. Other differences with Experiment 1 are detailed below. The hypothesis to be tested was: student performance in a CAPSI assignment, in which students are asked to supply higher level thinking responses to higher-level questions as a component of a blended designed course, develops student critical thinking skills in lieu of writing a third essay.

Method

Participants

The participants were 276 University of Manitoba students enrolled in two lecture sections of *Introduction to University* for the Fall 2007 academic term consisting of 13 weeks. As in Experiment 1, there were no academic requirements for admission to the course other than having been admitted to the University of

Manitoba; however, the course is only open only to students who have completed less than twelve credit hours.

Procedure

The procedure, materials, course, and type of students used in Experiment 2 were similar to those used in Experiment 1, therefore a separate ethical committee approval was not necessary. See Appendix H for a copy of the final consent as presented to the students in the first week of classes

One lecture section was designated the CAPSI section and the other the non-CAPSI section. The number of students originally registered in the CAPSI section was 137 and the number of students originally registered in the non-CAPSI section was 139. As in Experiment 1, students signed up for the lecture sections without prior knowledge of the different teaching methods used in the different lecture sections.

The instructor in this experiment was the Instructor 1 from Experiment 1 who had previous knowledge of the CAPSI program. Eight TAs were selected by the department in charge of the course and matched with lab sections based on availability. The TAs did not have prior knowledge of which section would be using the CAPSI assignment and which section would be assigned the third essay.

The textbook used for the course was *ARTS 1110 Introduction to University: Custom Edition for the University of Manitoba* (2007). The book was a compilation of chapters from McWhorther (2006) and Browne and Keeley (2007). All students used the same textbook and were given the same course outline, which detailed the course requirements.

All students were required to write two 1,000-word APA-style essays. However, students in the non-CAPSI section were required to write a third 1,000-word APA-styled essay while students in the CAPSI section had CAPSI as a required assignment. In addition, students in both lecture sections were given an ACTM prior to the beginning of the CAPSI assignment and at the end of the term. Different questions were used in the two administrations of the ACTM. This was done to avoid a practice effect from writing answers to the same questions in the pre-ACTM assessment.

The CAPSI program was set up in a similar manner to that of Experiment 1. It was again set up so that when two peer reviewers assessed the answers, both had to independently agree that the unit test was a pass in order for the program to record it as a pass. There was also a built-in appeal process for arguing the validity of a given answer.

Two lecture sections were used, with the same instructor being responsible for each lecture section and for overseeing the TAs in all 16 lab sections. The two lecture sections were scheduled for either 2:30 to 3:45 on Tuesdays or Thursdays. Through a simple random method of the instructor flipping a coin and having a third party call the toss, it was determined that the Thursday lecture section would receive the CAPSI assignment.

There were eight lab sections associated with each of the two lecture sections, with each lab section having a maximum enrolment of 20 students. The labs were scheduled once a week on the Tuesday or Thursday that was not being used for its associated lecture. The lab sections were again taught by TAs. As a result of the

Thursday lecture section being the CAPSI section, the TAs associated with the Tuesday labs would be the TAs for the CAPSI lab sections.

CAPSI Assignment

In Experiment 1, the CAPSI program amalgamated all students in the CAPSI sections and all TAs as part of the same large class. In Experiment 2, the CAPSI program treated the eight lab sections as separate classes. That is, each lab section was a separate CAPSI class. TAs only assessed answers from students in their lab section; and, students only peer reviewed other students in their lab section. The reasons for this were to increase the efficiency of certain aspects of the program and to build a greater feeling of responsibility and communication within the sections. Students could ask the TA directly about progress through CAPSI or other specific questions related to the assignment.

Prior to the beginning of the CAPSI assignment, the researcher and the instructor gave the TAs a demonstration of the CAPSI program, with instructions regarding proper feedback procedures. In addition, the instructor maintained weekly contact with the TAs either by email or in person in order to address any concerns, and to provide general feedback and motivation.

Whereas in Experiment 1, students could access the program from the first week of classes, in Experiment 2, students were able to take unit tests only in the last third of the academic term. The purpose of this was to correspond to the timing of the required third essay assignment of the non-CAPSI section and to the material as it was presented in the lectures. As in Experiment 1, in the week prior to the beginning of the CAPSI assignment, students in the CAPSI section were shown and directed to the CAPSI website and given a demonstration of the program.

The design of the questions for the CAPSI assignment followed Keller's (1999) model of attention, relevance, confidence, and satisfaction (ARCS). In following this model, small modules were set up to help students' master one unit at a time, provide students with clear criteria, and deliver effective and timely feedback for motivation and confidence.

Only the final eight chapters of the textbook were used. These were the chapters that focused on the aspect of critical thinking and critical assessment. These eight chapters were then condensed into five units.

The CAPSI assignment was designed for students to pass 5 unit tests and to peer review 5 times. Each unit test consisted of 1 question pseudo-randomly selected by the CAPSI program from a database of questions. Each question was worth 2% of the final grade and each instance of peer reviewing was worth 1%. Therefore, if a student completed all 5 unit tests and peer reviews, he or she would have earned 15% of his or her grade. A student received a final grade on the assignment associated with the number of units passed and peer reviews completed.

The questions for each unit were designed to correspond to the appropriate chapters from the course textbook. The answers could not be found in the chapters themselves; however, information from the chapters could be incorporated into the answers. The questions were originally selected by the instructor, which were then submitted to the researcher for approval prior to inclusion in the CAPSI program. The questions were either taken from letters sent into a daily newspaper or were designed by the instructor for the course. All references to the source of the questions, as well as authors of the letters were removed from the original letter. Example questions for each of the units are provided in Appendix E.

Data Collection and Assessment

Students in both lecture sections were given a pre-ACTM prior to the beginning of the CAPSI assignment in the CAPSI section or third essay assignment in the non-CAPSI section. The ACTM was administered to both lecture sections in the last week of classes. The pre-ACTM contained problems similar to but different from those the students received on the ACTM at the end of the term. The same marking rubric as in Experiment 1 was used to assess the responses to the ACTM scenarios. Three critical thinking questions were assigned in the Pre-ACTM, while a different set of three questions assigned in the Post-ACTM. The questions used on the Pre-ACTM and ACTM were the same ones that were used in Experiment 1.

Inter-rater reliabilities

Reliability of assessing the Pre-ACTM and the ACTM were calculated in the same manner as in Experiment 1. The reliability measures obtained were .74 and .83, respectively.

Results

As seen in Table 13, the mean pre-ACTM score for the CAPSI section was only slightly higher than the non-CAPSI section, indicating that the two sections were approximately equal in critical thinking prior to the CAPSI assignment. However, the difference between the non-CAPSI and CAPSI sections increased on the ACTM by a factor of over 41 (.0174 vs. .7295).

Table 13: Populations, means and standard deviations for CAPSI and non-CAPSI sections on the Pre-ACTM and ACTM

		N	Mean	Mean Difference	Std. Deviation
Pre-ACTM	CAPSI	88	3.8295	.0174	1.9371
	Non-CAPSI	80	3.8121		1.7979
ACTM	CAPSI	66	5.3636	.7295	1.7856
	Non-CAPSI	61	4.6341		1.9615

For reasons of confidentiality there was no personal identification information on the ACTM answer sheets; therefore independent samples *t*-tests for the equality of means were conducted on the differences between lecture sections on the pre-ACTM and the ACTM. As seen in Table 14, the scores on the pre-ACTM did not differ significantly between the two lecture sections. However, the scores on the ACTM did differ significantly at the $p = .03$ level using a two-tailed test. Figure 1 depicts the results graphically.

Table 14: Independent samples t-tests for the equality of means between the two lecture sections using a two-tailed test.

	t	df	Sig.
Pre-ACTM	.060	166	.952
ACTM	2.194	125	.030

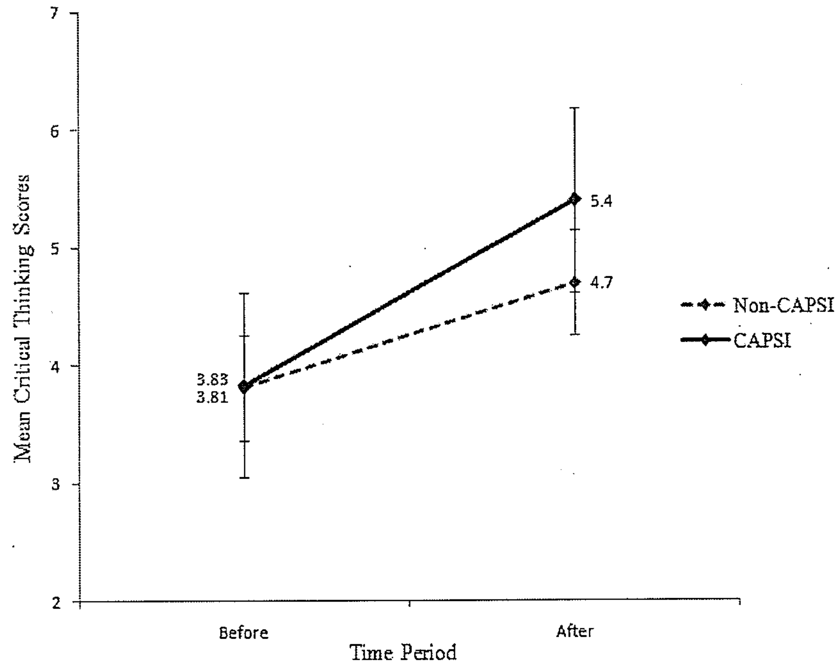

Figure 1. *Mean scores for the pre-ACTM (Before) and ACTM (After).*

Discussion

The results indicate that the CAPSI program was effective in increasing critical thinking skills as measured by the ACTM. The results on the pre-ACTM indicate that the students in both lecture sections were at approximately the same level of critical thinking ability at the start of the CAPSI assignment or third essay assignment. While students in both lecture sections did show an increase in critical thinking ability, the CAPSI section showed greater improvement than the non-CAPSI section did.

The results support Grant and Spencer's (2003) point that PSI is applicable to teaching critical thinking in post-secondary courses. Also, the outcome is consistent with Reboy and Semb's (1991) findings of improved higher-order cognitive skills in students enrolled in PSI courses.

The superiority of the CAPSI section on the ACTM could have resulted from several factors:

1. Feedback received from peers could have promoted students examining their answers for deficiencies in critical thinking.
2. The requirement of students to assess their peer's answers, attuning the reviewers to deficiencies and also providing them with examples of good critical thinking (cf. Alonso et al., 2005).
3. The opportunities for students to appeal restudy results, which may have helped to develop their critical thinking skills because of the level at which they generally had to couch their arguments in order to be successful.
4. The type of questions used, which were designed for higher-learning thinking skills according to Bloom's Taxonomy.

EXPERIMENT 3

This experiment was designed to in order to gain a further understanding of CAPSI in a blended design. According to Brothen and Wambach (2000); "PSI is especially useful in introductory psychology because the course focuses heavily on terminology and basic principles" (p. 253). Brothen and Wambach (2000) also point out that PSI should be particularly effective with introductory psychology students because PSI makes use of written materials, frequent testing, and feedback. Therefore, to follow up on these observations of PSI effectiveness, and to build on previous experiments involving the use of an Internet assignment for blended designs, CAPSI was incorporated as an assignment for an *Introduction to Psychology* class.

Two primary differences between this experiment and the previous experiments were introduced to further the generalization and transferability of CAPSI in a blended design for university courses. The first difference was the use of CAPSI in a different course. The second difference was the use of CAPSI in an introductory course placing a greater emphasis on knowledge, comprehension, and application. Therefore, because of focusing on the lower levels, only development of course knowledge was assessed in this experiment, while the critical thinking assessment was excluded. This was in accordance with the general design of the study, which involved measuring both knowledge and critical thinking together in Experiment 1, then assessing these separately in follow-up experiments.

To further assess the transferability of CAPSI as part of a blended design for university education, this experiment also differed from Experiments 1 and 2 in five other ways: (a) there were no TAs or labs, (b) the experiment took place in a different university, (c) the length of the course was 26 weeks as opposed to 13 weeks, (d) multiple-choice quizzes were used as the assignment to be compared with CAPSI, and (e) the dependent variable was performance on four multiple-choice exams. Other differences are detailed below. The hypothesis to be tested was that CAPSI assignments as a component of a blended designed course develops student course knowledge more than do multiple-choice quizzes.

Method

Participants

The participants were 194 University of Winnipeg students enrolled in two lecture sections of *Introduction to Psychology*. The course consisted of 26 weeks extended over two academic terms. The first term began in September 2007 and ended before the Christmas break in December; the second term began in January and ended in April 2008.

Procedure

The procedure, and materials used in this course were similar to those used in Experiments 1 and 2. However, this experiment was conducted in a different university; therefore, a discussion with the head of the Psychology Department at the University of Winnipeg was necessary in order to discuss ethical approval. It was decided after the meeting; based on ethical approval being granted by the University of Manitoba, additional approval by the University of Winnipeg was unnecessary. See Appendix I for a copy of the consent form as presented to the students in the first week of classes

As in Experiments 1 and 2, students signed up for a particular lecture section without prior knowledge of the teaching methods to be used in that section. Each lecture section met for a fifty-minute lecture three times a week. The CAPSI section met from 9:30 to 10:20 AM. The non-CAPSI section met from 10:30 to 11:20 AM. Each section was taught by the same instructor (who was also the author), in the same room, using the same materials. Assignment of these sections to the instructor was done by the psychology department at the University of Winnipeg based on availability of the instructor.

