University of Nevada, Reno

THE EFFECTS OF WEB-BASED SUPPLEMENTAL MODULES ON EXAM PERFORMANCE IN AN UNDERGRADUATE ECONOMICS STATISTICS COURSE

A thesis submitted in partial fulfillment of the
requirements for the Degree of Master of Arts in
Psychology

by

Thomas Wade Brown

Dr. Mark P. Alavosius/Thesis Advisor

May, 2010

University of Nevada, Reno
Statewide • Worldwide

THE GRADUATE SCHOOL

We recommend that the thesis
prepared under our supervision by

THOMAS WADE BROWN

entitled

The Effects Of Web-Based Supplemental Modules On Exam Performance in an Undergraduate Economics Statistics Course

be accepted in partial fulfillment of the
requirements for the degree of

MASTER OF ARTS

Mark P. Alavosius, Ph.D., Advisor

Ramona Houmanfar, Ph.D., Committee Member

Victoria Follette, Ph.D., Committee Member

William Cathey, Ph.D., Graduate School Representative

Marsha H. Read, Ph. D., Associate Dean, Graduate School

May, 2010

ABSTRACT

Behavior analysis has had great success in designing effective and efficient methods of instruction at multiple different levels. Despite large amounts of empirical evidence, these methods are not widely adopted in terms of general educational approaches. The present study describes the use of web-based supplemental modules in an undergraduate Economics course in Inferential Statistics and their effect on exam performance. Specifically, the present study asks what are the effects of installing supplements into a lecture-based course, and does attaching a point contingency directly impact contact with the supplements. While differentiation was observed in the two experimental conditions of this study, effects may have been diminished by general course design. Future areas of research are discussed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

More and more, university instructors are beginning to embrace supplemental materials, technological service programs, and outside tutoring as ways to compliment their classroom-based courses. Methods of teaching and definitions of effective instruction are incredibly variable, however, and often times, inefficient for producing learning by students. Teachers may design objectives that are too difficult and unrealistic for students to meet, given the wide range of personal histories that students bring to college courses (Fox, 2004). Student failure is commonly viewed as a behavior problem rather than issues with the instructional variations present (Michael, 1991). Instruction can be viewed as the application of technology derived from evidenced-based educational practice with careful design placed on the facilitated learning environment (Vargas & Vargas, 1991). In this regard, simple modifications in instructional design could potentially alter the outcome associated with a course and maximize student-passing rate.

Web-based instruction is an innovative and emerging method of instruction that promises efficiency in its application. The end of the 20th century was characterized by many exciting innovations regarding technology. The Internet became more widespread and efficiently used, especially when compared to adoption trends with older technologies such as television and radio (Naughton, 2000). In recent years, many organizations have begun using distance-learning techniques involving the Internet for training of employees (Goldstein & Ford, 2002). Such methods prove both effective and low cost in terms of achieving skilled workers and consultants. It is evident that such methods reach multiple viewers in multiple locations. This trend is also seen in academia. By the late 1990s, the United States and other developed countries witnessed a

surge in commerce of technology, further causing a majority of universities to adopt some form of a computer-based learning program to stay competitive in attracting students and generating support from granting agencies (Curran, 2008). Relatively few universities today do not initialize some distance education program, with the majority realizing the economical benefits of doing so. Enrollments into these courses have also increased over the last decade. For instance, Waits, Lewis & Greene (2003) report that 1.3 million students were enrolled in a distance education course between 1997 and 1998. However, the number of student enrollments in distance-based courses doubled in the following years to approximately 2.9 million between 2000 and 2001. Without question, distance education and web-based instruction are becoming an intriguing option in terms of higher education.

While the majority of college classes offered on a regular basis are done in traditional formats, many professors have begun to incorporate various streams of technology, in particular elements from web-based instruction, into their lectures as a technique to communicate more efficiently with their students. While technology is advancing everyday, little evaluation is given to these resources, as the advancement of technology moves rapidly. It is evident, especially when dealing with education, that the methods approached and used in teaching should be evaluated constantly. The present study evaluates the use of web-based course supplements on student learning within an undergraduate course on Inferential Statistics. The methods of evaluation are drawn heavily from Behavior Analysis (Skinner, 1938; 1958).

Behavior Analysis & Programmed Instruction

Behavior analysis defines behavior as a living organism's interaction with the environment and conducts functional analyses to further assess antecedent and consequential variables present in order to predict and, at times, control behavior (Skinner, 1938). The application of behavior analysis to social contingencies has had great success in multiple different environments and populations (Austin & Carr, 2000; Azrin, 1977; Baer, Wolf, & Risley, 1968). In particular, the use of behavior analysis in education demonstrates data-driven methods and results (Austin, 2000; Keller, 1968; Lindsley, 1991; Skinner, 1958; Vargas & Vargas, 1991). Despite decades of evidence confirming the use of behavior analytic methods in educational settings, most are not fully adopted or integrated within the normal educational curriculum (Austin, 2000; Boyce & Hineline, 2002; Sherman, 1992). Instead, most educators rely on their anecdotal assessments of instruction, without further consideration of the learning gains by students. Most course evaluations in universities do not progress beyond consumer satisfaction inventories. The application of well-tested behavioral methods, such as Programmed Instruction (Vargas & Vargas, 1991), Personalized System of Instruction (Keller, 1968), Precision Teaching (Lindsley, 1991), and Interteaching (Boyce & Hindline, 2002) could have significant consequences. The evidence supporting them includes data on learning gains that far surpass lecture type courses.

Programmed instruction refers to the foundation of behavioral education. Mainly, programmed instruction brings the instructor under the control of the student. The student's interaction and learning from the instructional materials guides and refines the instructor's teaching (Vargas & Vargas, 1991). Student success in this context is contingent upon the instructor's ability to design efficient and effective methods of

delivering the material within the course. Student failure then, is not a result of a lack of "understanding" or "motivation" by the student, but rather the ineffective methods the instructor used to teach the course (Keller, 1968; Vargas & Vargas, 1991). Typical models of programmed instruction have five main components, focusing on behavioral objectives, reinforcement, activity rate, successive approximations, and programming for mastery of the material.

Prior to the start of the course, behavioral objectives are set that state exactly what the student should be able to accomplish at the culmination of the course (Vargas & Vargas, 1991). Given that every student brings a wide variety of behavior into the instructional environment, every student must be handled on an individual level. Setting behavioral objectives allows a minimum mastery criterion to be set. In this regard, setting behavioral objectives are similar to setting goals in organizational settings (Fellner & Sulzer-Azaroff, 1984; Locke & Latham, 2002). In their review of goal setting and application, Locke & Latham (2002) summarized the basic foundations of setting goals. They point out that higher performance is correlated with more difficult explicit goals and that feedback helps the learners monitor their behavior and make adjustments when necessary. By making the goals explicit, ambiguity is reduced, and performance can be measured in terms of its desired outcome. Further, behavioral accounts of goals have identified them as discriminative stimuli, meaning that the goal sets the occasion for behavior to occur, further increasing the likelihood of goal directed behavior (Fellner & Sulzer-Azaroff, 1984). Such ideas also function within the educational environment. The instructor is responsible for designing clear, precise, and explicit objectives for the course as to communicate with the students what is expected of them.

As an organism behaves, changes in the environment occur. The changed environment produces consequential effects that potentially serve as a reinforcer for the behavior of an organism. A reinforcer is a stimulus presented following a particular response that further increases the likelihood of the response occurring again in the future (Skinner, 1938; Vargas & Vargas, 1991). This at times is a problem in traditional classrooms as most classroom teachers deliver reinforcers non-contingently, which causes difficulty in identifying the response that the teacher wishes to maintain at high levels. Teachers inadvertently reinforce problem behaviors of elementary students by providing attention, or professors could award homework points for incorrect responses. The most reinforcing consequence that can be delivered in a classroom for a student is getting the right answer (Michael, 1991; Vargas & Vargas, 1991). Activity rates of students are important and instructors should provide multiple different opportunities to respond to increase reinforcement of learning. Low activity rates, meaning few opportunities to actively respond, reduce opportunities to reinforce appropriate or desirable behaviors (Vargas & Vargas, 1991). Frequent testing, however, is labor intensive. Instructors provided with few resources may select lectures over testing and feedback. Students often fear testing and report a preference for lectures. In a context that uses customer satisfaction as an index of teacher effectiveness, lectures are embraced as the optimal teaching method. A review of the literature finds no data on learning outcomes to support this preference for lectures.

Programmed instruction reinforces successive approximations of mastering the course material, moving as rapidly as possible to help students meet the objectives set by the course instructor (Vargas & Vargas, 1991). In programmed instruction, the

instructor breaks units into small steps that further assist the student in completing the final objectives. The general assumption is that students enroll in a course to seek knowledge in an area that is unfamiliar to them. The materials presented at the front end of the course are shaping necessary skills that are required at the conclusion of the course. Giving multiple opportunities to respond sets the occasion for the instructor to reinforce successive approximations, which further assist in the mastery components built into a course.

In the purest form of Programmed Instruction, progression in the course only occurs when mastery is demonstrated (Keller, 1968; Vargas & Vargas, 1991). Frequent examinations are a metric used to demonstrate mastery throughout the course. In general, traditional instruction uses infrequent examinations and relies heavily on lectures to disseminate information. From a behavioral perspective, students are only given a small number of opportunities to demonstrate mastery, in addition to relying heavily on large quantities of lecture materials to supplement their learning environment. The lives of typical college students revolve around entertaining events and stimuli that compete for the student's time (Michael, 1991). Unfortunately, with this model, if students are not engaged or challenged on a regular basis, the distracting stimuli within the environment control the behavior of the student, and insufficient time is allocated to their studies.

Personalized System of Instruction

Keller (1968) articulated a completely mastery-based program in the Personalized System of Instruction, or PSI. PSI is a form of teaching that allows the student to take approximate steps towards the completion of a course. PSI has become the most researched and effective technique of instruction in Behavior Analysis (Austin,

2000). Despite documentation, PSI has not been adopted fully into our educational system, mainly because of resistance from the teaching community to reform their original methods (Sherman, 1992; Sulzer-Azaroff, 1986).

In an original PSI course, learning is self-paced, meaning that students advance only when they engaged with the material. There are no time limits for completions, and no contingencies are in effect for procrastination. Students who do not complete the course are awarded a grade of "Incomplete" until the class is completed. Keller also installed a mastery criterion necessary to advance within the course. In other words, students must perform at high levels before given materials for the following section. Lectures in this model only provide a forum to ask questions for clarification on material or motivate students to complete the course. Undergraduate proctors facilitate section quizzing and tutor the students via peer-to-peer coaching. In utilizing proctors, PSI provides for immediate feedback and reinforcement of successive approximations (Keller, 1968).

The success of PSI has been established through rigorous evaluations of student learning (Austin, 2000; Kulik, Kulik, & Cohen, 1979). Despite the empirical evidence that supports PSI, it is not embraced by the modern educational system. Several explanations have been explored as to the reasons why. First, PSI is very labor-intensive to operate and maintain (Austin, 2000; Boyce & Hineline, 2002; Pear & Crone-Todd, 1999). An effective PSI system requires the efforts of multiple different facilitators, including the instructor, teaching assistant and proctors. Many educators consider it an aversive event to construct. Second, PSI yields negatively skewed grade distributions (Austin, 2000; Boyce & Hineline, 2002; Fox, 2004). University administrators typically

expect a more normal distribution in large enrollment introductory courses and become suspicious of grade inflations or suspect faulty administration of the class when the data are released (Boyce & Hineline, 2002; Muchado & Silva, 1998). Given the self-paced nature of the true PSI model, there are a high proportion of students that receive a grade of Incomplete or Failure, depending on the sponsoring university's policy on grades, (Fox, 2004; Saville, Zinn, & Elliott, 2005). The negatively skewed distribution often becomes a bi-modal distribution, as the surge in grades lie on opposite ends of the grading distributions (Boyce & Hineline, 2002). Despite beneficial effects of implementing this teaching technology into traditional classrooms, these reasons have caused a shortage of support for a modern adoption of PSI.

Precision Teaching

Precision Teaching, like PSI, is a form of programmed instruction also shown to be effective. Precision teaching developed out of free operant research and is more focused on fluency of behavior rather than knowledge (Austin, 2000; Binder, 1996; Johnson & Layng, 1992; Lindsley, 1991). Fluency from the behavior analytic perspective involves the combination of speed in addition to accuracy as a metric of fluent performance (Johnson & Layng, 1992). In this regard, when a student scores 100% on a test designed to measure knowledge, the experimenter can infer no more than the student demonstrates adequate knowledge of the presented material. Knowledge of the material is generally a poor indicator of other academic measures, such as retention (Johnson & Layng, 1992). By focusing on rate of response, not just accuracy, the educator better tracks learning and becoming fluent of material.

In a typical precision teaching model, objectives and small components are identified and targeted for rate building (Austin, 2000). Progress is charted on a standard celeration chart and graphed frequently to monitor performance (Binder, 1996; Pennypacker, Koenig, & Lindsley, 1972). If the component targeted is too difficult for the learner, or it is evident that more basic skills should be approached first, the instructor can readily identify the low rate performance and adjust the curriculum accordingly. Once fluency is established, the student moves on to more complex components (Johnson & Layng, 1992). Thus, progression is managed with considerable precision.

Like PSI, Precision Teaching is not widely adopted by the educational system. Though individual programs of instruction exist across the country, most traditional educators are resistant to reform, especially in regards to fluency. Some educators think fluency is a difficult metric to hold students responsible for (McDade, 1993). Most instructors design questions that ask students to think carefully, and take large amounts of time to complete. Machado and Silva (1998) argue that most instructors, including behavior analysts, are not creating educational environments that allow students to emit complex ideas and exercise critical thinking skills. Fluency may teach key concepts in a course so these critical skills can be exercised. Like PSI, some design and knowledge of behavioral principles are vital to successfully administering a Precision Teaching model. This skill set is not taught in many education departments and this may explain the resistance of teachers. More empirical work needs to be done to explore barriers to integrate evidenced-based practice with traditional methods of instruction.

Interteaching

The lack of adoption of behavioral methods inspires formation of alternative, user-friendlier methods of instruction in behavior analysis. Boyce and Hineline (2002) introduced Interteaching as a method of instruction that requires less management skills to facilitate and gives more flexibility to instructional staff. Interteaching contains elements of PSI and Precision Teaching, however, it also incorporates aspects that may be intriguing to educators. In a typical Interteaching session, students are paired in groups of two to three, and given a prep guide to complete for a portion of the class. The prep guide contains materials and questions that resemble exam materials that the student may be held accountable for at a later date. During class time, students engage with the prep guide and the instructor moves from group to group and initiates discussion with group members. Students then fill out an evaluation of the session, which further allows the student to pinpoint potential problem areas they may be experiencing with the exam material. The instructor gathers these evaluations, and prepares a short lecture aimed at the problem areas identified by the students (Boyce & Hineline, 2002).

An “Interteach” is an interaction among students that involves probing and informing content from required readings (Boyce & Hineline, 2002). The instructor prepares a preparation guide beforehand that outlines readings and concepts discussed. During these interactions, students actively communicate the information and clarify key terms with one another. Time should not be spent doing the required reading or looking over notes (Boyce & Hineline, 2002).

Though anecdotal support exists for Interteaching, few empirical studies examine its effectiveness. Saville, Zinn, & Elliott (2005) examined different styles of teaching techniques in a laboratory setting with the students randomly assigned to four different

experimental groups. In the Interteaching condition, participants read a short article, and then engaged in an Interteaching session with a group of other students. Students in the lecture condition listened to a lecture on the article. Students in the reading condition read the article quietly to themselves without any discussion. A week later, students from the three experimental conditions, and a control group that did not have any exposure to the article, took a quiz on the readings. Students in the Interteaching condition scored significantly better than students in the other three conditions by at least 15 percent points (Saville, Zinn, & Elliot 2005).

Saville et al. (2006) examined Interteaching in a graduate level course. The instruction of the course alternated between traditional formats and Interteaching formats. All students performed better following the Interteaching sessions when compared to the lecture sessions. A second study examined Interteaching in two undergraduate courses. Interteaching sessions alternated with traditional sections across the two courses in counterbalanced delivery. Results showed that students performed better after the Interteaching sessions than the lectures. On a cumulative final, students answered more questions correctly from the Interteaching sessions than the lecture sessions. The authors call for more research on the controlling contingencies present within this model of instruction, in particular, use of prep guides to assist and shape study preparation by students (Saville et al. 2006).

Some potential limitations of Interteaching revolve around the number of students in a class. In smaller classrooms, Interteaching could be more effective as the instructor could spend more time with each individual group to reinforce successive approximations. In larger classrooms where attendance is beyond 50 students, instructors

may have difficulty distributing their time between high and low performing students. If students do not prepare or read required material prior to coming to the class meeting, they are not able to properly contribute to their individual group discussion (Boyce & Hineline, 2002; Saville et al., 2006). Clearly, more work needs to be done to answer some of these problems. The present study uses elements from the Interteaching model to automate instructional sessions by posting and allowing students access to web-based prep guides, or supplemental modules, and structured a contingency for accurate guided feedback on incorrect responses. Though not a direct replication of the Interteaching strategy, the use of interactive prep guides creates a virtual environment that enables student engagement while not directly interfering with traditional classroom instruction.

Supplemental Instruction

Many different disciplines outside of behavior analysis are beginning to see the benefit of designing supplements to assist instruction in both general education and college courses (Aberson et al. 2000; Baker, Wnetz, & Woods, 2009; Johnson & Kiviniemi, 2009; Jones & Fileds, 2001; Ketterlin-Geller, Chard, & Fien, 2008). Various different approaches have been made, and with the advancements in technology still to come, the possibilities seem endless. For example, Baker, Wentz, and Woods (2009) describe how some instructors use computer programs to create virtual worlds where instructors and students interact without face-to-face interaction. These programs allow for instructors to hold office hours, deliver lectures, and create an innovative learning environment for their students. Perhaps this is insight as to where education is heading.

For the most part, effective supplementation of classroom instruction involves some technological aspect, whether it is computerized or innovative practice within the

classroom. To supplement teaching of mathematical concepts, different approaches have been made. Jones & Fields (2001) report supplemental instructional sessions are effective for students in an introductory large enrollment accounting course for business majors. Fourteen individuals identified as experienced with the course materials led supplemental group sessions to assist others with the course material. Their data suggest that these sessions were effective at increasing academic performance. The methods they used were labor intensive, and did not use web-based interventions as an independent variable.

Aberson et al. (2000) designed a tutorial to teach the Central Limit Theorem for students in multiple sections of a statistics course. Students in the tutorial group received a packet of information and worked independently on the assignment with the assistance of the tutorial. Students in the lecture group attended lectures on similar materials. Students were evaluated via a pretest and posttest that measured knowledge of procedures, theory, definitions, and application. The results were not statistically significant but the tutorial offered more flexibility in terms of learning, and consumer satisfaction was high. Perhaps these types of tutorials should be designed to supplement courses rather than replace.

Over the past two decades behavior analysts embraced technology on multiple levels to design and implement effective interventions and effective supplements to learning (Pear & Crone-Todd, 1999; Ray, 1995; Tudor, 1995; Tudor & Bostow, 1991). Areas targeted include the teaching of spelling (Mayfield, Glenn, and Vollmer, 2008), training parents that have children with Autism (Sulzer-Azaroff et al., 2008), and prompting day care employees to greet parents by name (Ingvarsson & Hanley, 2006).

The present study designed supplemental modules derived from behavior analytic methods in a course on inferential statistics.

Internet Interventions

Using the Internet to deliver interventions is an emerging practice gaining wide support from multiple disciplines (Ritterband et al., 2003). In the last decade, many studies have examined various health issues and delivered interventions via an Internet connection, ranging from smoking cessation (Dallery & Glenn, 2005; Glenn & Dallery, 2007; Lenert, Munoz, Perez, & Bansod, 2004; Schneider, Walter, & O'Donnell, 1990) to treating eating disorders (Tate, Jackvony, & Wing, 2003; Winett et al., 1999; Winzelberg et al. 2000). With technology rapidly expanding and bandwidth capabilities increasing, the use of the Internet to deliver treatment becomes exceedingly popular, with such agencies as the National Institute of Health paying millions of dollars of support for such interventions (Ritterband et al., 2003; Ritterband, et al., 2006).

Only few studies exist that have evaluated Internet-based contingencies. Dallery & Glenn (2005) and Glenn & Dallery (2007) research using behavioral technology to reduce smoking in adults. Through use of a voucher program that offers cash rewards, the researchers report an effective program completely delivered through means of an Internet connection. Participants who are heavy smokers are required to take a reading of their carbon dioxide level present in their system by blowing into a detector. Readings are recorded and done in front of a web-cam, which allows the researchers an objective way to monitor progress. By setting goals, measuring progress, and earning rewards, heavy smokers reduce their smoking. This work demonstrates the potential for web-based interventions to change behavior. Little attention has been given to the design of

Internet interventions delivered to non-health specific behaviors. Porritt, Burt, & Poling (2006) examined the effects of a non-health related intervention delivered over the Internet. The productivity of fiction writers was increased by means of online goal setting and graphic feedback, all of which was delivered via the Internet.

