DOCUMENT RESUME

ED 443 847	TM 031 492
AUTHOR	Geske, Jenenne A.; Mickelson, William T.; Bandalos, Deborah L.; Jonson, Jessica; Smith, Russell W.
TITLE	Predicting Acquisition of Learning Outcomes: A Comparison of Traditional and Activity-Based Instruction in an Introductory Statistics Course.
PUB DATE	2000-04-00
NOTE	35p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 24-28, 2000).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Academic Achievement; College Faculty; *College Students; Curriculum; *Educational Change; Higher Education; Instructional Effectiveness; Outcomes of Education; *Statistics; *Teaching Methods

ABSTRACT

The bulk of experimental research related to reforms in the teaching of statistics concentrates on the effects of alternative teaching methods on statistics achievement. This study expands on that research by including an examination of the effects of instructor and the interaction between instructor and method on achievement as well as attitudes, classroom environment, and statistics self-efficacy. Data were gathered from 156 students aged 19 to 43 years in an undergraduate introductory statistics course. Two doctoral level graduate student instructors taught these students in four course sections. Results indicate that the anticipated benefits of statistics education reform may be affected by the instructor. Appendixes contain the alternative class course outline, a random rectangles problem, self-efficacy measurement items, and the take-home final examination. (Contains 28 references.) (SLD)

Predicting Acquisition of Learning Outcomes: A Comparison of Traditional and Activity-Based Instruction in an Introductory Statistics Course

Jenenne A. Geske, William T. Mickelson, Deborah L. Bandalos, Jessica Jonson, Russell W. Smith

University of Nebraska-Lincoln

Paper presented at the 2000 annual meeting of the American Educational Research Association in New Orleans. LA.

Abstract

Statistics education researchers are increasingly calling for reforms in the procedures used to teach introductory statistics classes. The bulk of experimental research in this area concentrates on the effects of alternative teaching methods on statistics achievement. The current study expands on this research by including examination of effects of instructor and the interaction between instructor and method on achievement as well as attitudes, classroom environment and statistics self-efficacy. Results indicate that the anticipated benefits of statistics education reform may be affected by the instructor.

It is often argued that the traditional lecture format hinders students' development of statistical reasoning abilities, perhaps because this traditional presentation style distances students from the dynamic nature of data collection and statistical analysis. As such, numerous statistics educators have been advocating the need for dramatic changes to the introductory statistics course. Essentially, the message has been to reorganize the introductory course around activity-based learning using real data that focus on promoting the learning of statistical concepts and the development of statistical reasoning skills. In particular, the recommendations include: (a) consciously develop course objectives based on the needs of student and of future employers, (b) utilize experiential learning more and lecture less, (c) teach scientific inquiry first and analysis tools afterward, (d) point out common misuses of statistics, and (e) recognize and confront common errors in students' thinking (Bradstreet, 1996; Cobb, 1993; Garfield & Ahlgren, 1988; Garfield, 1993; Hogg, 1991; Konold, 1995; Moore, 1997). These recommendations stem from the basic tenets of the constructivist theory of learning and a pragmatic belief that active learning experiences consistent with a professional statistician's methods are beneficial in aiding students in making statistical concepts and reasoning personally relevant. The implication is that modifications in instructional method will lead to improved student outcomes such as higher achievement, more positive attitudes toward their statistics class and toward statistics as a subject.

Experimental research on teaching method reform, however, is often hampered by constraints of the academic educational system. The typical study compares two classes taught by the same instructor with different methods (Giraud, 1997; Keeler & Steinhorst 1995). While these studies have their merit, the instructor is fully confounded with method so an estimate of the instructor's impact separate from method is not possible. The present study employed two instructors, each of whom taught two statistics sections. Each instructor taught one section in what is typically thought of as a "traditional" lecture-style format. Teaching practices in the other section taught by each instructor (the alternative section) incorporated many of the suggestions emerging from the statistics education reform movement. By exploring the interaction between teacher and teaching method, commentary can be given on curriculum and "best practices" for teaching introductory statistics. Specifically, our study incorporates experimental factors (type of teaching method, instructor, and their interaction) in an exploratory attempt to examine the acquisition of pre-specified learning goals. In addition, we examined the effects of method, instructor and their interaction on student perceptions about their statistics class and statistics in general.

In addition, while much of the previous research and commentary in this area focuses on student achievement and attainment of instructional goals (Keeler & Steinhorst, 1995; Smith, 1998; Moore, 1997), less emphasis is placed on student attitudes toward statistics as a subject and toward their statistics classroom environment in particular (Becker, 1996). We are therefore including measures of attitudes toward statistics, self-efficacy and classroom environment measures in our study.

Curricula

Much of the literature addressing statistics education reform emphasizes the need for instructors to examine and articulate the goals they have for their students (Garfield, 1995; Rinaman, 1998; Roiter & Petocz, 1996). The curricula for courses used in our study were developed based on goals set forth by a team of faculty members and graduate student course instructors at the beginning of the study. In particular, we wanted to ensure that students understood (1) how to collect data effectively, (2) how to summarize data, (3) how to interpret data in context and draw meaningful conclusions, (4) how to critique, value and evaluate the numerical arguments and data presented by others, and (5) the process of how to ask and answer appropriate descriptive and inferential statistical questions.

Course context and sequencing differed for the two teaching methods, but the content remained the same. The traditional course format had materials presented in a traditional lecture format with an organizational structure similar to that found in the introductory text, Statistics for the Behavioral Sciences by Gravetter & Wallnau (1988). Lectures were complemented by a course packet of lecture notes compiled jointly by the two instructors involved in the study. The instructors were at liberty to augment the text as they saw fit while keeping a predominantly lecture format. Student assessment measures included assignments based on lecture presentations and multiple choice in-class examinations at the end of each unit.

The alternative class curriculum was activity-based and designed with the goal of bringing about conceptual understanding of statistical principles and procedures. The five step process of posing a research question, designing a study to examine the question, collecting data, summarizing/analyzing the data, drawing a conclusion and communicate findings is very well-known (Graham, 1987; Kader & Perry, 1994) and parallels the scientific method. For the activity class, our framework was for these five steps to be explicitly modeled and employed during instruction (see Course Outline, Appendix A). Activities were presented with the goal of leading the students into their own construction of statistical content.