Both lecture sections used the same textbook: *Psychology in Context* 3rd edition by Stephen M. Kosslyn and Robin S. Rosenberg (2006). This book was selected because of a familiarity with the text material from previous experiences in teaching introductory psychology courses. The book was originally chosen because it was a well written with well designed pedagogical features, such as an opening illustrative story in order to connect all new concepts for a particular chapter.

The two lecture sections differed in an assignment worth a potential 15% of the total grade. The assignment for students in the non-CAPSI section was to write 13 multiple choice quizzes at designated dates throughout the term. The assignment for students in the CAPSI section was to work through 15 CAPSI units. The difference between the number of CAPSI units and multiple-choice quizzes was due to the two mid-term exams that were given on days on which quizzes would otherwise have been given in the non-CAPSI section.

Quizzes

Each of the 13 quizzes contained 10 questions from a chapter that was just discussed in the previous classes. Each quiz was worth 1.15% of the final grade. Questions for each of the quizzes were selected from the *Pearson Education: TestGen version 7 Test Bank* accompanying the course textbook.

The quizzes were given to the students in the non-CAPSI section on specific dates selected by the author. The quiz dates corresponded with the first class after completing the discussion for the chapter on which the quiz was based. Students were given 10 minutes to complete each quiz. They were allowed to use information they had written down on one side of a full page of standard size paper. The quizzes were scored and returned during the next class, thus approximating the quick turnaround and feedback on submitted answers in the CAPSI assignment.

Students did not need to write all thirteen quizzes to pass the course, nor were they required to write quizzes in order to write an exam. If a student missed a quiz, an opportunity was given for that student to write the quiz at another time. If a student received a low grade on a quiz he or she would not be allowed to repeat the quiz as rewriting quizzes would probably not have helped in exam performance (Brothen & Wambach 2001). The primary aspects of the quizzes that provided for a comparison with CAPSI were (a) the process of studying for the quizzes, (b) writing the quizzes, and (c) gaining feedback from the quiz results.

CAPSI

During the first class of the term, students in the CAPSI section were shown and directed to the CAPSI website at http://home.cc.umanitoba.ca/~capsi/. After this, they were ready to enter the program and start on the CAPSI unit tests. Students had access to the CAPSI assignment from the first day of classes to the last day of classes in the academic term.

The CAPSI program required that unit tests be marked by either the instructor or by two peer reviewers. When assigning peer reviewers, the CAPSI program selected two peers with the lowest number of peer-reviewing points. When two peer reviewers assessed the answers, both had to independently agree that the unit test was a pass in order for the program to record it as a pass. There was also a built-in appeal process for arguing the validity of a given answer. The appeal was submitted to the principal investigator, who then judged the strength of the argument and decided whether or not to accept the appeal.

The CAPSI assignment was designed for students to pass 15 unit tests and to peer review 10 unit tests. Each CAPSI unit corresponded to one chapter from the course textbook. Each unit test was worth 0.66 % of the final grade, while each peer review was worth 0.5%. For each unit, students were required to answer 3 randomly generated questions from the CAPSI unit corresponding to the appropriate chapter from the textbook.

The questions for each unit test were composed according to the requirements for the six levels of Bloom's taxonomy. The questions were designed in order for students to test their knowledge of the textbook material. The question level definitions for this design were the same as the definitions used for Experiment 1.

While it would likely be to a student's advantage to complete the four CAPSI units prior to one of the four multiple-choice exams on those units, there was no requirement that this had to be done. Students were also not required to complete all unit tests in order to pass the course. A student who did not complete all of the units would receive a percentage of what was completed toward the final grade.

Exams

Students in each section received four exams throughout the year. These were given after every fourth chapter was discussed in the lectures. Exams 1-3 were worth 20% of the total course mark, Exam 4 was worth 25%. Exams 1 and 3 were written during designated class periods; Exams 2 and 4 were written in the exam periods following the end of lectures for that term.

The four exams each comprised 60 multiple-choice questions chosen from the *Pearson Education: TestGen Version 7 Test Bank* accompanying the course textbook.

Each exam covered the previous four chapters of the textbook that were discussed in class. The exam questions were the same for each lecture section; however, the order of the questions in the exams was different for each section. The change in the order of the questions was done to reduce the possibility of students who had written the exam from relaying answers to the students who had not yet written the exam.

Data Collection and Assessment

The primary data for this experiment was student performance on the exams. Since the exams were multiple choice with clearly defined answers, a measure of the reliability of the scoring was unnecessary. Other data of interest were the number and percentage of students completing the various assessments of the course, and the course as a whole.

Results

Table 15 shows the number and percentage of students of the non-CAPSI section who wrote each of the quizzes; an *N* of 64 was used because this was the number of students who completed all four exams, and therefore, completed the course. Quizzes 1-3 were given prior to Exam 1, quizzes 4-7 were given prior to Exam 2, quizzes 8-10 were given prior to Exam 3, and quizzes 11-13 were given prior to Exam 4.

Table 15: Number and percentage of students of students who wrote each of the quizzes

Quiz	1	2	3	4	5	6	7	8	9	10	11	12	13
N	64	63	62	58	57	56	55	53	56	55	53	55	51
%	100	95.32	93.75	90.62	89.07	87.5	85.94	82.82	87.5	85.94	82.82	85.94	76.69

Table 16 shows the number and percentage of students of the CAPSI section who wrote tests on each of the units; an *N* of 82 was used because this was the number of students who completed all four exams, and therefore, completed the course. Note that student participation in 13 of the 15 units of the CAPSI tests was greater than or equal to student participation in the quizzes. Also note that the percentage of students writing the quizzes became higher over the semester.

Table 16: Number and percentage of students who completed the CAPSI unit tests

Unit	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
N	82	81	80	77	75	72	69	66	65	61	59	55	51	48	37
%	100	99.0	97.6	94.0	91.5	87.8	84.1	80.5	79.3	74.4	72	67.1	62.2	58.5	45.1

An unexpected result of this experiment was that although both sections started with the same number of students (N=97), more than twice as many dropped out of the non-CAPSI section (33 vs. 15). A two-tail chi-square analysis ($\chi^2 = 8.97$, df = 1) indicated that the probability of this difference occurring by chance was $p < .01$.

As can be seen in Table 17, the CAPSI section scored higher on all four exams. The mean overall difference over all four exams was 4.88 for all students who wrote each exam. The table includes students who finished the course, and students who dropped out of the course.

Table 17: Mean differences between groups of all students who wrote each exams

	Exam 1	Exam 2	Exam 3	Exam 4	Mean over 4 exams
CAPSI	63.5	63.6	69.3	71.65	**67.01**
Non-CAPSI	57.5	57	65.3	68.72	**62.13**
Mean differences	**6**	**6.6**	**4**	**2.93**	**4.88**

Table 18 shows the mean differences on exam scores between the sections just for the students who completed the course. Note that the CAPSI section still repeatedly scored higher, with a mean overall difference over all four exams of 3.319. Of note is the second largest mean difference was in the final exam, which occurred after the final date for withdrawing from the course.

Table 18: Mean differences on exam scores between groups of students who completed the course

	Exam 1	Exam 2	Exam 3	Exam 4	Mean over 4 exams
CAPSI	63.75	65.11	67.82	71.65	**67.08**
Non-CAPSI	61.2	59.24	65.91	68.72	**63.77**
Mean differences	**2.55**	**5.87**	**1.91**	**2.93**	**3.31**

A repeated measures ANOVA was conducted on these data and is shown in Table 19. As can be seen in Table 19, the difference between the sections in exam scores did not reach statistical significance. The table also shows that the interaction between the sections and exam scores was not statistically significant. However, there was a statistically significant increase in student performance on the exams over the term. Figure 2 depicts a graphical representation of the data in Table 17.

Table 19: Repeated measures ANOVA for exams between and within sections

	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Sections	1583.562	1	1583.562	2.002	.159	.014
Error (Between subjects)	113891.285	144	790.912			
Exam	6328.128	3	2109.376	33.609	.000	.189
Exams within sections	332.940	3	110.980	1.768	.152	.012
Error (Within Subjects)	27112.994	432	62.762			

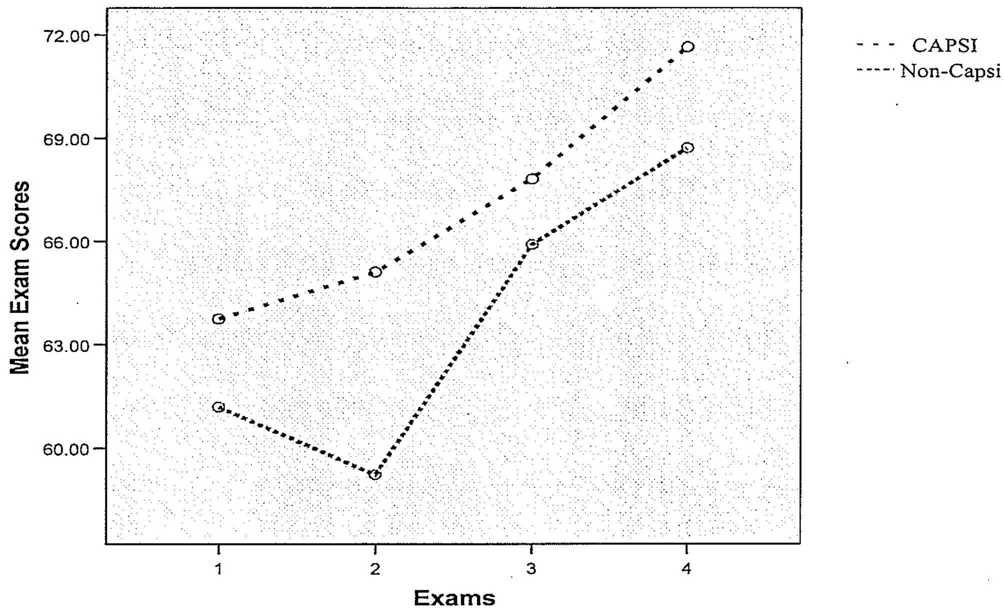

Figure 2. Mean scores for CAPSI and non-CAPSI sections over the 4 exams

Discussion

Students in the CAPSI section scored higher on all four exams, which is consistent with previous studies showing that PSI produces higher examination scores than are obtained with the lecture method (Eppler & Ironsmith, 2004: Ironsmith & Eppler, 2007; Kulik, Kulik, & Bangert-Drowns, 1990). However, this difference was not statistically significant in the present experiment.

Both sections started with 97 students. The CAPSI section ended with 82 students, and the non-CAPSI section ended with 64. Therefore, the CAPSI section had 15 dropouts and the non-CAPSI section 33 dropouts, which was a statistically significant difference. The significant differential dropout rate could have been partially responsible for the lack of statistical significance between the sections. The experiment did not include a questionnaire regarding student responses to the course and possible reasons for dropping out; however, the data indicate that students with low examination scores tended to drop out (see Tables 17 and 18). If this was the case, then if these students had stayed in the course, or were included in the statistical analysis for the exams they did write, this would have lowered the mean scores of the non-CAPSI section and possible produced a significant difference.

The lower dropout rate of the CAPSI section is not consistent with Born, Gledhill, and Davis (1972), who found a higher dropout rate in their PSI section which they attributed to academically poorer students tending to withdraw from classes taught using PSI. However, the present study is consistent with Dzubian, Hartman, and Moskal (2004) and Pereira et al. (2007) who had fewer students dropping out in blended designed classes than in lecture–based classes; and Hobbs (1987) who had 49 of 51 students completing a PSI course in behavioural psychology. Thus, PSI does not necessarily lead to a higher dropout rate.

One possible confound to consider, given the higher dropout rate in the non-CAPSI section, is that higher achieving students may have selected the earlier class time. However, there was only a one hour difference between class times. If the difference was an early morning class and a late afternoon class, time of day might be more strongly considered as a confounding variable.

It can be argued that the procedure was weighted in favour of the students in the non-CAPSI section. First, they had practice opportunities simulating the exam situation. Second, they received feedback indicating right and wrong answers from the quizzes, which could have aided them in recognizing areas of weaknesses. Third, the time limits on writing the quizzes could have favoured the non-CAPSI section by better preparing those students to write time-limited exams.

The students in the CAPSI section did not have the same opportunity to write multiple-choice tests. Instead, the students in the CAPSI section wrote out answers to the various questions. The results from this study; however, suggest that writing answers to questions is more effective than writing quizzes in preparation for exams, even for multiple choice exams.

General Discussion

This study demonstrated that using CAPSI as an assignment within a blended design can be effective for developing student knowledge and critical thinking skills. The study included three experiments, with results consistently showing the CAPSI sections scoring higher on knowledge of course content and critical thinking in all assessments. While the results were not always statistically significant, their consistency indicates a reliable effect.

Of note are the multiple blended learning designs that CAPSI was used in. First, CAPSI was shown to be effective with designs involving a number of instructors, mentors (TAs), and students. In Experiment 1, CAPSI was used with 2 instructors, 16 TAs, and 193 students. In Experiment 2 CAPSI was used with one instructor, eight TAs, and eight CAPSI sections of less than 20 students. Experiment 3 had one instructor, no TAs, and 194 students. Second, CAPSI was effective with different textbooks. Third, CAPSI was effective using three different sets of questions. Fourth, CAPSI was used in three different lengths of time: 5 weeks in Experiment 2, 13 weeks in Experiment 1, and 8 months in Experiment 3.

