AN INVESTIGATION INTO THE EFFECTIVENESS OF
INTELLIGENT TUTORING ON LEARNING
OF COLLEGE LEVEL STATISTICS

DISSERTATION

Presented to the Graduate Council of the
University of North Texas in Partial
Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

By

Nanta Palitawanont, B.Ed., M.Ed.
Denton, Texas
May, 1989
Palitawanont, Nanta, An Investigation into the 
Effectiveness of Intelligent Tutoring on Learning of College 
Level Statistics. Doctor of Philosophy (College Teaching), 
May, 1989, 90 pp., 11 tables, 6 figures, bibliography, 37 
titles.

The present research incorporated the content of basic 
statistics into the Artificial Intelligence Physics Tutor 
(ARPHY), which was used as the expert system shell, and 
investigated the effects of the Artificial Intelligent 
Statistics Tutor (ARSTAT) as a supplement to learning 
statistics at the college level.

Two classes of an introductory educational statistics 
course in the Department of Educational Foundations, 
University of North Texas, were used in the study. The day-
time class was used as the experimental group and the 
evening class was used as the control group. The 
experimental group's lecture/discussion was supplemented 
with ARSTAT, and the control group received only 
lecture/discussion. A one-way analysis of covariance was 
used to compare students' test scores. No significant 
difference was found; however, the adjusted mean score of 
the experimental group was slightly higher than that of the 
control group. A two-way analysis of covariance showed no 
significant main effect or interaction between gender and
study technique. A second two-way analysis of covariance showed no significant interaction between the students' attitude toward statistics and the study technique used. However, the students with a statistics-positive attitude scored significantly higher on the test than students who had a negative attitude toward statistics.

This study concluded that the ARSTAT can be used effectively as a tutor for students taking an introductory course in educational statistics.

The following recommendations for further study were made: incorporate more advanced topics of statistics into the ARPHY teaching model; incorporate the ARPHY learning theory and statistical content using another version of LISP language or another programming language such as PROLOG; and compare the ARSTAT tutor to some other kind of supplement to lecture/discussion.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>vi</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF ILLUSTRATIONS</td>
<td>vii</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td></td>
</tr>
<tr>
<td>Major Purposes of the Study</td>
<td></td>
</tr>
<tr>
<td>Specific Purposes of the Study</td>
<td></td>
</tr>
<tr>
<td>Hypotheses</td>
<td></td>
</tr>
<tr>
<td>Significance</td>
<td></td>
</tr>
<tr>
<td>Definition of Terms</td>
<td></td>
</tr>
<tr>
<td>Limitations</td>
<td></td>
</tr>
<tr>
<td>Delimitations</td>
<td></td>
</tr>
<tr>
<td>Assumptions</td>
<td></td>
</tr>
<tr>
<td>Development of the Software</td>
<td></td>
</tr>
<tr>
<td>Procedure for Collection of Data</td>
<td></td>
</tr>
<tr>
<td>Procedures for Data Analysis</td>
<td></td>
</tr>
<tr>
<td>Chapter Bibliography</td>
<td></td>
</tr>
<tr>
<td>II. REVIEW OF RELATED LITERATURE</td>
<td>16</td>
</tr>
<tr>
<td>Computer-Assisted Instruction</td>
<td></td>
</tr>
<tr>
<td>Use of the Computer in the Teaching of Statistics</td>
<td></td>
</tr>
<tr>
<td>Computer-Assisted Instruction of Statistics</td>
<td></td>
</tr>
<tr>
<td>Intelligent Computer-Assisted Instruction</td>
<td></td>
</tr>
<tr>
<td>The Components of ICAI Systems</td>
<td></td>
</tr>
<tr>
<td>Knowledge Representation in Expert System</td>
<td></td>
</tr>
<tr>
<td>Conclusions</td>
<td></td>
</tr>
<tr>
<td>Chapter Bibliography</td>
<td></td>
</tr>
<tr>
<td>III. PROCEDURES OF THE STUDY</td>
<td>33</td>
</tr>
<tr>
<td>Incorporation of the Content of Statistics into the ARPHY Teaching Module</td>
<td></td>
</tr>
<tr>
<td>Expert Module</td>
<td></td>
</tr>
<tr>
<td>Student Module</td>
<td></td>
</tr>
<tr>
<td>Tutorial Strategy Module</td>
<td></td>
</tr>
<tr>
<td>Chapter</td>
<td>page</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>Communication Module</td>
<td></td>
</tr>
<tr>
<td>Completion of Tutorial</td>
<td></td>
</tr>
<tr>
<td>Equipment</td>
<td></td>
</tr>
<tr>
<td>Procedures for Data Collection and Analysis</td>
<td></td>
</tr>
<tr>
<td>Selection of the Subjects</td>
<td></td>
</tr>
<tr>
<td>Description of Data Collection</td>
<td></td>
</tr>
<tr>
<td>Variables</td>
<td></td>
</tr>
<tr>
<td>Statistics Procedures Used in</td>
<td></td>
</tr>
<tr>
<td>Data Analysis</td>
<td></td>
</tr>
<tr>
<td>Chapter Bibliography</td>
<td></td>
</tr>
<tr>
<td>IV. RESULT OF DATA ANALYSIS</td>
<td>59</td>
</tr>
<tr>
<td>Description of the Population</td>
<td></td>
</tr>
<tr>
<td>Results of the Hypotheses</td>
<td></td>
</tr>
<tr>
<td>Additional Results</td>
<td></td>
</tr>
<tr>
<td>V. DISCUSSION, SUMMARY, RECOMMENDATIONS</td>
<td>68</td>
</tr>
<tr>
<td>Discussion of the Hypotheses</td>
<td></td>
</tr>
<tr>
<td>Discussion of Additional Results</td>
<td></td>
</tr>
<tr>
<td>Summary and Conclusions</td>
<td></td>
</tr>
<tr>
<td>Recommendations</td>
<td></td>
</tr>
<tr>
<td>Chapter Bibliography</td>
<td></td>
</tr>
<tr>
<td>APPENDICES</td>
<td>76</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>84</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table | Page
--- | ---
1. First Generation Prerequisites of the Four Statistics Problems | 41
2. Second Generation Prerequisites of the Four Statistics Problems | 42
3. Third Generation Prerequisites of the Four Statistics Problems | 43
4. Fourth Generation Prerequisites of the Four Statistics Problems | 44
5. Summary of Analysis of Covariance Achievement Test Scores by Study Technique | 62
6. Mean Scores and Adjusted Mean Scores on Achievement Test | 62
7. Summary of Two-Way Analysis of Covariance Achievement Test Scores by Study Technique and Sex | 63
8. Mean Scores and Adjusted Mean Scores on Achievement Test | 64
9. Summary of Two-Way Analysis of Covariance Achievement Test Scores by Study Technique and Attitude Toward Statistics | 65
10. Mean Scores and Adjusted Mean Scores on Achievement Test | 66
11. Summary of Students Evaluation of the Artificial Intelligent Statistics Tutor | 83
LIST OF ILLUSTRATIONS

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Networked hierarchy of the four problems of statistics building in the expert module</td>
<td>35</td>
</tr>
<tr>
<td>2.</td>
<td>Initial depth-first pass in recursion for building the statistics net-archy for Problem 1</td>
<td>37</td>
</tr>
<tr>
<td>3.</td>
<td>Initial depth-first pass in recursion for building the statistics net-archy for Problem 2</td>
<td>38</td>
</tr>
<tr>
<td>4.</td>
<td>Initial depth-first pass in recursion for building the statistics net-archy for Problem 3</td>
<td>39</td>
</tr>
<tr>
<td>5.</td>
<td>Initial depth-first pass in recursion for building the statistics net-archy for Problem 4</td>
<td>40</td>
</tr>
<tr>
<td>6.</td>
<td>High-level structure of the tutorial strategy module (TSM)</td>
<td>46</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

The rapid growth and acceptance of microcomputers are creating an enormous impact on instruction in colleges and universities. Teaching by computer has been introduced with both good and questionable results in classrooms from kindergarten to graduate school (Elliott, 1985, p. 39). There have been three general approaches in the use of computers as devices which interact directly with the students, rather than as assistants to the human teacher. First, the environmental approach allows the student to use the machine more or less free style. The second approach uses games and simulations as instructional tools. The third computer application in education is computer-assisted instruction (CAI). Carbonell defines a second type of CAI that is known today as knowledge-based or Intelligent CAI (ICAI). ICAI systems and the earlier CAI systems both have representations of the subject matter they teach, but ICAI systems also carry on a dialogue with the student and use the mistakes of the student to diagnose his knowledge deficits (Barr, & Feigenbaum, 1982).

Intelligent Computer-Assisted Instruction (ICAI) attempts to provide a natural (computer-based) environment
for the student and tutor in a one-on-one situation. Artificial Intelligence (AI) systems differ from conventional scientific or data processing computing systems in several ways. In conventional computing, a program embodies an algorithm that provides a solution to the problem at hand. Given the input data, the program executes a fixed sequence of operations to produce the correct output. Languages such as FORTRAN, BASIC, or PASCAL are commonly used. In contrast, AI systems manipulate and reason using characteristics and/or attributes expressed in concepts and establish relationships among concepts. As with complex human reasoning, the data available are often in error, uncertain, or changing during the course of problem solving. There are no known algorithms that produce "guaranteed" or "best" solutions. AI systems are programmed in LISP or PROLOG languages, which permit a facile representation of logical relationships among complex knowledge structures (Hayes-Roth and Thorndyke, 1985, p. 231).