Because of their prevalence in reports found in the general media, the first unit focused on simple proportions. In-class group activities were used to demonstrate the statistical concepts of validity, randomness, variability, the impact of random versus nonrandom sampling, bias, sampling distributions, the logic of hypothesis testing and confidence intervals. The five-step process of reasoning with data was modeled or employed with each activity. For example, during the 4th class session, students were introduced to the concept of random sampling and bias, using the "Random Rectangles" activity suggested by Scheaffer, Gnanadesikan, Watkins & Witmer (1996). Students were given a copy of the Random Rectangles page of the Student Guide (see Appendix B), face down. Students were advised of the goal of the exercise (to estimate the average area of the sample of rectangles on the page) and were then asked to view the Random Rectangles for a brief period of time and then again turn the page face down. Data were collected from the students regarding their estimates, and a histogram was constructed for the class to see. The students were then asked again to look at the page and select five rectangles that they feel are representative of the population of rectangles on the sheet, and calculate mathematically the average area of the chosen rectangles (incorporating the definition of the Mean, and the formula for calculating it). Next, students used a random number table to randomly select samples of 5 and 10 rectangles, and calculated the average area for both samples. Histograms were conducted after each sampling. Students were then told that the actual average area of the rectangles on the page is 7.3. Students were asked to compare this information with the information found in each of the histograms. The histograms differed, allowing for the opportunity to introduce the concept of variability to the class. Additionally, bias is demonstrated in the histograms constructed from data chosen via the non-random methods. The class ended with a discussion of how the concepts of sampling and bias fit into the five-step process of reasoning with data. Remaining topics were presented utilizing projects and activities similar to the ones presented in Activity-Based Statistics (Scheaffer, Gnanadesikan, Watkins, Witmer, 1996).

Upon completion of the first unit, the five-step data reasoning process was again modeled and employed as students formed groups and conducted survey projects on research questions they posed. The second unit consisted of data analysis issues intended to supplement the students' own data collection and analysis. Topics covered included exploratory data analysis, t-tests, chi-square tests, and correlation and regression. Class time was allotted during the second unit to allow groups to formulate their research questions, design their studies and plan their analyses. Thus, students were working together and constructing their own knowledge of statistical methods, with guidance provided by an instructor.

Students in the alternative classes were assessed via written 1-2 page activity reports which summarized the activity and reaffirmed its learning goals. Students also received an in-class essay examination at the end of the first unit. Additionally, student groups were asked to give a presentation and write a paper on their research findings.

While the sequence and context of the material varied across the two instructional methods, every attempt was made to ensure that the content of the material presented in both classes was comparable and that it addressed the established learning goals, which were identical for both methods. At the end of the semester, all of the students in the study (in both the traditional and the alternative classes) were asked to complete the same take-home essay final examination, which was used to assess student achievement of the learning goals. Along with this examination, data were collected from the students on measures of attitudes toward statistics, self-efficacy and perception of classroom environment.

Research Questions / Hypotheses

Our primary research goal focused on the exploration of interactions between instructional method and instructor on their effect on the statistics achievement, attitude and perceptions of classroom environment. While there is a great deal of literature that advocates reform in instructional practices (Bradstreet, 1996; Garfield, 1995; Hogg, 1991; Moore, 1997), there is little available information concerning the role the instructor plays in the process of change. Clearly, the instructors involved have different teaching philosophies and different theories about learning. The question becomes "are student variables affected by the possibility that instructors react in a differential manner to instructional method?"; it is therefore our goal to expand upon previous research which examined instructional method alone. Because there are no previous data concerning the instructors involved in this study, or concerning instructor variables in general, no predictions are made regarding the nature of any suspected interactions. While we expect to find differences on the outcome variables across instructors, no specific directional hypotheses are formulated.

In addition, we are interested in the main effects of method on statistics achievement, self-efficacy and attitudes. The presumed benefits of statistical education reform include a deeper understanding of statistical concepts and stronger statistical reasoning abilities (Moore, 1997) along with higher self-efficacy and a more positive attitude toward statistics (Garfield, 1995; Davidson & Kroll, 1991). Thus, in this study, we predict that students in the alternative classes will have higher scores on the final exam and the statistics self-efficacy scale. These students should place a higher value on statistics and feel an increased sense of cognitive competency toward the subject matter. The activity-based nature of the alternative class should result in higher perceptions of student involvement, cohesiveness, individualization and classroom innovation. Because most of the students who participated in this study are probably very well accustomed to a lecture-style classroom, we feel that the students will feel more comfortable in this type of learning environment. Therefore, we believe that students in the traditional classroom will have higher affect and satisfaction scores than will be obtained from students in an alternative learning environment.

Method

Participants

Data were gathered from 156 students, ranging in age from 19 to 43 years (mean age=21.24, sd=3.7644). Students were enrolled in one of four sections (approximately 40 students each) of an undergraduate introductory statistics course at a large Midwestern university. Of those who indicated their gender, fifty-six were males and eighty-nine were females. Sixteen of the participants indicated that they were freshman, 48 were sophomores, 47 were juniors, 33 were seniors and 1 was a graduate student (the remaining students did not indicate their year in school). The vast majority (88%) indicated that they had no prior statistics or research methods courses. Participation in this study was voluntary.

Instructors

Two doctoral level graduate student instructors were each responsible for teaching two sections of the course with supervision from a faculty member. Both students (one male, one female) had a minimum of 2 semesters prior experience teaching this course, utilizing primarily lecture-style formats.

Materials

The participants were asked to complete several instruments throughout the course of the semester. In addition to completing a short demographic questionnaire, they were asked to complete a Survey of Attitudes Toward Statistics (Schau, Stevens, Daupinee & Delvecchio, 1995), 10 items concerning statistical self-efficacy which were written by the researchers, and the College and University Lecture Classroom Environment Inventory (Schuh, 1996). At the end of the semester, all students completed a non-comprehensive, take-home final with short essay questions.

The Survey of Attitudes Toward Statistics (SATS) (Schau, et. al, 1995) is designed to measure the attitudes and beliefs that students have about statistics and measures four dimensions: Affect (positive and negative feelings toward statistics), Cognitive Competence (attitudes about intellectual knowledge and skills applied to statistics), Value (attitudes about the usefulness, relevance and worth of statistics in personal and professional life), and Difficulty (attitudes about the difficulty of statistics as a subject). Students indicated their level of agreement using a 7-point Likert-type scale (1=Strongly Disagree, 4=Neither, 7=Strongly Agree). Each factor consisted of 6-9 items.

Ten statistical self-efficacy items were created by the researchers. No existing statistical self-efficacy instruments could be located, and research indicates that content-specific measures of self-efficacy are preferred to generalized measures (Pajares, 1996). The items that were created were specific to the subject of statistics. This unidimensional scale contains items which were measured on the same 7-point scale as the SATS. See Appendix C for the self-efficacy items that were used. A summated self-efficacy score was computed by reverse coding items 2, 4, and 8 and summing the scores.