The positive outcomes in all three experiments indicate that CAPSI is an effective program for learning and critical thinking development in multiple blended designs. The prediction would be that no matter whom the instructor is, or what the course content and specific objectives are, a CAPSI assignment can aid that instructor in developing critical thinking and course knowledge for his or her students.

Critical Thinking

Contrary to the criticism that behaviourist methods are not effective for developing higher level thinking skills (Budd, 2002), scores on critical thinking measures, while not statistically significant, showed a consistently positive difference in favour of the CAPSI sections in two experiments. The results from these experiments support the findings of Reboy and Semb (1991) that PSI can be used to teach critical thinking at levels of achievement that exceed those attained in a lecture-discussion format. The positive effect of PSI on critical thinking could be due to, as stated by Reboy and Semb (1991), early diagnosis of inadequate answers, which increased the probability of detecting and correcting errors in reasoning. The early diagnosis and correction in CAPSI is attained by receiving feedback on written answers and on appeals that argue the merits of specific answers.

Other factors that could have contributed to this result are the types of questions asked, especially in Experiment 2, and the inclusion of peer reviewing. As stated earlier, Reboy and Semb (1991) point out that the components of the PSI delivery system – in this case the questions designed for higher-order thinking according to Blooms Taxonomy – can be focused on developing higher-order cognitive skills. That peer reviewing might be a factor contributing to the development of critical thinking is consistent with Terenzini et al. (1995), who stated that interactions among students seems to have a positive influence on critical thinking. For CAPSI, the interactions involving peer reviewing could have produced a similar enhancement of critical thinking. Peer reviewing can be viewed as a highly systematic written dialogue between students on specific course information (Carr-Chellman & Duchastel, 2001).

Course Content

The consistent superiority of CAPSI is in accordance with the findings of greater college-level exam scores for PSI type learning (Kulik et al., 1979, 1990; M. Skinner, 1990). It is also consistent with the findings of Dzubian, Hartman, and Moskal (2004), and Pereira et al. (2007) who showed that students in blended learning courses achieved higher scores than students in lecture-based courses. The small effect sizes found in the present study are not consistent with the larger effect sizes found in previous studies on PSI. However, most of the prior studies compared PSI with a straight lecture method. In the present study, steps were taken in an attempt to ensure that the methods used in the comparison sections were as effective as possible; e.g., assigning essays, providing TAs that assisted students in their learning (Experiments 1 and 2, and giving, frequent quizzes with rapid feedback in Experiment 3). In addition, in all three of the present experiments, CAPSI was blended with lectures that were identical across sections. All of these factors should have decreased the differences between the CAPSI and non-CAPSI sections. That the CAPSI sections still consistently outperformed the sections with which they were compared with in regard to course content learning and critical thinking supports the view that CAPSI is a highly effective empirically based educational methodology.

Decrease in dropouts

An unexpected result of this study was the significantly lower dropout rate in the CAPSI section in Experiment 3. This result is not consistent with previous studies of PSI, and should be looked at more carefully in future studies. If the finding turns out to be reliable it should be of interest to instructors and administrators who are concerned about high dropout rates, especially in first year university courses.

As previously discussed, there has been a change in student population demographics; students are now, on average, older, and with varied interests and responsibilities, making it more difficult to attend classes on campus. The use of an online component in a blended learning design, therefore, could allow these students to stay in a course, rather than having to drop out (Dzubian, Hartman, & Moskal, 2004).

Limitations

Care was taken in designing the research for each experiment; however, the research involved departments and faculties with other objectives, which decreased the amount of experimental control that could be implemented. In Experiments 1 and 2, the research was adapted to a course design after the decision for the design was made and a description of the course put into the student undergraduate calendar. For Experiment 3, the author had more control for the design of the class, but was still confined to offering the course as it was described in the course calendar – with three face-face-face classes per week.

An additional limitation was the 0.74 inter-observer correlation for the pre-ACTM scores in Experiment 2. This score is below the usual standard of 0.80. While this was the only situation of this occurring in the four assessments, greater care would be taken to insure this would not occur. To do this, two research assistants would initially be given 5% of the assessments. If a score of 0.80 was not attained, then the assistants would be provided with additional 5% of the assessments until a consistent score of 0.80 was established.

ACTM

Obtaining significant differences between CAPSI and non-CAPSI sections might have been difficult because there were multiple issues that needed to be considered in determining any improvement with critical thinking skills. In this regard, Halpern (2001) makes three important points: First, subtle, yet positive, changes in critical thinking might be difficult to be "pick up" or detect. Second, the responses to hypothetical situations could suffer from reliability. Third, it may be unrealistic to expect the occurrence of a significant gain in thinking abilities within just one semester. Instead, it is "more realistic to expect modest improvements in thinking abilities, a fact that makes assessment all the more difficult" (p. 273). In Experiment 1, CAPSI was available to students for only 13 weeks, while in Experiment 2 CAPSI was available for only five weeks. Therefore, the possibility exists that the time available for developing critical thinking skills was too short for the effectiveness of CAPSI in this regard to be fully realized.

In addition, accommodations for implementing the ACTM assessments in the labs were done in order for the TAs and instructors to fulfill their course objectives and requirements. Therefore, there was a time limit of 15 minutes for students to complete the ACTMs. As a result, responding to three scenarios within that time limit might have limited the opportunity for students to fully engage with the assessments.

CAPSI

Not all students in the CAPSI sections completed the unit tests. Therefore, the consistently positive results of the CAPSI groups are based on a smaller amount of CAPSI than would be indicated by a superficial look at the procedure. This could have decreased the probability of obtaining a significant effect of CAPSI and the effect size. In fact, on the basis of the positive results from this study, and the literature on PSI, an assumption could be made that greater student participation in CAPSI would have resulted in greater positive results.

The possibility could exist that the positive results that occurred were due to extraneous variables rather than CAPSI; however, the consistent positive outcomes in favour of CAPSI over three experiments, and the prior research demonstrating the effectiveness of PSI, would argue against that possibility (Brothen, 1996; Brothen & Wambach, 1998; Crosbie & Kelly, 1993; Hambleton, Foster, & Richardson, 1998; Price, 2000; Pear & Crone-Todd, 1999; Lloyd & Lloyd, 1986 Sherman, 1992).

Instructor Effects

Differences between instructors in Experiment 1 or instructor bias in Experiments 2 and 3 could have been confounding variables. However, differences between instructors were controlled for as a confounding variable in Experiments 2 and 3 because the same instructor was used for the CAPSI and non-CAPSI sections. Instructor bias was controlled for because different instructors taught in the courses in those experiments. Instructor 1 from Experiment 1 taught both sections in Experiment 2, while Instructor 2 from Experiment 1 taught both sections in Experiment 3. In both cases, the instructor did not alter his teaching methodology regarding the presentation of information to the students. In addition, reliability measures on the dependent variable were used in Experiments 1 and 2 and multiple-choice questions, which are more objective than essay type questions, were used in Experiment 3.

Differences between TAs might have been a confounding variable in Experiments 1 and 2. However, this could not have been the case in Experiment 3 because there were no TAs in that experiment. In Experiment 2, the potential effect of TAs was decreased due to the overall decrease of TA participants from 16 to 8, while in Experiment 3 the effect of TAs were removed completely.

Implications

As stated earlier, when discussing the present study, there is little theoretical discussion in the literature on blended learning's pedagogical and epistemological basis (Mortera-Guitierrez, 2006). In addition, there is little in the literature regarding important design features in blended learning classrooms (Ausburn, 2004). Moreover, there is limited empirical evidence to support many of positive claims being made for blended learning (Matheos et al., 2006). An important focus, therefore, of this study was to address this lack of information on the effects of blended learning, and more specifically, on online learning systems within a blended design.

Blended learning allows for incremental inclusions of Internet based learning methods into traditional classrooms (Matheos et al. 2006). In this regard, using CAPSI as a learning assignment in three lecture courses was a first step to discovering the effectiveness of an online assignment as an addition to face-to-face lectures. Progressing in this manner will allow for greater understanding on the effects of CAPSI as opposed to a design including limitation of classroom interactions and lectures. The effectiveness of CAPSI may be such as to allow it to be included as an online assignment for blended courses with decreased lecture hours.

There are also practical implications for instructors using CAPSI in any blended design. There could be a situation where a large amount of information needs to be covered in a limited amount of time. In this case, instructors could be confident that students would learn the required material by completing the CAPSI tests that are designed by the instructor, leaving classroom time available for discussion of the material and for student motivation. Additional implications are that instructors can use CAPSI for any class size, length of time, type and number of questions in a unit test, or the type of assessment used for mid-term tests or final exam.

With the increase of technology in the classroom (Dzubian, et al. 2004), the evidence presented in this study of an Internet-based learning management tool will likely be a welcome addition for any course design. This research has added to the earlier findings of PSI as an effective pedagogical design (Born, Gledhill, & Davis, 1972; Kulik, Kulik, & Cohen, 1979), and to the effectiveness of CAPSI (Pear & Crone-Todd, 1999; Pear & Martin, 2004). This study has also added to the literature regarding higher exam and tests scores when comparing student performance in a PSI section to completely lecture based section (Eppler & Ironsmith, 2007; Ironsmith & Eppler, 2004).

Future Directions

ACTM

In future research, regarding the use of the ACTM, three changes should occur. First, the time to write the ACTM should be increased to one hour in order to complete a minimum of five scenarios, as this would provide better evidence for changes in critical thinking. Therefore, instead of measuring averages out of a score of 27, averages could be rated on scores of 36, 45 and so on. Overall, this would increase the amount of data from each student, which should increase the reliability of the measure. Second, to provide greater consistency, it would be necessary to have one or two people provide the ACTM instructions, as opposed to the 16 or 32 separate individuals (the TAs) from these experiments. Third, further research should include an assessment of the ACTM in relation to other critical thinking measures, such as the Ennis-Weir Critical Thing Essay Test, the Watson-Glaser Critical Thinking Appraisal, and the California Critical Thinking Skills Test.

CAPSI

Seven future directions to consider regarding future research on CAPSI are as follows. First, CAPSI's inclusion in these experiments was part of courses that had full compliments of face-to-face hours per week. Further research should progress to include CAPSI in course designs with reduced classroom time.

Second, if research is to be conducted across multiple sections of the same course, then a procedure would be put into place to maintain instructor consistency in all sections. It is possible that, unknown to the instructor, certain behaviours could have been emitted to motivate greater student learning in the CAPSI section. Therefore, videos of the live lectures could be conducted and shown to observers who are not aware of which lecture belongs to which section. This could also be done in a double-blind design where the videos are numbered by a research assistant, then randomly given to a second assistant, who then delivers these to the independent observers. These observers would then score the lectures along various dimensions such as content delivery, pacing, repition of descriptions, and number of questions to the students in order to assess any possible differences between the lectures.

Third, student activity in regard to the CAPSI assignment in all three experiments was done without deadlines or requirements of students to complete unit tests. Future research might include the use of deadlines and requirements for student involvement in the CAPSI assignment. Following up on earlier research (Born et al., 1972; Crosbie and Kelly, 1993; Hambleton, Foster, & Richardson, 1998; Price, 1999; Ross & McBean, 1995), one method to encourage students to move faster through a CAPSI course would be to give bonus marks to students for completing units before a test or a specific date and apply penalties to students who did not reach the deadline date. However, the punishment should not be so harsh as to have students dropping the course. Punishment could be mitigated by allowing students to offset it by achieving bonuses later in the course.

Fourth, in Experiment 3 the quizzes were done in class. A future study could look at having student's complete quizzes as in a CAPSI design, but the questions would be multiple-choice, with mastery set at 90%.

Fifth, in regard to efficiency of questions, a study could look at introductory psychology students receiving only specific levels of Bloom's taxonomy. For example, one group might receive questions only from the first three levels; the second group would receive only questions from the top three levels. This design could also include a group of students only participating in multiple-choice unit tests. Therefore, the experiment would involve three groups: (a) multiple-choice, (b) Blooms levels 1-3, and (c) Blooms levels 4-6.

Sixth, a more powerful test of CAPSI that was done in Experiment 3 should be done in a follow-up experiment, thus increasing the probability of obtaining statistical significance. This study would involve two sections of *Introduction to Psychology* over two terms, as in Experiment 3. The difference between groups would be in-class quizzes and CAPSI; however, the assignment would change after the Christmas break. The group receiving the quizzes in the first half of the course would receive CAPSI in the second half, while the opposite would occur for the other section. A more complicated, but possibly more effective, design would be to have half of each section of two sections receive either the quiz assignment or the CAPSI assignment. Again switching would occur after the first half of the course, with changes occurring within sections. The analysis would then look at differences within sections and between sections.

A further manipulation would be to have three groups within one section. One group in both sections would continue with one assignment through two semesters, while the other two groups would switch mid-way through the course. If two sections were used, the continuous assignment would be quizzes in one group and the CAPSI assignment in the other group.

References

Abrami, P. C. (2001). Understanding and promoting complex learning using technology. *Educational Research and Evaluation, 7*(2), 113-136.

Abrahmov S. L., & Ronen, M. (2008). Double blending: online theory with on-campus practice in photography instruction. *Innovations in Education and Teaching International* 45(1), 3–14.

Adams, A. (2004) Pedagogical underpinnings of computer-based learning. *Journal of Advanced Nursing*, 46(1), 5-12.