Previous Work in Instructional Design

The experimental team in the present study has done extensive work in instructional design of math-based university courses (Brown & Alavosius, 2008; Dagen & Alavosius, 2006; Leeming & Alavosius, 2007; Leeming & Alavosius, 2009). Their work focused on multiple interventions to improve exam performance and course mastery in a psychology statistics course. The course uses a modified version of PSI to administer instruction. Historically, students that enroll in the course usually perform well, with 80% or higher passing with a grade of A or B. However, a sub-population in this course struggles. The experimental team has targeted improving performance in this sub-population through varying interventions. Dagen & Alavosius (2006) used a math tutorial to teach order of operations and basic mathematical skills. The results identified that low performing students could perform basic math skills, but still struggled with the course material, which might be attributed to lack of engagement and failure to apply skills. Leeming & Alavosius (2007) utilized peer coaching and exam retakes to guide the learning of students. While effective in increasing performance of low performers, peer coaching was labor intensive and required staff to work long hours to meet the demands for retakes. Brown & Alavosius (2008) designed web-based instructional modules as a pre-test to focus the study efforts of the students prior to exams. While not as effective as the peer coaching, the pre-test produced similar results. These required significantly less

labor to install and manage. Leeming & Alavosius (2009) used both color-coded and black and white web-based fluency modules as a supplement to the same course. While the black and white modules produced improvements in performance, the color-coded modules actually inhibited student performance on exams. These results suggest that relatively subtle changes in instructional design (e.g., color-coding) can affect student performance.

The previous work by this team has identified several different approaches for supplemental instruction. While peer coaching was the most effective intervention installed (Leeming & Alavosius, 2007), the labor correlated with maintaining such a program is not realistic in thinly resourced settings. Supplements should be user-friendly, efficient, effective, and designed so instructors can easily replicate them. The present study is the first effort by this team to design and install supplements for a course outside of their courses and home department.

Replication in a Business School

The present study evaluates the effects of web-based supplemental instruction modules on performance in an Inferential Statistics course for Business Majors. This study was within a traditional course (lectures with few exams). History shows that behavioral methods of instruction, through effective, are not adopted into the mainstream educational system. Therefore, perhaps it would be best to introduce components of behavioral education into the educational system to supplement instruction rather than reform traditional methods. In the present study, course supplements that resemble the prep guides described by Boyce & Hineline (2002) within the Interteaching paradigm were installed in a lecture-based course. As Saville et al. (2006) point out, “the similarity

between prep guide questions and test items is likely to exert stimulus control that may affect test-taking performance (p. 58)." An automated, programmed guide that supplements regular instruction could potentially provide a resource to instructors that seek ways to articulate effective methods of education and initiate student contact with course materials outside of the classroom. Unlike Boyce & Hineline's (2002) model, however, the course in the present study was taught from a relatively traditional perspective, meaning the supplemental instructional modules was completed outside of the classroom, and in class time was used for lectures and discussion of the material.

Two questions are asked in this research. The first is what are the effects of installing web-based supplemental instruction modules on exam performance in an undergraduate economics course on inferential statistics? The second asked if attaching a point contingency to the completion of such modules affect students' contact with supplemental materials?

METHOD

Participants

Fifty-eight students enrolled in three sections of Economics 262 were the participants for this study. The course is a prerequisite for higher division courses in the business curriculum, and is also the second course on statistics required for students enrolled into this program. All students followed standard procedures to register for the course, and did not have advanced knowledge of which section received which independent variable.

There were 18 participants in Section 002 (Experimental Group). Of those 18, 5 were sophomores, 7 were juniors, and 6 were seniors in class standing. Of the 18, 10 were Business Administration Pre-Majors, while the rest varied across different business majors or other programs on camps. There were 10 males and 8 females that chose to participate in this section.

There were 23 participants in Section 003 (Control Group). Of those 23, 8 were sophomores, and 15 were juniors in class standing. Of those 23, 16 were Business Administration Pre-Majors, 2 were Finance Majors, and the rest were varied across the business college and other programs on campus. There were 16 males and 7 females that chose to participate in this section.

There were 17 participants in Section 004 (Choice Group). Of those 17, 6 were sophomores, 9 were juniors, and 2 were seniors in class standing. Of the 17, 12 were Business Administration Pre-Majors, while the rest varied across the business college and other programs on campus. There were 8 males and 9 females that chose to participate in this section.

Personnel

The experimental team consisted of the Principal Investigator and the Student Investigator. The Principal Investigator is an Assistant Professor of Psychology at the sponsoring university, and supervised all experimental procedures. The Student Investigator is a graduate student in the Psychology Department at the same university, and completed this study to partially satisfy the requirements for the Degree of Master of Arts in Psychology. The Instructor of the course is a lecturer in the Economics Department at the sponsoring university. The university also employs a technical support staff to assist in the event of technical malfunctions or questions regarding WebCampus.

Participant Protection

Descriptions of all experimental procedures were submitted and approved by the sponsoring university's Institutional Review Board (IRB) prior to any participant recruitment. The consent protocol for this study was explicit in nature and informed students of the experimental variables of interest, and assured them that their participation in the course would in no way affect their final grade. Once data were collected, the name of the student, student number, and any additional personal information were converted to a three digit number series. The Principal and Student Investigators did not disclose the identity of any of the students within the course. The instructor had no knowledge as to which students choose to participate in the study. The students were informed that their participation in the study had no effect in how the course was taught or managed. Thus consent only granted permission for the experimental team to monitor and analyze student progress within the course.

Consent was gathered the first day of the class. The instructor left the room, and the Student Investigator read the following statement:

> "We are conducting a study to help us develop the instruction used in this course. As you just heard, all students in this class take regular scheduled exams and also complete supplemental homework. We wish to evaluate how supplements affect exam scores. We invite you to participate in this study by allowing us to examine how the supplements affect your test performance. The instructor will not know if you are participating or not, and your participation or decline will have no impact on the instruction you receive and the grade you earn. Please review the informed consent form. Are there any questions?"

Students were then given the opportunity to ask questions, and signed a prepared consent form, indicating their voluntary participation in the study. Consent forms are kept in a locked drawer for two years after the culmination of the study. After this time period, all documents involving consent will be destroyed as per university regulations. A copy of the consent is displayed in Appendix A.

Setting

This study took place in Land-grant University in the Western United States. The university enrolls over 16,000 students and employs 1,000 faculty members, and offers a variety of different programs, ranging from Mining Engineering to Business to Psychology. In particular, the undergraduate business curriculum offers training in a variety of different majors including Information Systems, Finance, Marketing,

Accounting, Management, & Economics. Before students are permitted to take higher division courses and declare a major, all business students are required to complete a core curriculum. The core courses consist of introductory survey classes in different disciplines, including math and statistics. Economics 262 is one of two statistics classes required by the College of Business as part of the core curriculum, and focuses on the principles of Inferential Statistics. All students must average at least a 2.75 Grade Point Average (GPA) in order to be recognized as successfully completing the core curriculum. A list of the core curriculum for the College of Business is listed in Appendix B. Each Econ 262 course section meets two times a week for 75-minutes each class. The instructor used *Business Statistics for Contemporary Decision Making, 5th edition* by Ken Black for the text. A copy of the course syllabus from is included in Appendix C.

This study emerged after a committee formed by the Provost explored options for supplemental instruction at the sponsoring university. The committee investigated high DFW rates in "gateway" courses and potential out-of-class supplements in preparation for potential cuts to the university's budget. Supplemental instructional resources at the university were eliminated in cost cutting including a Writing Center and a Math Center where struggling students could seek additional help. Funding for these programs was cut, forcing such programs to be eliminated or scaled down.

The university provides a number of "gateway" courses with a high level of either student failure or withdrawal rates. The present study was conducted in the context of these budget issues and unacceptable student failure in math-rich courses. Efforts to identify and evaluate an alternative form of supplemental instruction should involve

collaboration of multiple different departments on campus and improve the success of students.

All course supplements were designed and facilitated using WebCampus, an online instructional program that centralizes assessments, learning modules, and creates a template for threaded discussion. All supplements were accessed out of class and required an Internet connection to use the materials.

Procedure

All students in Econ 262 go through the same basic course format. The course is broken down into three different units. The first unit focuses on Inference and Hypothesis Testing. There are four weeks of general instruction, followed by an exam on this material. The second unit focuses on Chi-Square and Analysis of Variance (ANOVA). Again, students go through four weeks of general instruction, followed by an examination. The last unit is on Correlation and Regression Analysis, which has 5 weeks of general instruction and culminates with the last exam. All together, students have three opportunities to demonstrate mastery in the context of this course. For each exam, students have access to a calculator, the inside cover of their textbook which contained a copy of the Unit Normal Table, and a prepared "cheat sheet" that the student brings in to reference for the exam. In addition, each exam contains a "take-home" portion where the student is required to complete a computer output problem. Students also complete weekly take-home assignments that are assigned by the instructor, in addition to a case study.

Each of the three course sections was randomly assigned to either an experimental or control condition. Two sections (Sections 002 and 004) received a variation of the

Independent Variable, while one section (Section 003) served as a control group. Students enrolled in the control group took the course as stated above without any deviation or required interaction with any supplement. Students in the experimental sections were to complete a supplemental instructional module. These modules were similar in format to the prep guides described by Boyce & Hineline (2002) within the Interteaching paradigm, in that the material on these modules resembles that of the exam materials. Guided instructional feedback was contingent upon completion of the module. In this design, student interaction with course material is with the interactive module and not peer-to-peer as to control for accurate feedback and active participation.

Each module consists of a variety of multiple-choice questions. The Student Investigator designed these modules for functionality on the Web and aligned the content within these modules to relate to the course instructor's learning objectives. After students complete the module, they receive immediate feedback on their performance. Students are informed of their correct and incorrect responses. Written descriptive feedback explaining the logic behind the answer and the page number in the textbook that corresponds to the material are also displayed.

Students enrolled in Section 002 were assigned into the experimental group with the supplement as a homework assignment given one week prior to the exam. Students had until the class period before the exam to complete the supplement, but could interact with it after the deadline passed. In order to interact with the supplement, students logged onto WebCampus with their personalized username and password, and clicked on their appropriate course number. Students enrolled in Section 002 received the following instructions prior to the completion of the supplement:

> "Open the file below to engage with a web-based supplement. These supplements have been designed to help you prepare for the exams given in this course. **THE COMPLETION OF THESE SUPPLEMENTS WILL COUNT TOWARDS YOUR HOMEWORK POINTS FOR THIS COURSE. IN OTHER WORDS, NOT COMPLETING THE SUPPLEMENT WILL IMPACT YOUR FINAL GRADE. "**

See Appendix D for a digital image taken from the course webpage for Section 002.

Students enrolled in Section 004 were assigned to the Choice Group and had access to the supplements, but no points were given for completion of the supplement. In other words, completion of the supplement was available but optional. Like the Experimental Group, the Choice Group accessed the supplements from WebCampus, and were given the following instructions:

> "Open the file below to engage with a web-based supplement. These supplements have been designed to help you prepare for the exams given in this course.

See Appendix E for a digital image taken from the course webpage for Section 004.

Students enrolled in Section 003 were assigned to the Control Group. No webpage was designed for their course, and no web-based supplements were available to the students outside of standard practices of the course.

Dependent Variables

This course was taught following the established format and their exam scores were the primary measures taken by the experimental team. The instructor designed all testing answer keys and grading criteria used on the exams. The Student Investigator designed and managed the web-based modules. Inter-observer agreement (IOA) was gathered for 30% of two of the three exams to ensure an objective grading criteria was in effect. IOA was calculated by sampling 30% of all exams from each section, and using the formula:

(total number of questions – number of deviations)/total number of questions x 100%.

For section 002, 003, and 004 on Exam 1, IOA was found to be 97.9%, 98.2%, and 97.5%, respectively. For sections 002, 003, and 004 on Exam 2, IOA was found to be 99.1%, 100%, and 100%, respectively. No IOA was calculated on Exam 3.

The instructor approved all materials put on the module to ensure they were relevant course materials. Exams were administered three times during the course of the semester. The first exam was worth 110 points and the other two exams were worth 90 points toward the final grade. All exams can be seen in Appendix F, G and H respectively. Time measures were taken on each exam in order to measure the time taken to complete all examination materials. All timings were done with a Lathem 1000e time clock. Prior to handing out exams, the Student Investigator took an initial time measure to notate the start of the exam. As students handed in their exam, they would stamp their test with the time clock. The Student Investigator then subtracted the student's time from the original start time to calculate how long the student had the exam in their possession.

Prior to the beginning of instruction, students received a pre-test to assess equivalence of all three sections. Given that this class is the second course in Statistics these students take, the pre-test measured abilities on Descriptive Statistics and basic mathematical computations. This pre-test identifies low, midlevel, and high performers to compare their data to other individuals whose scores were similar.

Module performance was another dependent measure for evaluation. WebCampus tracks how long each individual student took to interact with the supplement, how many attempts were made to achieve mastery, and also how accurate their performance was. These measures were evaluated within and between experimental sections to detect difference between the two sections and individual students.

Independent Variable

The independent variable was access to the supplemental instruction modules. Students in the control section never had access to the modules. All students in the control section still took and participated in all graded work in terms of exams, case studies, and weekly assignments.

Students in the experimental sections had access to the supplemental instruction modules throughout the semester. These consisted of a series of multiple-choice questions that gave guided feedback once the module has been completed. Students had unlimited attempts on the module, which was to ensure mastery of the material. Students completed the module, reviewed their answers, reviewed guided feedback on questions, and took again if mastery was not achieved (mastery level on module set at 100%). One module was designated per unit in this class. Overall, there were three supplemental modules presented in the experimental sections. In order to control for potential

cheating, or students not directly interacting with the material and submitting blank modules, each module had a time limit set to three minutes, meaning that the students must spend at least three minutes interacting with the module. Also, the list of potential answers was randomized, as no two modules looked exactly the same, but all covered the same material.

Students in the first experimental section, or the Experimental Group, were required to complete the supplements as per a point requirement in their syllabus. Failure to complete a supplement resulted in a point deduction towards the final grade.

The other experimental section (Choice Condition) had regular access to the same supplemental materials. However, the completion of the module was not contingent towards the final grade of the student. No point value was placed on the completion of the supplemental modules, and further, the completion of such was in the control of the student. Verbal prompts by the instructor were administered throughout the semester to encourage students to interact with the module as a form of exam preparation. Doing so did not affect the final grade of the student.

Experimental Design

A between group design evaluates the differences in performance of students in the experimental sections and the control section. All in class course materials that the groups received were held constant, meaning the modules were the only difference on how the courses were taught.

Data Analysis

A combination of statistical analysis and visual inspection of data (Baer, 1977; Michael, 1974) were used to detect variations of group analyses in addition to individual

student performance. An analysis of variance (ANOVA) test was run to detect any effect of supplements on grades for exams and assignments. Representative students were evaluated in regards to their individual performance. The pre-test identified potential low performers in all three sections. Their progress through the course was examined compared to representative mid-level and high performers.

Social validity measures (Hawkins, 1991; Wolf, 1978) were gathered at the culmination of the course to measure satisfaction with the modules and students' comments on the course. Students completed a brief evaluation that asked them to rate their satisfaction and recommend how future modules should be operated.

Effort measures were taken to calculate the amount of time to program and design these supplemental modules. This time included initial contact with material, the design of questions, the design of guided feedback, programming onto WebCampus, testing prototypes, and meetings to discuss supplemental content. These effort measures estimate the amount of work necessary to replicate these modules.

RESULTS

Global Summary

The research questions for this study asked: 1) what were the effects of installing supplements in an economics statistics course? 2) Did attaching a point contingency effect contact with the supplement?

Figure 1 provides a general summary of the results of the present study. The pretest given before the release of any supplement yielded scores for the three sections (Experimental, Choice, Control) of 79.56%, 77.18%, and 72.14% respectively. On the three exams, Section 002, or the Experimental Group, averaged 84.09%, 75.74%, and a 77.22%, respectively. Section 004, or the Choice Group, averaged 79.95%, 79.35%, and 73.53%, respectively. Section 003, or the Control Group, averaged 72.61%, 73.00%, and 69.95%, respectively. As evident by the averages, there was a clear distinction between the two experimental conditions and the control condition. Figure 2 depicts the same graph as Figure 1 with the choice condition divided, with one bar representing students who took the supplement and the other for students that did not. Students in the choice condition that took the supplement performed at a higher average when compared to the rest of the class (Supplement=80.91%, 85.11%, 74.67%; No Supplement=79.82%, 76.94%, 73.06%).

Figure 3 shows the average time spent on each exam. As depicted in the graph, the experimental condition averaged a faster completion time in comparison to the other two sections. Figure 4 shows a similar graph, but separates the choice condition between students that took the supplement and students that did not. Little differentiation is seen within the choice condition, with the exception of Exam 3 where the students that

engaged with the supplement took longer to complete when compared to the rest of the class.

Final grade distributions are shown for all sections in Table 1 and graphically in Figure 5. The experimental and choice condition had higher percentages of students in passing range (A, B, & C) when compared to the control condition. It is important to note that exam scores were not the only index used to evaluate students, especially in calculating final grades. Figure 6 presents the final course grade if only calculated by exam scores (excludes non-targeted course materials; i.e. in-class quizzes and case studies). As is evident by this graph in comparison to Figure 5, it appears that grades fluctuated more towards the center of the distribution, thus other assignments and case studies used in the course helped low performers and hurt high performers.

A one-way analysis of variance (ANOVA) compares scores on all three exams, and the final grade distribution. Differences on Exam 1 were statistically significant (α=.05, df= 2, 55, F=5.119). No other statistically significant effects were found, although significance was approached in the final grade distribution, especially when calculating final grades with exam scores only. Table 2 shows the ANOVA table and *F* values for all exams and the final grade distribution. Scheffe's Post Hoc analysis was run to detect which group deviated from the others. Table 3 displays the output. Note the asterisks, as they indicate statistically significant effects. The experimental group differed from the control group on the first exam. This was the only exam where significant differences between groups are noted.

The second question asked whether attaching a point contingency to supplemental modules was necessary for students to contact the modules. Table 4

displays general statistics run from WebCampus that convey activity levels on the webpage for the entire semester. The differences between the two sections are notable, clearly showing that the supplement was accessed more by students in the experimental section compared to those enrolled in the choice section. Table 5 show the number of students that interacted with the supplement for the particular exam, the compliance percentage ((number of students that took supplement/number of student enrolled in course) x 100%), and the number of retakes exercised. Again, the differences between the two sections are large, with the Experimental Section having high compliance and a large number of retakes.

Rate Measures

Figure 7 depicts the correlation between rate of response and exam percentage in Exam 1 for all three sections. Rate was calculated by taking the students raw point score and dividing it by their completion time. The unit of measurement is therefore points per minute. Each quadrant is meaningful in this depiction. For example, Quadrant I (upper right) represents high rate and high accuracy. Quadrant II (upper left) represents high rate and low accuracy. Quadrant III represents (lower left) represents low rate and low accuracy. Quadrant IV (lower right) represents low rate and high accuracy. For this exam, the cutoffs on the x and y-axes were made based on identified performance benchmarks. For example, the x-axis cuts off at 70%, as 70% is the lowest passing score a student can obtain. The y-axis cuts off at 1.03, as this is how fast students need to be answering in order to complete the exam in the allotted time (110 points in 75 minutes). The experimental group performed faster, more accurately, and with less variability when compared to the other two sections. Pearson correlation coefficients (r) were equated and

show a positive correlation between rate and accuracy (r=.587(exp), .585(choice), and .538(control)). The red squares in the graph depict individuals in the choice condition. Squares without fill indicate that the student completed a supplement prior to the exam. Figure 8 depicts the same graph, but separates students based on contact with the supplement, and not by course section. In other words, the experimental section and students that completed the supplement in the choice section are grouped together ("Supplement Rate"), and the control and the remainder of students in the choice section are grouped together ("No Supplement Rate"). As evident in this graph, there is a clear distinction between the two groups, suggesting a relationship exists between contact with the supplement and rate of accurate responding.

Figure 9 shows the same correlation with Exam 2. Quadrant assignments are the same, but the axes were adjusted to accommodate the different point structure in this exam compared to Exam 1. In this depiction, the cutoff on the y-axis is set at 0.84 (90 points, 75 minutes). Again, there is a clear separation with respect to rate, accuracy, and variation between the experimental sections and the other two sections. Pearson correlation coefficients show another positive correlation (r=.574 (exp), .269 (choice), .706 (control)). Figure 10 shows the same graph, but separates students based on contact with the supplement. In both graphs, the students who contacted the supplement responded quickly and accurately.

Figure 11 shows the correlation for Exam 3. Quadrant assignments and axes were based on the same assignment as Figure 9 (x-y intersection at (70%, 0.84)). Although there was more variation in the experimental section than previous exams, the experimental section still performed with a high degree of rate and accuracy in

comparison to the other sections. Pearson correlation coefficients reveal another positive correlation between rate and exam percentage (r= .526 (exp), .276 (choice), .734 (control)). Figure 12 shows the same graph, but differentiates based upon supplement contact. There is a clear separation between students that contacted the supplement with respect to rate, accuracy and variation in comparison to the other students, suggesting a relationship between the supplement and rate of accurate responding.

