ISBN 0-315-55649-8

THE UNIVERSITY OF ALBERTA

APPRENTICESHIP AND COMPUTER MANAGED LEARNING
--A TRIAL RUN BY MICROCOMPUTER

BY

ROLAND RUMBOLT

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF EDUCATION

IN

VOCATIONAL EDUCATION

DEPARTMENT OF INDUSTRIAL AND VOCATIONAL EDUCATION

EDMONTON, ALBERTA

(FALL, 198[illegible])

Dedication

I dedicate this work to my three girls--

Marion, Nancy, and Susan.

ABSTRACT

The purpose of this study was to investigate the feasibility of using microcomputer based on-line testing as the evaluation component of computer-managed learning (CML) systems in self-paced, mastery learning, postsecondary vocational programs. To do this thirty-six second year electrical apprentices at the Northern Alberta Institute of Technology were invited to participate in a study to compare the effects of taking summative exams on-line by microcomputer as compared to off-line by computer generated and scored paper and pencil exams.

This quasi-experimental study was based on 1) a posttest-only randomized control group design, with the independent variable being the testing method and the dependent variables being achievement and testing time, and 2) a randomized control group pretest-posttest design with the treatment again being the testing method and the dependent variables being test anxiety and computer attitudes. Random assignment was used to divide the sample into an experimental group and two control groups. The experimental group and one control group exchanged roles in the second half of the study to provide a type of counterbalanced design.

Statistical analysis of the data resulted in a failure to reject the four null hypotheses of the study, thus suggesting that student achievement levels, testing times,

changes in attitudes toward computers, and increases in test anxiety levels do not differ significantly whether students complete tests on-line by microcomputer or off-line by computer generated and scored paper and pencil tests. Based on analyses of qualitative data collected from students and instructors, a fifth hypothesis regarding equal attempts at cheating by either testing method was not rejected. Qualitative data collected from instructors and students supported certain aspects of the statistical findings, pinpointed advantages and disadvantages of each testing method, and provided suggestions that could make the on-line tests by microcomputer a realistic contender in CML.

It appears that on-line testing by microcomputer may be feasible for single classroom settings, specific dedicated test purposes, and some aspects of formative evaluation. However, answers to logistical questions, such as machine to student ratios and time requirements, are required before one can consider using them in larger integrated CML programs.

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation and gratitude to the many people who have encouraged and assisted me throughout this study, without whose support it could not have taken place.

I thank my advisor, Dr. Fred Ilott, for his consistent support, advise, and encouragement; and my committee members, Dr. David Collett and Dr. Craig Montgomerie, for their encouragement and assistance, not only for this research, but throughout my year of studies.

I thank Mr. Ed Kuefler and his collegues from the Electrical Department of the Northern Alberta Institute of Technology, for their permission and support enabling the study to take place.

Likewise, I extend sincere appreciation to Mr. Don Bell and his staff at the Apprenticeship and Trade Certification branch of Alberta Career Development and Employment, for their help in insuring that the study could take place and their cooperation throughout the study.

I would also like to thank and acknowledge Dr. Milton Petruk and the Apple Canada Education Foundation for generously providing microcomputers and peripherals for the duration of the study.

I thank Wally Latto for his interest and assistance, and especially the hours he spent modifying the original QUIZMASTER program to make it more closely match the requirements needed for the study.

I thank my wife, Marion, and daughter, Nancy, for their typing efforts.

In addition, I would like to thank Jim Sanders and the instructors, Art deBoer, Roger Gagnon, George Harris, Bob Jacobs, Harold Pearson, Doug Pederson, Joe Schlosser, Bill Yaremko, and Jerry Wright, of the first and second year electrical apprenticeship programs at NAIT, for their cooperation, assistance, and really making me feel at home.

I would like to give a special thank you to the second year apprentice students who took part in the study in recognition of their participation and cooperation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: THE PROBLEM AND ITS SETTING

Introduction

The general instructional design principles of self-paced, competency-based instruction have been promoted by proponents over the last two decades, and there have been many individual efforts at implementation. In the past, however, there have been implementation problems, a topic that has received little attention and has not been well documented. (Suzuki & Finch, 1987, p. 47) In recent years, a shift to this type of delivery method has started to gain momentum, particularly in vocational and technology programs. (Blank, 1987, p. 45; Fretwell, 1987, p. 47) This change to self-paced, individualized instruction has increased, considerably, the demand for one-to-one student-instructor consultation. This need for individual student-instructor interaction is magnified even more in vocational and technology programs, because of the demands of practicum--the practical hands-on component of the programs.