Akkoyunlu, B. & Yılmaz-Soylu, M. (2008). Development of a scale on learners' views on blended learning and its implementation process. *Internet and Higher Education, 11*, 26–32.

Alonso, F., Lopez, G., Manrique, D., & Vines, J. (2005). An instructional model for web-based e-learning education with a blended learning process approach. *British Educational Communications and Technology, 36(2)*, 217-235.

Allan, I. E., & Seaman, J. (2006). Making the grade: Online education in the United States, 2006. Needham, MA: Retrieved July 10, 2007 from http://www.sloan-c.org/resources/making_the_grade.pdf

Aretz, A. J., Bolen, M. T., & Devereux, K. E. (1997). Critical thinking assessment of college students. *Journal of College Reading and Learning*, *28*(1), 12-22.

Arts 1110: Introduction to University. (2007). New York, New York: Longman

Astleitner, H. (2002). Teaching critical thinking online. *Journal of Instructional Psychology*, *2*, 53-77

Ausburn, L. J. (2004). Course design elements most valued by adult learners in blended online education environments: An American perspective. Retrieved July 28, 2006, from http://www.tandf.co.uk/journals

Aycock, A., & Kaleta, R. (2002). Lessons learned from the Hybrid Course Project. Retrieved May 23, 2007, from http://www.uwsa.edu/ttt/articles/garnham2.htm

Baker, W., Hale, T., & Gifford, B. R. (1997). From theory to implementation: The mediated learning approach to computer-mediated instruction, learning and assessment. Retrieved September 22, 2004, from http://www.educause.edu/pub/or/review/review.Articles/33542.html

Balram, S. & Dragićević, S. (2008). Collaborative spaces for GIS-based multimedia cartography in blended environments. *Computers & Education*, *50*, 371–385

Bellefeuille, G. L, (2006). Rethinking reflective practice education in social work education: A blended constructivist and objectivist instructional design strategy for web-based child welfare practice course, *Journal of Social Work Education. 42*(1), 85-103

Ben-Jacob, M. G. (1999). A workshop for distance learning educators of the 21st century. *Journal of Educational Technology Systems, 27*(2), 117-131.

Bernard, R. M., Abrami, P. C., Lou, Y., Borokhovski, E., Wade, A., Wozney, L. (2004). How does distance education compare with classroom instruction? A meta-analysis of the empirical literature. *Review of Educational Research, 74*(3), 379-439.

Binder, C., & Watkins, C. L. (1990). Precision teaching and direct instruction: Measurably superior instructional technology in schools. *Performance Improvement Quarterly, 3*(4), 74-96.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of Educational Objectives*. New York: Longmans, Green and Company.

Bloomfield, J. G., While, A. E., & Roberts, J. D. (2008). Using computer assisted learning for clinical skills education in nursing: Integrative review *Journal of Advanced Nursing. 63*(3), 222-235.

Bodie, G. D., Powers, W. G., & Fitch-Hauser, M. (2006). Chunking, Priming and Active Learning: Toward an innovative and blended approach to teaching communication-related skills. *Interactive Learning Environments, 14*(2), 119-135.

Born, D. G., Gledhill, S. M., & Davis, M. L. (1972). Examination performance in lecture-discussion and personalized instruction courses. *Journal of Applied Behavior Analysis*, *5*(1), 33-43.

Borreson, C. J., & Salaway, G., (2008). The ECAR study of undergraduate students and information technology, 2008 roadmap, *Educause Roadmap*. EDUCAUSE Center for Applied Research, 2008, Boulder, CO: Retrieved, December 15 from http://www.educause.edu/ecar.

Botsch, C. S., & Botsch, R. E. (2001). Audiences and outcomes in online and traditional american government classes: A comparative two-year case study. *Political Science & Politics*, *34*(1), 135-141.

Brothen, T. T. (1996). Comparison of non-performers and high-performers in a computer-assisted mastery learning course for developmental students. *Research & Teaching In Developmental Education, 13*(1), 69-74.

Brothen, T., & Wambach, C. (1998). An evaluation of lectures in a computer-based, PSI introductory psychology course. *Journal of Educational Technology Systems, 27*(2), 147-155.

Brothen, T., & Wambach, C. (2000). The effectiveness of computer-based quizzes in a PSI introductory psychology course. *Journal of Educational Technology Systems, 28*(3), 253-261.

Brothen, T., & Wambach, C. (2001). Effective student use of computerized quizzes. *Teaching of Psychology, 28*(4), 292-294.

Brothen, T., & Wambach, C. (2004). The value of time limits on internet quizzes. *Teaching of Psychology, 31*(1), 62-64.

Browne, M. N., (2007). *Asking the Right Questions: A Guide to Critical Thinking* (Eighth ed). Upper Saddle River, New Jersey: Prentice Hall

Budd, J. W. (2002). Teaching labor relations: Opportunities and challenges of using technology. *Journal of Labor Research, 23(3)*, 355-370.

Caldwell, E. C. (1985). Dangers of PSI. *Teaching of Psychology*, *12*(1), 9-12.

Carman, J. M. (2005). Blended learning design: Five key ingredients. Retrieved April 24, 2007, from http://www.agilantlearning.com/pdf/Blended%20Learning%20

Carr-Chellman, A., & Duchastel, P. (2001). The ideal online course. *Library Trends, 50*(1), 145-158.

Chickering, A. W., & Kytle, J. (1999). The collegiate ideal in the twenty-first century. *New Directions For Higher Education, 105*, 109-120.

Christensen, T. K. (2003). Finding the balance: Constructivist pedagogy in a blended course. *The Quarterly Review of Distance Education, 4*(3), 235-243.

Chyung, S.-Y. Y., & Stepich, D. (2003). Applying the "congruence" principle of Bloom's taxonomy to designing online instruction. *The Quarterly Review of Distance Education, 4*(3), 317-330.

Collucciello, M. L. (1997). Critical thinking skills and dispositions of bacculaureate nursing students-a conceptual model for evaluation. *Journal of Professional Nursing, 13*(4), 236-245.

Concannon, F., Flynn, A., & Campbell, M. (2005). What campus based students think about the quality and benefits of e-learning. *British Journal of Educational Technology, 36*(3), 501-512.

Cooner, T. S. & Hickman, G. (2008). Child protection teaching: Students' experiences of a blended learning design. *Social Work Education, 27*(6), 647-657.

Cracolice, M. S., & Roth, S. M. (1996). Keller's old personalized system of instruction: A new solution for today's college chemistry students. *The Chemical Educator, 1*(1), 1-18.

Crone-Todd, D.E., & Pear, J.J. (2001). Application of Bloom's Taxonomy to PSI. *The Behavior Analyst Today*, *3*, 204-210.

Crone-Todd, D. E., Pear, J. J., & Read, C. N. (2000). Operational definitions for higher-order thinking objectives at the post-secondary level. *Academic Exchange Quarterly, 4*(3), 99-106.

Crosbie, J., & Kelly, G. (1993). A computer-based personalized system of instruction course in applied behavior analysis. *Behavior Research Methods, Instruments, and Computers, 25*(3), 377-370.

Daley, B. J., Watkins, K., Williams, S. W., Courtenay, B., Davis, M., & Dymock, D. (2001). Exploring learning in a technology-enhanced environment. *Educational Technology & Society, 4*(3), 126-138.

Dalsgaard, C. & Godsk, M. (2007). Transforming traditional lectures into problem-based blended learning: Challenges and experiences. *Open Learning, 22*(1), 29–42.

Dettori, G., Giannetti, T., & Persico D. (2006). SRL in online cooperative learning: Implications for pre-service teacher training. *European Journal of Education, 41*(3/4), 398-414

Dewhurst, D. G., Macleod, H. A., & Norris, T. A. M. (2000). Independent student learning aided by computers: An acceptable alternative to lectures? *Computers & Education, 35*, 223-241.

Dziuban, C. D., Hartman, J. L., & Moskal, P. D. (2004). Blended learning. Retrieved May 23, 2007, from http://www.educause.edu/ir/library/pdf/ERB0407.pdf

Ennis, R. H. (1987). *Teaching Critical Thinking Skills*. New York: W. H. Freeman and Company.

Eppler, M. A., & Ironsmith, M. (2004). PSI and distance learning in a developmental psychology course. *Teaching of Psychology, 31*(2), 131-133

Eynon, R. (2008). The use of the world wide web in learning and teaching in higher education: Reality and rhetoric. *Innovations in Education and teaching International, 45*(1), 15-23

Eyre, H. L. (2007). Keller's personalized system of instruction: Was it a fleeting fancy or is there a revival on the horizon. *The Behavior Analyst Today*, 8(3), 317-324

Facione, P. A. (1986). Critical thinking assessment. *Theory Practice, 32,* 179-186.

Falconer I., & Littlejohn A. (2007). Designing for blended learning, sharing and reuse. *Journal of Further and Higher Education, 31* (1), 41-52

Gallup, H. F., & Allen, R. W. (n.d.). Concerns with some recent criticisms of the personalized system of instruction (PSI). Retrieved February 19, 2004, from http://ww2.lafayette.edu/~allanr/concerns.html

Garnham, C., & Kaleta, R. (2002). Introduction to hybrid courses. Retrieved May 23, 2007, from http://www.uwsa.edu/ttt/articles/garnham.htm

Garrison, D. R. (1991). Critical thinking and adult education: A conceptual model for developing critical thinking in adult learners. *International Journal of Lifelong Education, 10*(4), 287-303.

Garrison D. R., Kanuka H. (2004). Blended learning: Uncovering its transformative potential in higher education. *The Internet and Higher Education 7*, 95–100.

Gerber, M., Grund, S., & G. Grote G. (2008). Distributed collaboration activities in a blended learning scenario and the effects on learning performance. *Journal of Computer Assisted Learning, 24*, 232–244

Giancarlo, C. A., & Facione, P. A., (2001). A look across tour years at the disposition toward critical thinking among undergraduate students. *The Journal of General Education, 50*(1), 29-55

Giouvanakis, T., Samaras, H., Konstantinos, T. (2001). Designing a pedagogically sound web-based interface: The critical role of prior knowledge, *World Conference on*

Educational Multimedia, Hypermedia, & Telecommunications, Tampere, Finland, June 25-30, 2001

Grant, L. K., & Spencer, R. E. (2003). The personalized system of instruction: Review and applications to distance education. Retrieved December 3, 2004, from *http://www.irrodl.org/content/v4.2/grant.spencer.html*

Gresh, K. S. & Mrozowski, S., (2000). Faculty/student interaction at a distance: Seeking balance. *Educause*, Nashville, Tenn. October, 10-13, 2000

Guskey, T. R. (2001). Benjamin S. Bloom's contributions to curriculum, instruction, and school learning, Paper presented at the Annual meeting of the American Educational Research Association, Seattle, WA. April 10-14, 2001

Gustafson, K. (2004). The impact of technologies on learning. *Planning For Higher Education, December 2003-February 2004*, 37-43.

Halpern, D. F. (1998). Teaching critical thinking for transfer across domains. *American Psychologist, 53*(4), 449-455.

Halpern, D. F. (2001). Assessing the effectiveness of critical thinking instruction. The *Journal Of General Education, 50*(4), 270-286.

Hambleton, I. R., Foster, W. H., & Richardson, J. T. E. (1998). Improving student learning using the personalized system of instruction. *Higher Education, 35,* 187-203.

Hobbs, S. H. (1987). PSI: Use, misuse, and abuse. *Teaching of Psychology, 14*(2), 106-107.

Hosie, P., Schibeci, R., & Backhaus, A. (2005). A framework and checklists for evaluating online learning in higher education. *Assessment & Evaluation in Higher Education, 30*(5), 539-553.

Hughes, G. (2007). Using blended learning to increase learner support and improve retention. *Teaching in Higher Education. 12*(3), pp. 349-363

Ironsmith, M., & Eppler, M. A. (2007). Mastery learning benefits low-aptitude students. *Teaching of Psychology, 34*(1), 28-31.

Jeris, L. & Poppie, A. (2002). Screen to screen: A study of designer/instructor beliefs and actions in internet-based courses. Adult and Education Research Conference, Raleigh, N.C., May 24-26, 2002

Kaczynski, D., Wood, L., & Harding, A. (2008). Using radar charts with qualitative Evaluation: Techniques to assess change in blended learning. *Active Learning in Higher Education, l9*(1), 23–41

Keller, F. S. (1968). "Good-bye teacher ..." *Jounal of Applied Behavior Analysis, 1(Spring)*, 79-89.

Keller, J. M. (1999). Motivation in cyber learning environments, *International Journal of Educational Technology, 1* 7-30.

Kim, K. & Bonk, C. J. (2006). The future of online teaching and learning in higher education: The survey says….*Educause Quarterly 29(4)*. Retrieved February 26, 2007 from http://www.educause.edu/apps/eq/eqm06/eqm0644?bhcp=1

King, A. (1995). Designing the instructional process to enhance critical thinking across the curriculum. *Teaching of Psychology, 22*(1), 13-17.

Kulik, J. A.; Kulik, C-L. (1987). Review of recent research literature on computer-based instruction. *Contemporary Educational Psychology, 12*(3), 222-230

Kulik, C.-L., Kulik, J. A., & Bangert-Drowns, R. L. (1990). Effectiveness of mastery learning programs: A meta analysis. *Review of Educational Research, 60*(2), 265-299.

Kulik, J. A., Kulik, C-L., & Cohen, P. (1979). A meta-analysis of outcome studies of Keller's personalized system of instruction. *American Psychologist, 34*(4), 307-318.