The purpose of this research was to incorporate the content of basic statistics into an artificial intelligence learning-based system shell. The Artificial Intelligence Physics tutor (ARPHY) was used as the artificial intelligence learning-based system shell, a system developed in 1985 by Stephen F. Brown to teach non-calculus physics at the University of North Texas. The present study attempted
to investigate the effectiveness of intelligent tutoring in learning statistics at the college level.

Statement of the Problem

The problem of this research was to determine whether there were significant differences in the learning achievement of college students in descriptive statistics when the lecture/discussion was supplemented with the Artificial Intelligent Statistics Tutor.

Major Purposes of the Study

The major purposes of this study were as follows:

1. to incorporate the basic content of statistics (Statistical Concepts, Frequency Distributions and Graphs, Measures of Central Tendency, and Measures of Variation) into the teaching module of the ARPHY tutor; and

2. to determine the feasibility of an artificial intelligent statistics tutor in promoting and enhancing learning.

Specific Purposes of the Study

The specific purposes of the study were as follows:

1. to determine if students who learned by lecture/discussion plus the Artificial Intelligent Statistics Tutor could correctly answer a significantly greater number of questions on a test than students who learned by lecture/discussion only;

2. to determine if there was a significant difference
in interaction between gender and treatment (study technique) in achievement test scores;

3. to determine if students' attitudes toward statistics were a significant factor in learning statistics; and

4. to determine if the Artificial Intelligent Statistics Tutor could be used effectively as a tutor for students taking EDUC 5210, Educational Statistics.

Hypotheses

To carry out the purposes of this study, the following hypotheses were tested:

1. Students who learn by lecture/discussion plus the artificial intelligent statistics tutor will answer correctly a significantly greater number of questions on a test than students who learn by lecture/discussion only.

2. There will be a significant difference in interaction between gender and treatment (lecture/discussion plus the artificial intelligent statistics tutor and lecture/discussion only) when using an achievement test as the dependent variable.

3. Students who have a positive attitude in statistics and who learn by lecture/discussion plus the artificial intelligent statistics tutor will answer correctly a significantly greater number of test items than students with a statistics-positive attitude who learn by lecture/discussion only.
Significance

This study incorporated the content of statistics into the teaching module of the ARPHY tutor. The significance of this study is that it determined whether the artificial intelligent statistics tutorial system could be used effectively as a supplement in the teaching/learning of statistics. It determined whether students who learn by lecture/discussion presentation plus tutoring with the artificial intelligent statistics tutorial system could answer correctly a significantly greater number of questions on the test and have more knowledge retention than students who learned by lecture/discussion presentation only. It determined if the students' attitude toward statistics was a factor in learning statistics. It determined if there was a significantly different interaction between gender and treatment (study technique) in the achievement test scores. Perhaps the most significant result of this study was the production of a truly learner-based software program for statistics instruction.

Definition of Terms

The following terms are defined for the purposes of this study.

Artificial Intelligence (AI). The area of study or discipline which seeks the development of computer software which will allow the computer to perform tasks which, at the time, people are better at performing (Rich, 1983). Some
fields or specializations which are developing artificial intelligence include speech recognition, pattern recognition, deep space probes, advanced chess programs, numerical integration programs, and other applications where human reasoning is the goal of the program.

Intelligent Computer-Assisted Instruction (ICAI). Instruction through the use of computer software which has elements of artificial intelligence incorporated into its modules.

LISP. A loose acronym for LISt Programming. LISP is the premier language for AI programming due to its ability to adapt to concept and symbolic manipulation. The specific version of LISP used in this dissertation was COMMON LISP. This version is the most universally used of all LISP dialects.

Software. The intangible portion of a computer system which includes the logic guiding the control of the operation of the computer and the program.

Limitations

ICAI systems represent the state-of-the-art in computer-based instruction. However, the systems are not yet readily available and are still being developed. The limiting factors are the huge amounts of time and effort required to build an ICAI system, the costly hardware requirements, and the narrow range of content domains for which ICAI systems have been built. During the course of
this study, the content which was incorporated into the ICAI system shell was limited to the four chapters of the course materials are: statistical concepts, frequency distribution, measures of central tendency, and measures of variation.

Delimitations

The population in this study consisted of the students in two sections of the Educational Statistics class (EDUC 5210), Department of Educational Foundations, College of Education, University of North Texas.

Assumptions

In the development of this study, several assumptions were made. First, it was assumed that students would respond honestly to the Attitude Toward Statistics Test. Secondly, a basic assumption was made that the students would not put "bugs" into the ICAI software. These bugs can be introduced by typing errors, or entering incorrect answers.

Development of the Software

To develop the Artificial Intelligent Statistics Tutor, the Artificial Intelligent Physics Tutor (ARPHY) was used as a system shell. The ARPHY was developed by Steven F. Brown for his doctoral dissertation in 1985. The system was developed to teach non-calculus physics at the first semester college level at University of North Texas. The system incorporates Ausubel's advance organizer, Gagne's
level of learning, and Bloom's classification system for cognitive behaviors into an ICAI system.

The structure of the ARPHY was constructed in four sections corresponding to the modularity of ICAI systems. The four sections were expert, student, tutorial strategy, and communications modules. The first module, the expert module, contained the domain of knowledge. The learning theories were integrated and the learning types were represented as a semantic net within this module. The second module, the student module, kept all the records specific to each student which were accessed by the other three modules. Information kept by this module included the student's reading speed, topics completed, areas of misconceptions, and the path instruction had taken through the topics. The third module, the tutorial strategy module, was the controller of the system, using forward chaining as a mechanism to drive the system. The last module, the communication module, contained two basic types of communication: direct and indirect. This module was a translator between the other three modules.

The ARPHY was used as a system shell by removing its domain-physics knowledge in the expert module and replacing it with the domain-statistics knowledge. The inference engine and support facilities of the other modules retained their identity.

Since the Artificially Intelligent Statistics Tutor was
a generation of the ARPHY, it was named ARSTAT, which is an acronym for ARTificially intelligent STATistics tutor.

More details for the structure of the ARPHY system and the incorporation of the content of statistics into the ARPHY system will be discussed in Chapter 3.

Procedure for Collection of Data

The following is an explanation of the procedure used to investigate the effectiveness of the ARSTAT as a supplement to learning statistics.

Student Population

The population used for the development of this study was seventy students from two classes of the Educational Statistics course (EDUC 5210), Fall 1988, Department of Educational Foundations, College of Education, University of North Texas.

Instruments

The Attitude Toward Statistics Test is a seven-point scale of ten pairs of questions designed to measure students' attitude toward statistics. It was developed by the Department of Educational Foundations (Educational Research), College of Education, University of North Texas. The internal reliability coefficient of the Attitude Toward Statistics Test was found with $r = .884$. The reliability coefficient was determined in the 1988 spring semester with 39 students, using Cronbach's alpha formula (SPSS, 1983).
Experiment

A pilot study was conducted during the 1988 spring semester. The subjects in the pilot study were students enrolled in two sections of the EDUC 5210 class, Department of Educational Foundations, College of Education, University of North Texas. In this pilot study, it was found that there were some "bugs" in the ARSTAT system and some errors in the content which had been incorporated into the system. The ARSTAT system was verified with the advice of Stephen F. Brown. The statistics content was verified with the advice of William K. Brookshire, Associate Professor in the Department of Educational Foundations, College of Education, University of North Texas. A second pilot study was administered during the first summer session of 1988. The ICAI software was validated in regard to the statistical content, graphics, and the ARPHY system. Also, the validated statistics content was approved by William K. Brookshire.

For the actual experiment, two classes of the Fall 1988 Educational Statistics course were used: A daytime class taught by Jon I. Young and an evening class taught by L. Fred Thomas. The students in the evening class were used as the control group because most of them lived off campus and it was too inconvenient for them to use the computer laboratory. The students in the daytime class were used as
the experimental group. Although the instructor for each group was different, the two men had similar credentials. Both were full professors who were experienced in teaching statistics and had taught this course for several years. Also, they co-authored the textbook used in both classes, *An Introduction to Educational Statistics: The Essential Elements*.

The experimental group received normal class lecture/discussion plus a tutorial with the artificial intelligent statistics tutoring system. The control group received only class lecture/discussion. In the first session of each class, students were administered a statistics attitude test which included a seven-point scale of ten pairs of questions. A pre-test of the content of the first four chapters was given to the students. A letter (Appendix A) to the students explained the study and noted that information collected would be recorded by code number and kept confidential. In the second class meeting of the experimental group, a second letter of consent (Appendix B) was handed to the students by the instructor. The letter asked students to participate in the tutorial session. Each student signed the letter indicating whether or not he or she would participate. Then the students studied the content of these chapters, (1) Statistical Concepts, (2) Frequency Distributions and Graphs, (3) Measures of Central Tendency, and (4) Measures of Variation, by
lecture/discussion. After the lecture/discussion of the first four chapters, the students in the experimental group who had agreed to participate in the tutorial session were asked to complete the statistics tutorial lesson by using the artificial intelligent statistics tutoring system. The students were informed that the ARSTAT took about four hours to complete. The content in the artificial intelligent statistics tutoring system corresponded to the first four chapters of the textbook and class discussion. The students were shown how to access the programs, and the objectives of the artificial intelligent tutoring system were explained. At the end of the tutorial session, each student completed a six-item five-point Likert scale of student evaluation (Appendix C).