Individual Performance

Pretest scores were used to identify low, mid-level, and high performers for the duration of the course. More specifically, the top three, the low three, and the middle three scores were used for individual analysis. Note the participant number when comparing the data within and across these graphs. Numbers beginning with "1" refer to students in the experimental group (supplement interaction was required). Those beginning with "3" indicate students in the choice condition (supplement optional). Participant numbers beginning with "2" refer to students enrolled in the control group (supplement withheld). Figure 13 depicts the individual performance for the identified low performers in the course. The asterisks indicate that the supplement was completed at least one time by the student. Comparison of exam averages for low performers across the three sections indicates an average of 75.17% for the experimental section, 67.06% for the choice, and 75.86% for the control, suggesting that the supplement did not have a strong effect with low performers.

Data from the mid-level performers are depicted in Figure 14. Again, an asterisk directly above the bar graph indicates that a supplement was accessed and completed at least one time. There is some difference between the experimental participants and the

other two groups. However, a comparison of the means reveals averages of 74.85% for the experimental section, 80.80% for the choice, and 77.83% for the control, again suggesting that the supplement did not have a strong effect with mid-level performers.

Figure 15 shows the results from the high performers. As seen in comparison to the other performers, the students that accessed the supplement generally performed at a higher level, and were distinct from the other high performers. Comparison of exam averages for high performers across the three sections indicates an average of 79.31% for the experimental section, 78.57% for the choice, and 78.94% in the control.

Choice Condition: Individual Performance

Students enrolled in Section 004 were assigned to the Choice Condition. No student interacted with the supplement in every exam, but different combinations were exercised. Figure 16 depicts the exam scores for every student that interacted with at least one supplement for the duration of the semester. An asterisk directly above the bar indicates that a supplement was accessed prior to the exam. Note in this graph the different combinations used by students. For instance, Participant 303, 306, and 319 interacted with the supplement solely on Exam 2. Further, this pattern of responding resembles a Baseline-Intervention-Baseline experimental design.

Social Validity

At the culmination of the project, students filled out a consumer satisfaction survey that asked them to rate their opinions of the course and supplementation. Three questions asked about Professionalism, and six questions asked about the Supplemental Modules. Table 6 shows the responses gathered for students enrolled in both the

experimental and choice groups. Overall, the majority of students reacted favorably to the supplements, with a majority suggesting that they be kept in the design of the course.

Effort Measures

The student investigator tracked the time (in minutes) required to design and launch the supplemental modules. Figure 17 shows the self-reported data from the student investigator depicting effort placed in the completion of the supplements. The graph shows three categories, in addition to a sum of all three for a grand total per supplement. The three categories are: meetings between the student investigator and the instructor discussing the contents of the supplement; the design of questions and the act of posting them onto WebCampus; and the design of feedback. The supplement for Exam 3 by far took the most time to complete, given the complexity of the content and different format when compared to the other two areas of content.

DISCUSSION

General Comments

The supplement affected the performance of the students, especially in Exam 1. Students in the experimental group performed higher and with less variability in the first exam, but then as content became more difficult, exam performance in the experimental group resembled the performance of the other two conditions. The supplement seemed to have variable effects on the individual performance of students. A clear benefit was most evident in the high performer category, compared to the other performers.

Comparisons with our results to other Interteaching interventions reveal notable differences. Saville et al. (2005) reported high differences (15 points) between intervention conditions and control conditions when using an Interteaching design. Web-based supplements, while attempting to automate the Interteach sessions, appear to be less effective than the social methods used in previous studies (Saville et al., 2005; Saville et al., 2006). A survey of our previous work in instructional design shows that our most effective intervention involved peer-coaching (Leeming & Alavosius, 2007). While attempting to replicate this via web-based supplements has been successful in PSI-type courses (Brown & Alavosius, 2008), the effect was less potent in the current study. The social interactions in the Interteach designs may provide elaborate and extended feedback in comparison to more general feedback provided directly after a multiple-choice assessment. Analyses of the content, duration, timing, schedules of feedback, and other supplement operations may reveal directions for refinement of supplementation.

Further analyses of the supplemental effects reveal some promising results. The most obvious conclusion is that these modules are used most when a clear point value is

assigned for completion. While the supplement was interacted with in the Choice Condition, it was not universally used as an instructional tool, nor was the supplement used to its utmost capability. Students that would possibility benefit from the supplement chose not to interact with the module. Figure 13 validates this statement by looking at the supplements accessed out of 3 low-performing students in the Choice Condition. Out of the three exams, only two supplements were accessed by the three low performers. Michael (1991) notes the importance of the point contingency in course design with respect to modifying the behavior of students. Assigning points for academic work provides a powerful way for the instructor to shape studious behavior, in addition to promoting opportunities for active responding outside of the course. The optimal assignment of points is an area of future investigation.

Computed correlations between rate and exam performance provide another promising result from this study. There was a clear distinction across all three exams with the experimental section responding at highly accurate rates in comparison to the other two sections. These data allow us to infer that there is a clear relationship between rate of response and performance, and that the supplement may have allowed for the more fluid responding observed in the experimental and choice sections. Future implications could target better supplemental design, with particular emphasis on training for fast and accurate responding. Point values could be manipulated to promote high rate and accuracy in completing modules. Supplemental modules that interlock these two features of responding may better approximate the interactions students find with peer tutors.

The mission of this study was to introduce potentially effective methods of supplementation to a traditional lecture based course. It is worthwhile to discuss some of the obstacles in undergoing such a project. While the largest effect was shown in Exam 1, similar results might have been evident had more opportunities for active responding been present, in terms of more exams. We know of no published studies showing reliance on lectures as the only method for disseminating complex materials and an effective way at boosting exam scores. In taking advice from the leaders in Behavioral Education (Austin, 2000; Boyce & Hindline, 2001; Vargas & Vargas, 1991), using exams as a teaching tool rather than just a learning assessment could have significant implications for the successful completion of coursework. Short, successive approximations via frequent examinations allow the student to demonstrate the required skills necessary to succeed in a course on a regular basis. When infrequent examinations are used as the primary metric to evaluate students, students are less likely to be actively engaged with materials, and more likely to lack component skills required to complete more complex materials.

This leads to larger issues in higher education. The present state of the United States economy suggests educational reform on multiple levels. Lecture-based courses are not found to produce high levels of student success. Perhaps this format needs to be re-evaluated in the context of the goals of the institution. Given the obtained grade distributions in this study, simple modifications can be made in courses that boost low performers to higher levels of academic achievement. Grade distributions can be altered by low cost supplements but reform of basic course operations may be necessary before supplementation reaches the maximum potential. Instructional designs that replace

lectures with frequent opportunities to actively respond and inevitably lead to detailed and timely feedback are found to be optimal models, especially when such activities result in consequential points toward the final grade. Anecdotally, this appears to be met with resistance from instructors as preference for lecture methods of instruction and cognitions towards student performance are generally not data-based.

The students positively rated the supplemental modules in both the experimental and choice groups. What is particularly interesting is the number of positive responses taken from the choice condition, considering the number of interactions with the supplement in the choice condition (see Table 4). It appears that students emitted preference for the option of extra help, even though the supplements were not routinely engaged with in the choice condition. In our previous work, we uniformly find students to positively rate course supplements and frequent quizzing.

Limitations

Although the obtained results are a promising stepping-stone for more research in classroom techniques, there are some limitations to this study worthy of discussion. The number of students enrolled in this study was significantly lower than was previously estimated. Although all experimental groups were roughly equivalent in size, the actual number of participants differed greatly from the original proposal by close to 100 students. Far fewer students enrolled than were expected, perhaps due to the economic situation at the sponsoring university. Had more students enrolled in the study, the power of the significance tests may have yielded better detection of effects.

Second, the Pretest scores indicated that there were differences between Experimental Groups and the Control Group prior to any intervention. An argument can

be made that the supplements may have had little effect on the overall performance of the student, and that the individual history of each student is potentially an explanation of the obtained results. However, given the significant effect shown in Exam 1 and replications in the data of the individual performers, it is likely that the supplement had some impact with respect to exam preparation for all students that contacted them, and that the predictive value of the pretest was overestimated. Individual data clearly depicted the variability of performance.

When conducting this project, there were several instances where the experimental staff sacrificed experimental factors to adapt to the context. For instance, due to budget restrictions, it was impossible to hire an independent grader for all exams like originally proposed. All experimental manipulations were conducted outside of the classroom to not conflict with the established routine of the course. The experimental staff are grateful towards the Economics Department for hosting this project. Our previous research in course development has been done in Psychology courses. We find students in Psychology to be more willing to participate in research on course development. Business students were more reluctant to participate, which could be attributed to the different environmental variables in the individual departments. Psychology students are frequently recruited for participation in research studies, whereas Business students are not, and may have altered cognitions towards being a research subject. Further analyses should investigate resistance of students and faculty in adoption of evidenced-based practices. Argyris (1968) recommends a model of research where participants (ie. students, faculty, administrators, etc.) actively engaged in the planning, design, and evaluation of a research study. This idea is provocative, especially for

applied researchers in university settings during economic recessions. Students might advocate for better instructional design, but this demand would likely require a shift in technique commonly used by most instructors, in addition to a redefinition of the student's role as a consumer of university offerings. Lectures would shift from students passively listening to vocal instructions to students actively shaping their own educational environment to a way that best fits their individual repertoires. An environment where the students are active in the course design presents a challenge worthy of future investigation.

This was our first study of course operations outside our home department. The opportunities for refinement are many. While the *n* size for this study was sufficient, but not ideal, future research should attempt to replicate these methods with larger enrollment courses using enhanced materials, and perhaps try to branch outwards into other subject matters like social sciences or other business curriculum. Another area for improvement would fall on the development of software that accompanies the goals of the instructors and better supplements their teaching methods. For instance, while computational questions would have been beneficial to add to the supplements, the amount of variability in responding by the student would have altered answers by decimal points, and further would have been very difficult to program into WebCampus. Future research should evaluate software that can account for the variability of student responses on mathematical problems.

In addition, a replication of this study that transferred all in-class exams to computer-based assessments would potentially allow for more meaningful metrics to be applied to analysis. While rate of exam completion was computed for each student, this

is not a true fluency reading, and further, these data should be interpreted with caution. Rate was calculated by taking the raw point score from the exam, but different questions are weighted differently. Accounting for this by breaking questions down into component skills and measuring active response rate on computerized exams would allow for more in-depth analysis rather than percent correct, as was the key metric in this study.

The most interesting area for future research is analyzing the institutional variables operating in the context of university instruction. Given that higher education is moving more towards large class enrollments with minimal resources, the instructional design exercised by instructors refers to a variety of different behaviors that all operate on interlocked contingencies throughout the academic environment. Research might seek to identify the goals of the institution, and whether individual departments are in compliance with those goals. For instance, an instructor can design a course where high levels of mastery are achieved across the majority of students, or can view the course as more of an academic selection that isolates an elite few as passing and fails those that cannot keep up. Instructor and administrator values play an interesting role in the analysis of these institutional variables, as well as identifying the source of resistance to adopt mastery methods of instruction. Such research might result in the development of successful rubrics that instructors could cross reference with their individual teaching methodology and alter accordingly. For example, most lecture-based courses lack clear behavioral objectives and opportunities for active responding, and further do not design the course with successive approximations to some terminal set of skills and abilities. Identification of these institutional variables would provide a way to assess how

educators and officials are regulating and operating the academic environment and integrate evidenced-based practices into the college classroom.

Statistics courses teach well articulated and straight-forward analyses used in a variety of applications and settings. Many students are ill-prepared for instruction in statistics due to a variety of reasons, such as prior history with poor instructional designs or negative cognitions towards math courses. The design of university-level statistics courses can be made to accommodate the faulty preparation and remediation of students without sacrificing academic integrity. This would require effort in terms of the instructional design, but potential outcomes outweigh this concern. Such redesigned courses would involve the assessment of the entering skills of students, the presentation of opportunities for active responding to shape component skills, and creating links to these skills to composite applications and meaningful utility. This goes beyond simply supplementing conventional lectures and delves into the fundamental design of the learning environment, but clearly presents an enticing alternative to the normal practices observed presently.

REFERENCES

Aberson, C.L., Berger, D.E., Healy, M.R., Kyle, D.J., & Romero, V.L. (2000). Evaluation of aninteractive tutorial for teaching the central limit theorem. *Teaching of Psychology,* 27, 289-291.

Argyris, C. (1968). Some unintended consequences of rigorous research. *Psychological Bulletin, 70(3*), 185-197.

Austin, J. (2000). Behavioral approaches to college teaching. In J. Austin & J.E. Carr (Eds.), The Handbook of Applied Behavior Analysis. (449-472). Reno, NV: Context Press.

Austin, J. & Carr, J. E. (2000). *Handbook of Applied Behavior Analysis*. Reno, NV: Context, Press.

Azrin, N. (1977). A strategy for applied research: learning based but outcome oriented. *American Psychologist, 32,* 140-149.

Baer, D.M. (1977). Perhaps it would be better not to know everything. *Journal of Applied Behavior Analysis,* 10, 167-172.

Baer, D.M., Wolf, M.M., & Risley, T.R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis,* 1, 91-97.

Baker, S.C., Wentz, R.K., & Woods, M.M. (2009). Using virtual worlds in education: Second Life as an educational tool. *Teaching of Psychology,* 36, 59-64.

Binder, C. (1996). Behavioral fluency: Evolution of anew paradigm. *The Behavior Analyst,* 19, 163-197.

Brown, T.W., & Alavosius, M.P. (2008). *The effects of an online pre-test on exam performance in an undergraduate statistics course*. Unpublished honor's thesis, University of Nevada, Reno, Reno, NV.

Boyce, T.E. & Hineline, P.N. (2002). Interteaching: A strategy for enhancing the user-friendliness of behavioral arrangements in the college classroom. *The Behavior Analyst,* 25, 215-226.

Curran, C. (2008). Online learning and the university. In W.J. Bramble & S. Panda (Eds.) *Economics of Distance and Online Learning.* New York: Routledge.

Dagen, J.C., & Alavosius, M.P. (2006). *Assessing the potential effects of a mastery-based mathematics program on exam performance in an undergraduate psychology statistics course*. Unpublished master's thesis, University of Nevada, Reno, Reno, NV.

Dallery, J. & Glenn, I.M. (2005). Effects of an internet-based voucher reinforcement program for smoking abstinence: A feasibility study. *Journal of Applied Behavior Analysis,* 38, 349-357.

Fellner, J.D. & Sulzer-Azaroff, B. (1992). A behavioral analysis of goal setting. *Journal of Organizational Behavior Management,* 6, 33-51.

Fox, E.J. (2004). The personalized system of instruction: A flexible and effective approach to mastery learning. In D.J. Moran & R.W. Malott (Eds.), *Evidence-based educational methods* (p. 201-221). San Diego: Elsevier Academic Press.

Glenn, I.M. & Dallery, J. (2007). Effects of internet-based voucher reinforcement and a transdermal nicotine patch on cigarette smoking. *Journal of Applied Behavior Analysis,* 40, 1-13.

Goldstein, I.L. & Ford, J.K. (2002). *Training in Organizations*, 4th Edition. Belmont, CA: Thompson Learning, Inc.

Hawkins, R.P. (1991). Is social validity what we are interested in? Argument for a functional approach. *Journal of Applied Behavior Analysis*, 23, 205-213.

Ingvarsson, E.T., & Hanley, G.P. (2006). An evaluation of computer-based programmed instruction form promoting teachers' greetings of parents by name. *Journal of Applied Behavior Analysis*, 39, 203-214.

Johnson, B.C. & Kiviniemi, M.T. (2009). The effect of online chapter quizzes on exam performance in an undergraduate social psychology course. *Teaching of Psychology*, 36, 33-37.

Johnson, K.R. & Layng, T.V. (1992). Breaking the strucuralist barrier: Literacy and numeracy with fluency. Special issue: Reflections on B.F. Skinner and psychology. *American Psychologist*, 47, 1475-1490.

Jones, J.P. & Fields, K.T. (2001). The role of supplemental instruction in the first accounting course. *Issues in Accounting Education*, 16, 531-547.

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

Ketterlin-Geller, L.R., Chard, D.J., & Fien, H. (2008). Making connections in mathematics: Conceptual mathematics interventions for low performing students. *Remedial and Sepcial Education*, 29, 33-45.

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

Leeming, E., & Alavosius, M.P. (2007). *The effects of peer-coaching and exam retakes on performance in an undergraduate statistics course*. Unpublished honor's thesis, University of Nevada, Reno, Reno, NV.

Leeming, E. & Alavosius, M.P. (2009). *The effects of color-coded and black and white fluency modules on exam performance on an undergraduate statistics course*. Unpublished master's thesis, University of Nevada, Reno, Reno, NV.

Lenert, L., Munoz, R.F., Perez, J.E., & Bansod, A. (2004). Automated e-mail messaging as a tool for improving quit rates in an internet smoking cessation intervention. *Journal of the American Medical Informatics Association*, 11, 235-240.

Lindsley, O.R. (1991). Precision teaching's unique legacy from B.F. Skinner. *Journal of Behavoral Education*, 1, 253-266.

Locke, E.A. & Latham, G.P. (2002). Building a practical useful theory of goal setting and task motivation. *American Psychologist*, 57, 705-717.

Machado, A. & Silva, F.J. (1998). Greatness and misery in the teaching of the psychology of learning. *Journal of the Experimental Analysis of Behavior*, 70, 215-234.

Mayfield, K.H., Glenn, I.M., & Vollmer, T.R. (2008). Teaching spelling through prompting and review procedures using computer-based instruction. *Journal of Behavioral Education*, 17, 303-312.

McDade, C.E. (1993). Computer-based precision learning: Achieving fluency with college students. *Education & Treatment of Children*, 63, 290-305.

Michael, J. (1974). Staistical inference for individual organism research: Mixed blessing or curse? *Journal of Applied Behavior Analysis*, 7, 647-653.

Michael, J. (1991). A behavioral perspective on college teaching. *The Behavior Analyst*, 14, 229-239.

Naughton, J. (2000). *A brief history of the future: The origins of the Internet.* London: Phoenix.

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

Pennypacker, H.S., Koenig, C.H., & Lindsley, O.R. (1972). *Handbook of the Standard Celeration Chart*. Kansas City: Precision Media.

Porritt, M., Burt, A., & Poling, A. (2006). Increasing fiction writers' productivity through an internet-based intervention. *Journal of Applied Behavior Analysis*, 39, 393-397.

Ray, R.D. (1995). A behavioral systems approach to adaptive computerized instruction. *Behavior Research, Methods, Instruments, & Computers*, 27, 293-296.

Ritterband, L.M., Andersson, G., Christensen, H.M., Calbring, P., & Cuijpers, P (2006). Directions for the international society for research on internet interventions. *Journal of Medical Internet Research*, 8, 60-65.

Ritterband, L.M., Gonder-Frederick, L.A., Cox, D., Clifton, A.D., West, R.W., & Borowitz, S.M. (2003). Internet interventions: In review, in use, and into the future. *Professional Psychology*, 34, 527-534.

Saville, B.K., Zinn, T.E., & Elliot, M.P. (2005). Interteaching vs. traditional methods of instruction: A preliminary analysis. *Teaching of Psychology*, 30, 133-136.

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

Schneider, S.J., Walter, R. & O'Donnell, R. (1990). Computerized communication as a medium for behavioral smoking cessation treatment: Controlled evaluation. *Computers in Human Behavior,* 6, 141-151.

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

Skinner, B.F. (1958). Teaching machines. *Science,* 128, 969-977.

Skinner, B.F. (1938). *The Behavior of Organisms: An Experimental Analysis.* Oxford: Appleton Century.

Sulzer-Azaroff, B., Fleming, R., Tupa, M., Bass, R., & Hamad, C. (2008). Choosing objectives for a distance learning behavioral intervention in autism curriculum. *Focus on Autism and Other Developmental Disabilities,* 23, 29-36.

Tate, D.F., Jackvony, E.H., & Wing, R.R. (2003). Effects of an Internet behavioral counseling on weight loss in adults at risk for type 2 diabetes: a randomized trial. *Journal of the American Medical Association,* 14, 1833-1836.

Tudor, R.M. (1995). Isolating the effects of active responding in computer-based instruction. *Journal of Applied Behavior Analysis,* 28, 343-344.

Tudor, R.M. & Bostow, D.E. (1991). Computer-programmed instruction: The relation of required interaction to practical application. *Journal of Applied Behavior Analysis,* 24, 361-368.

Vargas, E.A. & Vargas, J.S. (1991). Programmed instruction: What it is and how to do it. *Journal of Behavioral Education,* 1, 235-251.

Waits, T., Lewis, L., & Greene, B. (2003). *Distance education at degree-granting post-secondary institutions:2000-2001.* Washington DC: U.S. Department of Education, National Center for Education Statistics. (NCES 2003-017).

Winett, R.A., Roodman, A.A., Winett, S.G., Bajzek, W., Rovniak, L.S., & Whiteley, J.A. (1999). The effects of the Eat4Life internet-based health behavior program on the nutrition and activity of high school girls. *Journal of Gender, Culture, and Health,* 4, 239-254.

Winzelberg, A.J., Eppstein, D., Eldredge, K.L., Wilfley, D., Dasmahapatra, R., Dev, P., & Taylor, C.B. (2000). Effectiveness of an Internet-based program for reducing risk factors for eating disorder. *Journal of Consulting and Clinical Psychology,* 68, 346-350.

Wolf, M.M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis,* 11, 203-214.