Kosslyn, S. M., & Rosenberg, R. S., (2006). *Psychology In Context,* (3rd ed.). Boston, MA: Allyn & Bacon

Lee, Y.-F., & Nguyen, H. (2007). Get your degree from an educational ATM: An empirical study in online education. *International Journal on E-Learning, 6*(1), 31-40.

Levy, S. (2004). Six factors to consider when planning online distance learning programs in higher education. Retrieved December 23, 2004, from *http://www.westga.edu/~distance/ojdla/spring61/levy61.htm*

Lim, C. P., & Chai, . S. (2008). Rethinking classroom-oriented instructional development models to mediate instructional planning in technology-enhanced learning environments, *Teaching and Teacher Education. 24*, 2002-2013.

Lim, D. H, Morris, M. L, Kupritz, V. W. (2006). Online vs. blended learning: Differences in instructional outcomes and learner satisfaction. *Online Submission; Paper presented at the Academy of Human Resource Development International Conference (AHRD) (Columbus, OH, Feb 22-26, 2006) p809-816 (Symp. 39-1) Retrieved April 28, 2008 from http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/1b/df/c3.pdf*

Lloyd, M. E., & Lloyd, K. E. (1986). Has lightning struck twice? Use of PSI in college classrooms. *Teaching of Psychology, 13*(3), 149-151.

Ma, Y., & Runyon, L. R. (2004). Academic synergy in the age of technology-a new instructional paradigm. *Journal of Education for Business, July/August*, 367-371.

Mackey, P. T., & Ho, J. (2008). Exploring the relationships between Web usability and students'

perceived learning in Web-based multimedia (WBMM) tutorials. *Computers & Education 50*, 386–409.

Manton. E., Turner, C., & English, D. (1999). Testing the level of student knowledge. *Education*, 124(4), 682-687.

Martin, T. L., Pear, J. J., & Martin, G. L. (2002). Analysis of proctor marking accuracy in a computer-aided personalized system of instruction course. *Journal of Applied Behavior Analysis, 35*(3), 309-312

Matheos, K., Daniel, K., McCalla, G. I. (2005). Dimensions for blended learning technology: learners' perspectives. *Journal of Learning Design*, 1(1), 56-76.

McWhorter, K. T. (2006). *Study and Critical Thinking Skills in College*, (Sixth ed). New York, New York: Longman

Mehlenbacher, B., Miller, C. R., Covington, D., & Larson, J. S. (2000). Active and interactive learning online: A comparison of web-based and conventional writing classes. *IEEE Transactions On Professional Communication, 43*(2), 166-184.

Merisotis, J. P. (2001). Quality and equality in internet-based higher education: Benchmarks for success. *Higher Education in Europe, 26*(4), 489-497.

Merisotis, J. P., & Phipps, R. A. (1999). What's the difference? Outcomes of distance vs. traditional classroom-based learning. *Change, May/June*, 13-17.

Minke, A. (1997) Conducting Repeated Measures Analyses: Experimental Design Considerations. Paper presented at the annual meeting of the Southwest Educational

Research Association, Austin. Retrieved January 20, 1997 from http://ericae.net/ft/tamu/Rm.htm on April 25

Moallem, M. (2001). Applying constructivist and objective learning theories in the design of a web-based course: Implications for practice. *Educational Technology & Society, 4*(3), 113-125.

Moore, G. S., Perlow, A., Judge, C., & Koh, H. (2006). Using blended learning in training the public health workforce in emergency preparedness. *Public Health Report, 121*(2), 217-221

Morss, D. A. (1999). A study of student perspectives on web-based learning: WebCT in the classroom. *Internet Research: Electronic Networking Applications and Policy, 9*(5), 393-408.

Mortera-Gutierrez, F. (2006). Faculty best practices using blended learning in e-learning and face-to-face instruction. *International Journal on E-Learning, 5*(3), 313-337.

Mueller, J., Wood, E., Willoughby, t., Ross, C., & Specht, J. (2008). Identifying discriminating variables between teachers who fully integrate computers and teachers with limited integration. *Computers & Education, 51,* 1523-1537

Olson, T. M., & Wisher, R. A. (2002). The effectiveness of web-based instruction: An initial inquiry. Retrieved January 23, 2004, from "http://www.irrodl.org/content/v3.2/olson.html

Osguthorpe, R. T., & Graham, C. R. (2003). Blended learning environments: Definitions and directions. *The Quarterly Review of Distance Education, 4*(3), 227-233.

Pahinis, K. P., Stokes, C. W., Walsh, T. F., & Cannavina, G. (2007). Evaluating a blended-learning course taught to different groups of learners in a dental school. *Journal of Dental Education, 71*(2), 269-278.

Pahinis, K. P., Stokes, C. W., Walsh, T. F., & Cannavina, G. (2008). A blended learning course taught to different groups of learners in a dental school: Follow-up evaluation. *Journal of Dental Education, 72*(9), 1048-1057.

Pascarella, E. T., & Terenzini, P. T. (1991). *How college affects student: Findings and insights from 20 years of research.* San Francisco: Jossey-Bass.

Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students. Volume 2: A third decade of research.* San Francisco: Jossey-Bass.

Pear, J. J. (2002). Teaching and researching higher-order thinking in a virtual environment. In J. A. Chambers (Ed.), *Selected papers from the 13th International Conference on College Teaching and Learning* (pp. 143-150).Jacksonville, FL: Florida Community College at Jacksonville.

Pear, J. J., & Crone-Todd, D. E. (1999). Personalized system of instruction in cyberspace. *Journal of Applied Behavior Analysis*, *32*, 205-209.

Pear, J. J., & Crone-Todd, D. E. (2002). A social constructivist approach to computer-mediated instruction. *Computers and Education, 38*(1), 221-231.

Pear, J.J., Crone-Todd, D.E., Wirth, K., & Simister, H. (2001). Assessment of thinking levels in students' answers. *Academic Exchange Quarterly*, *5*(4), 94-98.

Pear, J. J., & Kinsner, W. (1988). Computer-aided personalized system of instruction: An effective and economical method for short-and long-distance education. *Machine-Mediated Learning, 2*, 213-237.

Pear, J. J., & Martin, T. L. (2004). Making the most of PSI with computer technology. In D. Moran J., & Malott, R. W. (n.d) *Evidence-based educational methods* (pp. 223-243). San Diego, California: Elsevier & Academic Press.

Pear, J. J., & Novak, M. (1996). Computer-aided personalized system of instruction: A program evaluation. *Teaching of Psychology, 23*, 119-123.

Pereira, J. A., Pleguezuelos, A. M., Molina-Ros, A., Molina-Tomas, M. C., & Masdeu, C., (2007). Effectiveness of using blended learning strategies for teaching and learning human anatomy. *Medical Education. 41*, 189-195.

Power, M. (2008). The emergence of a blended online learning environment. *MERLOT Journal of Online Learning and Teaching, 4*(4), 503-514.

Price, R. V. (1999). Designing a college based course using a modified PSI model. *Tech Trends. 43*(5), 23-28.

Quitadamo, I. J., & Brown, A. (2001). *Effective teaching styles and instructional design for online learning environments.* Paper presented at the National Educational Computing Conference, Chicago, Illinois.

Reboy, L. M., & Semb, G. B. (1991). PSI and critical thinking: Compatibility or irreconcilable differences. *Teaching of Psychology, 18*(4), 212-215.

Regehr, K. (2003). A new measure of critical thinking: The construct validity of the renaud critical thinking measure. Unpublished Honour's Thesis, University of Manitoba, Winnipeg

Renaud, R. D., & Mandzuk, D. (2006). Case studies in teacher education: Do they really enhance critical thinking. *Paper presented at the Canadian Society for the Study of Education*, Toronto, Ontario.

Robertson, C. (2006) Development and transfer of higher order thinking skills in pilots. *Humanities and Social Sciences*, 66(8-A), 2896

Robertson, M. (2002). Quality and university teaching: Juggling competing agenda. *Quality in Higher Education*, 8(3), 273-286.

Ross, L. L., & McBean, D. (1995). A comparison of pacing contingencies in classes using a personalized system of instruction. *Journal of Applied Behavior Analysis*, *28*, 87-88.

Rovai, A. P. (2002). Sense of community, perceived cognitive learning, and persistence in asynchronous learning networks. *The Internet and Higher Education, 5*(4), 319-332.

Salinas, M. F. (2008). From Dewey to Gates: A model to integrate psycoeducational principles in the selection and use of instructional technology. *Computers & Education, 50*, 652-660.

Salter, D., Richards, L., & Carey, T. (2004). The 'T5' model: An instructional model and learning environment to support the integration of online and campus-based courses. *Educational Media International*, Retrieved September 9, 2007 from http://www.tandf.co.uk/journals

Saville, B. K., Zinn, T. E., Neef, N. A., Norman, R. V., & Ferreri, S. J. (2006). A comparison of interteaching and lecture in the college classroom. *Journal of Applied Behavior Analysis, 39*(1), 49-61.

Saville, B. K., & Elliott, M. P. (2005). Interteaching versus traditional methods of instruction: A preliminary analysis. *Teaching of Psychology, 32*(3), 161-163.

Schmitt, D. R. (1998). The introductory class in higher education: Some old problems and new considerations. *The Behavior Analyst, 21*(2), 281-287.

Scriven, M., & Paul, R. (2004). *Defining critical thinking*. Retrieved August 9, 2006, from http://www.criticalthinking.org/aboutCT/definingCT.shtml

Selwyn, N. (2008). An investigation of differences in undergraduates' academic use of the internet, *Active Learning in Higher Education, 9*(11), 50-61

Sherman, J. G. (1992). Reflections on PSI: Good news and bad. *Journal of Applied Behavior Analysis, 25*(1), 59-64.

Skill, T. D., & Young, B. A. (2002). Embracing the hybrid model: Working at the intersections of virtual and physical learning spaces. *New Directions For Teaching and Learning, 92*, 23-32.

Skinner, B. F. (1958). Teaching Machines. *Science*, *129*, 969–977.

Skinner, B. F. (1968). *The Technology of Teaching*. New York: Appleton-Century-Crofts.

Skinner, M. (1990). The effects of computer-based instruction on the achievement of college students as a function of achievement status and mode of presentation. *Computers in Human Behavior, 6*, 351-360.

So, H. (2009) Is blended learning a viable option in public health education? A case study of student satisfaction with a blended graduate course. *Journal of Public Health Management Practice*, *15*(1), 59-66.

So, H., & Brush, T. A. (2008). Student perceptions of collaborative learning, social presence and satisfaction in a blended learning environment: Relationships and critical factors. *Computers & Education 51*, 318–336.

Spence, L. D. (2001). The case against teaching. *Change, November/December*, 11-19.

Springer, C. R., & Pear, J. J. (2008). Performance measures in courses using computer-aided personlaized system of instruction. *Computers & Education, 51*, 829-835.

Stoney, S., & Oliver, R. (1999). *Can higher order thinking and cognitive engagement be enhanced with multimedia?* Retrieved July 8, 2003, from http://imej.wfu.edu/articles/1999/2/07/index.asp

Svinicki, M. D. (2007). Moving beyond "It worked": The ongoing evolution of research on problem-based learning in medical education. *Educational Psychology Review, 19*, 49–61.

Svinicki, M. D. (1999). New directions in learning and motivation. *New Directions for Teaching and Learning, 80*, 5-27.

Szulewski, A. & Davidson L. K. (2008). Enriching the clerkship curriculum with blended e-learning, *Medical Education. 42*(11), 1114–1114.

Tabachnick, B. G., & Fidell, L. S., (2007). *Using Multivariate Statistics*, (Fifth ed). Boston, MA: Allyn and Bacon

Terenzini, P. T. (1999). Research and practice in undergraduate education: And never the twain shall meet. *Higher Education, 38*, 33-48.

Terenzini, P. T., Springer, L., Pascarella, E. T., & Nora, A. (1995). Influences affecting the development of students' critical thinking skills. *Research in Higher Education, 36*(1) 23-39.

Theall, M. (1999). New directions for theory and research on teaching: A review of the past twenty years. *New Directions for Teaching and Learning, 80*, 29-52.

Vaughn, N. (2007). Perspectives on blended learning in higher education. *International Journal on E-Learning, 6*(1), 81-94.

Vaughn, N. & Garrison, D. R. (2005). Creating cognitive presence in a blended faculty development community. *Internet and Higher Education,* 8, 1-12

Waschull, S. B. (2001). The online delivery of psychology courses: Attrition, performance, and evaluation. *Computers in Teaching, 28*(2), 143-147.

Weis, P., & Guyton-Simmons, J. (1998). A computer simulation for teaching critical thinking skills. *Nurse Educator, 23*(2), 30-33.

Welker, J., & Berardino, L. (2006). Blended learning: Understanding the middle ground between traditional classroom and fully online instruction. *Journal of Educational Technology Systems, 34*(1), 33-55.

Williams, R. L. (1999). Operational definitions and assessment of higher-order cognitive constructs. *Educational Psychology Review, 11*(4), 411-427.

Wingard, R. G. (2004). Classroom teaching changes in web-enhanced courses: A multi-institutional study. *Educause Quarterly, 27*(1), 1-14.