Seven dimensions of students' perceptions of their classroom environment were measured using the College and University Lecture Classroom Environment Inventory (CULCEI) (Schuh, 1996). This 49-item instrument is a modification of an earlier instrument, the College and University Classroom Environment Inventory (Fraser & Treagust, 1986). Dimensions were measured using seven items each and consisted of Personalization (student has opportunities to interact with the instructor), Involvement (students participate in class discussions and activities, Student Cohesiveness (students in the class know each other and are helpful toward each other), Satisfaction (students enjoy the class), Task Orientation (class projects are clear and wellorganized), Innovation (the instructor uses a variety of teaching methods and assessments), Individualization (students' individual differences are incorporated in the class). Responses were gathered on a 4-point scale ranging from Strongly Agree (4) to Strongly Disagree (1).

Course instructors along with faculty advisors collaborated on writing eight short essay items for the course final exam. Students were allowed to complete this exam on their own outside of class, and were instructed not to work together or obtain help from anyone other than their instructor. Items on the exam were constructed with the intention of assessing the learning goals established at the beginning of the semester, including the ability to evaluate published research and to employ statistical reasoning skills in contextual situations. Students were first presented with a scenario in which the announcers from a local radio station discussed the results of a published study. Students were asked questions pertaining to the conclusions that the announcers reached. The second part of the exam consisted of an abridged version of an actual research article taken from a journal. Students were expected to evaluate the research and interpret the presented results. A copy of the final exam can be found in Appendix D.

Procedures

Instructors and methods were randomly assigned to the four sections three days prior to the beginning of the semester. Although it was not possible to randomly assign students to sections, no prior knowledge regarding instructor or teaching method was available to the students prior to the first day of class.

Course instructors meet weekly with their faculty supervisor. The purpose of these meetings was to coordinate the curriculum. In addition, these sessions enabled the instructors to interact with each other and share ideas, experiences and obstacles. Problems were discussed and resolved with input from all three members.

Data were collected from the students at three points in the semester. During the second week of classes, the researchers went into the classes and explained the purpose of the study. Informed consent forms were presented and signed, and the students were given identical packets containing a demographic questionnaire along with the SATS. Midway through the semester, the students completed the CULCEI. The take-home final exam was given to the students approximately two weeks before the end of the semester and students had one week to complete them. After the final exams were collected, students were asked to again complete the SATS and the CULCEI.

The process of scoring the final exams was completed by one of the graduate student instructors and a faculty advisor. The instructor first developed a scoring rubric detailing the points to be assigned for certain types of responses to the questions on the exam. A detailed training session followed between the instructor and the faculty advisor, with goal of maximizing inter-rater consistency. The exams were then divided among the two scorers for final grading. If there was uncertainty regarding the score for an item, the scorers collaborated to assign a score on that item. Scores on the exam ranged from 8-28 points; the maximum score was 28 points.

Results

Initial Equivalence

The purpose of this study is not to investigate changes over time; rather, the primary goal is to simply explore the possible interactions between instructional method and instructor. In order to determine that the four groups were similar at the beginning of the semester, analyses of variance were conducted on the SATS subscales and the self-efficacy scales that were administered during the second week of class. Results can be found in Table 1. Instructor B's scores on the initial SATS difficulty scale were significantly higher (mean=27.31, sd=4.20) than instructor A's (25.67, sd=5.39). Due to the random procedures used for assigning instructors and methods to each class, and because this survey was administered very early in the semester, it is thought that this significance is due to high statistical power. The four groups were statistically equal on the remaining SATS subscales. Additionally, as can be seen in Table 2, of the students who have not previously taken any statistics courses, a total of 8 of them were in Instructor A's classes and 10 were in instructor B's classes (9 each in the alternative and traditional sections). Thus, it would appear that the groups are equal on prior statistics ability.

Table 1: ANOVA Results, Effect Sizes on Second-week Measures

	Method x Instructor Interaction		Method			Instructor		
Variable	F	p	F	p	d	F	p	d
Self-Efficacy	2.371	.13	3.123	.08	.28	.191	.66	.05
SATS Subscales								
Affect	1.012	.32	1.000	.32	.16	2.271	.13	.25
Cognitive Competency	2.146	.15	.997	.32	.16	.555	.48	.12
Value	3.023	.08	.004	.95	.19	.384	.54	.27
Difficulty	2.352	.13	.884	.36	.14	4.167	.04	.34

Table 2: Distribution of Students With Prior Statistics Courses

	Traditional Method	Alternative Method	Total
Instructor A	4 (10%)	4 (11%)	8 (10%)
Instructor B	5 (12%)	5 (13%)	10 (13%)
Total	9 (11%)	9 (12%)	18 (11%)

Note: Numbers in parentheses indicate the percentage of students in each class.

Internal Consistency

Measures of internal consistency reliability were obtained for all of the semester-end dependent variable measures. Results can be found in Table 3.

Table 3: Reliability of Measures

Dependent Variable	Coefficient Alpha	Dependent Variable	Coefficient Alpha
CULCEI Subscales		SATS Subscales	
Personalization	.83	Affect	.80
Involvement	.69	Cognitive Competency	.81
Student Cohesiveness	.46	Value	.86
Satisfaction	.86	Difficulty	.65
Task Orientation	.78		
Innovation	.56	Self Efficacy	.84
Individualization	.52		

Final Exam

Results of a two-factor ANOVA indicate that all classes scored approximately equally on the final exam. Table 4 indicates the means, standard deviations and ANOVA results. Figure 1 graphically displays the effects of method and instructor on final exam results. There is no significant interaction between instructor and method on the final exam scores ($F(1,126)=.065$, $p=.80$). Additionally, there are no main effects for instructor ($F(1,126)=.271$, $p=.60$) or for instructional method ($F(1,126)=.438$, $p=.51$).

Table 4: Means, Standard Deviations and Two-Way Analyses of Variance for the Effects of Instructor and Instructional Method on the Final Exam Scores

Effect	Mean	Standard Deviation	n	F	p	d
Method				.438	.51	.11
Traditional	19.19	4.38	66			
Alternative	18.71	4.30	64			
Instructor				.271	.60	.09
Instructor A	18.75	4.03	59			
Instructor B	19.13	4.59	71			
Method x Instructor Interaction	18.95	4.33	130	.065	.80	