At the same time, under this type of instructional design, the load on the instructor for the administration of tests, referral to resources, and record keeping has also increased significantly. According to a 1985 state-funded report on the development of a computerized data management system for competency-based vocational education (CBVE) in Florida, "the paper work generated by CBVE complete with open entry/open exit is voluminous and demanding" (p. 10). In order to reduce the clerical tasks as much as possible some institutions have attempted to institute computer-managed learning (CML) systems, or some subset of them, to take care of the testing and record keeping components of the program. Primarily, these institutions have run CML systems on large mainframe computers, or minicomputers equipped with a lot of memory, employing off-line paper and pencil generated tests. Thus, most of the descriptions of CML systems reported in the literature deal with those larger hardware configurations. This statement was substantiated by a review of various meta-analytic studies, such as the one by Kulik, Kulik & Bangert-Drowns (1985) in which they noted the limitation that no micro-computer based studies were included in their review.

Many of the small institutions, however, do not have the resources to invest in mainframes or minicomputers to handle CML programs. Thus, in order for these smaller institutions to have access to CML systems, they must become generally available for the microcomputer. According to a report published by Oregon State University (1984), "extensive commercial and teacher written software is available for the vocational setting" for CML applications. (p. 27) However, it appears that many of these programs have addressed only selected aspects of CML, have not gained widespread use, and have resulted in few documented evaluations of them in the literature. Roblyer, Castine and King (1988), assessing the state of research regarding the impact of computers, particularly microcomputers, in education, observed:

> Unless there is a rapid increase in research to accompany computer implementation in education, future reviews will have to contend with the same problems as those of the past: few available studies and many studies with inadequate data and/or flaws which prevent their inclusion in a review. Also, since past reviews included almost no microcomputers, future reviews should emphasize this technology and compare it with findings from mainframe studies. (p. 39)

To facilitate the development of practical and useful microcomputer based CML systems, various institutions and/or researchers need to experiment with various microcomputer programs in this CML genre to determine their capabilities and limitations. This study was one such attempt.

Since record keeping, grade book, and prescription components of CML are primarily a function of available memory, and that problem is already being addressed with the ever expanding memory capability with each new model change; those particular aspects were not addressed in this study. Rather, this particular study focused on the testing component of CML; in particular, on the viablity of administering summative evaluation on-line by microcomputer.

Since it is felt, generally, that the CML mainframe programs have a "proven track record", it gives a benchmark against which to measure similar application programs for the microcomputer. (Romaniuk & Lee, 1984, p. 4) By using one of those larger systems as a reference for comparison, investigations were made to see if it were possible to set up a CML system on microcomputer such that student achievement levels, student anxiety, time, and test security were not significantly compromised. In addition, the study tried to detect any potential weaknesses, strengths, and required improvements needed to make the delivery of CML by microcomputer a realistic alternative.

This investigation was carried out in a program composed of postsecondary adult students at the vocational trades level, an area in which this type of methodology is in need of research because of the increasing trend toward individualized, self-paced instruction. In contrast, most of the published studies concerning CML were made using undergraduate and graduate university students as subjects.

Statement of the Problem

The primary focus of this study was to investigate the following general problem: What are the differential effects of testing CML program modules on-line by microcomputer as compared to testing the modules off-line by paper and pencil tests generated and scored by minicomputer?

In order to examine this general question the following research questions were developed:

Research Question 1: Is there a significant difference in achievement when students are tested on-line by microcomputer as compared to off-line by computer generated paper and pencil tests?

Research Question 2: Is there a significant change in test anxiety when students are tested on-line by microcomputer as compared to off-line by computer generated paper and pencil tests?

Research Question 3: Is there a significant change in attitudes of students toward computers when tested on-line by microcomputer as compared to off-line by computer generated paper and pencil tests?