TABLES

Table 1

Grade Distribution: Sections 002, 003 & 004

FINAL GRADE	EXP	CONROL	CHOICE
A	2	0	4
B	10	11	7
C	3	8	4
D	2	3	0
F	1	1	2
TOTAL	18	23	17

Table 1: Final grade distribution for all sections.

Table 2

One Way Analysis of Variance (ANOVA) Output

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Pretest	Between Groups	2228.700	2	1114.350	3.322	.043
	Within Groups	18447.524	55	335.410		
	Total	20676.224	57			
Exam1	Between Groups	1687.554	2	843.777	5.119	.009
	Within Groups	9066.050	55	164.837		
	Total	10753.603	57			
Exam2	Between Groups	319.409	2	159.705	1.103	.339
	Within Groups	7963.487	55	144.791		
	Total	8282.897	57			
Exam3	Between Groups	433.797	2	216.899	1.697	.193
	Within Groups	7027.927	55	127.780		
	Total	7461.724	57			
Final Grade	Between Groups	9007.469	2	4503.734	2.037	.140
	Within Groups	121597.135	55	2210.857		
	Total	130604.604	57			
Final Grade (Exam Only)	Between Groups	5407.612	2	2703.806	3.102	.053
	Within Groups	47946.474	55	871.754		
	Total	53354.086	57			

Table 2: ANOVA output for all three exams and the final grade distribution. Exam 1 was the only found to be statistically significant.

Table 3

Multiple Comparisons

Scheffe

Dependent Variable	(I) Section	(J) Section	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Exam1	Exp	Control	12.63043*	4.04035	.011	2.4651	22.795
		Choice	4.55882	4.34211	.579	-6.3657	15.483
	Control	Exp	-12.63043*	4.04035	.011	-22.7958	-2.4651
		Choice	-8.07161	4.10648	.155	-18.4033	2.2601
	Choice	Exp	-4.55882	4.34211	.579	-15.4833	6.3657
		Control	8.07161	4.10648	.155	-2.2601	18.403
Exam2	Exp	Control	2.47101	3.78671	.809	-7.0561	11.998
		Choice	-3.24510	4.06952	.729	-13.4838	6.9936
	Control	Exp	-2.47101	3.78671	.809	-11.9982	7.0561
		Choice	-5.71611	3.84868	.339	-15.3992	3.9670
	Choice	Exp	3.24510	4.06952	.729	-6.9936	13.483
		Control	5.71611	3.84868	.339	-3.9670	15.399
Exam3	Exp	Control	6.54348	3.55733	.194	-2.4066	15.493
		Choice	3.32353	3.82301	.687	-6.2950	12.942
	Control	Exp	-6.54348	3.55733	.194	-15.4935	2.4066
		Choice	-3.21995	3.61555	.675	-12.3165	5.8766
	Choice	Exp	-3.32353	3.82301	.687	-12.9420	6.2950
		Control	3.21995	3.61555	.675	-5.8766	12.316
Final Grade	Exp	Control	29.34879	14.79695	.150	-7.8795	66.577
		Choice	11.38562	15.90207	.775	-28.6231	51.394
	Control	Exp	-29.34879	14.79695	.150	-66.5771	7.8795
		Choice	-17.96317	15.03910	.494	-55.8008	19.874
	Choice	Exp	-11.38562	15.90207	.775	-51.3944	28.623
		Control	17.96317	15.03910	.494	-19.8744	55.800

Table 3: Scheffe's Post Hoc Analysis. Asterisks indicate statistical significance.

Table 4

Supplement Data from Sections 002 & 004

Statistic	Exp	Choice
Total user sessions:	338	163
Average user session length:	0:10:52	0:03:07
Average user sessions per day:	5	3
Average user sessions per day on weekdays:	5	4
Average user sessions per day on weekends:	5	2
Most active day:	2-Nov-09	21-Sep-09
Least active day:	12-Nov-09	19-Sep-09
Most active hour of the day:	2:00-3:00PM	11:00AM-12:00PM
Least active hour of the day:	4:00-5:00AM	5:00-6:00AM

Table 4: Data taken from WebCampus measuring activity within the course webpage for the entire course.

Table 5
Supplement Data: Sections 002 & 004

	Exam 1 (Exp)	Exam 1 (Choice)	Exam 2 (Exp)	Exam 2 (Choice)	Exam 3 (Exp)	Exam 3 (Choice)
Number of Students	18	2	16	5	16	5
Compliance (%)	100%	11.76%	88.89%	29.41%	88.89%	29.41%
Number of Retakes	27	0	24	1	16	1

Table 5: Data taken from WebCampus depicting the number of students that engaged with the supplement, number of retakes, and the percentage of compliance with the assignment for the experimental and choice groups.

Table 6

Social Validity Results: Section 002 & 004

Professionalism

1. The assignments & materials described in the syllabus were available when indicated.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Exp	12	9	0	0	0
Choice	12	6	0	0	0

2. I found the materials helpful for preparing for exams.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Exp	11	8	1	1	0
Choice	6	7	5	0	0

3. Course assistant(s) were helpful and courteous.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Exp	7	5	7	1	0
Choice	7	3	4	0	0

Supplemental Program

1. The supplemental modules ran smoothly and did not crash.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Exp	12	8	1	0	0
Choice	7	3	4	0	0

2. I had already mastered the content within the supplemental modules and did not need them.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Exp	0	4	8	9	0
Choice	1	4	3	4	2

3. The modules improved my basic math abilities and understanding of order of operations.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Exp	1	11	2	4	2
Choice	2	7	2	2	0

4. The appearance of the modules was attractive and easy to read/navigate.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Exp	7	13	1	0	0
Choice	5	5	3	0	0

5. The feedback component showed me my level of mastery of material before the exam.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Exp	9	7	4	1	0
Choice	4	5	4	0	0

6. I recommend that this supplemental program continue to be an assignment.

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Exp	8	8	3	1	0
Choice	4	4	5	0	0

Table 6: Social Validity survey and results for Experimental and Choice Group.

FIGURES

Figure 1

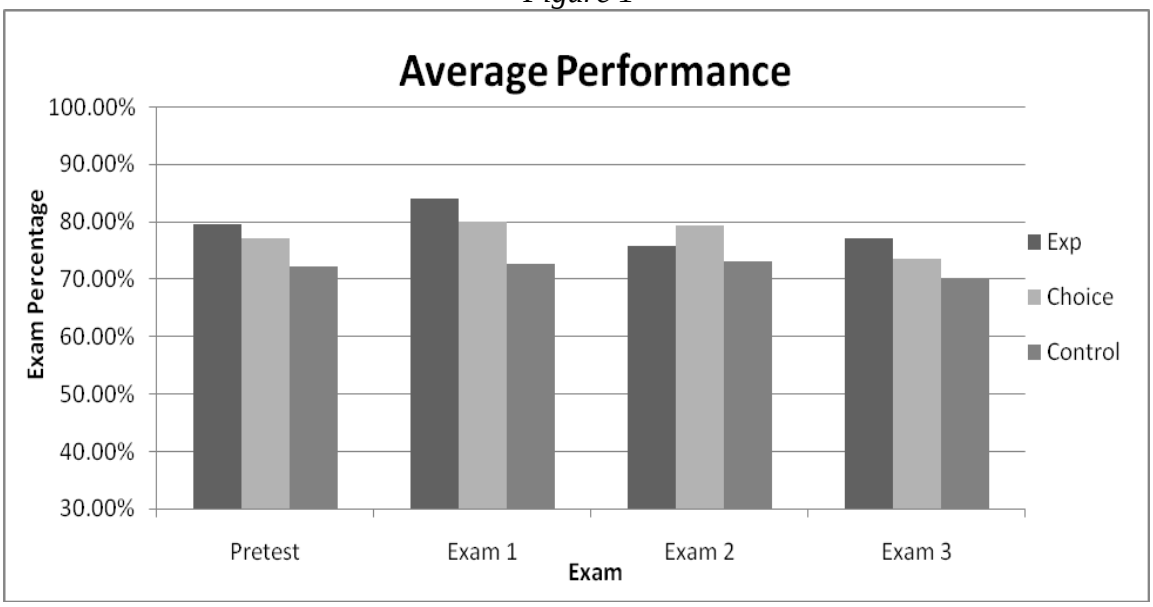


Figure 1: Pretest and exam averages from all three sections.

Figure 2

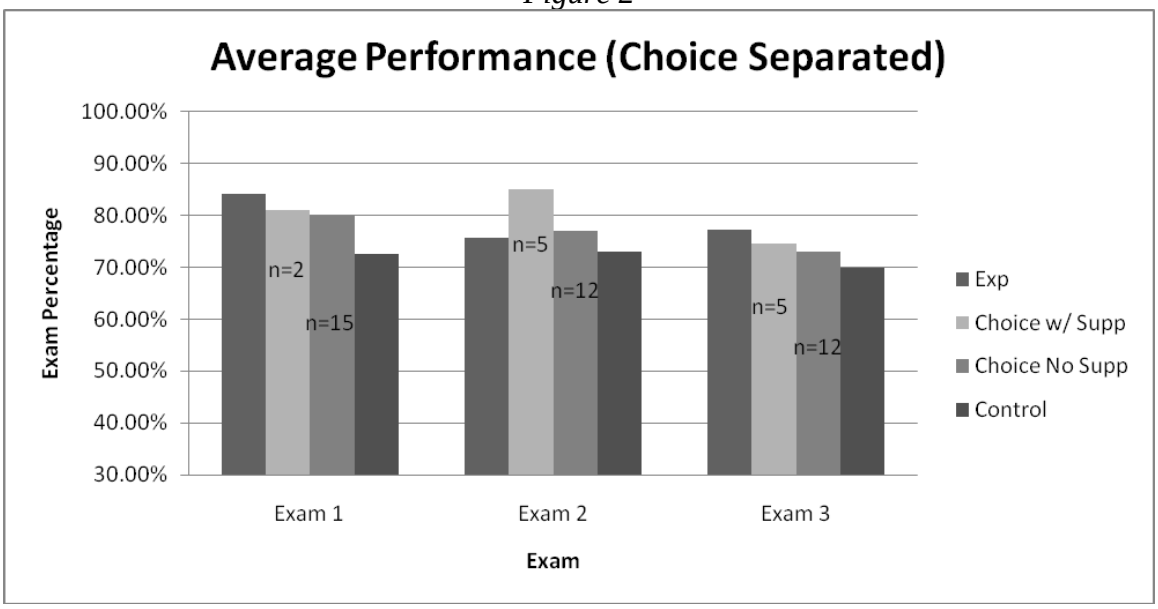


Figure 2: Exam averages from all three sections. Choice Condition is separated between students that accessed the supplement and students that did not.

Figure 3

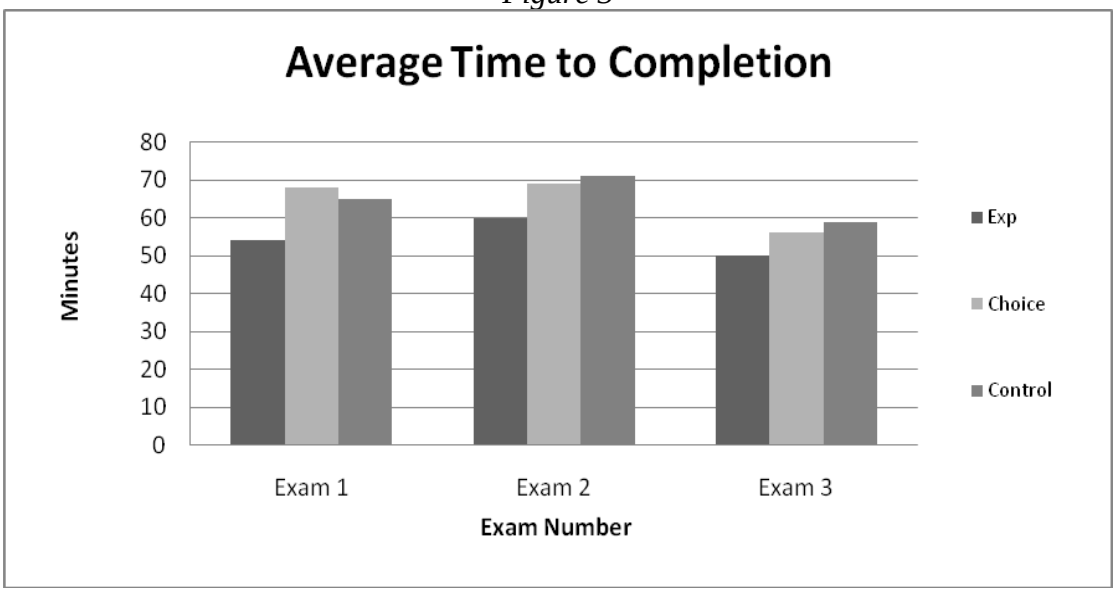


Figure 3: Average time to completion for all three sections. Time is measured in minutes.

Figure 4

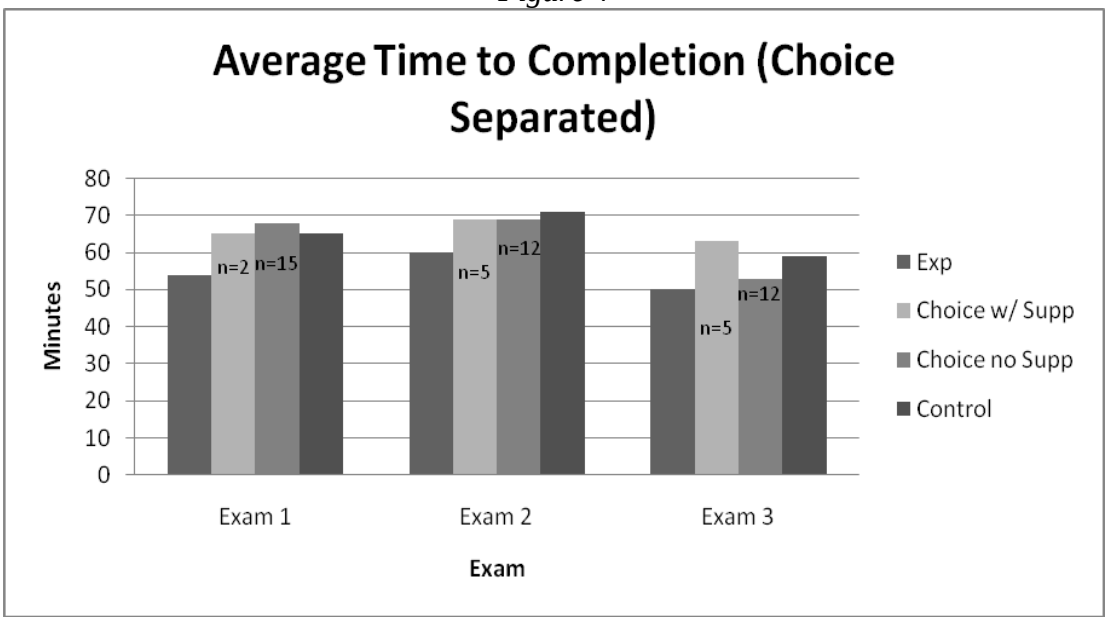


Figure 4: Average time to completion for all three sections. Choice condition is separated between students that completed supplements and students that did not. Time is measured in minutes.

Figure 5

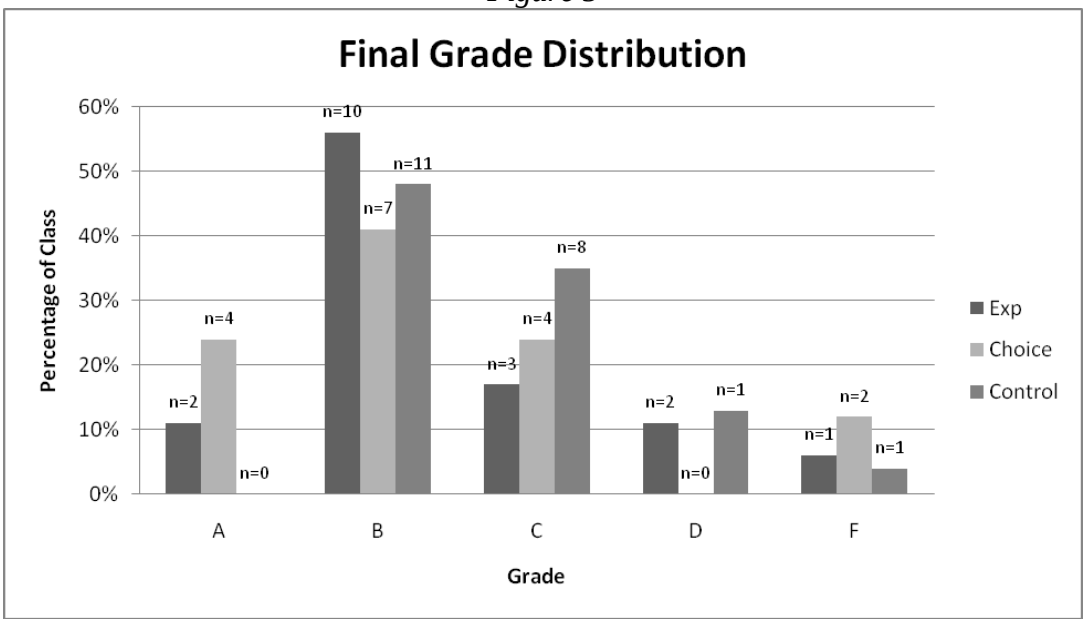


Figure 5: Final grade distributions for all sections. Final grades are computed based off all course materials.

Figure 6

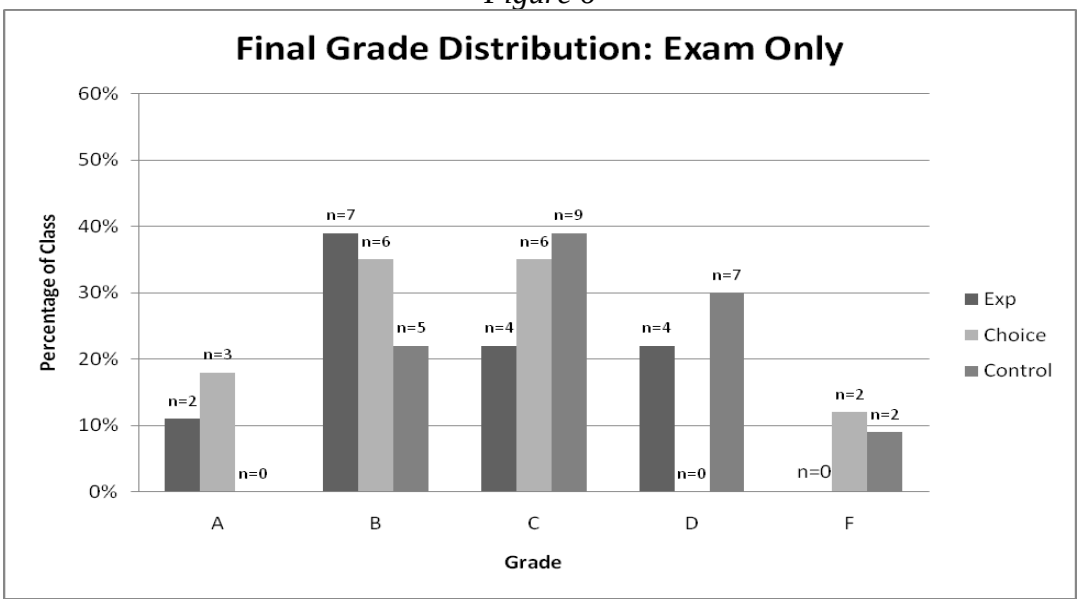


Figure 6: Final grade distributions for all sections using only exam scores to compute final grades.