Appendix A – Final Exam – Introduction to University

THE UNIVERSITY OF MANITOBA
FINAL EXAMINATION - 099:111W *INTRODUCTION TO UNIVERSITY*
06R – Term 2

April 19, 2006
Location: 210-224 University Centre

Time: 2 hours

This exam consists of three parts. The weighting of each question within the examination is indicated. Ensure that you read all questions carefully. Some questions offer you a choice. Allocate your time, taking account of the weighting. Be sure to write your name, student number, instructor's name and section number on your answer booklet. Answer all items in the answer booklet. Ensure that your responses are all double-spaced. Pay attention to form as well as spelling, grammar, punctuation, etc.

********Answers to all questions must be written in INK*******

EXAM TOTAL --100 Points

I. IDENTIFICATION (30 Points).

Define 15 of the following terms and explain the significance of each.

analytical lecture format
marginal annotations
active learning
APA
revising/editing
Cornell method
course syllabus
distributed practice
selective attention
experiential learning
goal setting
self-monitoring
study group
metacognition
performance goal
retrieval
SEND
spatial learners
SQ3R
stress
cyber plagiarism

II. APPLIED QUESTION (30 Points). ***Answer A or B.***

A. Bloom's Taxonomy:

i. Using Bloom's/Anderson's model, identify and explain the six levels of thinking (12 points).

ii. Write a **minimum** of three questions or statements reflecting **each** of the six levels of thinking on **ONE** of the following topics (total of 18 questions or statements; 3 per level = 18 points). *(If you write questions you do not need to answer them.)*

1. Writing a research paper
2. Choosing a field of study
3. Taking notes from a chapter in our textbook
4. Orienting to university
5. Studying for a final exam

-OR-

B. Problem Solving:

Chris is a University 1 student and likes to read, follows written instructions and rules, works well with others and has good time management skills. Chris realizes that he/she is spending the money saved from working last summer more quickly than originally anticipated.

Based on the problem solving model identified in your textbook, discuss the steps Chris needs to take to decide whether or not to take on the part time job available in the local bookstore which would require working 20 hours/week.

III. CRITICAL THINKING ESSAY (40 Points)

*Choose **ONE** of the following for the introduction of your formal academic discourse. Using the standard essay format, clarify the issue and thesis, suggest one or two counterarguments (i.e., antithesis) and then make your case or present your argument (i.e., synthesis or conclusion). The essay should be 300-500 words in length.*

a. Computers have simplified our lives.
b. Universities prepare their graduates for their chosen careers.
c. Every successful student possesses good critical thinking skills.

Appendix B - Marking Key Final Exam 2006 – Fall Term

ARTS 1110 *INTRODUCTION TO UNIVERSITY*
Marking Key Final Exam 2006 – Fall Term

I. IDENTIFICATION (30 marks).

*Define **15** of the following terms and explain the significance of each. If a student supplies answers for more than fifteen of the terms, only the first fifteen will be marked. Each answer is worth 2 marks, with 1 mark awarded for definition and 1 mark awarded for the explanation of the term's significance or relevance to the first year student.*

APA – APA stands for the American Psychological Association. It is the writing protocol used by Social Sciences which dictates a standard format for citation of references, approach punctuation and grammar, and document layout. In APA, sources are noted in text with the year of publication (e.g. O'Brien-Moran 2006) and at the end of the paper on a Reference page. It is important that students understand that they are required to conform to the specific conventions of the respective discipline (eg. APA for Social Science, MLA for Literature, Chicago for History, etc)

Abstract – An abstract is a brief (100 - 200 word) summary of a written paper. It identifies the paper's key points and enables readers to decide quickly if it will serve their needs. Students may find it useful to review an abstract to determine whether it is relevant to their research.

Active learning - Active learning is a pro-active and inter-active approach to learning which includes thinking critically about new information, considering the purpose behind the information one learns, and making connections to those things one already knows.

In engaging in active learning, students bring higher order thinking to bear on any particular subject, which, in turn, facilitates greater understanding and greater recall. In contrast to passive learners, active learners take charge of their learning. Active learners:

- participate in class lectures;
- decide what is important to write down;
- ask questions to connect ideas;
- make outlines and study sheets;
- predict exam questions;
- look for trends & patterns;

- discover the significance of an assignment;
- looks for principles & concepts;
- try to expand their knowledge and experience with a topic, and try to connect it with course objectives or content.

Annotated Bibliography – An annotated bibliography is a list of cited references which are formatted in an appropriate style, with brief summaries of each of the itemized references. An annotated bibliography is both a useful record of one's research and a reference when writing a paper. The annotated bibliography can also be used purely as a literature review.

Cornell Method – The Cornell method is an effective three-stage system of note-taking that uses cue columns and summaries to facilitate comprehension and enhance the transfer of information to long-term memory. Successful students often find that this method improves their learning efficiency.

Course Syllabus - A course syllabus is the outline of all the requirements of a course. Given out to students in the first week of class, it details the material to be covered in a course, the required texts, additional readings, assignments & when they are due, the dates for test/exams, the value of all component of the course, any additional information about the course the instructors wishes to convey to the students, and a caution on the regulations concerning plagiarism and cheating. It is important for students to follow the syllabus carefully, noting any changes made during the course. It will help them to use time management for assignments, reading, and test or exams.

Critical Thinking – Critical Thinking is the mode of investigation in which an individual applies logical process and relevant evidence in the systematic exploration of a well-defined question. Successful students understand that they are expected to employ critical thinking skills in most university assignments. The exception to this rule may be assignments which require purely subjective responses.

Deductively Valid – Deductive validity refers to a relationship between premises and conclusion. A deductively valid argument is one in which true premises inevitably lead to a true conclusion. However, because an argument can be valid without being sound, false premises can lead to either a true or a false conclusion. Students should know that validity does not guarantee truth.

Distributed Learning – In distributed learning, the subject matter is divided into manageable parts and studied over time. This method of studying tends to be more effective for three reasons: 1) it prevents fatigue & keeps one working at peak efficiency;

2) since the mind keeps working on something once studying stops, the after effect or "thinking-on"occurs several times rather than

just once when studying over several blocks; 3) distributing the material over several sessions allows one to approach it in reasonable pieces that can be mastered more easily.

Encoding – Encoding is the first stage in the process of memory. It involves the process of rendering sensual data meaningful. It is important that students understand that, at this stage, memory is relatively unstable. Information that is not consolidated will not be recorded in long-term memory.

GPA – GPA is an acronym that stands for Grade Point Average. It is a point system used to calculate a student's overall standing and is based on the average of final marks. While all students need a minimum GPA of 2.00 to graduate from the University of Manitoba, the GPA required to enter specific Faculties is often much higher.

Goal Setting – Goal setting is a systematic strategy of identifying one's life targets and the steps necessary to reach those targets. Goals are not simply dreams or wishes. Goals should be achievable, realistic, time related (deadline/timeframe), measurable, documented, positive, worthwhile, and flexible.

Marginal Notation – Marginal notation is a note, word, symbol in the marginal of a text or set of notes that indicates the importance of material to be learned. It could be an arrow, line, star, word like "def" for definition, "ex" for example or "3 steps" for steps in a process etc. Active learners use this method so that they can go back to these pages and quickly review important points to be learned.

Metacognition - Metacognition is, most simply, "thinking about thinking." Metacognition is higher order thinking which allows students to analyze their own learning practices and adapt their learning strategies to maximize the chances for success.

Plagiarism - Plagiarism is the unacknowledged use of another's words or ideas. The failure to cite sources, regardless of intent, is a serious academic offence the consequences of which can include failure in the specific course or suspension from the university.

Retrieval – Retrieval relates to the third major stage in the Model of Memory. A student's ability to retrieve information is the true test of how well they have stored it. To improve retrieval rates, students can use a number of strategies: visualization, retrieval clues, rehearsal, and/or learning beyond mastery.

Retroactive Interference – Retroactive interference refers to the phenomena in new learning interferes with the recall of existing knowledge. In order to overcome this type of interference, students should continue to review previously learned material while they are acquiring new material.

SQ3R – SQ3R is an effective method of academic reading which increases the reading rate and improves retention. The acronym stands for: S= Survey, Q= Question, R= Read, R= Recite, R= Review. SQ3R is a strategy that promotes active learning.

Stress – Successful students are able to understand the causes and symptoms of stress, and take steps to reduce unmanageable stress in their lives. Stress builds up when people no longer feel in control of situations. Physical signs of stress include headaches, insomnia, irritability, and stomach aches. Some of the ways a student can reduce stress are: develop a healthy diet; maintain appropriate amounts of sleep; reduce employment hours; seek counseling; maintain an exercise regime; and practice good time management.

Student Advocacy – Student Advocacy is a university department that assists students in understanding their rights, and responsibilities. It provides guidance and support to students who require assistance with resolving problems related to academic/disciplinary decisions. At the University of Manitoba, Student Advocacy is part of the services offered by Student Affairs and is represented by an office on the 5th floor of the University Centre building. In addition the Student Advocacy office lobbies the appropriate University units for improvement in student matters and acts as an advocate for students on campus.

Thesis Statement – A thesis statement is a declarative statement that articulates and limits the direction and scope of an academic paper. It provides a student with focus and helps to guide the synthesis of many disparate ideas into a coherent argument. Without a strong thesis statement, papers tend to lack focus and purpose.

(Butterill, O'Brien-Moran, Siwak 2006)

II. (a) Problem Solving Rubric

Demonstrates an understanding of the problem-solving model:

1. Accurately and thoroughly identifies the various elements of the model: 1 2 3 4 5

Demonstrates an understanding of problem analysis by identifying the problem, identifying the problem's cause and identifying the problem's effect:

2. Accurately and thoroughly identifies the problem:
   1 2 3 4 5

3. Thoroughly defines the problem's cause and effect:
   1 2 3 4 5

Arrives at logical conclusions.

4. Identifies three possible solutions:
   1 2 3 4 5

5. Identifies advantages and disadvantages of each possible solution:
   1 2 3 4 5

6. Develops an accurate solution appropriate to all elements of the problem: 1 2 3 4 5

(O'Brien-Moran 2006)

II. (b) Bloom's Taxonomy

Demonstrates an understanding of the Bloom's Taxonomy:

1. Accurately identifies the various elements of the model:
   1 2 3 4 5 6

2. Accurately represents the types of thinking associated with each level. 1 2 3 4 5 6

Applies knowledge to learning environment:

2. Appropriately applies knowledge to specific field of study:
   1 2 3 4 5 6

3. Develops questions that are reflective of each of thinking within the context of the academic discipline (each question is worth two marks; to receive two marks, questions must reflect understanding of the sort of knowledge expected at the respective level of thinking):

   1 2 3 4 5 6 7 8 9
   10 11 12

(O'Brien-Moran 2006)

III. General Analytic Writing Rubric

1. Introduction 1 2 3 4 5

   1 = no or poor introduction.
   3= some introduction; nothing beyond a forecast.
   5= introduction grasps reader's attention (engages the reader) and forecasts major points.

2. Articulation of thesis 1 2 3 4 5

   1 = no or poor articulation of thesis.
   3 = some articulation of thesis
   5 = clear articulation of thesis or argument.

3. Paragraph development 1 2 3 4 5

   1.= poor paragraphs with no clear topic sentence; multiple topics; little or no development
   3 = some structure and development of paragraphs and/or some with clear topic sentences or focus, but not consistently.
   5 = paragraphs are consistently well developed, with a clear topic sentence and appropriate number of sentences that provide examples and develop points.

4. Use of examples 1 2 3 4 5

   1 = little or no use of examples.
   3= some use of examples or evidence, but not consistent; no examples or evidence in places where they are needed
   5 = frequent or consistent use of examples and evidence

5. Transitions 1 2 3 4 5

   1 = little or no transition between paragraphs; poor flow
   3 = some transition or flow between paragraphs; partial structure to argument
   5 = strong and/or consistent transition between points in essay; strong flow

6. Reasoning 1 2 3 4 5

   1 = no obvious rationale for position; serious flaws in logic.

3 = some evidence of reasoning; demonstrates ability to justify key components of a chosen position and/or solution

5 = reasoning is sound; evidence that contrary position has been considered fairly.

7. Coherence 1 2 3 4 5

1 = lack of coherence; i.e. mismatch between the thesis and the body; tangents
3 = occasional tangents; repetition
5 = every paragraph works to support the thesis; "linked" paragraphs

8. Conclusion 1 2 3 4 5

1 = no or poor conclusion or summary of argument
3 = some summary of points made, but nothing beyond summary; no broad conclusions/lessons
5 = a conclusion going beyond summary of what was written in the body of the essay.

Adapted from Dawn M. Zimmaro, Ph.D., Developing Grading Rubrics

Appendix C - Applied Critical Thinking Measure

Student Number: _______________

Applied Critical Thinking Measure
(99.111 06-07)

The purpose of this measure is to determine the degree to which your responses to various events reflect critical thinking. Critical thinking is defined, briefly, as the process of making a justified decision that involves (1) identifying the objective, and (2) either drawing upon relevant existing information to make a justified decision, or specifying what additional information would be needed to make a justified decision.

In this measure, you will be asked to respond to various events that most people are likely to be familiar with. More specifically, after reading the short description of a particular event, you will be asked to either (a) make a decision that can be justified on the basis of existing information (in the short description) and explain why your decision is most appropriate (i.e., justification), or (b) specify what information you would need in order to make a justified decision. To illustrate, a sample event is provided below:

Two candidates are competing for an upcoming election for the position of student council president in this university. Candidate A has served on the student council for three years and has outlined a firm plan to look at the most important issues that students are facing. Candidate B was the student council vice-president last year and appears to be very well liked by fellow students when delivering campaign speeches. Which candidate would you vote for?