The experimental group was asked to complete the tutorial program before the midterm examination. Only items on the midterm examination corresponding to the first four chapters were used for this study; the number of correct answers were recorded as achievement test scores.

Procedures for Data Analysis

To test the first hypothesis, a one-way analysis of covariance was used to analyze the difference in test scores between the students who learned by lecture/discussion plus the artificial intelligent tutoring system and the students who learned in the traditional lecture environment. Treatment (study technique) was used as the independent
variable. The students' pre-test scores were used as the covariate variable. The achievement test scores (midterm examination) were used as the dependent variable.

In the second hypothesis, a two-by-two analysis of covariance was utilized to compare the students. Study technique and sex were the independent variables. Students' pre-test scores were the covariate variable and the achievement test scores (midterm examination) were used as the dependent variable.

In order to test the third hypothesis, the students' attitude toward statistics was measured prior to the experiment by using a seven-point scale of ten pairs of questions from the Department of Educational Foundations, University of North Texas. A median was used to differentiate the students' attitudes toward statistics. The method of statistical analysis used for this hypothesis was a two-by-two analysis of covariance, with those who studied by lecture/discussion plus artificial intelligent statistics tutor and those who studied by lecture/discussion only in one dimension, and those who had a positive attitude and those who had a negative attitude toward statistics in the other dimension. Students' pre-test scores were the covariate variable and the achievement test scores (midterm examination) were used as the dependent variable.

Additional analyses of the student evaluation of the artificial intelligent statistics system (Appendix C) were
conducted using descriptive statistics. Student were asked to completed a six-item five-point Likert scale of student evaluation at the end of tutorial session, the results were described by frequency and percentage.
CHAPTER BIBLIOGRAPHY


CHAPTER 2

REVIEW OF RELATED LITERATURE

Computer-Assisted Instruction

The term "computer-assisted instruction" or "computer-aided instruction" (CAI) may refer to all the types of use that relate to instruction, such as drill and practice, tutorial programs, simulations, and educational games or it may simply refer to drill and practice or tutorial uses (Cornett, 1984, p. 2). Computer-assisted instruction (CAI) has been studied for more than two decades. Frequently these studies show that the learning generated by CAI is indistinguishable from that generated by other teaching techniques. However, many of these studies also report that CAI seems to accomplish these learning goals in significantly less time. The combination of these results—equal learning in less time—may be called the CAI phenomenon; it is an observation about the efficiency of instruction (Bright, 1983).

The documentation of the CAI phenomenon is taken from a variety of reviews and syntheses of the CAI literature. Molnar (1973, p. 68) stated that, according to a 1970 paper by Bunderson, "CAI permits students to achieve educational objectives in much less time and that saving in time of up
to 40% is not uncommon." Cody (1973, p. 25) agreed that "most of the studies indicate that material can be taught by computers in substantially less time than conventional techniques, with no loss in achievement."

Fletcher and Atkinson (1972) found that students who received supplementary CAI instruction in reading scored an average of .6 grade levels higher on a standardized test at the end of the year than students who received only traditional classroom instruction (Shively, 1984). At the college level, Kulik and Cohen (1980) reviewed eight CAI studies that included time as a measured variable. All eight reported that CAI students used significantly less time; on the average, CAI students used 2.25 hours per week and conventional students used 3.5 hours per week. The conclusion was that "there appears to be little doubt that students can be taught with computers in less time than with conventional methods of college teaching" (Bright, 1983).

The quality of educational computing in a school is going to depend on the quality of the software selected for use in that school and on the way in which the use of that software is integrated into the overall curriculum. Educational computing and education itself are currently at a critical juncture, and their combined fate depends to a large extent on the quality of educational software and on the quality of the uses to which it is put, both in school and at home (Komoski, 1984). The quality of software has
remained consistently low, while the sheer numbers of machines in the schools and the power of each machine have increased dramatically. The quality of educational software has simply not kept pace (Bork, 1984, p. 240).

According to Kenneth Komoski, executive director of the Educational Products Information Exchange (EPIE), many products are not meeting the needs of educators. Evaluations by the EPIE staff of software produced for schools have led to several observations: large scale software packages have not been developed for high schools; software seems to be designed predominantly for mathematics classes in the drill and practice mode; little emphasis is given to any higher order skills, such as analysis and synthesis of materials; and most programs provide little diagnostic help (Komoski, 1984).

Use of the Computer in the Teaching of Statistics

The use of computers has taken two forms in the teaching of elementary statistics:

1. Integrating the content of statistics in courses on computers. The introductory textbooks' attempt to integrate statistics with computers appears to take two forms. The objective of the first form is to learn a computer language, where statistics is used as a content area presenting program problems. The second form that many introductory statistics textbooks have attempted, focuses on learning statistics using the computer as an aid.
2. Using computers as a method of statistics instruction. The following are the three general methods of instruction:

"Hands-on." "Hands-on" methods of instruction of theoretical material are called Computer-Assisted Instruction (CAI).

Demonstrational. The instructor can do many things with the computer to provide useful information for the statistics classroom or laboratory.

Simulation. The computer can be used to eliminate the tedium of computation; it can also simulate the collection, input, storage, and manipulation of data. For example, STEXSIM, developed by W. Thomas (1972) generates data in FORTRAN for factorial designs having up to seven random or fixed main effects. Interactive vs. Non-interactive Statistics Packages, large statistics packages such as SAS (Barr and Goodnight, 1972)--Statistical Analysis System. SPSS (Nie, Bent, and Hull, 1970)--Statistical Package for the Social Science. Koh's (1975) TUSTAT-II (Tutorial system for statistics with time-sharing computers) and STRAP-II (Statistics system for research and production with time-sharing computers) provide 140 statistics routines in BASIC.

According to Goodman (1986), the microcomputer can be of great use in statistics because it can (a) operate as a powerful computational tool, (b) reinforce specific concepts and relationships, and (c) provide different settings in
which students can apply statistical concepts and techniques. The microcomputer can be used to help students learn fundamental concepts and techniques in statistics, as well as offering different opportunities for students to apply concepts and skills. Goodman illustrates three ways of using the computer to teach basic statistics concepts and techniques: (a) use of games, (b) use of simulations, and (c) use of computer problems and projects.

Computer-Assisted Instruction of Statistics

In 1969, Wassertheil successfully incorporated CAI into the laboratory portion of an introductory statistics course. In her study, the main purpose of CAI was to individualize instruction. Each student progressed at his own pace. The computer was programmed to interact with a student at a remote teletype. If a student's response to a programmed computer question was correct, the computer would go to the next question, present new material, skip simple material, or go to more complex matters based on the previous rate of progress of the student. On the other hand, if the student's response was incorrect, the computer would either "present tutorial material, review concepts, give hints, refer to text or tell the student to go home and study some more." Throughout the process, the computer recorded the performance of each student.

IBM'S T.J. Watson Research Center in Yorktown Heights, New York, developed The CAI Problem Laboratory in Statistics
and STATS Workbook of Problems. Thirteen of twenty-six students in Wassertheil's "Introductory Statistics" course volunteered to schedule one hour of terminal time in lieu of the usual 75-minute weekly laboratory. Wassertheil went over the homework problems and answered questions of the control group in the laboratory sessions. Both groups turned in weekly assignments from the workbook. From the grade point and examination score means and standard deviation of both groups, there were no statistically significant differences between the groups on any of the measures. The computer groups did somewhat better on the second examination and the final examination. The average amount of terminal time used to cover eight chapters was 10.1 hours with standard deviation of 3.9 hours. The most positive result of Wassertheil's study was that one 75-minute class period per week could be eliminated without deterioration of student performance. The benefit of CAI would be the freeing of the instructor for individual student contact or other duties (Tubb, 1977, p. 12).

Forsythe and Bleich (1973) conducted a similar study. Twenty-five students used computer terminals to interact with a CAI program that (a) helped them learn routine material, such as definitions and basic processes; (b) enabled each student to see if he had mastered the important content of the text by presenting short answer questions and responding to his replies with specific references to the
text; and (c) removed much of the drudgery of the statistics laboratory by doing the calculations once the student had demonstrated that he could do them. Forsythe and Bleich did not report any comparative evaluation other than that those who used the terminals were "overwhelmingly positive." The authors concluded that "the computer can be used beneficially to help the teaching of a 'service' statistics course."

Tsai and Pohl (1980) studied differences in student learning achievement and retention in a college-level statistics course taught in a variety of teaching/learning environments. The environments included in the study were (a) lecture/discussion; (b) lecture/discussion supplemented with planned teacher/student contacts; (c) programmed instructional texts; (d) programmed instructional texts supplemented with planned teacher/student contacts; and, (e) straight CAI. Although students experiencing a CAI environment performed no differently on achievement or retention tests than students experiencing a traditional lecture/discussion environment, students experiencing an "enriched" CAI environment (CAI plus teacher/student contacts) performed significantly better on achievement tests than students experiencing any of several other environments. One conclusion that can be drawn is that CAI may tend to be non-personal and some students may need a human touch to their instruction. Another related
conclusion is that the benefits of straight CAI are overshadowed by the fact that CAI systems lack the ability to respond to the misconceptions of the students, thereby being an inferior teacher compared to a human.

Intelligent Computer-Assisted Instruction

Artificial Intelligence is defined as the attempt to design computers to perform tasks that require human intelligence. AI has been applied in the field of education as well, with the result often referred to as intelligent tutoring systems or intelligent computer-assisted instruction (ICAI) (Roberts and Park, 1983, p. 7).