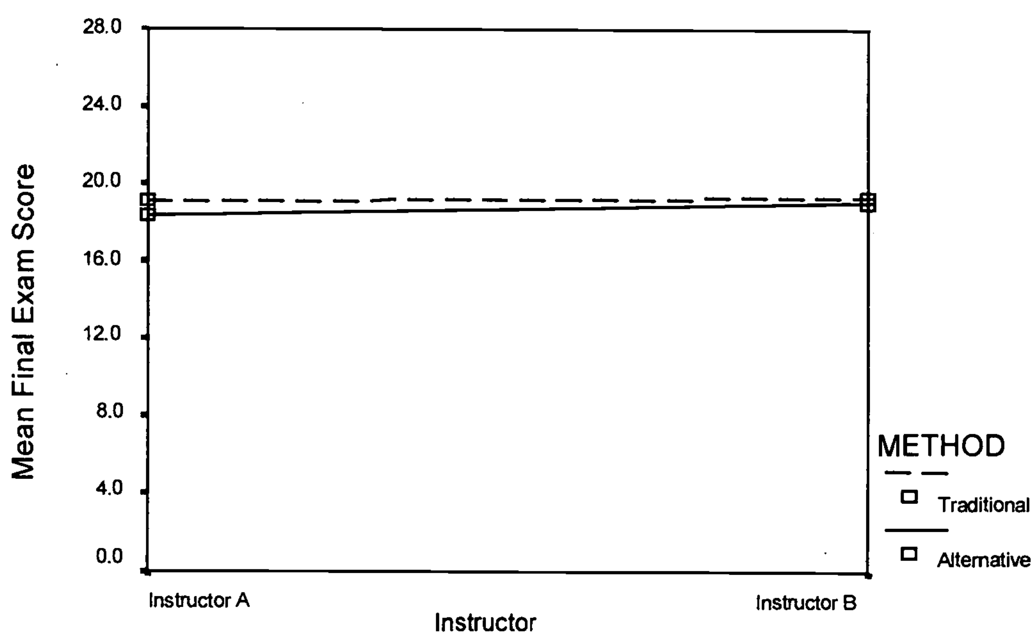


Figure 1: Effects of Method, Instructor and their Interaction on Final Exam Scores

Examination of the relationship of other dependent variable scores with scores on the final exam revealed that each of the SATS subscales, along with the self-efficacy measure were significantly correlated with performance on the final, while the CULCEI subscales were not related. Results provided in Table 5 indicate that those with more positive attitudes about their affect, cognitive competency, perceived value and the difficulty of the class did better on the final exam. Additionally, those with a higher self-efficacy toward statistics received higher exam scores. Scores on the CULCEI subscales did not correlate with final exam scores.

Table 5: Correlations with the Final Exam

Dependent Variable	r	p	n
SATS Subscales			
Affect	.317**	.001	105
Cognitive Competency	.326**	.001	104
Value	.282**	.004	104
Difficulty	.249*	.012	100
CULCEI Subscales			
Personalization	.032	.751	100
Involvement	- .067	.506	101
Student Cohesiveness	- .035	.728	100
Satisfaction	- .017	.870	99
Task Orientation	- .011	.911	97
Innovation	- .082	.413	101
Individualization	- .117	.256	97
Self-Efficacy	.315**	.001	104

* p<.05. ** p<.01.

Self-Efficacy

A method by instructor between-subjects factorial ANOVA was conducted on the results of the self-efficacy scale. As can be seen in Table 6 and Figure 2, There was not a significant main effect for instructor (F(1,102)=.045, p=.833, d=.15) or a significant method by instructor interaction (F(1,102)=.001, p=.979). However, while the main effect for method was not statistically significant (F(1,102)=3.718, p=.057, d=.43), there is a medium effect size, indicating that the alternative classes (51.333, sd=8.96) had self-efficacy scores almost a half of a standard deviation higher than the traditional classes (47.43, sd=8.99).

Table 6: Means, Standard Deviations and Two-Way Analyses of Variance for the Effects of Instructor and Instructional Method on the Self-Efficacy Scores

Effect	Mean	Standard Deviation	n	F	p	d
Method				3.718	.06	.43
Traditional	47.43	8.99	67			
Alternative	51.33	8.96	39			
Instructor				.045	.83	.15
Instructor A	49.48	8.70	58			
Instructor B	48.13	9.68	48			
Method x Instructor Interaction	48.87	9.14	106	.001	.98	

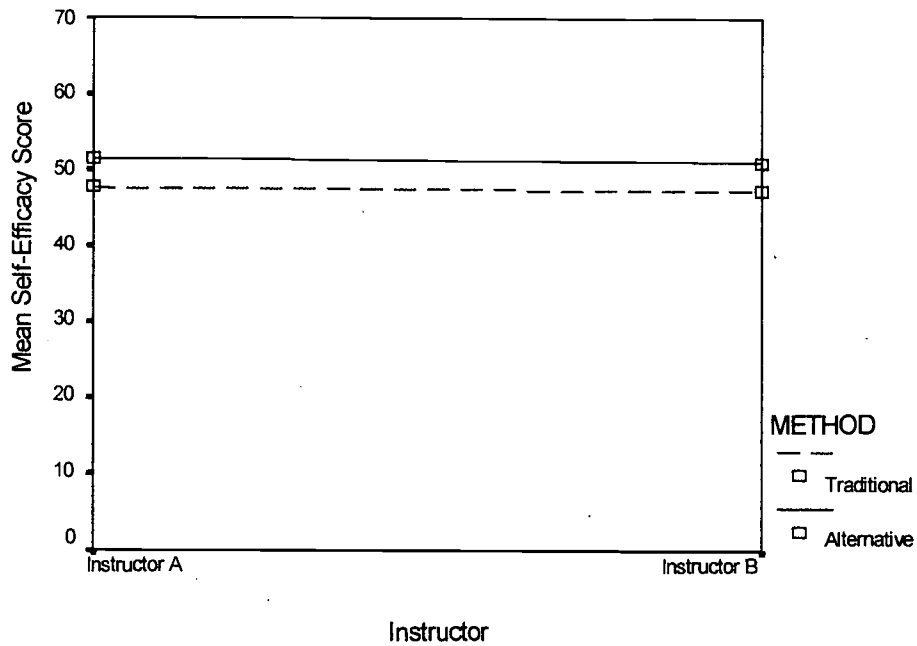


Figure 2: Effects of Method, Instructor and their Interaction on Self-Efficacy

Survey of Attitudes Toward Statistics

A 2 x 2 between-subjects factorial multivariate analysis of variance was performed using the four SATS subscales (affect, value, cognitive competence, difficulty) as dependent variables. The independent variables were method (traditional or alternative) and instructor (A or B). Fifty-six of the students did not have scores on all of the dependent variables and were not included in this analysis. Means and standard deviations of the subscale can be found in Table 7. A significant multivariate interaction was found between method and instructor, $F(4,93)=2.913$, $p=.026$. To further investigate the nature of the interaction, the simple main effects of instructor for each method were examined multivariately. For the traditional classes, there was a significant difference between instructors on the dependent variables, $F(4,58)=4.804$, $p=.002$. This same effect did not appear for the alternative classes ($F(4,58)=.886$, $p=.484$). As indicated in Table 8, results of discriminant function analyses suggest that students in instructor A's traditional class had higher affect, cognitive competency and value scores and lower difficulty scores, while the pattern of scores for instructor B's students were reversed.