(circle one) *CANDIDATE A* *CANDIDATE B* *NOT SURE*

In the boxes below, if you chose CANDIDATE A or CANDIDATE B, please give three distinct reasons why (one reason per box). If you answered NOT SURE, what would be the three most important and relevant things (one point per box) you want to know before deciding which candidate to vote for?

According to a recent report, in Canadian prisons, the sentences are too short and the time spent behind bars is too easy. In fact, a considerable number of those convicted of violent crimes are released before they have served even a third of their expected sentences.

Do you agree that the prison system in Canada is far too soft, especially with more serious offenders?

(circle one) **YES** **NO** **NOT SURE**

In the boxes below, if you answered YES or NO, please give three distinct reasons why (one reason per box). If you answered NOT SURE, what would be the three most important or relevant things (one point per box) you want to know before deciding whether or not the prison system in Canada is too soft?

To give students from low-income families the opportunity to attend better quality schools, the provincial government is considering the use of private school vouchers. These vouchers would be derived from taxpayer dollars that would pay for all or most of the cost for students from low-income families to attend private schools. With these vouchers, parents now have more choice over which school they would like their children to attend.

Do you feel these vouchers would help improve the overall level of educational achievement among all students from low-income families?

(circle one) **YES** **NO** **NOT SURE**

In the boxes below, if you answered YES or NO, please give three distinct reasons why (one reason per box). If you answered NOT SURE, what would be the three most important or relevant things (one point per box) you want to know before deciding whether or not these vouchers would be beneficial?

When we work out, our muscles are often working at high intensities, like football or basketball players. When our muscles are working, they're also creating what is known as lactic acid. As the amount of lactic acid in our bodies increase, our performance begins to decline. One type of drink that is becoming more known among athletes is a new oxygen-enriched water. Compared to regular water, the new oxygen-enriched water contains 20 times the normal amount of oxygen. Because oxygen breaks down lactic acid, this product is more in demand than ever.

At $3.25 per 500 ml bottle, would you buy this new oxygen-enriched water?

(circle one) **YES** **NO** **NOT SURE**

In the boxes below, if you answered YES or NO, please give three distinct reasons why (one reason per box). If you answered NOT SURE, what would be the three most important or relevant things (one point per box) you want to know before deciding whether or not to buy the oxygen-enriched water?

Recently, a leading medical journal published a report that found that talking on a cellphone while driving significantly increased a person's risk of an accident. As a result, the provincial government will have its members of the cabinet vote on whether or not talking on a cellphone while driving should be made illegal.

Do you agree that talking on a cellphone while driving should be made illegal?

(circle one) **YES** **NO** **NOT SURE**

In the boxes below, if you answered YES or NO, please give three distinct reasons why (one reason per box). If you answered NOT SURE, what would be the three most important or relevant things (one point per box) you want to know before deciding whether or not talking on a cellphone while driving should be made illegal?

According to a research firm in Vancouver, Ontario ranks dead last of all 10 provinces for spending on higher learning. This clearly shows a crisis in its ability to handle the projected increase in enrollment next year. With the largest university classes, second highest tuition fees, and the smallest operating grants in the country, Ontario has placed 10th in Canada for educational quality.

Based on this information, do you feel that the level of educational quality in Ontario universities is the lowest in Canada?

(circle one) **YES** **NO** **NOT SURE**

In the boxes below, if you answered YES or NO, please give three distinct reasons why (one reason per box). If you answered NOT SURE, what would be the three most important or relevant things (one point per box) you want to know before concluding whether or not the level of educational quality in Ontario universities is the lowest in Canada?

Last winter, many regions in Canada have endured record high temperatures. For example, during the month of January, the temperature went up to 18 degrees (Celsius) in some places. While much has been discussed about the effects of global warming, researchers have concluded that the record high temperatures that have occurred during this winter clearly indicate that global warming is definitely happening.

Do you feel that this information confirms or proves that global warming is definitely happening?

(circle one) **YES** **NO** **NOT SURE**

In the boxes below, if you answered YES or NO, please give three distinct reasons why (one reason per box). If you answered NOT SURE, what would be the three most important or relevant things (one point per box) you want to know before concluding whether or not global warming is actually happening?

Appendix D - Scoring Key for the Applied Critical Thinking Measure

Applied Critical Thinking Measure
Scoring Key 99.111 (August 2006)

Critical thinking can be defined as the process that leads one to making a justified choice. In order to make a justified choice, one must be able to (1) identify the objective and (2) either draw upon existing information or obtain new information that is relevant to the objective. The score of each question is determined by the number of points that are relevant to the objective.

Within each question, each statement or question should be compared to the main objective of the situation. For example, in the PowerFit question, the objective is to determine whether or not the device actually develops or strengthens muscles. Basically, each point either (1) refers directly, (2) refers vaguely or indirectly, or (3) is irrelevant.

So far, it appears that most students have little difficulty in identifying the main objective/question. This indicates that each situation is familiar to most students. The main source of variability in responses seems to be in the degree of justification or questioning that would enable justification.

Rather than asking students to simply list as many statements or questions as they can think of, students are asked to list three with one distinct statement/question per box. This should help elicit critical thinking even more closely as it asks students to evaluate and choose only the best reasons and enable scoring to be more valid as the number of distinct statements/questions will be fewer and in a clearer format (i.e., in the boxes).

Each question is designed to portray a product or a situation that most university students would likely be familiar with and to encourage the student to respond quickly with either a YES, NO or NOT SURE answer. In addition, each question is either worded quite vaguely or is missing vital information.

Therefore, a student is most likely to demonstrate his/her critical thinking skill not by quickly accepting or rejecting what is proposed in the question, but rather by asking questions that would clarify his/her understanding of the product or situation so that a more informed choice can be made. **In most cases, a student will obtain a higher score in a particular question by either selecting NOT SURE and asking three relevant and specific questions, or outlining three specific flaws or missing pieces of information to support a NO answer. Occasionally, there may be an exception where a student may select YES and has provided up to three relevant and specific points to support that choice.**

Scoring key for YES, NO and NOT SURE answers

YES

1. "This is beneficial..." - "this is important..." - "this is needed..." - "I agree..." (applies either to self or others)
    a. Reason is related to the objective (assumption, prediction, better than _____) **1 point**
    b. No additional reason provided - reason simply reiterates information in the question - nothing beyond "this is important..." **0 points**

2. Asks or implies a question
    a. Relevant and specific (i.e., requires a specific answer other than *yes* or *no)* **1 point**
    b. Relevant and vague (i.e., can be answered with either *yes* or *no*) - "need to know more..." **0 points**

3. "It is inexpensive..." - "nothing to lose..." - "easy to try..." **0 points**

4. "I have seen/heard of this (or something similar to this) before..." - "I have experienced this before..." - "everyone else has (is doing) it..." **0 points**

5. Indicates a NO response **0 points**

6. Unclear or irrelevant **0 points**
    a. Not clear if answer indicates either YES, NO, or NOT SURE
    b. Not clear how statement relates to the objective

NO

1. "Not interested..." - "I don't believe it:.." - "I disagree..." - "I don't like it..."
   a. Provides either possible reason or a seemingly feasible alternative **1 point**
   b. No reason or alternative provided, nothing beyond a negative reply **0 points**

2. "Does not apply to me..." - "Not my problem..." - (could be interested or agree if perceived as applicable)
   a. If person could easily avoid situation (e.g., buying a product) **do not score**
   b. If person may not be able to easily avoid situation (e.g., neighbourhood crime) **0 points**

3. "Not worth it..." - "Too expensive..." (could be interested, but perceived benefit is insufficient)
   a. Provides either possible reason or a seemingly feasible alternative **1 point**
   b. No reason or alternative provided, nothing beyond a negative reply **0 points**

4. "Not enough specific information or evidence to be convinced..." - "would need to know _____" - "we don't know _____"
   a. Indicates what specific information or evidence would be needed (i.e., requires a specific answer other than *yes* or *no)* - implies "NOT SURE" type question **2 points**
   b. Refers to needed information or evidence that is more vague (i.e., can be answered with either *yes* or *no*) - implies "NOT SURE" type question **1 point**
   c. Nothing beyond "not enough information", "need more evidence", etc. **0 points**

5. "This would be too difficult..." - "this cannot be done" (if possible, then might consider)
   a. Provides either possible reason or a seemingly feasible alternative **1 point**
   b. No reason or alternative provided, nothing beyond "too difficult...", "cannot be done..." **0 points**

6. "Probably won't work..." - "unlikely to be effective..." (although could be easy to try or might be interested)
   a. Provides either possible reason or a seemingly feasible alternative **1 point**
   b. No reason or alternative provided, nothing beyond "won't work..." **0 points**

7. Indicates a YES response **0 points**

8. Unclear or irrelevant **0 points**
   a. Not clear if answer indicates either YES, NO, or NOT SURE
   b. Not clear how statement relates to the objective

NOT SURE

1. Question or statement requires specific information other than *yes* or *no* and the question/statement is relevant to the objective in that the information could conceivably determine whether one would answer YES or NO **3 points**

2. Question requires a simple *yes* or *no* answer and the question is relevant in that the answer could conceivably determine whether one would answer YES or NO **2 points**

3. Question or statement that says nothing beyond something like "would want to know more..." **0 points**

4. Statement that indicates either YES or NO **0 points**

5. Unclear or irrelevant **0 points**
    a. Not clear if answer indicates either YES, NO, or NOT SURE
    b. Not clear how statement relates to the objective
    c. Asks about own values or feelings (e.g., "I don't know if I would want this.")

Appendix E - Detailed Description of Procedure for Experiment 1

The course in Experiment 1 contained 4 lecture sections and 32 lab sections.Two lecture sections were scheduled for an 8:30 to 9:45 time slot on Tuesdays and Thursdays, and two were scheduled for a 2:30 to 3:45 time slot on Tuesdays and Thursdays. The CAPSI sections were Thursday afternoon (Instructor 1) and Tuesday morning (Instructor 2), while the non-CAPSI sections were Thursday morning (Instructor 3) and Tuesday afternoon (Instructor 4).

After selecting one of the four lecture sections, students selected a lab section. Each student was given a list of eight lab sections associated with that particular lecture section and the student selected which one of the eight lab sections to be in, based on availability. Each lab section had a maximum enrolment of 20 students.

CAPSI Assignment

After registration and prior to the start of classes, all students registered in the 16 CAPSI lab sections were put into one CAPSI lecture section. As well, all the TAs were included as mentors – i.e., unit test markers - in that one CAPSI section. Therefore, TAs assigned to one of the 16 CAPSI labs assessed answers from students registered in all CAPSI lab sections. In addition, peer reviewers marked answers from students in all CAPSI lab sections.

To provide the TAs with a general overview of CAPSI and to help them in assessing answers and providing feedback, a demonstration of the CAPSI program was given prior to the start of classes. In addition to the initial training, weekly contact was maintained with the TAs either by email or in person in order to answer any concerns, maintain quality of responses, answer questions, provide general feedback and motivation, and to provide needed instruction.

Students in the CAPSI sections were able to take unit tests any time in the academic term from the second week of classes to the last day of classes. Students were able to access the CAPSI program through any computer connected to the Internet. Therefore, access location and time was at the students' choosing. The CAPSI program delivered and recorded student-requested unit tests, submitted answered unit tests to markers, returned the assessed answers back to the student, and managed student appeals.

The CAPSI assignment was designed for students to pass 20 unit tests and to peer-review 10 unit tests. When a student was prepared to write a test, he or she would access the program and click the test button. The student would then be presented with three questions pseudo-randomly selected by the CAPSI program from a database of questions for that unit. The student would then have one hour to complete the test and submit it for assessment.

The test would be sent to either a TA or to two peer reviewers, assessed, and returned to the student. If mastery was attained, the student progressed to the next unit. If the student received a "restudy" result, then he or she could either write a new test on the unit after one hour, or submit an appeal to the instructor.

The appeal contained the student's original answers, the assessment from the TA or the two peers, and the student's argument as to why the assessment was not accurate. If the instructor granted the appeal, the student could progress to the next unit. If the appeal was not granted, the student would need to restudy the chapter and write a new test for that chapter.

Questions for the unit tests were drawn from and designed to test students knowledge of the material in the course textbook (McWhorter, 2006). The study questions for each unit test were designed in accordance with the six levels of Bloom's taxonomy as modified by Crone-Todd, Pear, and Read (2000), Crone-Todd and Pear, (2001), and Pear, Crone-Todd, Wirth, and Simister (2001). Definitions and an example of a question for each level of the taxonomy that were used in the course are below:

Knowledge questions – the required answer is word-for-word or closely paraphrased from the study material.

"What is terminology? When, where, and why is terminology used?"

Comprehension questions – the required answer is in the student's own words.

"In your own words, describe the role of the professor and the purpose of class time in regard to taking responsibility for own learning."

Application questions – the required answer applies a concept to a problem or situation that student might experience.

"Describe the new demands University has placed on you and any plans you have made to deal with these demands."

Analysis questions – the required answer involves breaking down a concept into its individual parts (e.g., comparing and contrast two or more concepts).

"Compare the course outline from two different courses. How are they the same? How are they different?"

Synthesis questions – the required answer combines two or more concepts to form something new.

"What are the strengths and weaknesses of electronic sources? How do these relate to print resources?"

Evaluation questions – the required answer is a reasoned argument for or against a given position.

"How can a part-time job either help or hinder success in university?"