ICAI programs offer what Brown (1977) calls a reactive learning environment, in which the student is actively engaged with the instructional system, and his interests and misunderstandings drive the tutorial dialogue. This goal was expressed by other researchers trying to write CAI programs that extended the medium beyond the limits of frame selection (Koffman and Blount, 1975):

Often it is not sufficient to tell a student he is wrong and indicate the correct solution method. An intelligent CAI system should be able to make hypotheses based on a student's error history as to where the real source of his difficulty lies.

(Barr and Feigenbaum, 1982, p. 227)

AI developed from its roots in 1956 with heuristic
search as its defining characteristic. Problem solving was viewed as the search for solutions in a maze of possibilities, where the solution could not be found by blindly searching all possibilities. During the 1970s, attention shifted to the problem of how to represent the conceptual complexity inherent in real-world problem-solving situations. In general, there are now three main branches to artificial intelligence: (a) cognitive science, (b) intelligent software, and (c) autonomous machines. One major paradigm for intelligence is called the cognitive model. Within the branch concerned with intelligent software, Hayes-Roth & Thorndyke (1985), identify four major paradigms for intelligence: (a) the expert advisor, (b) the knowledge system, (c) the intelligent interface, and (d) the surrogate instruction. In the third branch, the single paradigm is called the autonomous system.

The Component of ICAI Systems

The ICAI system consists of four modules: (1) expert module; (2) student module; (3) tutorial module; and (4) communication module.

**Expert Module.** This module contains the domain of knowledge the system is trying to impart to the student. This knowledge includes the facts to be taught and procedural knowledge, which is the set of procedures used by experts in solving a particular problem.
**Student Module.** This module represents the student's understanding of the material being taught. When an ICAI system is spoken of as "knowing the student," it is due to developments in this module that the "knowing" is possible.

**Tutoring Strategy Module.** Also called the teaching module, this portion of the ICAI system is responsible for changing the student into an expert in the most efficient and effective manner possible.

**Communication Module.** The communication module is the vital link between human and some type of data storage, in order to understand student input, be it a question or an answer. Ideally, the communications module should be able to understand the utterance of the student.

Winograd (1977) notes that ICAI system need a communications module capable of complex interaction between the expert, student, and tutoring strategy modules. This communication must include the ability to understand the student's responses and respond to the student in a way that the student will understand what is said.

Knowledge Representation in Expert System

Waterman (1986) described three representation used in the expert system are rules, semantic nets, and frames.

**Rules** provide a formal way of representing recommendations, directives, or strategies; they are often appropriate when the domain knowledge results from empirical associations developed through years of experience solving
problems in an area. Rules are expressed as IF-THEN statements.

Semantic net is used to describe a knowledge representation method based on a network structure. Semantic nets were originally developed for use as psychological models of human memory but are now a standard representation method for AI and expert system (Brachman, 1979). A semantic net consists of points called nodes connected by links called arcs describing the relations between nodes.

Frame refers to special way of representing common concepts and situations. A frame is organized much like a semantic net. A frame is a network of nodes and relations organized in a hierarchy, where the topmost nodes represent general concepts and the lower nodes more specific instances of those concept.

Knowledge Engineering Language

A knowledge engineering language is a sophisticated tool for developing expert system, consisting of an expert system building language integrated into an extensive support environment. Knowledge engineering languages can be categorized as either: skeletal systems or general-purpose systems.

A skeletal knowledge engineering language is simply a stripped-down expert system -- that is, an expert system with its domain-specific knowledge removed, leaving only the
inference engine and support facilities.

A general-purpose knowledge engineering language can handle many different problem areas and types. It provides more control over data access and search than does a skeletal system but may be more difficult to use.

Examples of Skeletal Knowledge Engineering Languages

EMYCIN - the skeletal system EMYCIN derived from the MYCIN system. The MYCIN system was developed for diagnosing and treating bacterial infections at Stanford University. The implementation language used was INTERLISP. EMYCIN uses a rule-based knowledge representation scheme with a rigid backward chaining control mechanism that limits its application to diagnosis and classification-type problems. EMYCIN has been used to build diagnosis-type expert systems in medicine, geology, engineering, agriculture, and other areas.

EXPERT - this skeletal knowledge engineering language uses a rule-based knowledge representation scheme and has a limited forward chaining control. EXPERT has built-in explanation, knowledge acquisition, and consistency checking facilities to speed system development. This system shell developed at Rutgers University, and the implementation language used was FORTRAN. EXPERT has been used to build diagnosis programs in medicine, geology, and other areas.
Conclusions

The studies found that teaching by computer-assisted instruction (CAI) was indistinguishable from the results generated by other teaching techniques. Students can be taught by computer in substantially less time than conventional techniques with no loss in achievement. Generally, the conclusion was that teaching by computer depends on the quality of the software. Most commercial software did not emphasize the higher order skills such as analysis and synthesis of materials and provided little diagnostic help. In this research, the study incorporated the content of statistics into the ARPHY system shell and investigated to the effectiveness of the software as a supplement in enhancing learning. The next chapter will discuss the structure of ARPHY, the incorporation of the content of basic statistics into the teaching module of ARPHY, and procedures for data collection and analysis.
CHAPTER BIBLIOGRAPHY


CHAPTER 3

PROCEDURES OF THE STUDY

The first portion of this chapter is a discussion of the structure of the Artificially Intelligent Physics Tutor (ARPHY) and the incorporation of statistics content into the ARPHY teaching module. The second portion is a discussion of the procedures for data collection and analysis used to investigate the effectiveness of the Artificially Intelligent Statistics Tutor (ARSTAT) as a supplement to the learning of college level statistics.

Incorporation of the Content of Statistics into the ARPHY Teaching Module

The Artificially Intelligent Physics Tutor (ARPHY), which was used as a system shell in this study, was developed by Stephen F. Brown in 1985. The ARPHY was constructed in four sections corresponding to the modularity of ICAI systems. The four sections were expert, student, tutorial strategy, and communications modules. The ARPHY was used as a system shell by removing its domain-physics knowledge in the expert module and replacing it with domain-statistics knowledge. The inference engine and support facilities of the other modules retained their identity.
The four sections of the ARPHY and the incorporation of the domain-statistics knowledge into the expert module are discussed below.

**Expert Module**

The major theory which was used as the basis of the teaching strategy module was Ausubel's advance organizer. The organizer was constructed in a depth-first manner with each of the nodes being one of the Gagne learning types. The learning types fell naturally into the advance organizer net-archy. The pedagogical expert module was constructed by recursively asking the question "What prerequisite learning is necessary to learn this node?"

The knowledge domain is contained in this expert module. Once the physics knowledge domain was removed, the inference engine and support facilities of the ARPHY system retained their identity. To incorporate the content of statistics into the ARPHY tutor system, the domain-knowledge of statistics was placed in the expert module and programmed into the networked hierarchy. The domain-knowledge of statistics was presented in four problems.

- Problem type 1: Measurement Scales;
- Problem type 2: Frequency Distribution and Graphs;
- Problem type 3: Measures of Central Tendency;
- Problem type 4: Measures of Variation.

The four problems were built into the networked
hierarchy structure in the expert module as presented in Figure 1.

Figure 1. Networked hierarchy of the four problems of statistics building in the expert module.
The ARSTAT was constructed with 25 nodes, containing 4 problems, 6 principles, and 15 concepts. Based on the theory of the expert module of the ARPHY system, the net-archy was constructed in the depth-first manner and with prerequisites necessary for full understanding of the problem. The following figures show the depth-first path of each problem.

Problem Type 1: Measurement scales. The first problem is the least complex. There are several concepts required for full understanding and each of these concepts in turn will be a prerequisite for the other problems as shown in Figure 2. To understand the measurement scales requires the student to learn the level of measurement concept. To understand the level of measurement requires the student to learn the variables concept. To understand the variables concept requires the student to learn the concept of statistics, which in turn requires the concept of measurement.
Problem Type 2: Frequency Distribution and Graphs.

Figure 3 shows how the basic content of frequency distribution and graphs was incorporated in the learning strategy model of ARPHY. Also, in this problem there are several concepts required for full understanding, one of which is types of graph. The types of graph, in turn, requires the concept of frequency distribution. Frequency distribution requires the concept of class interval. And class interval requires the concept of level of measurement.
Problem Type 3: Measures of Central Tendency. To fully understand this problem, one of the principles to be learned is the comparison between mean, median, and mode. Due to the depth-first manner of construction, only one principle is initially placed in the net-archy. The comparison between mean, median, and mode, in turn, requires the concept of mean, which requires the concept of summation. Summation requires the concept of frequency distribution. The net-archy for Problem 3 is shown in Figure 4.
Problem Type 4: Measures of Variation. In this problem, there are several principles involved, one of which is mean deviation. The mean deviation, in turn, requires the concepts of mean, which requires the concept of summation. Summation requires the concept of frequency distribution, as shows in Figure 5.
Figure 5. Initial depth-first pass in recursion for building the net-archy for Problem 4.

There is a limited amount of networking within the learning net-archy for one problem; however, the networking becomes increasingly complex and networked upon the addition of other problems. The new problems usually call for some of the same principles and concepts as previous problems. As each problem type generates its own learning net-archy, the number of new prerequisites decreases due to the more complex network which is formed.