Table 7: Means and Standard Deviations on the SATS and CULCEI Subscales

	Traditional		Alternative		
	Instructor A	Instructor B	Instructor A	Instructor B	Total
Dependent Variable	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
SATS Subscales					
Affect	27.14 (6.31)	23.60 (6.80)	25.62 (6.71)	24.45 (7.22)	25.21 (6.74)
Cognitive Comp.	30.96 (5.22)	28.74 (6.62)	29.19 (7.12)	30.00 (7.13)	29.62 (7.03)
Value	40.50 (9.14)	38.03 (9.51)	41.58 (9.61)	41.45 (10.28)	40.02 (9.50)
Difficulty	25.04 (7.04)	27.26 (7.03)	25.58 (6.35)	23.45 (5.28)	25.78 (6.71)
CULCEI Subscales					
Personalization	24.44 (3.00)	23.56 (4.22)	24.42 (3.95)	25.30 (2.50)	24.23 (3.68)
Involvement	22.56 (2.49)	22.91 (3.17)	22.58 (3.48)	23.60 (2.17)	22.80 (2.97)
Cohesiveness	18.24 (3.67)	20.75 (6.14)	22.65 (3.57)	22.70 (2.71)	20.82 (4.85)
Satisfaction	19.08 (2.47)	18.91 (4.04)	16.23 (4.85)	16.90 (3.07)	17.99 (4.00)
Task Orientation	23.56 (2.95)	21.25 (2.93)	18.62 (4.39)	18.30 (3.40)	20.82 (3.96)
Innovation	15.96 (2.65)	20.03 (3.17)	20.27 (3.03)	23.50 (2.27)	19.34 (3.69)
Individualization	16.44 (1.90)	19.41 (2.84)	19.04 (2.89)	20.50 (1.96)	18.62 (2.87)

Table 8: Correlation of SATS Subscales with Discriminant Function (Function Structure Matrix) and Standardized Discriminant Function Coefficients for Students in the Traditional Classes

SATS Subscale	Correlation with Discriminant Function	Standardized Discriminant Function Coefficient
Affect	.472	1.250
Cognitive Competency	.323	.266
Value	.232	- .157
Difficulty	-.277	-1.302

Note: Group Centroid for Instructor A = .633; Group Centroid for Instructor B = -.507

College and University Lecture Classroom Environment Inventory

A 2 x 2 factorial MANOVA was conducted with the seven CULCEI subscales as dependent variables. While there was not a significant interaction between method and instructor ($F(7,83)=1.19$, $p=.319$), main effects for both method ($F(7,83)=14.14$, $p<.001$) and instructor ($F(7,83)=6.56$, $p<.001$) were statistically significant. (See Table 7 for group means and standard deviations.) Follow-up analyses using discriminant analysis (see Table 9) reveal that the alternative classes had higher cohesiveness and innovation scores and lower satisfaction and task orientation scores than did the traditional classes. Personalization had a high standardized discriminant function coefficient for the method analysis; however, because scores on this subscale correlate highly ($r=.780$) with scores on the involvement subscale, it is thought that scores on the personalization subscale do not contribute uniquely to the function. The scores from Instructor A's students were lower on innovation and individualization than were the scores from Instructor B's students.

Table 9: Correlation of CULCEI Subscales with Discriminant Function (Function Structure Matrix) and Standardized Discriminant Function Coefficients

CULCEI subscale	Correlation with Discriminant Function	Standardized Discriminant Function Coefficient
Method		
Personalization	-.110	.519
Involvement	.020	.100
Student Cohesiveness	.366	.286
Satisfaction	-.378	-.485
Task Orientation	-.595	-.716
Innovation	.481	.515
Individualization	.268	-.077
Instructor		
Personalization	-.115	-.955
Involvement	.157	.167
Student Cohesiveness	.139	.076
Satisfaction	.187	.146
Task Orientation	-.115	.187
Innovation	.730	.632
Individualization	.652	.593

Note: Group centroid for Traditional method = -.688; Group centroid for Alternative method = 1.089; Group centroid for Instructor A = -.484; Group centroid for Instructor B = .588

Discussion

The purpose of this study was to explore the effects of instructional method, instructor and their interaction on the achievement, attitudes, and self-efficacy of introductory statistics students. While we anticipated the presence of interactions on the outcome measures, we made no specific predictions as to the nature of the expected effects. Another objective of this study was to verify that alternative forms of statistics instruction resulted in more positive student outcomes. We expected to find that students who have been exposed to alternative, activity-based teaching methods would have higher achievement (as measured by the final exam) as well as higher self-efficacy and value for statistics, higher feelings of cognitive competency in statistics, a greater sense of student involvement and cohesiveness in their statistics class, and stronger feelings of innovation and individualization. We predicted that students in the traditionally-taught classes would feel more positive toward their statistics class and would be more satisfied with it.

The only significant instructor/method interaction we found was on the multivariate SATS subscales. Specifically, students in Instructor A's traditional class had higher scores on the SATS than the students in Instructor B's traditional class, with the affect and cognitive competency subscales being the primary contributors to this difference. Value (how worthy do the students find statistics as a subject) and difficulty (how difficult is the subject of statistics) did not contribute much to the differences and no differences were found on the SATS among students in the alternative classes.

Our results regarding our hypotheses about the main effects of instructor and of method were mixed. While there were no differences between the teaching methods on the final exam scores, student involvement or individualization, we did find that the scores of students in the alternative classes were higher on the CULCEI subscales than the scores of the students in the traditional classes, and that this difference resulted primarily from the alternative classes' higher scores on the cohesiveness and innovation subscales and their lower scores on the satisfaction and task orientation subscales. In addition, although the difference was not statistically significant, scores on the self-efficacy scale were higher for students in the alternative classes. We did not find the anticipated effects on the value and affect subscales of the SATS or on the student involvement and individualization subscales on the CULCEI. These results indicate that students in the alternative classes felt that the members of their class worked well together and that the instructor incorporated original tasks and activities; they also were less happy with their statistics class than the traditional students. Moreover, students in the alternative classes felt that the structure of the class was somewhat disorganized. However, students in both the alternative and traditional classes felt equally involved in the experiences of the class and both groups of students agreed on the extent to which their individual differences were acknowledged (as measured by the involvement and individualization subscales of the CULCEI, respectively).

Thus, contrary to much of the prevailing literature on statistics reform, incorporating activity-based, experiential learning reforms in the statistics curriculum did not result in higher achievement for the students in our study. The fact that there were no differences on the final exam is disappointing, but may be attributed to the exam itself. As Table 4 illustrated, the overall average score on the final exam was 18.95 points; the number of possible points was 28. Thus, the average percentage score for all students on the final exam was 68%. Thirty percent of all students failed the exam (received a score below 60%) and less than five percent of students received an "A" (90% or above). It seems evident that either the exam was an inadequate measure of students' achievement of learning goals or that the instruction did not adequately assist students in attaining these goals.