Figure 7

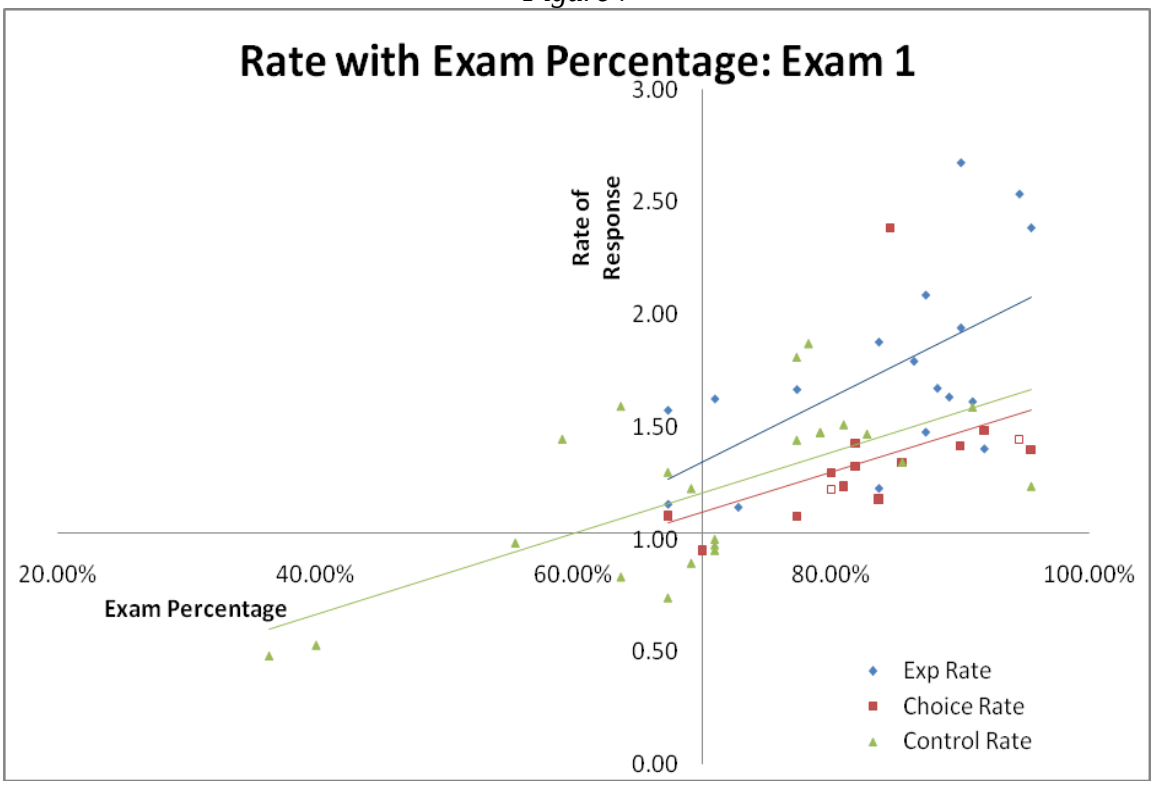


Figure 7: Rate of response with exam percentage for Exam 1. Red squares without fill indicate that a supplement was accessed in the Choice Condition. Pearson correlation coefficients (r) were equated and show a positive correlation between rate and accuracy (r=.587(exp), .585(choice), and .538(control)).

Figure 8

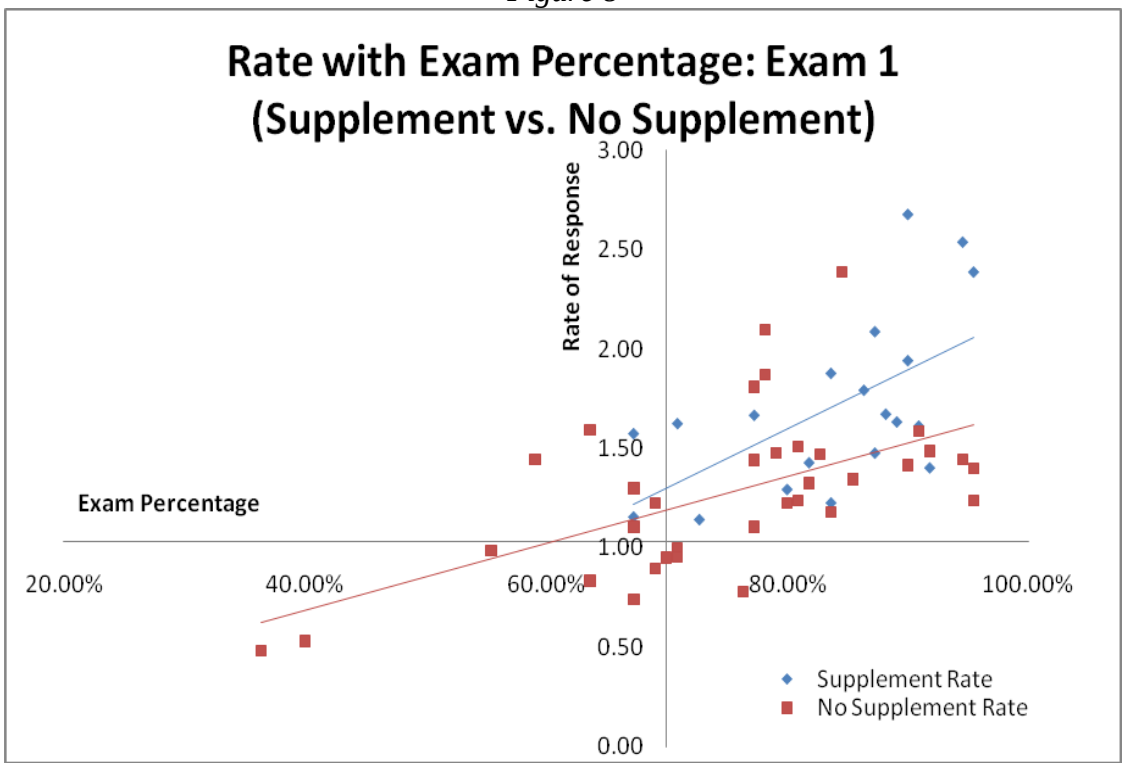


Figure 8: Rate with exam percentage for students who engaged with the supplement versus students that did not in Exam 1.

Figure 9

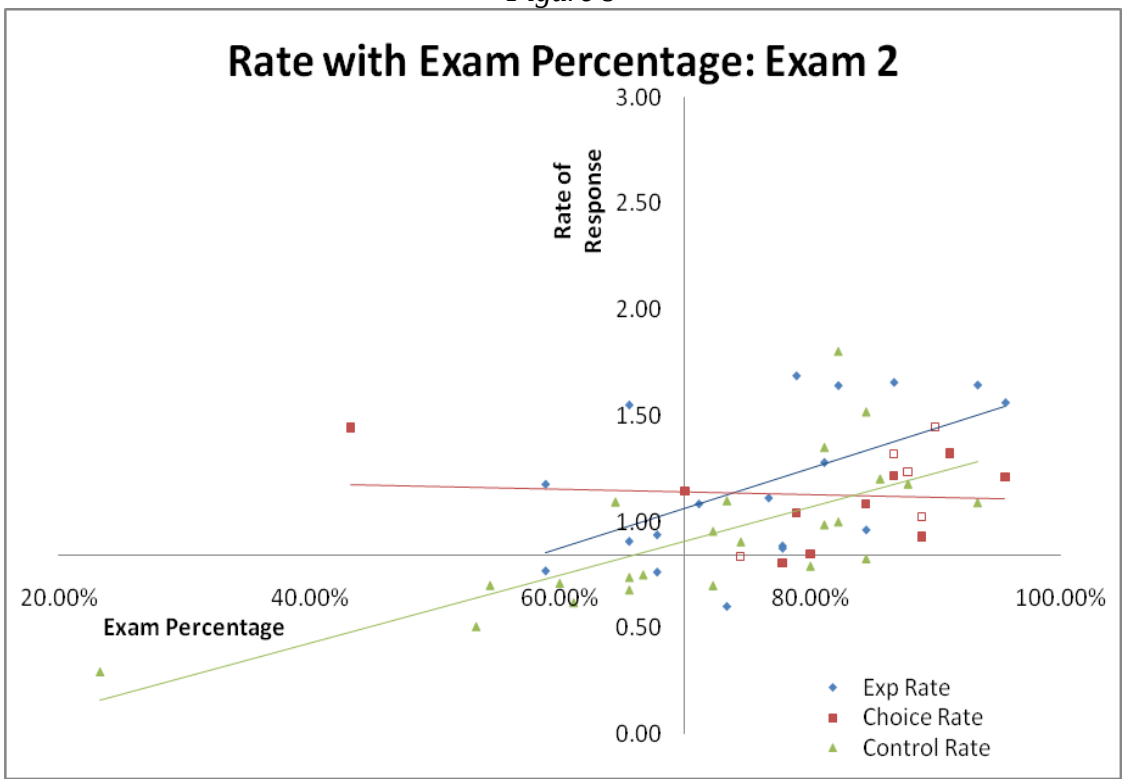


Figure 9: Rate of response with exam percentage for Exam 2. Red squares without fill indicate that a supplement was accessed in the Choice Condition. Pearson correlation coefficients (r) were equated and show a positive correlation between rate and accuracy (r=.574 (exp), .269 (choice), .706 (control)).

Figure 10

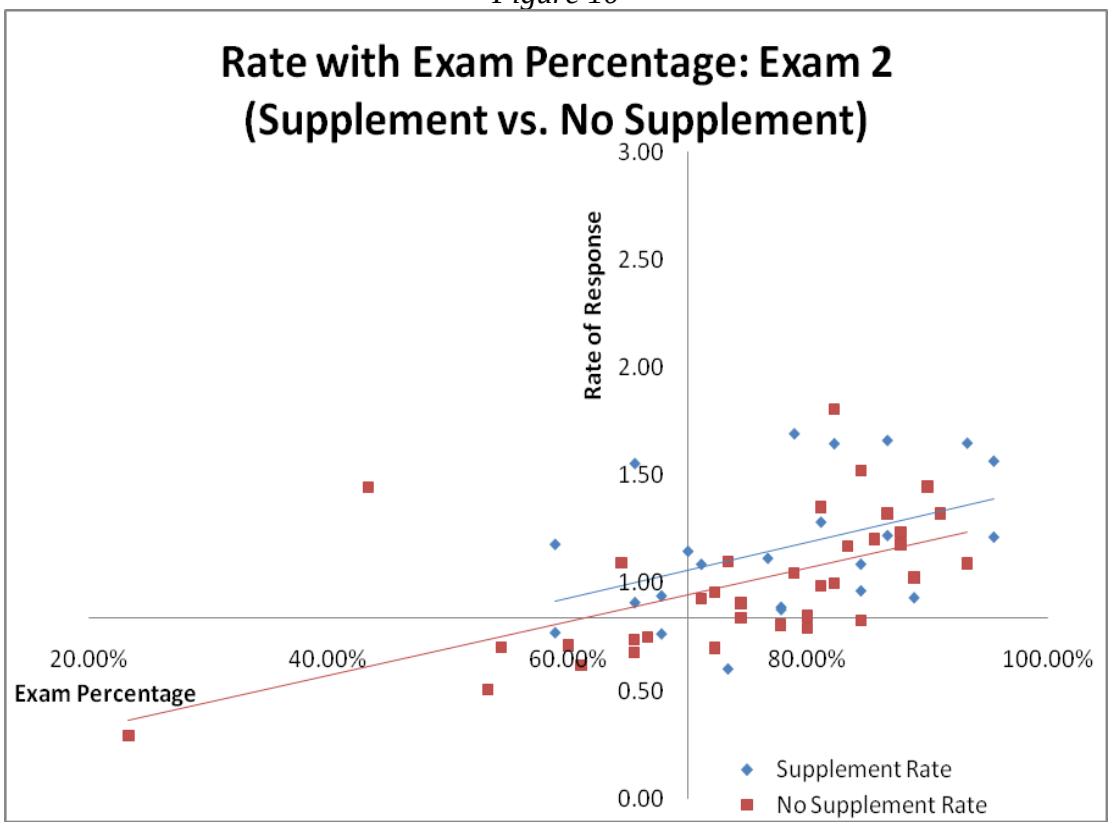


Figure 10: Rate with exam percentage for students who engaged with the supplement versus students that did not in Exam 2.

Figure 11

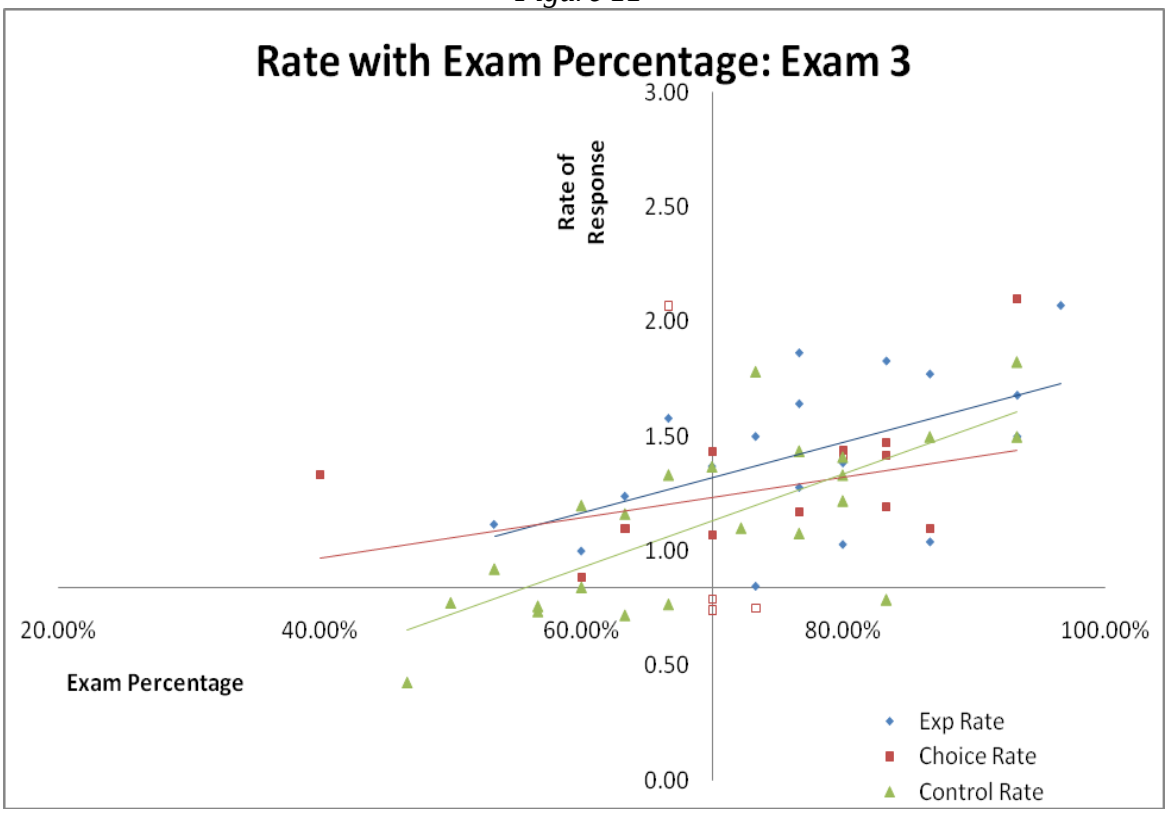


Figure 11: Rate of response with exam percentage for Exam 3. Red squares without fill indicate that a supplement was accessed in the Choice Condition. Pearson correlation coefficients (r) were equated and show a positive correlation between rate and accuracy (r= .526 (exp), .276 (choice), .734 (control)).

Figure 12

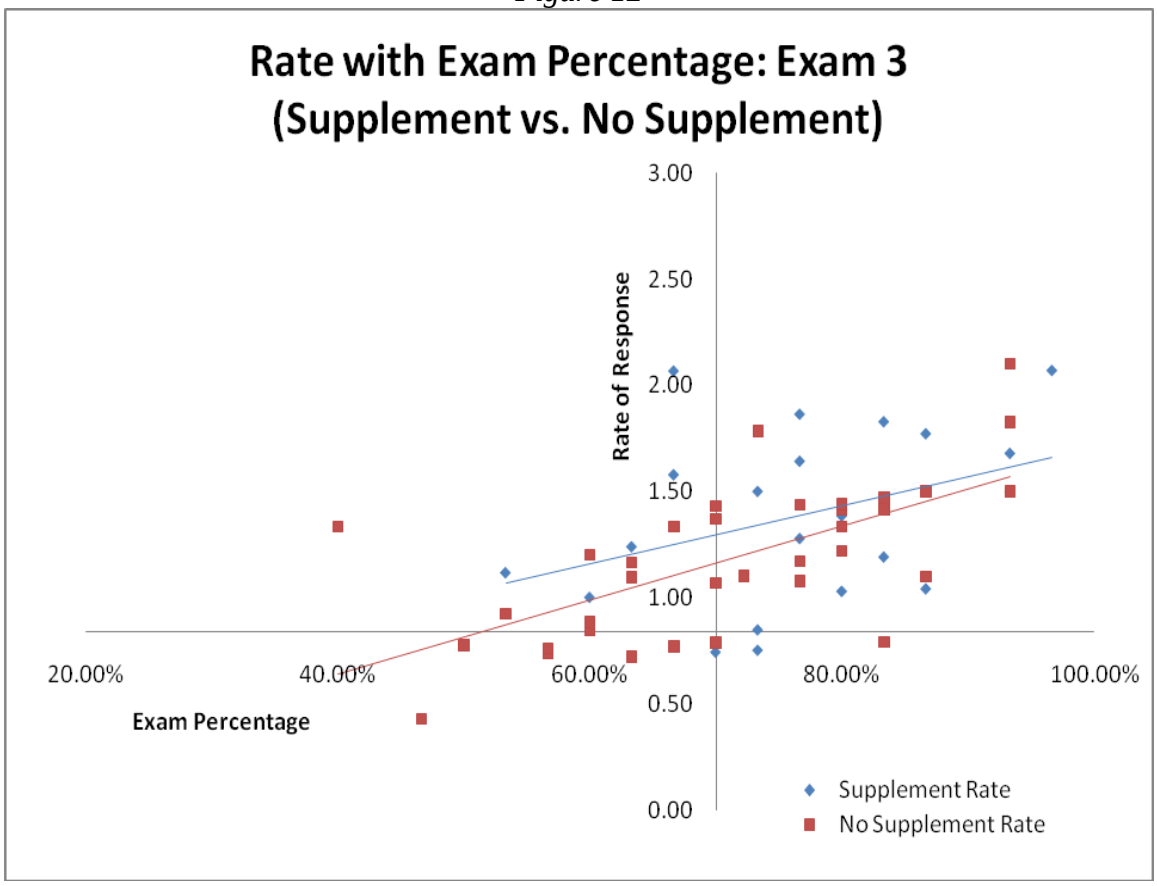


Figure 12: Rate with exam percentage for students who engaged with the supplement versus students that did not in Exam 3.

Figure 13

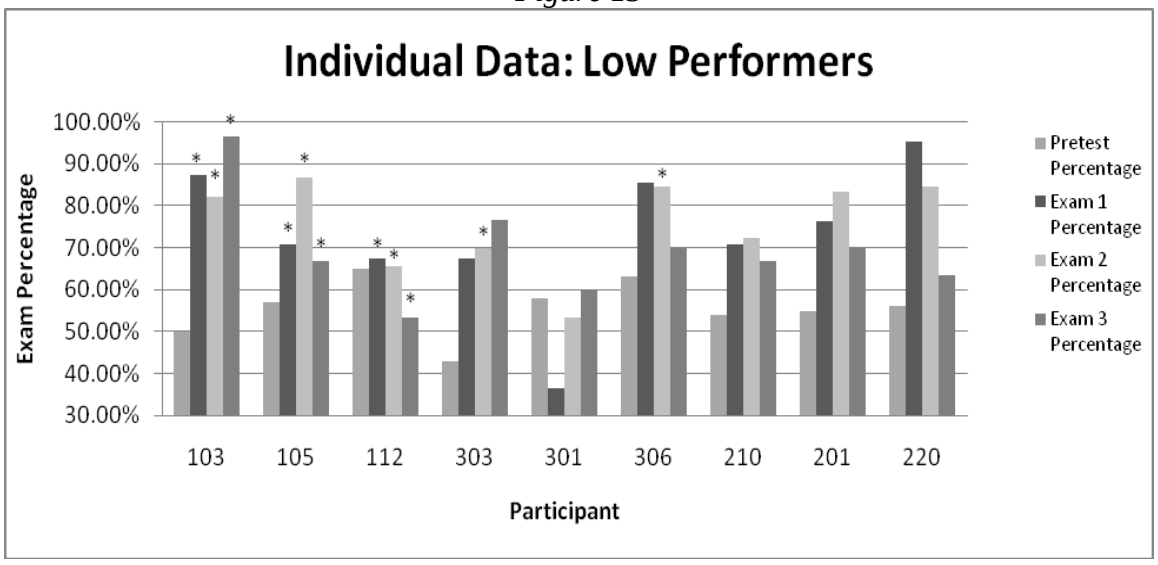


Figure 13: Individual data from low performers across all three sections. Asterisks above indicate that a supplement was completed. Comparison of exam averages for low performers across the three sections indicates an average of 75.17% for the experimental section, 67.06% for the choice, and 75.86% for the control.

Figure 14

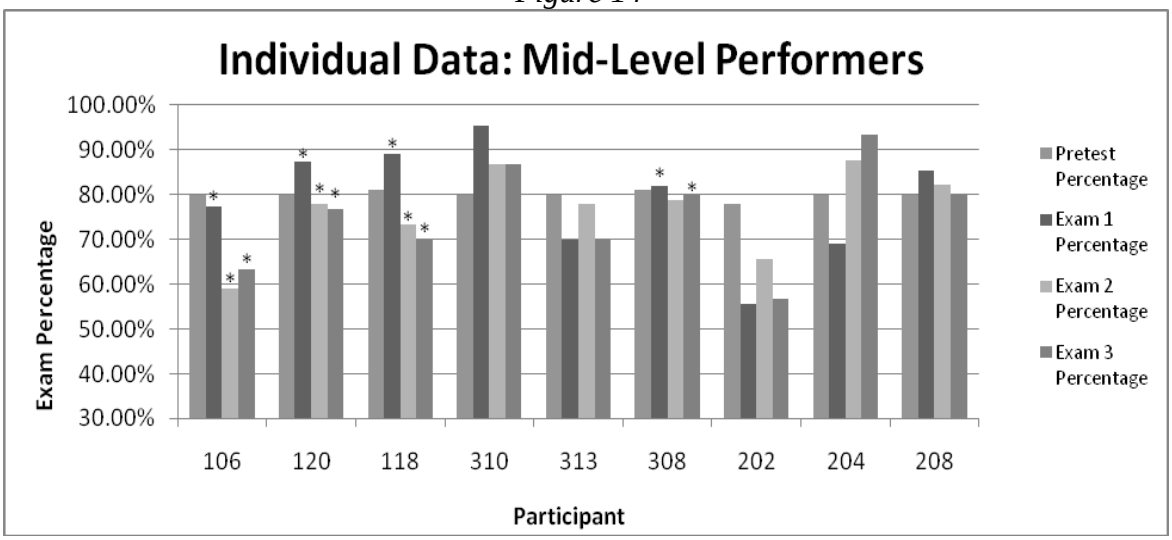


Figure 14: Individual data from mid-level performers across all three sections. Asterisks above indicate that a supplement was completed. Comparison of exam averages for mid-level performers across the three sections indicates an average of 74.85% for the experimental section, 80.80% for the choice, and 77.83% for the control.