All the prerequisite information for each problem is
represented in the form of parent-child relationships. Each problem has its associated children, or first-generation prerequisites. These children, in turn, have their children, second-generation prerequisites. The four problems of statistics are represented according to these relationships. Table 1 shows the first generation of four families, one family for each problem.

Table 1
First Generation Prerequisites of the Four Statistics Problems

<table>
<thead>
<tr>
<th>Parent</th>
<th>Children (prerequisites)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem 1 *</td>
<td>level of measurement (concept)</td>
</tr>
<tr>
<td>Problem 2 *</td>
<td>types of graph (concept)</td>
</tr>
<tr>
<td></td>
<td>shape of distribution (concept)</td>
</tr>
<tr>
<td>Problem 3 *</td>
<td>comparison between mean, median, and mode (principle)</td>
</tr>
<tr>
<td></td>
<td>deviation of the mean (principle)</td>
</tr>
<tr>
<td>Problem 4 *</td>
<td>standard scores (principle)</td>
</tr>
<tr>
<td></td>
<td>quartile deviation (principle)</td>
</tr>
<tr>
<td></td>
<td>mean deviation (principle)</td>
</tr>
</tbody>
</table>

Prior learning of many concepts is required for the solution of problems. The concepts are prerequisites for the principles, and will, therefore, appear in later generations of the family tree.
The second generation of the family tree is shown in Table 2.

Table 2

Second Generation Prerequisites of the Four Statistics Problems

<table>
<thead>
<tr>
<th>Parent</th>
<th>Children (prerequisites)</th>
</tr>
</thead>
<tbody>
<tr>
<td>level of measurement *</td>
<td>variables (concept)</td>
</tr>
<tr>
<td>shape of distribution *</td>
<td>frequency distribution (concept)</td>
</tr>
<tr>
<td>types of graph *</td>
<td>frequency distribution (concept)</td>
</tr>
</tbody>
</table>

Comparison between mean, median and mode:
- mean (concept)
- median (concept)
- mode (concept)

Deviation of the mean * | mean (concept)

Standard score (z-score, t-score) * | Variance & standard deviation (principle)

Quartile deviation * | range (concept)

Mean deviation * | deviation of the mean (principle)

The third generation of the family tree for the statistics problems contains the prerequisites for second generation children and is shown in Table 3.
Table 3

Third Generation Prerequisites of the Four Statistics Problems

<table>
<thead>
<tr>
<th>Parent</th>
<th>Children (prerequisite)</th>
</tr>
</thead>
<tbody>
<tr>
<td>variables *</td>
<td>population (concept)</td>
</tr>
<tr>
<td>*</td>
<td>statistics (concept)</td>
</tr>
<tr>
<td>frequency *</td>
<td>class interval (concept)</td>
</tr>
<tr>
<td>dis.</td>
<td></td>
</tr>
<tr>
<td>Mean *</td>
<td>summation (concept)</td>
</tr>
<tr>
<td>Median *</td>
<td>summation (concept)</td>
</tr>
<tr>
<td>Mode *</td>
<td>summation (concept)</td>
</tr>
<tr>
<td>variance &amp; standard</td>
<td>deviation of the mean (principle)</td>
</tr>
<tr>
<td>deviation *</td>
<td>Effect of adding and multiply constant to variate value (concept)</td>
</tr>
</tbody>
</table>

The remaining form for the Fourth generation and are shown with their children in Table 4.
Table 4

Fourth Generation Prerequisites of the Four Statistics Problems

<table>
<thead>
<tr>
<th>Parent</th>
<th>Children (prerequisite)</th>
</tr>
</thead>
<tbody>
<tr>
<td>population *</td>
<td>measurement concept (concept)</td>
</tr>
<tr>
<td>statistics *</td>
<td>measurement concept (concept)</td>
</tr>
<tr>
<td>class interval *</td>
<td>level of measurement (concept)</td>
</tr>
<tr>
<td>summation *</td>
<td>frequency distribution (concept)</td>
</tr>
</tbody>
</table>

The learning types are represented as a semantic net within the expert model. In a semantic net, information is represented as a set of nodes connected to each other by a set of labeled arcs, which represent relationships among the nodes (Tanimoto, 1987).

Assessing Student Competency

The expert module assessed each student's competency and also determined how well the student had learned the information transmitted. When the student had not learned a topic well, the expert module (EM) provided the tutorial strategy module with information about which misconceptions the student might have.

Assessing the level of learning. There were two general classifications of questions which ARSTAT posed to
the student. The first classification, called BLOOM-1, included questions which tested behaviors falling in the 1.00-2.30 range. This included the knowledge domains of terminology, facts, classifications, categories, methodology, principles, and structures. The comprehension domain included translation, interpretation, and extrapolation. The second classification called, BLOOM-2, included questions which tested behaviors in the 3.00-6.2 range. The domains included were application, analysis, synthesis, and evaluation. The most common BLOOM-2 questions were those at the application and analysis levels.

**Misconception Detection.** When a student answered a BLOOM-2 question incorrectly, ARSTAT passed control to the SAY-WRONG function located within the communications module. SAY-WRONG communicated the incorrectness of the response to the student and then passed control over to the misconception-detector function, MISCONDET, located within the expert module.

**Student Module**

The student module kept all the records specific to each student which were accessed by the other three modules. Information kept by the student module included the student's reading speed, topics completed, probable areas of misconceptions, and the path instruction had taken through the topics.
Tutorial Strategy Module

The pedagogical decisions are made by the tutorial strategy module (TSM). The TSM consists of a dual expert system in a nested configuration. The top-level expert system is the controller of the actions taken by ARSTAT.

```
+----------------------------------------+  
| +-----START LOOP THROUGH POSSIBLY-ACTIVATED PROCEDURES |  
| |  
| | | OPEN-SESSION  
| | | | CHOICE-A-TOPIC  
| | | | INTRODUCE TOPIC  
| | | | PRESENT-TOPIC-BACKGROUND  
| | | | EXPLAIN THEORY  
| | | | RELATE TOPIC TO OTHER LEARNING  
| | | | EXPLORE-COMPETENCY-BLOOM-1  
| | | | EXPLORE-COMPETENCY-BLOOM-2  
| | | | WRAP-UP  
| | | | CONTINUE-TUTORIAL  
| | | | CLOSE-UP-TUTORIAL  
| | | |  
| | | FORWARDS-CHAIN THROUGH EXPERT SYSTEM RULES  
| | +<-RE-ITERATE LOOP WITH NEW ASSERTIONS  
+----------------------------------------+
```

Figure 6, High-level structure of the tutorial strategy module (TSM). Procedures are activated ONLY by an appropriate assertion. Usually, only ONE procedure is activated during each pass through the loop.

OPEN-SESSION

In this procedure an introductory message is presented to the student who is then prompted for his name. If the student has not been met, the procedure (OPEN-NEW-STUDENT) will be called. This procedure first transfers control to the student module to assess the student's reading speed,
then gives an introduction to the structure and operation of
the tutorial system.

CHOOSE-A-TOPIC

This procedure is the most complex within the tutorial
system. The (CHOOSE-A-TOPIC) procedure is dynamic in its
approach to topic selection. It can also be passive and
allow the student to have total control over the
instructional process, but if the student exhibits poor
performance due to diagnosed misconceptions, (CHOOSE-A-
TOPIC) can take control and steer the student through a
period of remedial instruction. The student also has the
option of allowing the tutorial system to select the next
topic.

Presentation

The modularity of the presentation sequence is to
facilitate the implementation of future expert system rules
which may infer that the student may not need an
introduction or general overview of the topic.

Competency Determination

The control of the EXPLORE-COMPETENCY procedure rests
within the expert module. During this procedure there is a
constant communication between the expert module and the
TSM. During competency assessment, the two dynamic
thresholds: num-BLOOM-1 questions and num-BLOOM-2 questions
determined to ask questions at BLOOM-1 and BLOOM-2 level.
These thresholds are two of the eight dynamic pedagogical parameters. The expert system transfers control back to the TSM while adding the assertion: (COMPETENCY EXPLORE BLOOM-2). The TSM then invokes the WRAP-UP procedure.

WRAP-UP

The WRAP-UP procedure is essentially a final set of "house-cleaning" procedure designed to take care of all necessary tasks relating to the previous topic prior to a possible closing of the tutorial session which may subsequently be invoked. There are three routines in this procedure:

1. **MISCONDET Transfer.** All misconceptions are removed from the blackboard and placed in the student module.

2. **Threshold Transfer.** The modification of the eight pedagogical parameters' threshold values is based upon the actions of a nested expert system whose only function is to decide what modifications to the parameters need to be made. The six parameters governing the CHOOSE-TOPIC procedure and the two parameters governing the number of questions to ask during the EXPLORE-COMPETENCY procedure are only modified when the topic is chosen by the tutoring system.

3. **Topic completion.** A decision is made by the TSM whether the student has completed the topic satisfactorily. If completion is allowed, the topic name is added to the list stored as *TOPICS-COMPLETED* and transferred to the student module.
Quitting or Continuing

After wrapping-up the topic, the student is asked if he or she wants to continue the session. The student may choose to proceed with another topic, take a break, or terminate the session. If the student wants to terminate the session, the CLOSE-UP procedure will be invoked. The CLOSE-UP procedure will present the student with a report detailing the information contained in the current student module. The information will include a list of the problems completed and the prerequisite topics which the student has learned.

Communications Module

There are two basic types of communication: direct and indirect.