One possible explanation for these results may be that the instructors involved were very new to this style of teaching. It is acknowledged that constructivist-type teaching methods involve more planning and often result in less content coverage than do traditional lecture-style methods (Garfield, 1997; Moore, 1997; Steinhorst & Keeler, 1995). Additionally, it is clear that instructors using these methods must have clearly established goals in mind prior to teaching the course (Garfield, 1995; Hoerl, Hahn & Doganaksoy, 1997; Scheaffer, 1997). Hubbard (1997) acknowledges that the process of change for instructors is a gradual one. It is obviously a difficult task for an instructor with only a few semesters of any statistics teaching experience to implement an activity-based curriculum and to develop methods of assessment that accurately measure student achievement, and this difficulty may have resulted in less effective presentation. The implication for a statistics education is that changing one's teaching method to activity-based instruction is not a simple leap. Furthermore, this change does not guarantee an immediate increase in student achievement. Rather, as a statistics teacher, one needs to take the perspective that changing the teaching paradigm entails researching one's own teaching and assessment practices in order to gain insight to and experience with more authentic forms of instruction and assessment.

Students' self-efficacy (their confidence in their abilities to perform statistical tasks), however, does seem to have been positively affected by the alternative teaching format. This would be anticipated given the postulated benefits of activity-based learning. By the end of the semester, the constructivist style of instruction would be expected to result in better learning (Garfield, 1995) and better retention of the learned information they learned (Keeler & Steinhorst, 1995; Moore, 1997). These outcomes would therefore result in increased self-efficacy.

It is interesting to note that there were some differences between the students on the Survey of Attitudes Toward Statistics, but that these differences appeared only in the traditional classes. The expected improvements in attitudes did not materialize for students in the alternative classes. Instructor A in this study is a woman, while Instructor B is a man. This gender disparity may have contributed to some of the attitude differences; if the impression that women are more nurturing and caring holds true, it stands to reason that students who are taught by women will reflect higher scores on emotionally-related scales such as affect and cognitive competency. This theory, however, is partially contradicted when examining the main effect for instructor on the CULCEI. The individualization subscale, which reflects the degree to which students feel that their individual differences are incorporated into the class, seems to include an aspect of emotional comfort. However, Instructor A's students reflected a lower score on the individualization subscale than did Instructor B's students.

The scores on the CULCEI offer further evidence concerning potential difficulties that instructors who are inexperienced in alternative teaching methods may encounter. The nature of the alternative class curriculum (group-work, projects, activities) resulted in higher perceptions of student cohesiveness and innovation for students in those classes than for students in the traditional class. These results are not surprising; one would expect that classes that involve group-work would lead to the perception on the part of the students that the members of their class know each other and work well together. Similarly, the fact that students in the alternative classes are being exposed to a multitude of, perhaps novel, instructional tactics would intuitively lead one to believe that they would perceive more innovation in their classroom than would students who are taught via lecture. What is perhaps more interesting is the contribution made by the scores on the task-orientation subscale. Students in the alternative classes tended to score lower on this measure than did students in the traditional classes, indicating that the alternative class students felt a sense of disorganization or confusion regarding their projects. This again supports the belief that this type of instructional method may not have the anticipated benefits during the instructor's first few attempts at it.

The results of this study lend at least partial support to the notion that activity-based, constructivist styles of learning may benefit students. While an overall endorsement of alternative methods of teaching is not warranted, it seems clear that this type of instruction may help improve statistics students' attitudes and opinions regarding their statistics class. Caution should be used, however, in concluding that this method will benefit all students, all of the time. Instructors should consider the possibility that student characteristics may impact their success in alternative-style statistics courses. Additionally, it should be acknowledged that instructor characteristics are likely to have some impact on the effectiveness of reorganized statistics courses. It is important for the instructor to consider their abilities, interests and learning objectives along with their students' attitudes, beliefs, and intellectual development when planning and revising the statistics curriculum.

Further research in this area should continue to focus on the role that student and instructor characteristics play in the anticipated benefits of a refined curriculum. Prior measures of instructor attitudes toward change or instructor caring may be useful in predicting the success of the course, and potential differences between instructors who are experienced in teaching using these recommended methods versus novice constructivist-style instructors should be explored. Additionally, there may be differences based on the gender of the instructor and/or the students. The results of the interviews indicate that it is vital that instructors feel confident and prepared as they adopt this new method of teaching. The process of incorporating constructivist concepts is undoubtedly an evolutionary one, and instructors should continually monitor their teaching practices and their students' learning in context. Future researchers should also include more assessments of achievement of learning goals, as well as analyses of the assessment methods.

References

Becker, B.J. (1996). A look at the literature (and other resources) on teaching statistics. *Journal of Educational and Behavioral Statistics*, 21(1), 71-90.

Bradstreet, T.E. (1996). Teaching introductory statistics so that nonstatisticians experience statistical reasoning. *The American Statistician*, 50(1), 69-78.

Cobb, G.W. (1993). Reconsidering statistics education: A national science foundation conference. *Journal of Statistics Education*, [Online], 1(1), Available: http://www.amstat.org/publications/jse/v1n1/cobb.html

Davidson, N. & Kroll, D.L. (1991). An overview of research on cooperative learning related to mathematics. *Journal for Research in Mathematics Education*, 22(5), 362-365.

Fraser, B.J. & Treagust, D.F. (1986). Validity and use of an instrument for assessing classroom psychosocial environment in higher education. *Higher Education,* 15, 37-57.

Garfield, J. (1993). Teaching statistics using small-group cooperative learning. *Journal of Statistics Education* [Online], 1(1), Available: http://www.amstat.org/publications/ jse/v1n1/ garfield.html

Garfield, J. (1995). How students learn statistics. *International Statistical Review, 63*, 25-43.

Garfield, J. (1997). Discussion (Response to Moore, 1997). *International Statistical Review,* 65, 137-141.

Garfield, J. and Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal for Research in Mathematics Education,* 19(1), 44-63.

Giraud, G. (1997). Cooperative learning and statistics instruction. *Journal of Statistics Education*, [Online], 5(3), Available: http://www.amstat.org/publications/jse/v5n3/giraud.html

Graham, A. (1987). *Statistical Investigations in the Secondary School.* Cambridge, England: Cambridge University Press.

Gravetter, F.J. & Wallnau, L.B. (1988). *Statistics for the behavioral sciences: A first course for students of psychology and education.* St. Paul, MN: West Publishing Co.

Hoerl, R., Hahn, G. & Daganaksoy, N. (1997). Discussion: Let's stop squandering our most strategic weapon (Response to Moore, 1997). *International Statistical Review,* 65, 147-153.

Hogg, R.V., (1991). "Statistical education: Improvements are badly needed," *The American Statistician,* 45(4), 342-343.

Hubbard, R. (1997). Assessment and the process of learning statistics. *Journal of Statistics Education*, [Online], 5(1), Available: http://www.amstat.org/publications/jse/v5n1/hubbard.html

Kader, G., & Perry, M. (1994). Learning statistics with technology. *Mathematics Teaching in Middle Schools,* 1, 130-136.