Figure 15

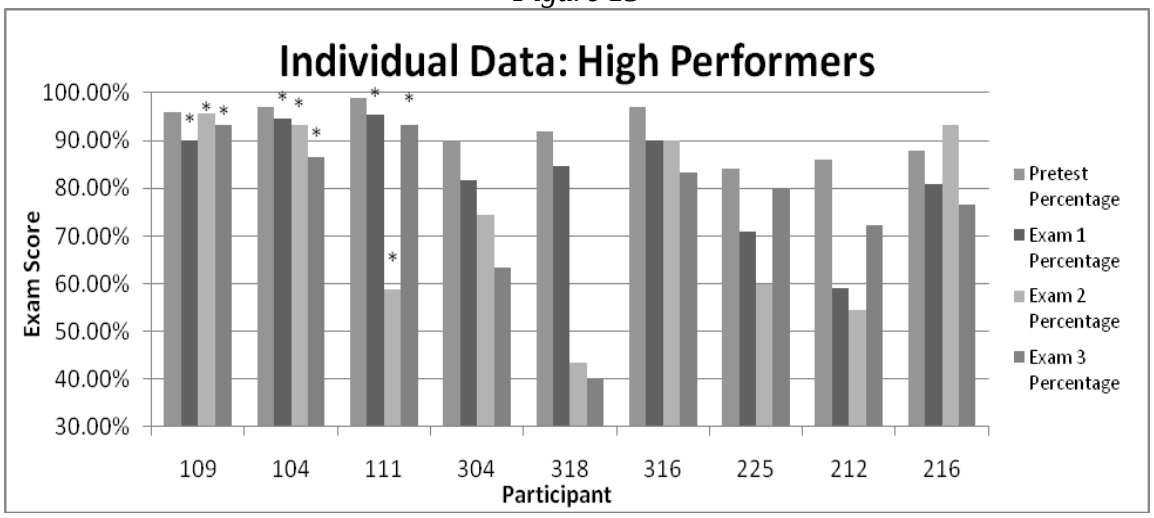


Figure 15: Individual data from high performers across all three sections. Asterisks above indicate that a supplement was completed. Comparison of exam averages for high performers across the three sections indicates an average of 79.31% for the experimental section, 78.57% for the choice, and 78.94% in the control.

Figure 16

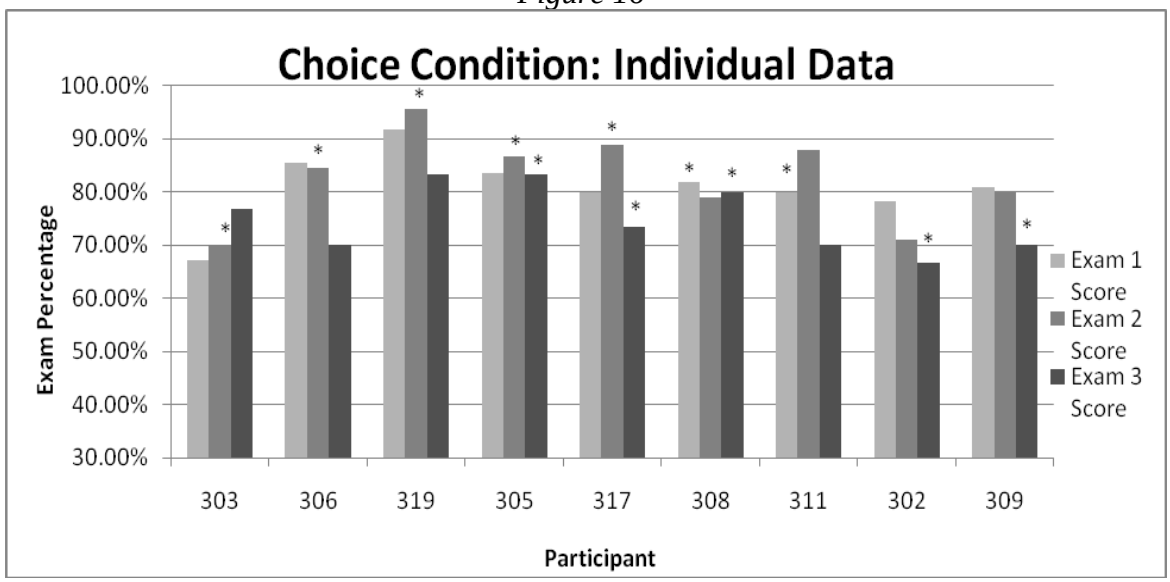


Figure 16: Individual data taken form the choice condition. Asterisks indicate whether a supplement was completed.

Figure 17

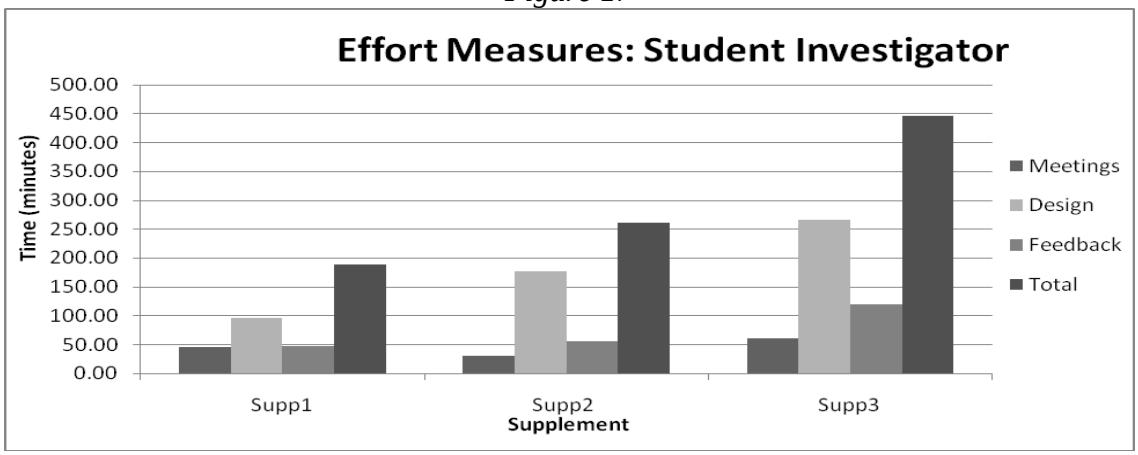


Figure 17: Self-reported effort measures.

APPENDIX A

UNIVERSITY OF NEVADA, RENO

Social Behavioral Institutional Review Board

CONSENT TO PARTICIPATE IN A RESEARCH STUDY

TITLE OF STUDY: Effects of web-based supplemental modules on exam performance in an undergraduate economics statistics course.

INVESTIGATORS: *Mark Alavosius, Ph.D. Principal Investigator*
Assistant Professor
Psychology Department
University of Nevada, Reno
(775) 682-8688

T. Wade Brown, Co-Investigator
Psychology Department
University of Nevada Reno

PROTOCOL NUMBER: SB08/09-172

PURPOSE

You are being asked to participate in a research study. The purpose of this study is to assess how web-based supplemental modules affect students' performance on exams in an undergraduate economics course on statistics. We wish to determine the effects of this individual training and assess how this component supports other class activities.

PARTICIPANTS

You are being asked to participate because as a student electing to take this Economics 262 course at the University of Nevada, Reno during the Fall 2009 term, you will be required to complete established course assignments designed to introduce you to inferential statistics as used in business and economics. Your mastery of the material will be tested, as is regular practice, during 3 exams**.** All instruction, including web-based supplements, are intended to support your success. As an undergraduate student at the University of Nevada, Reno, you are representative of all undergraduate students and we might generalize the results of this study to other students taking this course either here at UNR or at other universities.

PROCEDURES

Whether you choose to participate or not, you will be required to complete the established course requirements as specified in the course syllabus. These requirements include 3 examinations covering the course content and various exercises. Web-based supplements may help you study relevant materials and better prepare for exams.

The investigators will assess the effects of the supplements by examining students' test performance. Additionally, student satisfaction with the course content and procedures will be evaluated using the UNR course evaluations and a brief supplemental anonymous questionnaire surveying students' satisfaction with the supplemental training modules.

See the course syllabus for an itinerary of course activities. Note that participation in the study requires no additional work by you. All students who take this course are required to complete the course requirements within designated periods and your course grade is solely determined by your scores on the assignments. Your consent to participate merely allows the researchers to use your course grades to evaluate a component of the course.

RISKS

There are no known risks to working with web-based supplements during the course of an undergraduate class teaching business statistics. Deb Stiver, the course instructor, will not be informed of your choice about participation in the study and will submit grades based solely on your performance in the course requirements.

Your choice to participate in this study will not affect the grading procedures in any way. All reasonable efforts will be made to protect the confidentiality of student records. A slight and unknown risk of breach of confidentiality is a potential risk.

BENEFITS

There may be no direct benefits to you as a participant in this study. However, as instructors we hope to learn systematically how to improve our course and in subsequent classes we will incorporate what we learn about helping students master the course material. Benefits do **not** include payment/incentives to you as a participant in the study.

CONFIDENTIALITY

Your identity will be protected to the extent allowed by law. You will not be personally identified in any reports or publications that may result from this study**.** Dr. Mark Alavosius and T. Wade Brown Co-Investigators will have access to the data as will the UNR Social Behavioral Institutional Review Board. Course data, including all grades

will be stored securely in a locked office in the Mack Social Sciences building. The study involves no videotaping, audiotaping or photographs of the participants.

We might disseminate our results in professional meetings or journal publications. Your identity as a study participant will never be revealed and confidentiality of course data is assured. The Department of Health and Human Service, other federal agencies as necessary, and the University of Nevada, Reno Social Behavioral Institutional Review Board may inspect your study records

COSTS/COMPENSATION

There will be no cost to you nor will you be compensated for participating in this study.

RIGHT TO REFUSE OR WITHDRAW

You may refuse to participate or withdraw from the study at anytime without penalty. There will be no negative impact on your course requirements, grades, and instruction in this course. If the study design or use of the data is to be changed, you will be so informed and your consent re-obtained. You will be told of any significant new findings developed during the course of this study, which may relate to your willingness to continue participation.

QUESTIONS

If you have any questions, please ask us. If you have additional questions later, contact Dr. Mark Alavosius at MSS 403 (775) 682-8688 or T. Wade Brown at MSS 402.

You may report (anonymously, if you so choose) any complaints or comments regarding the manner in which this study is being conducted to the University of Nevada, Reno Social Behavioral Institutional Review Board at (775) 327-2368 or by addressing a letter to the Chair of the Board, c/o UNR Office of Human Research Protection, 205 Ross Hall / 331, University of Nevada, Reno, Reno, NV 89557.

CLOSING STATEMENT

I have read ( ) this consent form or have had it read to me ( ). [Check one.]

_____ has explained the study to me and all of my questions have been answered. I have been told of the risks or discomforts and possible benefits of the study.

If I do not take part in this study, my refusal to participate will involve no penalty or loss of rights to which I am entitled. I may withdraw from this study at any time without penalty.

I have been told my rights as a research subject, and I voluntarily consent to participate in this study. I have been told what the study is about and how and why it is being done. All my questions have been answered.

I will receive a signed and dated copy of this consent form.

Signature of Participant Date

Printed Name Date

Signature of Person Obtaining Consent Date

Signature of Investigator Date

APPENDIX B

College of Business Pre-Major

All students who are interested in graduating with a business degree start by being declared as "pre-majors." All business students must complete core curriculum requirements and the pre-business core.

Pre-Business Core

College of Business pre-majors are required to complete the following nine courses with a 2.75 GPA. Once you have completed the pre-business core, you will meet with an advisor to declare your business major.

ACC 201—Financial Accounting
ACC 202—Managerial Accounting
ECON 102—Principles of Microeconomics
ECON 103—Principles of Macroeconomics
ECON 261— Principles of Statistics I
ECON 262— Principles of Statistics II
IS 101—Introduction to Information Systems
MATH 176—Introductory Calculus
MKT 210—Marketing Principles

For a pre-business core GPA calculator, go to
http://www.coba.unr.edu/advisement/calc.htm

Pre-Business Core Course Descriptions
ACC 201 FINANCIAL ACCOUNTING
Purpose and nature of accounting, measuring business income, accounting principles, assets and equity accounting for external financial reporting.
Prerequisite: Sophomore standing.
ACC 202 MANAGERIAL ACCOUNTING
Forms of business organization; cost concepts and decision making; break-even analysis, fixed and variable costs, budgeting for internal reporting.
Prerequisite: ACC 201 and IS 101.

ECON 102 PRINCIPLES OF MICROECONOMICS
Introductory analysis of price determination, resource allocation, market structure, consumer behavior, producer behavior, market failure and government failure. Economic approaches to social issues and policy.
Co-requisite of MATH 126 or higher.

ECON 103 PRINCIPLES OF MACROECONOMICS
Introductory analysis of how markets and governments interact to determine an economy's output, employment, income, interest rate, wage, price, and trade balance levels over time.
Prerequisite: ECON 102.
ECON 261 PRINCIPLES OF STATISTICS I
Probability and major probability distributions; sampling theory; descriptive statistics; measures of central tendency and dispersion; index figures; time series.
Prerequisite: MATH 126; IS 101.

ECON 262 PRINCIPLES OF STATISTICS II
Statistical inference, estimation, hypothesis testing; simple linear regression and correlations; analysis of the variance.
Prerequisite: ECON 261.
IS 101 INTRODUCTION TO INFORMATION SYSTEMS
Introductions to microcomputers, computers in business and microcomputer software tools including word processors, spreadsheets and database management systems.

MKT 210 MARKETING PRINCIPLES
Objectives and policies of marketing managers as influenced by

marketing institutions, the functions performed and consumer wants and needs in a diverse culture.

(MATH 126 PRECALCULUS I)

Relations, functions, graphing; equations; linear, quadratic, polynomial systems; matrices and determinants; sequences, compound interest and amortization, binomial theorem; the complex numbers; logarithms; combinatorics.

This course DOES NOT satisfy the university core mathematics requirement.

Prerequisite: ACT score of 22 or SAT score of 520 or satisfactory score on readiness exam.

MATH 176 INTRODUCTORY CALCULUS FOR BUSINESS AND SOCIAL SCIENCES

Fundamental ideas of analytic geometry and calculus, plane coordinates, graphs, functions, limits, derivatives, integrals, the fundamental theorem of calculus, rates, extrema and applications thereof.

This course satisfies the university core mathematics requirement.

Prerequisite: ACT score of 27 or SAT score of 610 or satisfactory score on readiness exam.

MATH 176 INTRODUCTORY CALCULUS FOR BUSINESS AND SOCIAL SCIENCES

Fundamental ideas of analytic geometry and calculus, plane coordinates, graphs, functions, limits, derivatives, integrals, the fundamental theorem of calculus, rates, extrema and

APPENDIX C

SYLLABUS – ECON 262
Principles of Statistics II
Spring Semester, 2009

Instructor:	Debra Stiver	**Office:**	ABB 319G
Phone:	784-1124	**Fax:**	784-4728
E-mail:	stiver@unr.edu		

Required Prerequisites: Econ 261 is required prior to course enrollment. If you have not completed Econ 261, you should not be in this class.

During class ► Cell phone, blackberries, blueberries, raspberries etc. OFF !!!

Office hours: Tues-Thurs: 1:30 to 2:15 p.m.
Thurs: 9:45 to 10:45 a.m.
Wednesday: 2:30 to 3:45 p.m.
Monday: 5:30 to 6:30 p.m.
You may also make an appointment.

Class Schedule: ECON 262.002: TR 2:30 to 3:45 p.m. WBR 2006
ECON 262.004: TR 11 a.m. to 12:45 p.m. EJCH 205
ECON 262.005: MW 4 p.m. to 5:15 p.m. AB 101
BADM 700: MW 7 to 9:45 p.m. AB 201

***Business Statistics for Contemporary Decision Making*, 5th edition by Ken Black.** New texts include both a data CD and a student version of a Minitab CD. (The Minitab CD will be discussed in class). If you have purchased a used text, it may or may not have the data CD and will not have the Minitab CD that comes with the original package. It is important that you can access the material on the data CD because we will be using many of the problem sets found on the disk. If you do not have the data CD, you may access the data sets on the Internet or the COBA K-drive. There will be assigned readings from other sources during the semester.

A Few Words On Course Objectives and Goals: The College of Business Administration has identified 5 specific learning goals and several objectives for students seeking a business degree. The goals/objectives are as follows:

1. Learning Goal 1: Students will demonstrate mastery of fundamental business knowledge and practices.
2. Learning Goal 2: Students will possess analytical and critical thinking skills - **Objective 1: Students will demonstrate ability to apply key statistical tools.**

3. Learning Goal 3: **Students will effectively communicate information, concepts and ideas in writing.**
4. Learning Goal 4: **Students will be able to use information technology to support business analysis and operations.**
5. Learning Goal 5: Students will recognize, evaluate and resolve ethical issues, both personally and organizationally.

All of the above goals relate to the practice of statistical analysis, which is used extensively to aid in the decision-making process. In this second statistics course, you will be studying the field of statistics known as inferential statistics. The class is designed to familiarize students with techniques used to analyze data and gain insight into specific questions about populations. One of the major differences between EC261 and EC262 is the amount of emphasis placed on analysis in EC262. Most often, data is collected, tested and analyzed because it has managerial and/or research implications. There is plenty of math in the course, but statistics is not a math course. Statistics is more appropriately described as quantitative methods course that enhances one's ability to make decisions. The material in this course will be "applied"; that is, you will learn how to apply the appropriate statistical techniques to data with the intention of improving your decision making ability. As a conductor or user of statistical inference and hypothesis testing, you will be writing and verbalizing results to address managerial or research implications. Part of your grade will depend upon how well you are able to verbalize the analysis.

Computer Use: Due to the amount of data and complexity of math involved in this class, much of your computational work will be done using computer software. Both Minitab and Excel will be used in class. Students generally select one or the other to do the majority of their statistical analysis. However, you are not limited to using these packages. There are many statistical software packages available, and you may use another package if you are comfortable with that package. Please check with me if using different software, so that the applications may be assessed. Every enrolled statistics student will receive their own personal login and password. The login and password are assigned by the computer lab folks. You will be given instructions on accessing your login and password sometime early in the semester.

Steps to access class materials from any Internet browser:

1. Go to ► http://cobalabs.coba.unr.edu/
2. Select …. COBA Virtual Server – Access to K Drive Files (on the left side of the screen)
3. Select ….. Click Here to Launch CVSF……. Enter your UNR netID and password
4. Select K Drive (found on left hand side under Folders Tree)
5. Select Econ 262 on right hand side, then select Stiver folder.

Note ► There are 3 different file types from the Black text. The problem files contain each of the chapter exercises. The database files are for problems identified as database files. The case files are for end of chapter cases identified as cases.

Homework and Weekly Assignments: Weekly assignments (problems, cases, quizzes) will be given throughout the semester. These assignments gauge your daily progress in the course and will cover material presented in class or supplemental reading. Please note that **LATE WEEKLY ASSIGNMENTS WILL NOT BE ACCEPTED UNDER ANY CIRCUMSTANCES**. If you cannot attend class the day the assignment is due (professor frowns on this), you may e-mail or fax the assignment (attention to Stiver), or mail the assignment if out of town. Classmates may also bring the assignment to class. **If the answers to an assigned problem require calculations,**

please make sure that you show the required work for credit. Similarly, if an assignment requires the use of computer software, make sure that you turn in a hard copy of computer results with the assignment. Exams will have take-home problems that will require the use of the computer for calculations. Don't wait until the last minute to get on the computer and figure things out!

MAKEUP POLICY: ►THERE ARE NO MAKEUPS FOR WEEKLY ASSIGNMENTS/QUIZZES! ◄ Regardless of excuse – period. To mitigate this strict policy, I will drop the lowest weekly assignment grade, or alternately, you may skip one assignment during the semester. Exams can be made up IF I am contacted prior to the exam. I give one makeup exam after the second exam. This makeup exam is given to any student missing either the first or second exam and is comprehensive to the point of the makeup exam date. Please contact me if you know you are not going to be present on the day of a scheduled exam. I will make arrangements for you to take the exam early, depending upon the circumstances of your absence.

How Your Grade Will Be Determined: Grading is objective – Score Those Points!!!
Note: Exams may have take-home preparations given a week in advance.