**Direct communication** is used when the message is intended only for one module, when one module needs to communicate with another. For example, if the TSM needs the expert module to perform a task and return a value, the appropriate function within the expert module is called either by the TSM or by a high-level function within the expert module. Then one or more values will be passed back to the original calling procedure within the TSM. The other form of direct communication is to translate between the student and the tutorial system.

**Indirect communication** is used when the message is either of general interest to more than one module or if it
may need to be accessed at some later time. There are two types of indirect communication links. The first link is the blackboard. Messages will be placed on the blackboard by one module and be read by more than one other module. The other method of indirect communication is the assertion list used by the main and nested (threshold inspector) expert systems.

**Completion of Tutorial**

It is possible that some students may not be able to learn the concepts of statistics regardless of the delivery format. While the student solves each of the problems, the cognitive strategies and blatant hints is given; if the student still is not able to solve the problem, the instruction is terminated. Otherwise, instruction continues until the appropriate material is covered and understood.

**Equipment**

The equipment used for this research was an IBM PC-XT with 640 kilobytes of available internal RAM. Approximately 10 megabytes of storage were available on the fixed disk. The system was equipped with the color-graphics adapter and monitor. The software used was Golden Common LISP (GCLISP), a version of common LISP designed for the IBM PC by Gold Hill Computer Company.
Procedures for Data Collection and Analysis

Selection of the Subjects

Two of the four sections of EDUC 5210 (educational statistics for graduate students) being offered in the fall 1988 semester by the Department of Educational Foundations, College of Education, University of North Texas, were used in this study. These sections were used because they were offered on campus. One class was taught by Jon I. Young and the other class was taught by L. Fred Thomas. Thirty-seven students from Dr. Young's class were used as the experimental group. This class met two mornings a week, and each meeting took 1 hour and 20 minutes. Because using the computer laboratory was too inconvenient for students in the evening class, 33 students from Dr. Thomas' class were used as the control group. This class met three hours once a week in the evening. Although the instructor for each group was different, the two men had similar credentials. Both were full professors who were experienced in teaching statistics and had taught this course for several years. Also, they co-authored the textbook used in both classes, *An Introduction to Educational Statistics: The Essential Elements*.

Description of Data Collection

A statistics attitude test, which included a seven-point scale of ten pairs of questions, and a pre-test over
the content of the first four chapters of the textbook were administered to both classes during the first meeting of the class. The pre-test examination consisted of forty multiple-choice questions, selected from textbook exercises and sample questions to test each concept and knowledge of specified content of course material. The Artificial Intelligent Statistics Tutor (ARSTAT) was validated during the 1988 spring and summer semesters with regard to the statistical content, graphics, and the ARPHY system. The ARSTAT system was verified with the advice of Stephen F. Brown. The statistics content was verified with the advice of William K. Brookshire, Associate Professor in the Department of Educational Foundations, College of Education, University of North Texas. With the permission of the faculty and the cooperation of the lab assistants in the Department of Computer Education and Cognitive Systems, four IBM PC-XT with 640 kilobytes of available internal RAM and the hard disk in the Educational Media Center were used for the experiment. The ARSTAT were available during the fourth and fifth week of the class, before the midterm examination, which was scheduled in the seventh week. In the third week, the students who had agreed to participate in the tutorial session were asked to schedule times to work with ARSTAT during the fourth and fifth weeks. The appointed times were reserved with the lab assistants. A two-hour period was scheduled for each participant, but the participant could
work for longer or shorter periods. The participants were allowed to participate twice or more. If they did not finish or missed an appointed time, make up time was allowed in the sixth week.

In the tutorial session, each participant was shown how to access the tutorial program and given brief instructions about the content, the strategies, the goal, and the completion of the tutorial. After the tutorial session, each participant was asked to complete a six-item five-point Likert scale of student evaluation (Appendix C).

The midterm examinations were administered in the seventh week by the instructors. Both classes took the same examination. The midterm examination consisted of sixty multiple-choice questions, written by L. Fred Thomas, Professor of the Department of Educational Foundations, College of Education, University of North Texas. Only the forty items in the examination which corresponded to the first four chapters of the class lecture were used for analysis. The internal reliability coefficient of the examination was $r = .78$, found in the 1987 fall semester with 56 students. The reliability coefficient was found using Kruder-Richardson formula 20 by L. Fred Thomas.

**Variables**

**The Dependent Variable**

The dependent variable used in this study was the total
number of correct answers on the forty multiple choice questions from the midterm examination which students took during the seventh week of the class.

The Independent Variables

One of the independent variables used in this study was the study technique: by lecture/discussion only and by lecture/discussion plus ARSTAT. Another independent variable was sex: students were identified as male or female. The third independent variable was the students' attitude toward statistics, which was measured by the Attitude Toward Statistics Questionnaire, a seven-point scale of ten pairs of questions. It was developed by the Department of Educational Foundations, College of Education, University of North Texas. The seven-point scale was ranked with a score of 1 for the most negative attitude and a score of 7 for the most positive attitude. Each student's total score was found by adding the ten answers. The highest possible score attitude was 70 and the lowest possible was 10. To differentiate the students' attitude toward statistics, a median was used to divide the students' scores into two equal groups. The students who scored higher than the median were identified as having a positive attitude toward statistics, and those who scored below the median were identified as having a negative attitude toward statistics.
The Covariate Variable

The total number of correct answers on the forty-item multiple choice pretest, which was given to the students during the first class meeting, was used as the covariate variable.

Statistical Procedures Used in Data Analysis

All students who were in the evening class, filled out the Attitude Toward Statistics Questionnaire, answered the pretest questions, and took the midterm examination, were included in the control group. And all those who were in the day class, filled out the Attitude Toward Statistics Questionnaire, answered the pretest questions, participated in the tutorial session by using ARSTAT and took the midterm examination, were included in the experimental group.

To test the first hypothesis, a one-way analysis of covariance was used to compare the two groups (learning by lecture/discussion only and lecture/discussion plus ARSTAT) with respect to their adjusted midterm test scores on the forty items. In the second hypothesis, a two-by-two analysis of covariance was utilized to compare the students' adjusted test scores with the study technique and sex were the independent variables. And to test the third hypothesis, a two-by-two analysis of covariance also used to compare students' adjusted test scores with the study technique and the students' attitude toward statistics were the independent variables.
An analysis of covariance was used to statistically reduce the effects of initial group difference. This procedure made compensation adjustments for the final means on the dependent variable.

Descriptive statistics were conducted to describe the student evaluation of the artificial intelligent statistics system (Appendix C). The statistical data was analyzed using a National Advanced System 8043 computer at the University of North Texas and the SPSSX statistical software package (SPSS, 1983).
CHAPTER BIBLIOGRAPHY


CHAPTER 4

RESULTS OF DATA ANALYSIS

Description of the Population

Seventy students enrolled in the two EDUC 5210 classes in fall 1988. Thirty-three were enrolled in the evening class and 37 were enrolled in the day-time class. At the first meeting of the two classes, 63 students participated, 27 in the evening class and 36 in the day-time class. Of the 63 students, 39 were females and 24 were males, ranging in age from 20 to over 50. All were classified as graduate students. There were 33 master's students, 29 doctoral students, and 1 non-degree program student, who enrolled out of personal interest.

Of the original 27 students in the evening class, which was used as a control group, 26 completed the Attitude Toward Statistics Questionnaire and the pretest questions. This population was reduced to 24 students when one student dropped the course and one did not take the midterm examination. These 24 students consisted of 12 males and 12 females, with 8 master's students and 16 doctoral students.

Of these 24 students, the mean on the pretest was 18.88 and the mean of the postest was 33.54. The minimum score for the Attitude Toward Statistics was 21 and the maximum score was 70.
At the first meeting of the experimental group 36 students were present in the class. Of the 36 students, 35 completed the Attitude Toward Statistics Questionnaire and the pretest questions. In the second meeting, 33 students signed the agreement to participate in the tutorial session. 31 students were scheduled to work with the ARSTAT. Of these 31 students, four did not participate. The population of the experimental group was reduced to 27. Of these 27 students, one student did not complete the Attitude Toward Statistics Questionnaire and the pretest questions and one student did not take the midterm examination. The total population for the experimental group was 25 students. This group contained 8 males and 17 females, with 15 master's students and 10 doctoral students. The mean score of the pretest for this group was 17.74 and the mean score of the posttest score was 33.92. The minimum score for the Attitude Toward Statistics was 23 and the maximum score was 64.

Of these 25 students, 19 students completed the tutorial session and successfully solved the four problems in the tutorial program, 4 students completed the tutorial but did not successfully solve all four problems, and 2 students completed part of the tutorial which took at least two hours. Two of these 25 students completed ARSTAT within one visit of four hours. Twelve students spent two periods of two hours, ten successfully completed the ARSTAT, and two
completed the tutorial but did not successfully solve all four problems. Seven students spent three visits, 6 students successfully completed ARSTAT in approximately five to five and a half hours, and one completed without successfully solve all four problems. Two students spent four visits, one successfully completed the ARSTAT and the other did not successfully solve all four problems. Two students who completed part of ARSTAT took about two and a half hours.

The remaining 8 students in this class completed the Attitude Toward Statistics Questionnaire and the pretest questions and took the midterm examination, but did not participate in the tutorial session. This group consisted of 3 males and 5 females. Seven were master's student and one was taking the class out of personal interest.