Keeler, C. and Steinhorst, R.K. (1995). Using small groups to promote active learning in the introductory statistics course: A report from the field. *Journal of Statistics Education*, [Online], 3(2), Available: http://www.amstat.org/publications/jse/v3n2/keeler.html

Konold, C. (1995). Issues in assessing conceptual understanding in probability and statistics. *Journal of Statistics Education*, [Online], 3(1), Available: http://www.amstat.org/publications/jse/ v1n1/cobb.html

Moore, D.S. (1997). "New pedagogy and new content: The case of statistics." *International Statistical Review,* 65(2), 123-137.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research,* 66(4), 543-578.

Rinaman, W.C. (1998). Revising a basic statistics course. *Journal of Statistics Education,* [Online], 6(2), Available: http://www.amstat.org/publications/jse/v6n2/rinaman.html

Roiter, K. & Petocz, P. 91996). Introductory statistics courses – a new way of thinking. *Journal of Statistics Education*, [Online], 4(2), Available: http://www.amstat.org/publications/jse/v4n2/roiter.html

Schau, C., Stevens, J., Daupinee, T.L., & Delvecchio, A. (1995). The development and validation of the survey of attitudes toward statistics. *Educational and Psychological Measurement,* 55, 868-875.

Scheaffer, R.L. (1997). Discussion (Response to Moore, 1997). *International Statistical Review,* 65, 156-158.

Scheaffer, R.L., Gnanadesikan, M., Watkins, A., & Witmer, J.A. (1996). *Activity-based statistics,* New York: Springer-Verlag.

Schuh, K.L. (1996). The lecture classroom environment and its effects on change in computer anxiety of students taking computer proficiency classes. *Journal of Educational Computing Research,* 15(3), 241-259.

Smith, G. (1998). Learning statistics by doing statistics. *Journal of Statistics Education,* [Online], 6(3), Available: http://www.amstat.org/publications/jse/v6n3/smith.html

Steinhorst, R.K. & Keeler, C.M. (1995). Developing material for introductory statistics courses from a conceptual, active learning viewpoint. *Journal of Statistics Education*, [Online], 3(3), Available: http://www.amstat.org/publications/jse/v3n3/steinhorst.html

Appendix A

Alternative Class Course Outline

A. Introduction
   - Where does statistics fit into the world?
   - What does a statistician do?
   - Why study statistical methods?

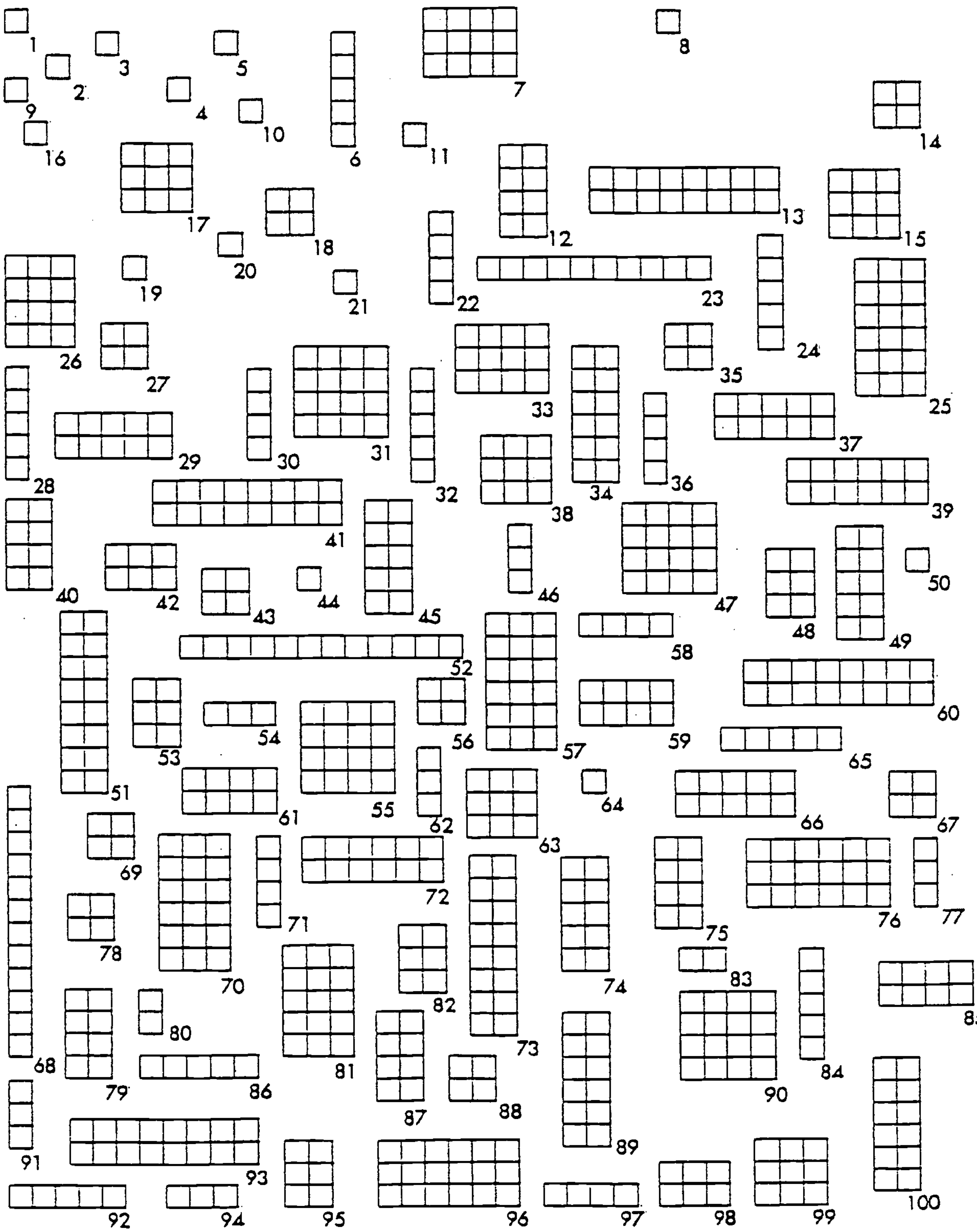
B. Understanding basic statistical concepts by way of proportions
   - Population, parameters, samples and statistics
   - Interpreting results: The role of prior experience and subjectivity
   - The "ins and outs" of sampling
   - Statistics, Estimation, and Bias
   - Sampling Distributions: What are they, and why are they important?
   - The role of probability and chance
   - The logic of hypothesis tests
   - Hypothesis tests for proportions
   - Confidence intervals for proportions
   - Putting it all together simultaneously is statistical reasoning
   - Critiquing numerical arguments