First Exam:	110 points
Second Exam:	90 points
Average of Weekly Assignments:	100 points
Third Exam (during finals week):	90 points
Case Study/Project:	30 points
Total Points Possible:	420 points

Students often ask if I curve the grading scale. Generally the class average in ECON 262 falls between 75% to 80%. If the average falls in this range, there will be a standard curve (A - 90%, B - 80%, C - 70%, C - 60%). If the average is lower than this range I may adjust the curve downward. There will be no "extra credit" in this class.

****** Special Note to Students: Econ 262 is one of the required courses for the lower division business core curriculum. Students wishing to declare a business major and take upper division courses must satisfy the 2.75 minimum GPA requirement in their lower division core courses. BE DILIGENT to avoid repeating courses.**

Class Expectations: The university classroom is an adult setting. Since the course can be challenging for many students, it is important to maintain an atmosphere conducive to learning and concentration. Walking in late can be very disruptive, not only to the instructor, but also to students. If you are late, enter quietly through the back or side door and take the first seat available. If you must leave early, take a seat next to the door. Please take care of necessities like getting drinks and using the restroom prior to the beginning of class. Lively discussions are encouraged - just remember to extend the same courtesies you expect.

Academic Dishonesty: UNR provides sanctions for academic dishonesty. It will not be tolerated in this class. Please see the following from the **University Code of Conduct and Policies**

SECTION IV: ACADEMIC STANDARDS

Subsection A: Definitions

Academic dishonesty is against university as well as the system community standards. Academic dishonesty includes, but is not limited to, the following:

Plagiarism: defined as submitting the language, ideas, thoughts or work of another as one's own; or assisting in the act of plagiarism by allowing one's work to be used in this fashion.

Cheating: defined as (1) obtaining or providing unauthorized information during an examination through verbal, visual or unauthorized use of books, notes, text and other materials; (2) obtaining or providing information concerning all or part of an examination prior to that examination; (3) taking an examination for another student, or arranging for another person to take an exam in one's place; (4) altering or changing test answers after submittal for grading, grades after grades have been awarded, or other academic records once these are official.

Disability Statement

Any student with a disability needing academic accommodations is requested to speak with me or contact the Disability Resource Center (Thompson Building, Suite 101), as soon as possible to arrange for appropriate accommodations.

READING, CLASS, AND EXAM SCHEDULE

SECTION I: INFERENCE AND HYPOTHESIS TESTING. In we examine the basics of estimation and hypothesis testing – the bread and butter of statistical research. We move beyond describing populations and now ask and answer specific questions. Sampling is required to answer the question and the statistical techniques introduced in this section address how we take a sample and apply the results to the population in general.

Week of:	Reading Assignment:
Jan 19	Monday, Jan 19 MLK Day – No Classes
	Review selected portions of chapter 3, 6, and 7. Chapter 8, pages 297 – 312.
Jan 26	Chapter 8, pages 314 – 323.
	Chapter 9, pages 336 – 355.
Feb 2	Chapter 9, pages Chapter 10: Selected portions as time allows.
Feb 9	Continue hypothesis testing.
Feb 16	Wrap-up and review hypothesis tests.
	First Exam: Thursday, Feb 19

SECTI II: CHI-SQUARE TESTS and ANOVA. We look at hypothesis tests that investigate specified distributions and relationships among categorical variables in Chi-Square techniques. In ANOVA we get our first good look at experimental designs and techniques involving multiple populations.

Week	Reading Assignment:
Feb 23	Chapter 12, pages 457 to 468.
Mar 2	Chapter 13, pages 490 to 502.
Mar 9	Chapter 13, additional material as time allows.

Mar 16	Spring Break Yippee – No Classes
Mar 23	Wrap-Up and Review **Second Exam: Thursday, Mar 26**

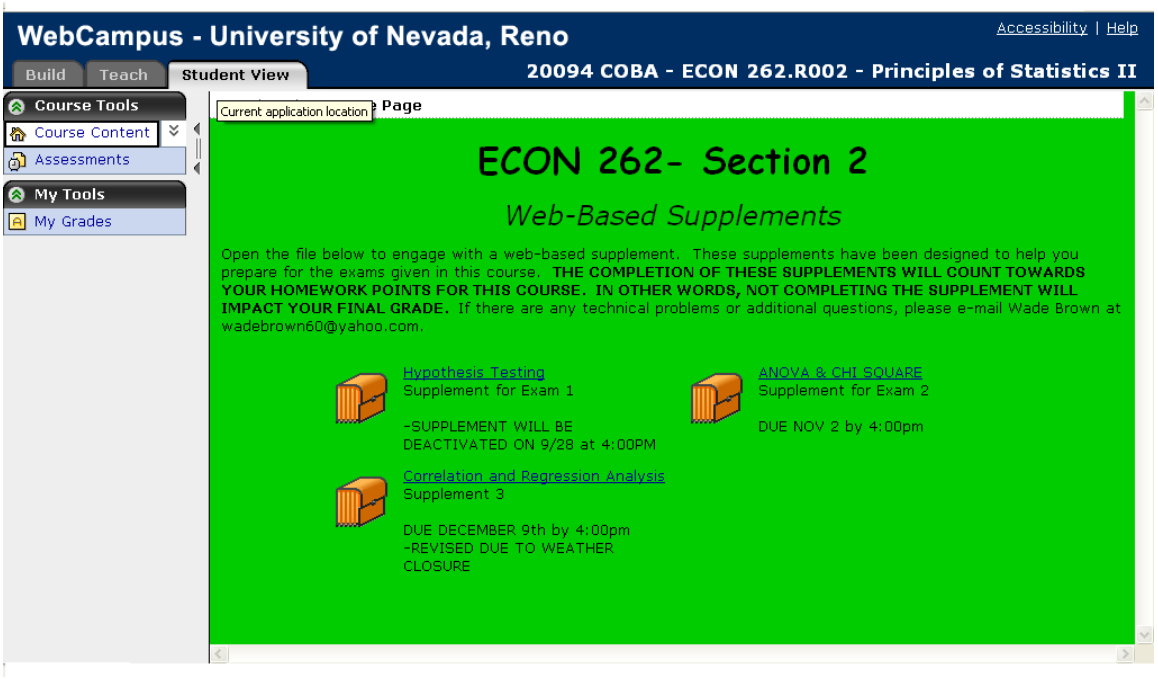
SECTION III: CORRELATION AND REGRESSION ANALYSIS. The techniques introduced in this section include some of the most important multivariate techniques in statistics. Government, business and economics researchers investigate the relationships among variables using regression and correlation techniques. Many of the previous concepts covered in class are used in regression and the related assessments associated with regression. You will learn how to develop mathematical models that describe and forecast.

Week	Reading Assignment:
Mar 30	Chapter 3, pages 105 to 112 and Chapter 14, page 572 and 573.
Apr 6	Chapter 14, pages 553 to 573.
Apr 13	Chapter 14, selected additional portions.
	Chapter 15, pages 634 to 665.
Apr 20	Multiple Regression Continued – selected portions from Chapter 16.
Apr 27	Multiple Regression Continued.
May 4	Last Day of Class: Wrap-Up and Review
	Final Case Study due Tuesday, May 5
May 6	Prep Day

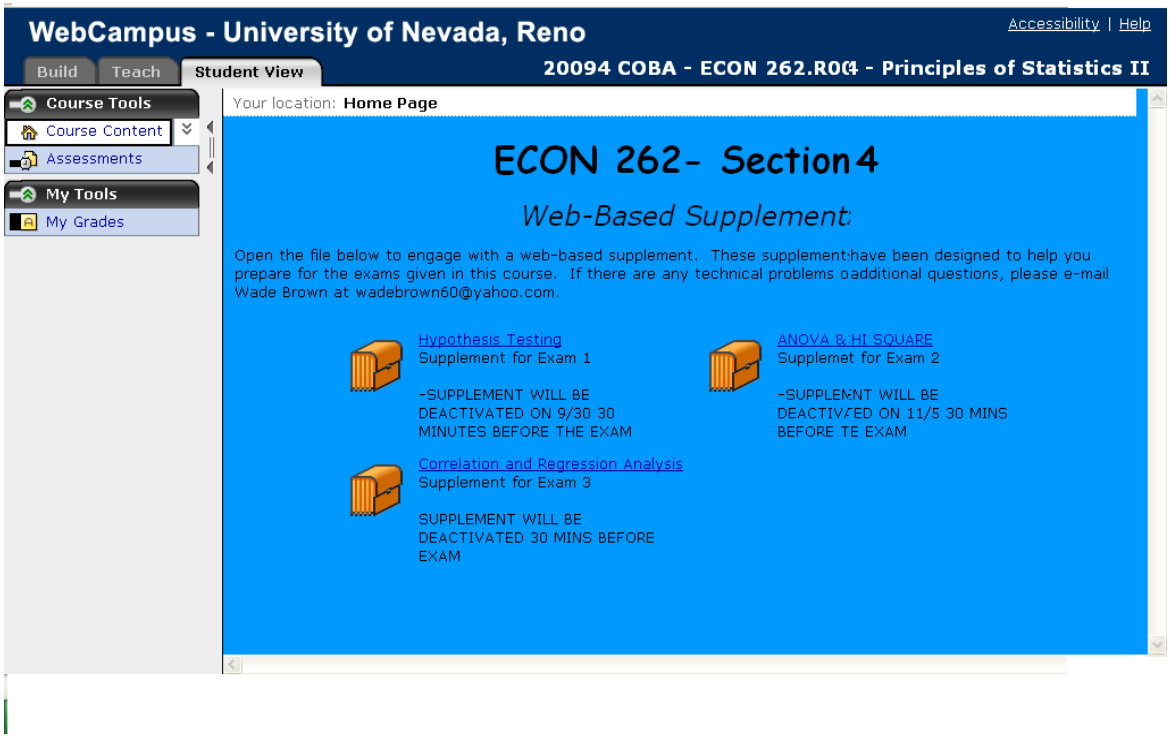
Last Exam (TR 11 a.m. Class), Thursday, May 7 at 9:45 to 11:45 a.m.

Last Exam (TR 2:30 p.m. Class) Thursday May 7 at 2:15 to 4:15 p.m.

APPENDIX D


WebCampus - University of Nevada, Reno
Accessibility | Help
Build
Teach
Student View
20094 COBA - ECON 262.R002 - Principles of Statistics II
Course Tools
Course Content
Assessments
My Tools
My Grades
Current application location
Page
ECON 262- Section 2
Web-Based Supplements
Open the file below to engage with a web-based supplement. These supplements have been designed to help you prepare for the exams given in this course. THE COMPLETION OF THESE SUPPLEMENTS WILL COUNT TOWARDS YOUR HOMEWORK POINTS FOR THIS COURSE. IN OTHER WORDS, NOT COMPLETING THE SUPPLEMENT WILL IMPACT YOUR FINAL GRADE. If there are any technical problems or additional questions, please e-mail Wade Brown at wadebrown60@yahoo.com.
Hypothesis Testing
Supplement for Exam 1
-SUPPLEMENT WILL BE DEACTIVATED ON 9/28 at 4:00PM
ANOVA & CHI SQUARE
Supplement for Exam 2
DUE NOV 2 by 4:00pm
Correlation and Regression Analysis
Supplement 3
DUE DECEMBER 9th by 4:00pm
-REVISED DUE TO WEATHER CLOSURE

APPENDIX E


WebCampus - University of Nevada, Reno
Accessibility | Help
Build
Teach
Student View
20094 COBA - ECON 262.R04 - Principles of Statistics II
Course Tools
Course Content
Assessments
My Tools
My Grades
Your location: Home Page
ECON 262- Section4
Web-Based Supplement:
Open the file below to engage with a web-based supplement. These supplement:have been designed to help you prepare for the exams given in this course. If there are any technical problems oadditional questions, please e-mail Wade Brown at wadebrown60@yahoo.com.
Hypothesis Testing
Supplement for Exam 1
-SUPPLEMENT WILL BE DEACTIVATED ON 9/30 30 MINUTES BEFORE THE EXAM
ANOVA &:HI SQUARE
Supplemet for Exam 2
-SUPPLEMNT WILL BE DEACTIVATED ON 11/5 30 MINS BEFORE TE EXAM
Correlation and Regression Analysis
Supplement for Exam 3
SUPPLEMENT WILL BE DEACTIVATED 30 MINS BEFORE EXAM

APPENDIX F

ECON 262 4 p.m. MW First Midterm NAME:

ANSWER ALL OF THE QUESTIONS ON THE EXAM. SHOW WORK, WHERE NECESSARY, TO RECEIVE PARTIAL CREDIT.

PART I. The questions in this part are based on the take-home preparation involving starting salaries for State U recruits.

A. What type of sample was taken in this problem?

Answer: _____

B. What is the sample estimate of the difference between the two means?

Answer: _____

C. Briefly interpret the value in part B above. Be specific.

D. What are the appropriate hypothesis statements for the test to determine if differences exist in the mean starting salary for State U recruits versus the mean starting salary for non State U recruits.

E. Write an appropriate conclusion for the test.

Part III. Suppose prior elections in a certain state indicated it is necessary for a candidate for governor to receive more than 70 percent of the vote in the northern section of the state to be elected. The incumbent governor is interested in assessing her chances of returning to office and plans to conduct a survey of 1,200 registered voters in the northern section of the state. A phone survey of the registered voters indicated 72.5 % would vote for the incumbent.

You are the statistician on staff and your task is to determine if the sample evidence indicates more than 70 percent of the registered voters in the north support the incumbent governor. You will perform the test at alpha = .05.

Circle the correct answer.

1. What is the appropriate alternative hypothesis for this test?

A. $P \geq .70$
B. $P = .70$
C. $P > .70$
D. $P > .725$

2. If the test is conducted using alpha = .01, what is the correct critical value (using the test statistic method) for the decision rule?

A. ± 1.96
B. 1.645
C. 2.33
D - 1.645

3. What does the standard error equal for this test?

A. .0132
B. ± .025
C. .045
D. ± .052

4. What is the appropriate conclusion for the test?

A. The proportion of registered voters supporting the incumbent is greater than 70%.
B. The proportion of registered voters supporting the incumbent is equal to 72.5%
C. The proportion of registered voters is not significantly more than 70%.
D. The proportion of registered voters is less than 70%.

5. The 99% lower bound for the proportion of voters who would vote for the governor is

Part III. A market research firm used a sample of individuals to rate the purchase potential of a particular product before and after the individuals saw a new television commercial about the product. The purchase potential ratings were based on a 0 to 10 scale, with higher values indicating a higher purchase potential. The researchers wished to determine if on average, there was an increase in the purchase potential ratings after the individuals saw the commercial. The data for the study is shown below.

	Purchase Rating	Purchase Rating
Individual	After	Before
1	6	5
2	6	4
3	7	7
4	4	3
5	3	5
6	9	8
7	7	5
8	6	6

The data was input into Minitab for analysis. The computer output for the analysis is as follows:

```
After - Before

            N   Mean  StDev  SE Mean
After       8  6.000  1.852    0.655
Before      8  5.375  1.598    0.565
Difference  8  0.625  1.302    0.460

95% CI for mean difference: (-0.464, 1.714)
T-Test of mean difference = 0 (vs not = 0): T-Value = 1.36  P-Value = 0.217
```

Answer the following questions for this test.

A. What type of sample was taken in this test?

Answer: _____

B. What is the appropriate null hypothesis for this test?

A. $\mu 1 - \mu 2 = 0$
B. $\mu 1 - \mu 2 \neq 0$
C. $D \leq 0$
D. $D \geq 0$

C. Were the results of the test statistically significant?

A. Yes, because the null hypothesis is rejected.
B. No because the null hypothesis is not rejected.
C. Yes, because it's very important for the company to know about product ratings.
D. No, because the means for the first sample does not equals the mean for the second sample.

Part IV. Most airlines allow passengers to carry luggage onto a plane. However, their studies show that the more carry-on bags passengers have, the longer it takes the plane to load and unload passengers. This is particularly important since tighter security procedures are now required. Because of falling revenue some airlines (like U.S. Air), began charging customers to check-in their luggage. This created a great deal of customer dissatisfaction so airlines have began dropping the charge for check-in luggage. These airlines are concerned, however, that dropping the charge will lead to too many carry-on bags and delay boarding time.

One airline will reinstate the charge if it determines the average number of bags per passenger exceeds 1.5 bags. The known standard deviation for number of carry-on bags from historical information = .5 bag. A random sample of 100 passengers was selected. The passengers were observed and the number of bags carried on the plane was noted. Out of the 100 passengers, the mean number of bags carried on = 1.6.

A. Perform the appropriate test using the P-VALUE method. Use alpha = .05. Be sure to state the appropriate hypotheses, decision rule, decision and conclusion.

B. Were the results of your test significant? Why or why not? (Be brief!)

C. Define a Type II error for this problem. What would be a possible consequence of committing a Type II error?

Part III. A university campus offers several different types of parking permits. The price differential between the silver zone and green zone permits is approximately $200 per year. The campus parking office has indicated the price differential reflects the difference in zone proximity to buildings, but does not reflect any differences in the number of available parking spaces per

permit holder. The student government on campus believes that the parking office oversells green zone permits so that it is very difficult to find a parking space in those areas during busy times of the day. They feel it may be intentional so that students will purchase the more expensive silver zone permits. The parking office will not release the statistics to show if the number of spaces per permit are equal in the two zones or whether the average number of spaces in the green zone have decreased since last year. However, the student government has a "mole" in Parking Services who discovered that last year's mean number of available spaces in the green zone = 12.8. To test whether the average number of spaces in the green zone has decreased, the student government conducts a survey to determine the number of available spaces in the green zone during a busy time of the day on campus. Students observe the number of empty spaces for 20 randomly selected days. The data was input into the computer with the following results:

N	Mean	StDev	95% Upper Bound	T
20	10.200	2.400	11.128	-4.84

1. What is the appropriate <u>alternative</u> hypothesis for this test?

E. $\mu < 10.2$
F. $\mu \geq 12.8$
G. $\mu < 12.8$
H. $\mu \neq 12.8$

2. If the test is conducted using alpha = .05, what is the correct critical value for the decision rule using <u>the test statistic method</u>?

A. 2.400
B. ± 2.093
C. 1.729
D - 1.729

3. What does the standard error equal for this test?

A. .537
B. ± 2.093
C. 2.400
D. ± .928

4. What is the appropriate conclusion for the test?

A. The mean number of available spaces is not significantly less than 12.8.
B. The mean number of available spaces is significantly less than 12.8.
C. The mean number of available spaces is not significantly less than 10.2.

D. The mean number of available spaces is significantly higher than 10.2.

Part IV. Blockbuster implemented a no late fee policy some time ago. It was hoped that the elimination of the late fee would increase the number of video rentals overall. To determine if the new "late fee" policy had the desired effect at one store, 10 randomly selected customers' accounts were sampled. In the sample, the number of rentals for a 2-month period for each customer was recorded prior to the implementation of the no late fee and then the number of rentals per customer for a 2-month period for each customer were again recorded after the no late fee policy was implemented. The data for the 10 customers is as follows:

	Number of Rentals	Number of Rentals
Customer	After the "No Late Fee" Policy	Before the "No Late Fee" Policy
1	14	10
2	12	7
3	14	10
4	13	13
5	10	9
6	13	14
7	12	12
8	10	7
9	13	13
10	13	9

```
"No Late Fee" Policy - Before the "No Late Fee" Policy

                          N    Mean  StDev  SE Mean
After the "No Late Fee   10  12.400  1.430    0.452
Before the "No Late Fee  10  10.400  2.503    0.792
Difference               10   2.000  2.211    0.699

95% lower bound for mean difference: 0.718
T-Test of mean difference = 0 (vs > 0): T-Value = 2.86  P-Value = 0.009
```

Answer the following questions for this test.

A. What type of sample was taken in this test?

Answer: _____

B. Briefly explain the statistical reason for taking this type of sample.

Yippee – almost done! One more page. Enjoy a stat-free weekend!!

C. What are the appropriate hypotheses for this test?

D. Write an appropriate conclusion for the test.

E. Briefly, how does the confidence interval value given support the results of the hypothesis test?

F. What is a Type II Error for this test? What would be a possible consequence of making a Type II Error?

APPENDIX G

Econ 262 Second Midterm Name:
Fall 2009 TR 11 a.m. class

ANSWER ALL OF THE QUESTIONS ON THE EXAM. SHOW WORK WHEN NECESSARY FOR PARTIAL CREDIT.

PART I. The questions below are based on the take-home preparation regarding instructor ratings.

A. What type of experimental design was used in this study?

Answer:
_____

B. What is the response variable in this study?

Answer:
_____

C. What is/are the treatments in this study? Answer:
_____

D. Write an appropriate conclusion for the test results.

E. Using your output, either describe any significant differences in the response variable or indicate where differences do not exist.

After performing the study, the school of business wanted to further investigate instructor ratings. Given possible differences in instructors, the study was redesigned to account for differences in mean ratings among the different instructors in the college. Ten instructors were selected. Each instructor teaches each of the 3 class levels. The revised data are shown below.