Results of the Hypotheses

According to the first hypothesis, the students who learned by lecture/discussion plus ARSTAT would have a significantly higher average score on the achievement test than the students who learned by lecture/discussion only. Results of the one-way analysis of covariance indicate that no significant difference existed between the adjusted mean scores on the achievement tests of the control group and the experimental group, as shown by the F value of .58 and the P value of .449 in Table 5. The adjusted mean score for the control group was 33.32, and the experimental group's
adjusted mean score was 34.14, as presented in Table 6.

Table 5

Summary of Analysis of Covariance:

Achievement Test Scores by Study Technique

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>DF</th>
<th>Mean Square</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariate</td>
<td>148.15</td>
<td>1</td>
<td>148.15</td>
<td>10.53</td>
<td>.002</td>
</tr>
<tr>
<td>Between Groups</td>
<td>8.20</td>
<td>1</td>
<td>8.20</td>
<td>.58</td>
<td>.449</td>
</tr>
<tr>
<td>Within Groups</td>
<td>647.20</td>
<td>46</td>
<td>14.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>803.55</td>
<td>48</td>
<td>16.74</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6

Mean Scores and Adjusted Mean Scores on Achievement test

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Pretest</th>
<th>Posttest</th>
<th>Adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Lecture</td>
<td>24</td>
<td>18.88</td>
<td>5.67</td>
<td>33.54</td>
</tr>
<tr>
<td>Lecture+AI</td>
<td>25</td>
<td>17.74</td>
<td>4.31</td>
<td>33.92</td>
</tr>
<tr>
<td>Total</td>
<td>49</td>
<td>18.25</td>
<td>5.01</td>
<td>33.74</td>
</tr>
</tbody>
</table>
The second hypothesis was tested using a two-by-two analysis of covariance with gender and study technique as the independent variables, the number of correct answers on the midterm examination as the dependent variable, and the number of correct answers on the pretest examination as the covariate variable. This test of the second hypothesis is summarized in Table 7. No statistically significant differences were found for differential main effects or interactions. A summary of the average scores and the adjusted average scores is presented in Table 8. The adjusted mean score was 33.08 for the male control group, 33.49 for the female control group, 33.33 for the male experimental group, and 33.46 for the female experimental group.

Table 7
Summary of Two-way Analysis of Covariance:
Achievement Test Scores By Study Technique and Sex

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>DF</th>
<th>Mean Square</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariate</td>
<td>148.15</td>
<td>1</td>
<td>148.15</td>
<td>10.19</td>
<td>.003</td>
</tr>
<tr>
<td>Study technique</td>
<td>5.52</td>
<td>1</td>
<td>5.52</td>
<td>0.38</td>
<td>.541</td>
</tr>
<tr>
<td>Sex</td>
<td>6.52</td>
<td>1</td>
<td>6.52</td>
<td>0.45</td>
<td>.506</td>
</tr>
<tr>
<td>Interaction</td>
<td>1.48</td>
<td>1</td>
<td>1.48</td>
<td>0.10</td>
<td>.751</td>
</tr>
<tr>
<td>Error</td>
<td>639.20</td>
<td>44</td>
<td>14.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>803.55</td>
<td>48</td>
<td>16.74</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 8
Mean Scores and Adjusted Mean Scores on Achievement test

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Pretest</th>
<th>Posttest</th>
<th>Adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Lecture Male</td>
<td>12</td>
<td>19.33</td>
<td>6.84</td>
<td>33.50</td>
</tr>
<tr>
<td>Lecture Female</td>
<td>12</td>
<td>18.42</td>
<td>4.46</td>
<td>33.58</td>
</tr>
<tr>
<td>Lecture+AI Male</td>
<td>8</td>
<td>16.88</td>
<td>2.42</td>
<td>32.88</td>
</tr>
<tr>
<td>Lecture+AI Female</td>
<td>17</td>
<td>18.00</td>
<td>4.99</td>
<td>34.41</td>
</tr>
<tr>
<td>Total</td>
<td>49</td>
<td>18.25</td>
<td>5.01</td>
<td>33.74</td>
</tr>
</tbody>
</table>

The third hypothesis states that students in the experimental group who had a positive attitude toward statistics would answer correctly a significantly greater number of test items than students in the control group who had a statistics-positive attitude. Results of the two-way analysis of covariance, as summarized in Table 9, indicate that a statistically significant differential main effect occurred with respect to attitude toward statistics. There was no significant difference in interaction. The adjusted mean score for students with a positive attitude in statistics who learned by lecture/discussion plus ARSTAT was 36.17. This score was significantly higher than the score for any of the other three groups; students with a
statistics-positive attitude who learned by lecture/discussion only (adjusted mean score = 34.70), students with a statistics-negative attitude who learned by lecture/discussion plus ARSTAT (adjusted mean score = 32.47), and students with a negative attitude in statistics who learned by lecture/discussion only (adjusted mean score = 31.84). A summary of the mean scores and the adjusted mean scores is presented in Table 10. The adjusted mean score was 34.96 for the positive attitude toward statistics group and was 32.22 for the statistics-negative attitude group. A Pearson product moment correlation was analyzed, a positive correlation between the attitude scores and the test scores was found with $r = 0.2805$, $p = 0.017$, $n = 57$.

Table 9

**Summary of Two-way Analysis of Covariance:**

**Achievement Test Scores by Study Technique and Attitude Toward Statistics**

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>DF</th>
<th>Mean Square</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariate</td>
<td>148.15</td>
<td>1</td>
<td>148.15</td>
<td>10.19</td>
<td>.003</td>
</tr>
<tr>
<td>Study technique</td>
<td>13.07</td>
<td>1</td>
<td>13.07</td>
<td>1.10</td>
<td>.301</td>
</tr>
<tr>
<td>Statistics Att</td>
<td>120.65</td>
<td>1</td>
<td>120.65</td>
<td>10.12</td>
<td>.003</td>
</tr>
<tr>
<td>Interaction</td>
<td>2.15</td>
<td>1</td>
<td>2.15</td>
<td>0.18</td>
<td>.673</td>
</tr>
<tr>
<td>Error</td>
<td>524.40</td>
<td>44</td>
<td>11.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>803.55</td>
<td>48</td>
<td>16.74</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 10

Mean Scores and Adjusted Mean Scores on Achievement test

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Pretest</th>
<th></th>
<th></th>
<th>Posttest</th>
<th></th>
<th></th>
<th>Adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Lecture Pos att</td>
<td>13</td>
<td>20.00</td>
<td>6.36</td>
<td>35.15</td>
<td>3.60</td>
<td>34.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lecture Neg att</td>
<td>11</td>
<td>17.55</td>
<td>4.66</td>
<td>31.64</td>
<td>3.80</td>
<td>31.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lecture+AI Pos att</td>
<td>11</td>
<td>19.36</td>
<td>4.72</td>
<td>36.45</td>
<td>1.97</td>
<td>36.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lecture+AI Neg att</td>
<td>14</td>
<td>16.29</td>
<td>3.56</td>
<td>31.93</td>
<td>4.50</td>
<td>32.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>49</td>
<td>18.25</td>
<td>5.01</td>
<td>33.74</td>
<td>4.09</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Additional Results

Twenty-seven students from the experimental group who participated in the ARSTAT tutorial session were asked to evaluate the tutorial program. They completed a 6-item 5-point Likert scale of student evaluation (Appendix C). Of the 27 students, 22 turned in the evaluation form. Students answered the evaluation form by indicating their agreement according to the 5-point Likert scale, ranking from 1 to 5 as strongly agree, agree, neutral, disagree, and strongly disagree. The first question of the evaluation form asked if the content of statistics in the tutoring system was relevant; of the 22 students, 16 (72.7%) strongly agreed, 5
(22.7%) agreed, and 1 (4.5%) disagreed. The second question asked if the content presented in the tutoring system was easy to understand; 8 (36.4%) students strongly agreed, 11 (50%) agreed, 2 (9%) were neutral, and 1 (4.5%) disagreed. The third question asked if the content presented in the tutoring system was well-organized; 9 (41%) students strongly agreed, 11 (50%) agreed, 1 (4.5%) was neutral, and 1 (4.5%) disagreed. The fourth question asked if it was easy to keep up with the content presented in the tutoring system; 9 (41%) students strongly agreed, 11 (50%) agreed, 1 (4.5%) was neutral, and 1 (4.5%) strongly disagreed. The fifth question asked if the tutoring system kept their attention; 9 (41%) students strongly agreed, 4 (18%) agreed, 5 (23%) were neutral, and 4 (18%) disagreed. The last question asked if the overall quality of the tutoring system was good; 6 (27%) students strongly agreed, 12 (55%) agreed, and 4 (18%) were neutral. A summary table for the student evaluation of the artificial intelligent tutoring system is presented in Appendix D.
CHAPTER 5

DISCUSSION, SUMMARY, RECOMMENDATIONS

This chapter presents a discussion of the hypotheses and the results of the ARSTAT evaluation, a summary of the study and conclusions derived from it, and recommendations for further study.

Discussion of the Hypotheses

The findings do not support the first hypothesis, that students who used the artificial intelligent statistics tutor would score significantly higher on the midterm test than students who learned by lecture/discussion only. The adjusted mean score of the students who learned by lecture/discussion plus ARSTAT was slightly greater than the adjusted mean score of the students who learned by lecture/discussion only, but the difference was not statistically significant.