C. Conducting larger statistical surveys and analyzing the data
   - The purpose of a survey
   - Questionnaire Design: The different kinds of possible variables
   - Survey Design: Are your final interpretations justified?
   - Finding patterns in the data: Graphical procedures
   - Summarizing the data: Descriptive statistics
   - Testing for a relationship between two categorical variables
     - The Chi-square test
   - Testing for a relationship between two continuous variables
     - Correlation
     - Simple Regression
   - Testing for a relationship between a categorical and a continuous variable
     - T-tests and/or ANOVA

D. Statistical Experimentation: What is it and how do I do it?
   - Designing an Experiment
   - Experimental Error vs. Sampling Variability
   - Working with human subjects (experimental design and randomization)
   - Analyzing the data
     - Paired data tests
     - Multiple independent sample tests

Appendix B

Random Rectangles



Appendix C

Self-Efficacy Items

	Strongly Disagree			Neither			Strongly Agree
1. I think I could use statistical information in making everyday decisions.	1	2	3	4	5	6	7
2. I would be confused if I had to decide what statistical test to use for a given research question.	1	2	3	4	5	6	7
3. I could recognize flaws in research studies.	1	2	3	4	5	6	7
4. I would have difficulty explaining the results of a correlational study.	1	2	3	4	5	6	7
5. I could understand information given in a statistical graph.	1	2	3	4	5	6	7
6. I am able to identify incorrect conclusions from research studies in newspaper articles.	1	2	3	4	5	6	7
7. I can explain the concept of variability.	1	2	3	4	5	6	7
8. I am not sure I could properly interpret the results statistical tests covered in this class.	1	2	3	4	5	6	7
9. I understand the advantages to using a random sample.	1	2	3	4	5	6	7
10. I could explain to my parents the logic of hypothesis testing.	1	2	3	4	5	6	7

Appendix D

Take-Home Final Exam

The following is an exam and is to be completed by you alone. Working with others is not allowed. You may use your book, your notes, and any other written resources you can find. If you have questions, do not ask your classmates. Do not ask your friends. Do not ask the CASI lab staff. Visit, call or e-mail your instructor. Violation of these instructions constitutes a violation of the Academic Honesty policy and could result in a failing grade. Your exam should be typed and double-spaced. Make sure your name is on every page of your write-up. This exam is due at the beginning of your scheduled final session.

Recently, The Point radio station reported on a study done at Yale University in which a negative correlation was found between the number of years a man lived and the physical beauty (or lack thereof) of his wife. The interpretation given by the disc jockey was that "marrying a really beautiful woman shaves years off of your life."

1) Based on the information you have, do you believe that the D.J. offered a valid conclusion? Why or why not?

2) What are three things you would want to know to determine if this were a valid research study? Why?

3) Are there other possible explanations for the observed differences between groups? If so, what are they? If not, why not?

4) In terms of this research scenario, what would constitute a Type II error? Don't give a definition of a Type II error; explain what a Type II error would be *in this particular situation.*

The following is an adaptation of a study found in an academic journal. Use this article to answer questions 5-8.
[*Students were presented with an excerpt from a study by Heckert, Mueller, Roberts, Hannah, Jones, Masters, Bibbs & Bergman (1999), entitled "Personality Similarity and Conflict among Female College Roommates" from The Journal of College Student Development, vol. 40, no. 1, pp. 79-81.*]

5) Discuss the interpretations the researchers could make based on the results of this study. Address all of the statistical tests that are included.

6) Evaluate the sample that was used in terms of the population that they would be able to represent. What are the strengths and/or limitations of the sample that was used?

7) Were the statistical tests that the researchers used appropriate for the type of data they collected? Why or why not?

8) What do the p-values actually measure in the context of these analyses?

AERA

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC®

TM031492

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: PREDICTING ACQUISITION OF LEARNING OUTCOMES: A COMPARISON OF TRADITIONAL AND ACTIVITY-BASED INSTRUCTION IN AN INTRODUCTORY STATISTICS COURSE

Author(s): JENENNE A. GESKE, WILLIAM T. MICKELSON, DEBORAH L. BANDALOS, JESSICA JONSON, RUSSELL W. SMITH

Corporate Source: UNIVERSITY OF NEBRASKA-LINCOLN

Publication Date: APRIL 28, 2000

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY ____ Sample ____ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY ____ Sample ____ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY ____ Sample ____ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
1	2A	2B
Level 1	Level 2A	Level 2B
[✓]	[ ]	[ ]
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, please

Signature: Jenenne Geske

Printed Name/Position/Title:

Organization/Address: University of Nebraska-Lincoln

Telephone: (402) 472-2580

FAX:

E-Mail Address: jgeske@unlserve.unl.edu

Date: 6/6/00

(over)

ERIC® Clearinghouse on Assessment and Evaluation

University of Maryland
1129 Shriver Laboratory
College Park, MD 20742-5701

Tel: (800) 464-3742
(301) 405-7449
FAX: (301) 405-8134
ericae@ericae.net
http://ericae.net

March 2000

Dear AERA Presenter,

Congratulations on being a presenter at AERA. The ERIC Clearinghouse on Assessment and Evaluation would like you to contribute to ERIC by providing us with a written copy of your presentation. Submitting your paper to ERIC ensures a wider audience by making it available to members of the education community who could not attend your session or this year's conference.

Abstracts of papers accepted by ERIC appear in *Resources in Education (RIE)* and are announced to over 5,000 organizations. The inclusion of your work makes it readily available to other researchers, provides a permanent archive, and enhances the quality of *RIE*. Abstracts of your contribution will be accessible through the printed, electronic, and internet versions of *RIE*. The paper will be available **full-text, on demand through the ERIC Document Reproduction Service** and through the microfiche collections housed at libraries around the world.

We are gathering all the papers from the AERA Conference. We will route your paper to the appropriate clearinghouse and you will be notified if your paper meets ERIC's criteria. Documents are reviewed for contribution to education, timeliness, relevance, methodology, effectiveness of presentation, and reproduction quality. You can track our processing of your paper at **http://ericae.net.**

To disseminate your work through ERIC, you need to sign the reproduction release form on the back of this letter and include it with **two** copies of your paper. You can drop of the copies of your paper and reproduction release form at the ERIC booth (223) or mail to our attention at the address below. **If you have not submitted your 1999 Conference paper please send today or drop it off at the booth with a Reproduction Release Form.** Please feel free to copy the form for future or additional submissions.

Mail to: AERA 2000/ERIC Acquisitions
The University of Maryland
1129 Shriver Lab
College Park, MD 20742

Sincerely,

Lawrence M. Rudner

Lawrence M. Rudner, Ph.D.
Director, ERIC/AE

ERIC/AE is a project of the Department of Measurement, Statistics and Evaluation at the College of Education, University of Maryland.