Instructor	Level 1 Rating	Level 2 Rating	Level 3 Rating
1	4.12	4.06	3.38
2	4.87	4.72	4.6
3	3.46	3.49	2.39
4	3.87	3.61	3.23
5	4.04	3.83	3.55
6	2.9	3.23	3.52
7	4.16	4.07	3.68
8	4.19	3.76	3.83
9	4.75	4.39	4.22
10	4.29	4.34	3.67

F. The data was input into a computer with the following partial results. **Complete the table by filling in the blanks.**

ANOVA TABLE

Source	DF	SS	MS	F
Rows	_____	6.7185	_____	_____
Columns	_____	1.1355	0.5677	_____
Error	_____	_____	_____	
Total	35	9.0605		

G. The second design employed blocking. Briefly explain why blocking is used in ANOVA.

H. Was there a block effect at alpha = .05? (Support your answer using the appropriate decision rule)

I. If there was a block effect – did it change the results of the test of treatment means? Explain briefly using the appropriate decision rule for the output shown in the redesigned test.

PART II. Market analysis often investigates the variables that are related to purchases. One national company that sells electronics on-line recently collected data on 400 orders placed over a period of several months. For each order, data was collected on the gender of the customer, the type of credit card used, the region of the country where the customer resides and other variables. To improve the marketing strategies developed to entice customers, type of credit card used to purchase the order was compared to past buying behavior. Buyer behavior is classified as high, medium or low volume. For example, a customer labeled as "high" behavior has purchased a high volume of products in the past. You are asked to perform the statistical analysis to determine whether a relationship exists between credit card used to purchase and buyer behavior. The cross-classification of buyer behavior data with credit card type is shown below.

	Past Buying Behavior		
Type of Card Used	High	Medium	Low

Proprietary (Electronics Co Card)	77	91	66
Other National Cards (Amex, Visa etc.)	43	69	54

A. What are the appropriate hypotheses appropriate for the test that should be conducted for this analysis?

B. What is the appropriate critical value if the test is conducted at alpha = .05? (Circle the correct answer)

A. 2.367
B. 4.329
C. 5.991
D. 7.815

The data was input into Minitab with the following results

	High	Medium	Low	Total
Proprietary	77	91	66	234
	70.20	93.60	70.20	
	0.659	0.072	0.251	
National	43	69	54	166
	49.80	66.40	49.80	
	0.929	0.102	0.354	
Total	120	160	120	400

C. What is the value of the observed sample test statistic?
(Circle the correct answer)

A. 2.367
B. 4.329
C. 5.991
D. 7.815

D. Write an appropriate conclusion for the test. If you found a relationship exists, briefly discuss its nature.

PART III. Although illegal, overloading (carrying too much weight) is common in the trucking industry. The Nevada Department of Transportation monitored the movements of overweight trucks on an interstate highway using an unmanned, computerized scale that is built into the highway. Unknown to the truckers, the scale weighed their vehicles as they passed over it. The sample data for three days of a selected week is shown below. The counts record the number of overweight trucks that were measured on Monday, Wednesday and Friday of the selected week.

	Monday	Wednesday	Friday
Number of trucks that were overweight	90	72	61

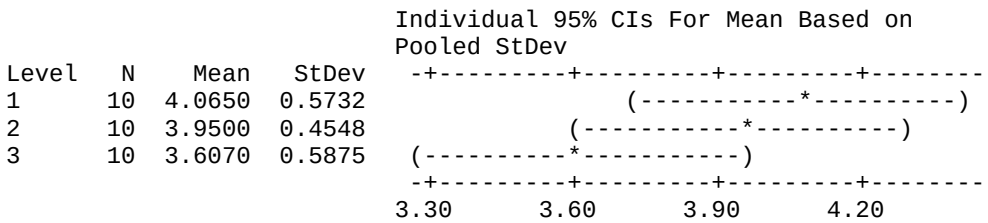
► The Department of Transportation would like to know whether the proportion of overweight trucks differs among the three days of the week. Perform the appropriate test using alpha = .05. Be sure to state the appropriate hypotheses, decision, decision rule and conclusion. If you find the distribution of proportions do differ, briefly describe where the differences exist.

O	E	O-E	O-E sq	O-E sq/E
50	60	-10	100	1.666667
70	50	20	400	8
35	60	-25	625	10.41667
45	30	15	225	7.5
				27.58333

```
         High  Medium    Low  Total
    1      77      91     66    234
        70.20   93.60  70.20
        0.659   0.072  0.251

    2      43      69     54    166
        49.80   66.40  49.80
        0.929   0.102  0.354

Total    120     160    120    400

Chi-Sq = 2.367, DF = 2, P-Value = 0.306
```

One-way ANOVA: Rating versus Class Type

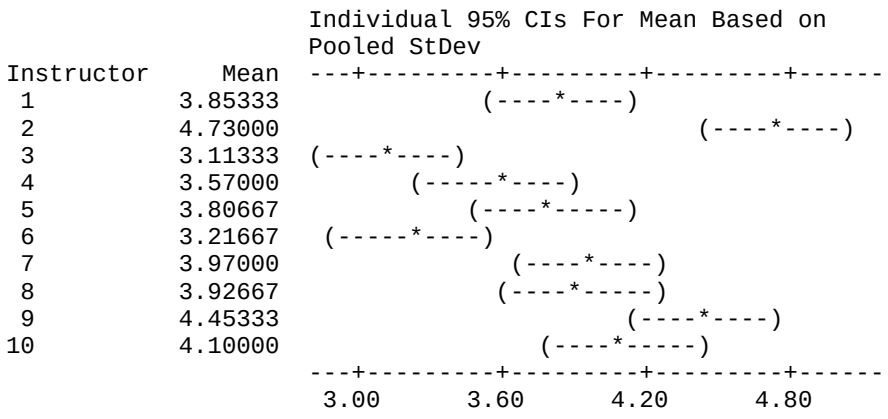
```
Source      DF     SS     MS     F      P
Class Type   2  1.135  0.568  1.93  0.164
Error       27  7.925  0.294
Total       29  9.061

S = 0.5418   R-Sq = 12.53%   R-Sq(adj) = 6.05%
```

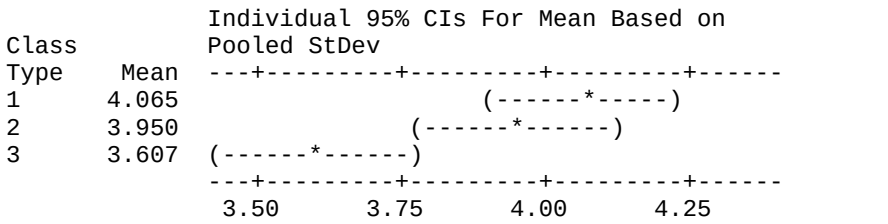



```
Pooled StDev = 0.5418
```

Two-way ANOVA: Rating versus Instructor, Class Type

```
Source      DF       SS        MS      F      P
Instructor   9  6.71852  0.746502  11.14  0.000
Class Type   2  1.13546  0.567730   8.47  0.003
Error       18  1.20654  0.067030
Total       29  9.06052

S = 0.2589   R-Sq = 86.68%   R-Sq(adj) = 78.55%
```


Individual 95% CIs For Mean Based on Pooled StDev
Instructor Mean
1 3.85333
2 4.73000
3 3.11333
4 3.57000
5 3.80667
6 3.21667
7 3.97000
8 3.92667
9 4.45333
10 4.10000
3.00
3.60
4.20
4.80


Individual 95% CIs For Mean Based on Pooled StDev
Class Type Mean
1 4.065
2 3.950
3 3.607
3.50
3.75
4.00
4.25

APPENDIX H

EC262 Third Exam, Fall 2009 Name:
9:30 Class **EXAM B**

ANSWER ALL OF THE QUESTIONS ON THE SCANTRON. PUT YOUR NAME ON THE EXAM AND THE SCANTRON. WRITE THE NUMBER OF THE EXAM ON YOUR SCANTRON. YOU MAY WRITE ON THIS EXAM. Each multiple choice question is worth 3 points.

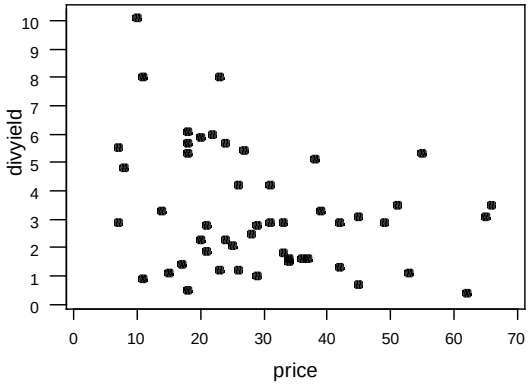
Questions 1 through 6 are based on the following problem: Suppose you are a financial analyst studying dividend yields (divyield) for various stocks. You believe that earnings per share (EPS) and the stock price (PRICE) are important explanatory variables for dividend yield. You randomly sample n = 50 firms from the Standard and Poor's Security Owner's Stock Guide. A sample of the data is shown below (note it does not include the entire sample).

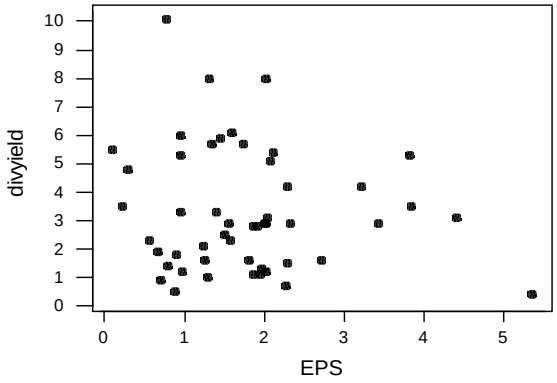
divyield	EPS	PRICE	
5.30	3.83	55	dividend yield is in percent
0.40	5.36	62	earnings per share is in dollars
5.50	0.10	7	price is in dollars
1.10	1.95	15	
3.10	4.41	65	
2.90	2.33	42	
4.20	3.22	31	etc …….

Shown below (and on the next page) are the correlation coefficient matrix and scatter plots for the variables being studied in this problem?

Correlations (Pearson)

	divyield	EPS
EPS	-0.130	
	0.368	
price	-0.302	0.647
	0.033	0.000





1. After examining the computer output, which statement below is correct?

A. Price should be the better predictor of dividend yield.
B. EPS should be the better predictor of dividend yield.
C. Both price and EPS should be good predictors of dividend yield.
D. EPS and price are not significantly correlated.

2. Which of the variable(s) is the response variable?

A. Both EPS and price.
B. EPS only.
C. Price only.
D. Dividend yield only.

3. The scatter plot for dividend yield versus price shows which of the following?

A. As dividend yield increases price also increases.
B. As price increases, dividend yield increases.
C. As price increases, dividend yield decreases.
D. A rather high coefficient of determination.

4. Suppose you wished to run a regression comparing dividend yield against price. Which of the following will occur in the regression output?

A. The regression equation will have an insignificant t-test.
B. The regression equation will have a significant F-test.
C. The slope of the equation will be positive.
D. The coefficient of determination will be extremely high.

5. The correlation matrix shows which of the following?

A. An insignificant regression model.
B. A problem with the residual assumptions.
C. The explanatory variables are significantly correlated with one another.

D. The explanatory variables are not significantly correlated with one another.

6. The p-values shown in a correlation matrix like the one above measures which of the following?

A. Whether the residual assumptions are violated.
B. Whether the regression equation is useful for predicting dividend yield.
C. Whether the correlation between variables is negative or positive.
D. Whether the correlation between variables is significantly different from 0.

Questions 7 through 13 are based on the following problem: The vice-president of marketing for a large firm is concerned about the effectiveness of advertising in generating sales of the firm's major product. To investigate the relationship between advertising and sales, data on the two variables were gathered from a random sample of 20 sales districts. The data are shown below:

sales	adv
4250	235
3700	210
2000	160
5800	345
6200	325
6500	365
7000	400
4900	370
6100	350
2900	200
3200	200
3500	210
4000	230
5175	210
5450	300
5900	325
7110	390
6500	375
7400	415
6600	380

sales measured in $100
adv (advertising expenditures) measured in $100

The vice president performed a regression analysis on the data with the following partial results:

Predictor	Coef	StDev
Constant	-57.3	509.8
adv	17.570	1.642

S = 594.8 R-Sq = 86.4% R-Sq(adj) = 85.7%

7. Which of the following is the correct interpretation for the slope?

A. For each $100 increase in advertising expenditures, sales are expected to decrease by $57.30.
B. For each $100 increase is sales, advertising expenditures are expected to increase by $1757.
C. For each $100 increase in advertising expenditures, sales are expected to increase by $1757.
D. For each $100 increase in advertising expenditures, sales are expected to decrease $1757.

8. Perform the test of significance on the slope at alpha equals .05. Which of the following statements is the correct conclusion for the result of the test?

A. Advertising expenditures is not useful for predicting sales because the t-value equals 1.642.
B. Advertising expenditures is useful for predicting sales because the t-value equals 10.7.
C. Sales is not useful for predicting advertising expenditures because the t-value equals 1.642.
D. Sales is useful for predicting advertising expenditures because the t-value equals 10.7.

9. The coefficient of determination indicates which of the following?

A. 86.4% of the total variability in sales is explained by advertising expenditures.
B. 86.4% of the total variability in advertising expenditure is explained by sales.
C. There is a weak positive correlation between sales and advertising expenditures.
D. There is a large amount of variability between the observed values of y and the predicted values of y.

The following Minitab printout shows the prediction for sales when advertising expenditure = $38,000.

Predicted Values for New Observations

New

Obs	*Fit*	*SE Fit*	*95% CI*	*95% PI*
1	6619	187	(6226, 7013)	(5309, 7929)

10. The value 6,619 represents which of the following?

A. The standard deviation of the residuals.
B. The observed value of Y for X = $38,000.
C. The predicted value of Y for X = $38,000.
D. The residual for Y versus Y-hat when X = $38,000.

11. The last sample observation in the data above shows sales = 6600 and adv = 380. What does the residual equal for this observation?

A. ± 38.6
B. $ 38,000
C. 19.3
D. – 19.3

12. The 95% confidence interval shown above represents which of the following?

A. The estimate of sales for all sales districts with an advertising expenditure of $38,000.
B. The estimate of sales for a single sales district with an advertising expenditure of $38,000.
C. A statistically significant regression model that is useful for predicting sales.
D. The estimate of advertising expenditures for the given value of Y.

13. The 95% prediction interval shown above represents which of the following?

A. The estimate of sales for all sales districts with an advertising expenditure of $38,000.
B. The estimate of sales for a single sales district with an advertising expenditure of $38,000.
C. A statistically significant regression model that is useful for predicting sales.
D. The estimate of advertising expenditures for the given value of Y.

You're almost half way home – keep on chugging along ……

Problems 14 – 16 are based on the following: In order to evaluate the effectiveness of a special energy saving package in new home construction, the Reno Builders' Association develop a regression model that estimated annual energy costs based on the square footage of the residence and whether or not the energy-savings package had been installed. The data appeared as follows:

Square Footage (in 1000s sq ft)	Insulation Package	Annual Energy Costs (in $100s)
1.5	Standard	9
2.0	Standard	12
2.2	Energy Saving	14
.	.	.
Etc.	Etc.	Etc.

For the Insulation Package, the following was used: 0 if the energy-savings package was used; 1 if a standard package was used. In total there were n = 12 homes that were tested. The following shows a partial Minitab output for this regression analysis:

Regression Analysis: Cost versus Sq Ft, Insul Pkg

```
Predictor    Coef  SE Coef     T      P
Constant    0.140    1.862  0.08  0.942
Sq Ft      5.7764   0.9206  6.27  0.000
Insul Pkg  0.8230   0.7971  1.03  0.032

S = 1.06630   R-Sq = 90.1%   R-Sq(adj) = 87.9%
```

14. The appropriate interpretation of the coefficient for the type of insulation is which of the following?

A. For each $100 increase in costs, the insulation package is expected to add $82.30 to the overall cost.
B. If the home used the energy savings insulation package, annual costs are expected to be $82.30 less versus a home with a standard insulation package.
C. If the home used the energy savings insulation package, annual costs are expected to be $82.30 more versus a home with a standard insulation package.
D. The type of insulation package used does not influence cost because the dummy variable is insignificant.

15. What is the appropriate null hypothesis to test the overall utility of the model?

A. $\beta_i = 0$
B. $\rho = 0$.
C. At least one $\beta \neq 0$
D. $\beta_1 = \beta_2 = 0$

16. What is the appropriate null hypothesis to test the significance of the type of insulation package?

A. $\beta_i = 0$
B. $\rho = 0$.
C. At least one $\beta \neq 0$
D. $\beta_1 = \beta_2 = 0$

17. Suppose Stiver wanted to determine if there was a significant correlation between the homework scores and test scores in her classes. She randomly samples 100 students and finds the correlation coefficient = .63. What conclusion should she reach regarding the significance of the correlation?

A. The correlation is not significant because the r is not close to 1.00.
B. The correlation is not significant because the t-test statistic to test for significance = 1.43.
C. The correlation is significant because the t-test statistic = approximately 8.036.
D. The correlation is significant because r is above .50.

18. What is the appropriate null hypothesis for the test conducted in question 17?

A. $\beta_i = 0$
B. $\rho = 0$.
C. At least one $\beta \neq 0$
D. $\beta_1 = \beta_2 = 0$

19. Outliers or unusual observations sometimes present difficulties for regression models. Which of the following is a problem that may be encountered when you have severe outliers?

A. Multicollinearity will be a problem.
B. The regression coefficients may not represent the true relationship between x and y.
C. The signs of the regression coefficients will be opposite of the correlation signs.
D. The Y-intercept will be un-interpretable.

20. Ordinary least squares regression is a technique that

A. Maximizes the coefficient of determination.
B. Minimizes the difference between the observed value of y and the predicted value of y.
C. Minimizes the t-test value on the slope.
D. Is both ordinary and square.

21. What is the primary purpose for a residual analysis?

A. It is used to determine the size of the residuals.
B. It is used to determine the significance of the regression equation.
C. It is used to determine whether the assumptions regarding the residuals are valid.
D. It is used to measure the “fit” of the line.

22. Which of the following is not one of the functions for constructing a scatter plot?

A. Determine whether the relationship between x and y is negative.
B. Determine whether the relationship between x and y is linear.
C. Determine whether the relationship between x and y is significant.
D. Detect the presence of outlier observations.

23. Which of the following is NOT an indication of multicollinearity in a model?

A. The F-test and the individual t-tests are insignificant.
B. A negative correlation sign switching to a positive coefficient in the regression equation.
C. Insignificant t-tests for explanatory variables that show significant correlation to the response variable.
D. A significant F-test when all the t-tests are insignificant.

Problems 24 - 27 are based on the following graphs:

Graph A

Graph B

Graph C

Graph D

24. Which graph shows a mis-specified model or a lack of independence?

A. Graph A
B. Graph B
C. Graph C
D. Graph D

25. Which graph shows a violation of the homoscedastistic assumption?

A. Graph A
B. Graph B
C. Graph C
D. Graph D

26. Which graph shows a correlation of approximately .021?

A. Graph A
B. Graph B
C. Graph C
D. Graph D

27. Which graph shows a coefficient of determination of approximately .852?

A. Graph A
B. Graph B
C. Graph C
D. Graph D

Problems 28 and 29 are based on the following: A large company is investigating variables that will help determine sales for the company. They have considered the following variables: advertising expenditures (adv), bonuses paid to the company's sales force (bonus), the company's market share in the industry (mktsh), and sales achieved by its closest competitor (compet). All four variables were included in a multiple regression model. The full model needed to be revised (see partial results below), so a step-wise regression was performed with the results listed below.

FULL MODEL REGRESSION RESULTS:

The regression equation is

sales = - 594 + 2.51 adv + 1.91 bonus + 2.65 mktshr - 0.121 compet

Predictor	Coef	SE Coef	T	P
Constant	-593.5	259.2	-2.29	0.033
adv	2.5131	0.3143	8.00	0.000
bonus	1.9059	0.7424	2.57	0.018
mktshr	2.651	4.636	0.57	0.574
compet	-0.1207	0.3718	-0.32	0.749

S = 93.7697 R-Sq = 85.9% R-Sq(adj) = 83.1%

Continued on next page ……

STEPWISE REGRESSION RESULTS:

Step	***1***	***2***
Constant	-157.3	-516.4
adv	2.77	2.47
T-Value	9.92	8.98
P-Value	0.000	0.000
bonus		1.86
T-Value		2.59
P-Value		0.017
S	101	90.7
R-Sq	81.06	85.49
R-Sq(adj)	80.24	84.18
Mallows C-p	5.9	1.6

28. The most likely reason market share was dropped from the regression is because

A. Market share reduced r-Sq when included in the model.
B. Market share is not linearly related to sales.
C. Market share should be negatively related to sales.
D. Market share reduced the standard error of the estimate when included in the model.

29. The value of the standard error dropped from 101 to 90.7 between steps 1 and 2 in the stepwise regression. Why is this reduction a desirable result?

A. A reduced standard error improves the model's precision when predicting sales.
B. A reduced standard error generally decreases the coefficient of determination.
C. A reduced standard error changes the value of the coefficients.
D. Reduced standard errors have nothing to do with a regression model.

30. Select the best answer. Which of the following is most likely now that you've finished this exam?

A. I will be partying to celebrate my victorious completion of this course.
B. I will be partying to alleviate the mental anguish I am currently experiencing.
C. I am so into statistics, I can't wait to go home and reread my text over the summer.
D. One would have to define "statistically" what is meant by "most likely" according to the rules of probability.

Have a restful and happy summer! ☺ Go Wolf Pack!!