These results are in accordance with the findings of Wassertheil (1969), who incorporated CAI into the laboratory portion of an introductory statistics course. There was no statistically significant difference between the experimental group and the control group on the three examinations, but the experimental group did somewhat better on the second and third examination. Tsai and Pohl (1980)
found that in a college-level statistics course, students experiencing a CAI environment performed no differently on achievement or retention tests than students experiencing a traditional lecture/discussion environment. However, they found that students who experienced an "enriched" CAI environment (CAI plus planned teacher/student contacts) performed significantly better on achievement tests than students who experienced any of several other environments, including lecture/discussion, lecture/discussion supplemented with planned teacher/student contacts, programmed instructional texts, programmed instructional texts supplemented with planned teacher/student contacts, and straight CAI.

The second hypothesis, that there would be an interaction between gender and study technique, was not found to be statistically significant on the achievement test scores. Also, there was no statistically significant difference on the main effect between male and female, or between students who learned by lecture/discussion only and students who learned by lecture/discussion plus ARSTAT, on the achievement test scores. Numerically, females tended to score higher than males, and females in the lecture/discussion plus ARSTAT group scored higher than females in the lecture/discussion group; also, males in the lecture/discussion plus ARSTAT group scored higher than males in the lecture/discussion group.
In the third hypothesis, a statistically significant difference in the achievement test scores was found in the main effect with regard to students' attitude toward statistics, but a significant difference was not found in the interaction between study technique and the students' attitude toward statistics. Students who had a positive attitude toward statistics had a higher score on the test than students who had a statistics-negative attitude. A statistically significant difference was not found in the interaction, but numerically the adjusted mean score of the students who had a positive attitude toward statistics in the lecture/discussion plus ARSTAT group was slightly greater than students who had a statistics-positive attitude in the lecture/discussion group. Likewise, the adjusted mean score of the statistics-negative attitude students of the lecture/discussion plus ARSTAT group was slightly greater than the statistics-negative attitude students in the lecture/discussion group. As one might expect, a Pearson product moment correlation revealed that students who had a positive attitude toward statistics (higher attitude score) were associated with a higher score on the achievement test ($r = 0.2805, p = .017, n = 57$).

Discussion of Additional Results

Twenty-two students evaluated the artificial intelligent statistics tutor on a 6-item 5-point Likert scale (Appendix D). It appeared that about 90% of the
students agreed with the content, the organization, and the presentation of the system. The graphics were adequate to explain the content. The major problem of the system was that the compiler time of the GCLISP (version 1) on the IBM PC-XT was too slow. The time required for the system to call up another procedure appeared to distract the students' attention.

Twenty-seven percent of the students strongly agreed that the quality of ARSTAT was good, 55% agreed, and the remaining 18% were neutral. One may conclude that, overall, the quality of the artificial intelligent statistics tutor was good. It was a good supplement to the learning of college-level statistics. Students who understood the statistics from the lecture/discussion found that the ARSTAT was a good review, and students who did not understand the material from the class lecture found that the tutor helped them to understand the concepts of the course materials.

Summary and Conclusions

The purpose of this study was to incorporate the content of elementary statistics into the ARPHY system shell and to investigate the effects of the artificial intelligent tutoring system as a supplement to learning statistics.

The results of this study indicate that the ARSTAT tutor can be used to supplement learning statistics at the college level, but there is no statistically significant difference between the achievement tests of two groups of
graduate students: one group learned in the normal lecture/discussion environment and the other group supplemented lecture/discussion with ARSTAT. But the students in the experimental group did slightly better on the test than the control group.

The findings also indicate that there is no statistically significant difference in interaction or main effect on the achievement test between gender and study technique. However, females in the lecture/discussion plus ARSTAT group did better on the test than any of the other three groups, and students in the lecture/discussion plus ARSTAT group did better on the test than students of the same sex in the lecture/discussion group. The significant finding in this study was that ARSTAT was viewed by students as being useful and that they had a positive attitude toward it.

From the results of the students' evaluations, one may conclude that ARSTAT can be used effectively as a tutor for students taking EDUC 5210, Educational Statistics. Also, the results confirm the advantages of the artificial intelligent tutor:

1. In this one-on-one situation, students can learn at their own pace. Faster students can move on to the next topics, while slower students can spend more time or repeat the instruction.

2. The ARSTAT frees the instructor for more individual
student contact.

One disadvantage of the artificial intelligent system is that the system cannot think like a human being and it tends to be impersonal; some students might need a human touch with their instruction. Another disadvantage is the development of the software, the amount of time and effort needed to implement the content into the system, even though the ARPHY system shell already existed. A minimum of 200 hours of programming time was needed to incorporate the 25 nodes of statistics content into the system. Each instructional node required about eight hours of programming time, using an intelligent GMACS editor resident in the GCLISP environment.

Recommendations

On the basis of the results of this study, the following recommendations for further study are made:

1. It is recommended that a similar study be conducted by incorporating more advanced topics of statistics, such as analysis of variance, factor analysis, etc., into the ARPHY teaching model.

2. It is recommended that a similar study be conducted by incorporating the ARPHY learning theory and the statistics content using another version of the LISP language or another programming language such as PROLOG.

3. It is recommended that other studies compare ARSTAT with some other kind of supplementary classroom
lecture/discussion.
CHAPTER BIBLIOGRAPHY


August 30, 1988

Dear Sir/Madam:

I am working on a doctoral dissertation entitled "An Investigation into the Effectiveness of Intelligent Tutoring on the Learning of College Level Statistics." My major professor is Dr. James R. Miller, Dean of the College of Education, University of North Texas.

The purpose of this study is to evaluate the effectiveness of the Artificial Intelligent Statistics Tutorial Program (ARSTAT) as a supplement to learning descriptive statistics. ARSTAT covers the content of statistical concepts, frequency distributions, measures of central tendency, and measures of variation.

Due to the advance of computer technology, teaching and learning by computer has an important role both today and in the future. The results of this study will be useful in the production of learner-based instruction for statistics.

I request your cooperation in answering the Attitude Toward Statistics Questionnaire and the pretest questions. In addition to this study I will need to look at part of your midterm score, the score on chapters 1-4. To preserve confidentiality, any information obtained in this study will be recorded by a code number.

Thank you in advance for your help and cooperation.

Sincerely,

Nanta Palitawanont
Researcher

Professor Dr. James R. Miller
Major Professor
APPENDIX B

SECOND LETTER TO THE STUDENTS
September 1, 1988

Dear Sir/Madam:

I am working on a doctoral dissertation entitled "An Investigation into the Effectiveness of Intelligent Tutoring on the Learning of College Level Statistics." My major professor is Dr. James R. Miller, Dean of the College of Education, University of North Texas.

I am requesting that you help me with my research by participating in a tutorial session using the Artificial Intelligent Statistics Tutor (ARSTAT). The content of the tutorial will correspond to the first four chapters (statistical concepts, frequency distribution, measure of central tendency, and measure of variation). ARSTAT will be available before the midterm exam, during Sept. 19-30, at the Media Center, Matthews Hall, room 349.

The ARSTAT tutorial will review the content of the first four chapters with the presentation and examples that will help you to understand the concepts of statistics. Also there are a series of exercises after each topic that will allow you to practice answering questions on statistics.

You may participate at any time when the lab is available. The tutorial session will be about four hours long. You may participate two times, and you are free to withdraw your consent and discontinue participation in this study at any time. Your decision whether or not to participate will in no way affect your course grade.

At the end of the tutorial you will be asked to evaluate the tutorial program.

Thank you in advance for your help and cooperation.

Sincerely,

Nanta Palitawanont
Researcher

I agree to participate in the tutorial program.

I prefer not to participate in the tutorial program.

Signature _______________________________________________________

Date __________________________________________________________________________
APPENDIX C

STUDENT EVALUATION QUESTIONNAIRES
For each of the following statements, please indicate your degree of agreement or disagreement by circling the appropriate number, based on the following scale:

1 - strongly agree
2 - agree
3 - neutral
4 - disagree
5 - strongly disagree

1 2 3 4 5 The content of statistics in the tutoring system was relevant.
1 2 3 4 5 The content presented in the tutoring system was easy to understand.
1 2 3 4 5 The content presented in the tutoring system was well-organized.
1 2 3 4 5 It was easy for me to keep up with the content presented in the tutoring system.
1 2 3 4 5 The tutoring system kept my attention.
1 2 3 4 5 The overall quality of the tutoring system was good.
APPENDIX D

SUMMARY OF STUDENT EVALUATIONS OF THE ARSTAT
### Table 11

Summary of Students Evaluation of the Artificial Intelligent Tutoring System

<table>
<thead>
<tr>
<th>Questions</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The content of statistics in the tutoring system was relevant.</td>
<td>16</td>
<td>5</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>72.7% 22.7% 4.5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. The content presented in the tutoring system was easy to understand.</td>
<td>8</td>
<td>11</td>
<td>2</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>36.4% 50% 9% 4.5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. The content presented in the tutoring system was well-organized.</td>
<td>9</td>
<td>11</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>41% 50% 4.5% 4.5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. It was easy for me to keep up with the content presented in the tutoring system.</td>
<td>9</td>
<td>11</td>
<td>1</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>41% 50% 4.5% 4.5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. The tutoring system kept my attention.</td>
<td>9</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>41% 18% 23% 18%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. The overall quality of the tutoring system was good.</td>
<td>6</td>
<td>12</td>
<td>4</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>27% 55% 18%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N = 22

1 - strongly agree  
2 - agree  
3 - neutral  
4 - disagree  
5 - strongly disagree
REFERENCES

Books


Articles


Elliott, A. C. (1985, February). The roll of call as a


**Dissertations**


Reports