Once a student completed the first unit, he or she was qualified as a peer reviewer for that unit. This pattern repeated, such that once a student mastered any unit he or she could be a peer reviewer for that unit, or any preceding unit. Therefore, if a student had passed units 1 through 10, he or she could receive tests for any of those units.

Once a student completed the required 10 peer reviews, he or she could then go into the program and click on a button to opt out of receiving any more tests. This option was also available if students did not want to receive tests during busy times or when they know they would be unavailable.

Unit tests were first assigned to the TAs on a rotating basis from the list of TAs within the CAPSI program. This continued until two or more students were available to peer review a particular unit. When assigning peer reviewers, the CAPSI program selected two peers with the lowest number of peer-reviewing points. When two peer reviewers assessed the answers, both had to independently agree that the unit test was a pass in order for the program to record it as a pass.

Students were advised to check into the program at least once a day to see if they had been issued a test to assess. If a test appeared, they would have 24 hours from the time the test was submitted to assess the test. If they did not complete the assessment within 24 hours, they would have .5% deducted from their grade.

Students continued with this process of writing unit tests and conducting peer reviews until all units were mastered and peer reviews completed, or until the end of the term. If a student did not complete all units or peer reviews, then a percentage of the total mark for the assignment would be calculated and submitted for the final grade.

ACTM

After the first class and in the first week of labs, the TAs had students complete the Applied Critical Thinking Measurement (ACTM) (Appendix B). This same measure was assigned and completed in the last week of classes for later analysis of any changes in critical thinking skills over the term. Each ACTM assessment took approximately fifteen minutes to complete. Three critical thinking questions were assigned in the pre-test, while a different set of questions were assigned in the post-test.

Appendix F - Detailed Description Questions Answers For Experiment 2

Unit 1

In a paragraph of approximately 200 words, assess the following argument, identifying, specifically, the proposition being argued. In your assessment, determine whether or not there is sufficient reason to support the proposition, whether there are ambiguous terms that need to be clarified, and whether the argument relies upon value assumptions that have either not been articulated (i.e. had not been made clear) or have not been proven.

Euthanasia, assisted suicide or whatever you want to call it, I think it makes perfect sense.

If I'm terminally ill or become severely disabled, I want to have the choice to die when I want. I don't want to be fed with tubes or with a spoon and I don't want to wear diapers or have a bedpan. I don't want my family members to remember me as incapable of taking care of myself. I have no relationship with God and I don't want people to tell me when it's "my time." I know that better than anyone else. That's my wish and it should be my choice. I want to die in dignity, my way.

Unit 2

Review the following passage. In a paragraph of approximately 200 words, assess the argument, identifying again, specifically, the proposition being argued. In your assessment, determine whether or not the argument is structurally flawed and, if so, what fallacy is evident. If the argument employs more than one fallacy, identify the one you believe to be most significant. Explain why you believe the argument to be *fallacious (i.e. why are you not convinced that the reasons that have been offered do not support the conclusion).*

Don't Criticize RCMP

Re: To serve and Taser.

Mr. ---------, how dare you. As I sit here and read over your letter, I can't help but be appalled by the ignorance you've set forth on paper. I shouldn't even need to mention our brave RCMP officers who have recently perished in the line of duty, but it is clear I need to seeing as how you've obviously neglected to consider this. These RCMP officers are also citizens of this country and have a right to protect themselves from harm. If those four members of Canada's elite police service acted inappropriately, I have no doubt they will receive punitive measures as a result. To go as far as to criticize their training and refer to them as cartoon characters is downright disgraceful and you should be ashamed.

Unit 3

Review the following passage, identifying the arguable proposition. In a paragraph of approximately 200 words, assess the argument, determining the sort of evidence that is being provided and whether the evidence offered is sufficient to accept the proposition. If it is not, explain why it is not.

Immigrants to Canada complain about nothing. My parents were immigrants from England and they say that they had no problems adjusting. They say people were very supportive and welcoming. Therefore, all these people who claim that immigrants are suffering during their transition to a new country are wrong.

Unit 4

In a paragraph of approximately 200 words, assess the following argument, identifying, specifically, the proposition being argued. In your assessment, determine whether or not there is sufficient reason to support the proposition. In your answer, suggest one alternate explanation for the phenomena being studied.

Global warming is a result of human activity. Over the last 200 years, industrial activity and carbon emissions have increased dramatically, with the result that the average temperature in Canada has risen 1 degree Celsius. In northern Canada, the permanent ice has begun to recede and sea ice has begun to break up earlier in the spring. There must be a connection between these things.

Unit 5

In a paragraph of approximately 200 words, assess the following argument, identifying, specifically, the proposition being argued. In your assessment, determine whether or not the evidence that has been provided provides sufficient reason to accept the proposition. In your answer, identify whether the statistics that have been provided are flawed, and, if so, in what way.

The number of students cheating at university cannot be as high as people believe. Despite the estimate that approximately 30% of students engage in academically dishonest practices, a recent study, conducted in universities across the country, showed that only 4.1 % of students admitted ever cheating in assignments or tests.

Appendix G - Examples of Questions Used At Each Level in Experiment 3

Knowledge question –Define and provide the goals of structuralism.

Comprehension question –In your own words, describe "resting potential," and the process that occurs in order to change a neuron from its resting potential to its "action potential."

Application question –In your own words, describe Ebbinghaus's forgetting curve, encoding failure, and decay, in relation to forgetting information. How can this information help you in studying for your next psychology test?

Analysis question –In your own words, describe cognitive and latent learning, and why this type of learning is significantly different from either classical or operant conditioning. Provide your own examples to help illustrate your answer.

Synthesis question –In your own words, compare and contrast the four biologically based theories of personality.

Evaluation question –Is dissociative identity disorder is a verifiably distinct diagnosis? Use information from this chapter to support your answer.

Appendix H- Consent Form for Experiment 1

Consent Form

Research Project Title: Assessment and Comparison of Critical Thinking Development of Students in a First Year Course With a CAPSI Assignment

Researcher: Louis Svenningsen M.Sc.

Purpose

The purpose of the research is to analyze the effectiveness of a specific learning management system to increase understanding of course material, and in developing critical thinking skills. To do this, the research will use data from the Computer-Aided Personalized System of Instruction (CAPSI) assignment within the course, an APA styled paper assignment, scores on a critical thinking assessment tool, and scores on the final exam. For specific information on CAPSI, please refer to "http://www.webcapsi.com"

Procedures

Within the first two weeks of class, all students will be asked to complete the Applied Critical Thinking Measure (Renaud's Critical Thinking assessment tool). This same assessment will also be given in last week of classes. This is a paper and pen assessment to be completed within ten to fifteen minutes during class time.

During the semester, as part of the requirements for the course, students will do either the CAPSI assignment or to write an APA formatted academic research paper, no less than fifteen pages long. A full description of both assignments will be given on the first day of class, with further information provided as necessary, throughout the term.

Once the course is completed and final grades are submitted, scores from the Applied Critical Thinking Measure, marks from the assignments, and marks from final exam will be used to assess if any differences in critical thinking occurred between students choosing the CAPSI assignment or the paper assignment.

Risk

Participation in the research study will not effect, nor be effected by marks attained during the course. This includes marks for all assignments, and the final exam. All research associated with the obtained data will start once the course is completed and marks are submitted.

Confidentiality

Please note that course information (progression through the course, and grades) will be used once the course is complete. Personal information such as student number or name of individual students will not be linked to the material used for research, but is assigned a number generated through the CAPSI program, eliminating potential student confidentiality problems. Throughout the term, the CAPSI system logs transaction records of student behaviour and tests. In addition, only archived data (data used after the course is completed), will be used, further reducing the potential for confidentiality conflicts.

Feedback

Please check the box below and provide an email address to indicate interest in receiving information on this research.

____

Your signature on this form indicates that you have understood to your satisfaction the information regarding participation in the research project and agree to participate as a subject. In no way does this waive your legal rights nor release the researchers, sponsors, or involved institutions from their legal and professional responsibilities. You are free to withdraw from the study at any time, and /or refrain from answering any questions you prefer to omit, without prejudice or consequence. Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation.

For clarification and information throughout the study, please contact Louis Svenningsen by email at svenning@ms.umanitoba.ca" or by phone, at 474-7948

This research has been approved by the Psychology/Sociology Research Ethics Board. If you have any concerns or complaints about this project you may contact any of the above-named persons or the Human Ethics Secretariat at 474-7122, or e-mail margaret_bowman@umanitoba.ca. A copy of this consent form has been given to you to keep for your records and reference.

____

Participant's Signature Date

____

Researcher and/or Delegate's Signature Date

Appendix I- Consent Form for Experiment 2

Consent Form

Research Project Title: Assessment and Comparison of Critical Thinking Development of Students in a First Year Course With a CAPSI Assignment

Researcher: Louis Svenningsen M.Sc.

Purpose

The purpose of the research is to analyze the effectiveness of a specific learning management system in developing critical thinking skills. To do this, the research will use data from the Computer-Aided Personalized System of Instruction (CAPSI) assignment within the course, and scores on a critical thinking assessment tool. For specific information on CAPSI, please refer to http://www.webcapsi.com

Procedures

In the eighth week of classes, students will be asked to complete the Applied Critical Thinking Measure (Renaud's Critical Thinking assessment tool). This same assessment will also be given in last week of classes. This is a paper and pen assessment to be completed within ten to fifteen minutes during class time.

During the last third of the semester, as part of the requirements for the course, students will do either the CAPSI assignment or to write a 1,000 word APA formatted academic research paper. The option will be made available for students in all sections of the course except for the English Second Language section, in which case the CAPSI assignment is mandatory.

Once the course is completed and final grades are submitted, scores from the Applied Critical Thinking Measure, will be used to assess if any differences in critical thinking occurred between students doing the CAPSI assignment or the paper assignment.

Risk

Participation in the research study will not effect, nor be effected by marks attained during the course. This includes marks for all assignments, and the final exam. All research associated with the obtained data will start once the course is completed and marks are submitted.

Confidentiality

Please note that course information (progression through the course, and grades) will be used once the course is complete. Personal information such as student number or name of individual students will not be linked to the material used for research, but is assigned a number generated through the CAPSI program, eliminating potential student confidentiality problems. Throughout the term, the CAPSI system logs transaction records of student behaviour and tests. In addition, only archived data (data used after the course is completed), will be used, further reducing the potential for confidentiality conflicts.

Feedback

Please check the box below and provide an email address to indicate interest in receiving information on this research.

____

Your signature on this form indicates that you have understood to your satisfaction the information regarding participation in the research project and agree to participate as a subject. In no way does this waive your legal rights nor release the researchers, sponsors, or involved institutions from their legal and professional responsibilities. You are free to withdraw from the study at any time, and /or refrain from answering any questions you prefer to omit, without prejudice or consequence. Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation.

For clarification and information throughout the study, please contact Louis Svenningsen by email at svenning@ms.umanitoba.ca or by phone, at 474-7948

This research has been approved by the Psychology/Sociology Research Ethics Board. If you have any concerns or complaints about this project you may contact any of the above-named persons or the Human Ethics Secretariat at 474-7122, or e-mail margaret_bowman@umanitoba.ca. A copy of this consent form has been given to you to keep for your records and reference.

____

Participant's Signature Date

____

Researcher and/or Delegate's Signature Date

Appendix J- Consent Form for Experiment 3

Consent Form

Research Project Title: Assessment and Comparison of Critical Thinking Development of Students in a First Year Course With a CAPSI Assignment

Researcher: Louis Svenningsen M.Sc.

Purpose

The purpose of the research is to analyze the effectiveness of a specific learning management system to increase understanding of course material, and in developing critical thinking skills. To do this, the research will use data from the Computer-Aided Personalized System of Instruction (CAPSI) assignment within the course, scores on multiple-choice quizzes, and scores from four exams. For specific information on CAPSI, please refer to http://www.webcapsi.com

Procedures

During the semester, as part of the requirements for one section of the course, students will do either the CAPSI assignment or write 13 multiple-choice quizzes. Once the course is completed and final grades are submitted, scores from the CAPSI assignment, quizzes, and exams will be used to assess if any differences in test scores between students being assigned the CAPSI assignment or the quizzes.

Risk

Participation in the research study will not effect, nor be effected by marks attained during the course. This includes marks for all assignments, and the final exam. All research associated with the obtained data will start once the course is completed and marks are submitted.

Confidentiality

Please note that course information (progression through the course, and grades) will be used once the course is complete. Personal information such as student number or name of individual students will not be linked to the material used for research. Only archived data (data used after the course is completed), will be used, further reducing the potential for confidentiality conflicts.

Feedback

Please check the box below and provide an email address to indicate interest in receiving information on this research.

Your signature on this form indicates that you have understood to your satisfaction the information regarding participation in the research project and agree to participate as a subject. In no way does this waive your legal rights nor release the researchers, sponsors, or involved institutions from their legal and professional responsibilities. You are free to withdraw from the study at any time, and /or refrain from answering any questions you prefer to omit, without prejudice or consequence. Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation.

For clarification and information throughout the study, please contact Louis Svenningsen

This research has been approved by the Psychology/Sociology Research Ethics Board. If you have any concerns or complaints about this project you may contact any of the above-named persons or the Human Ethics Secretariat at 474-7122, or e-mail margaret_bowman@umanitoba.ca. A copy of this consent form has been given to you to keep for your records and reference.

____

Participant's Signature Date

____

Researcher and/or Delegate's Signature Date