Research Question 4: Is there a significant difference in the time required to complete a test on-line by microcomputer as compared to off-line using computer generated paper and pencil exams?

Research Question 5: Are there perceived any increased attempts or opportunities for cheating or trying to beat the system when using on-line computer tests as compared to off-line tests?

Hypotheses

To try to establish answers to the stated research questions the following null hypotheses were developed:

Hypothesis 1: There will be no significant difference in achievement between students who are tested on-line by microcomputer and students who are tested off-line by computer generated paper and pencil tests.

<u>Hypothesis 2</u>: There will be no significant difference in test anxiety levels between students who are tested on-line by microcomputer and students who are tested off-line by computer generated paper and pencil tests.

<u>Hypothesis 3</u>: There will be no significant difference in attitude toward the role of computers and CML in training between students who are tested on-line by microcomputer and students who are tested off-line by computer generated paper and pencil tests.

<u>Hypothesis 4</u>: There will be no significant difference in test completion times between students who are tested on-line by microcomputer and students who are tested off-line by computer generated paper and pencil tests.

<u>Hypothesis 5</u>: There will be no perceived increase in attempts at cheating or increased opportunities for cheating when using on-line computer tests as compared to using off-line tests. This hypothesis will not be treated as a null hypothesis and its support or rejection, in a general way, must rely on qualitative data based primarily on the observations and reported perceptions of students, instructors, and the investigator.

Limitations

The following limitations of this study may influence the degree to which the results might be generalized to a larger population and are a threat to the external validity of the experiment.

1. The size of the sample was relatively small. The experimental and control groups were formed by dividing the accessible sample (36 second year electrical apprenticeship students returning to class at the Northern Alberta Institute of Technology (NAIT) for theoretical training), randomly, into control and experimental groups.
2. The accessible sample was not selected at random from the population. The sample was formed from the population (the eligible second year electrical apprentices from central and northern Alberta) according to the normal selection process used by the Apprenticeship and Trade Certification Division of Alberta Career Development and Employment. Bias under these circumstances is highly unlikely, but this factor was not investigated.

Delimitations

1. The testing was restricted to measuring achievement in two laboratory modules.
2. To make the on-line and paper and pencil tests as equivalent as possible the study did not attempt to evaluate the effect of immmediate and delayed knowledge of results on achievement, a feature available on the on-line testing program.

Assumptions

1. It was assumed that the available sample was representative of the larger population of second year electrical apprentices in Alberta.
2. It was assumed that a one hour practice session on microcomputer made the students familiar enough in the operation of the microcomputer for testing purposes so that they suffered no disadvantage using it for taking tests as compared to using the hardcopy terminals.
3. It was assumed that test items which are randomly selected, by objective, from a test bank will produce unique tests of equal difficulty.
4. It was assumed that the random assignment process used to establish the control and experimental groups did indeed create equivalent groups.

Importance of the Study

The accelerating shift to self-paced, performance-based instruction has nearly overwhelmed instructors of vocational programs because of the combined requirements of many more clerical and management type duties on the one hand and more individual attention in practicum on the other. Any help that computers can give in simplifying and better managing this type of environment would be welcome. With the cost of microcomputers and networks coming within reach of most institutions, any studies to determine the practicality of using them as full CML systems--or some useful subset thereof, such as, testing and record management--will be beneficial and could serve to motivate instructors and institutions to try such systems.

Definition of Terms

General Terms

Computer-based education (CBE) is a term referring to any general instructional application of computers in education.

Computer-assisted instruction (CAI) is an application of computer-based education that "involves an on-line interactive process between a learner and a computerized delivery system, in which the computer assumes a direct instructional role" (Gery, 1987, p. 214).

Computer-managed instruction (CMI) is an application of computer-based education which may be "defined as the procedure employed to manage the instructional process by assisting teachers in diagnosing instructional needs, testing and monitoring student achievement, prescribing learning activities, and matching instructional materials to teaching/learning activities" (Bluhm, 1987, p. 153).

Computer-managed learning (CML) is the term used in Alberta in place of CMI. See the definition for computer-managed instruction.

Computer-assisted testing (CAT) is an application of computers to aid instructors to construct, administer, score, analyze, and record tests. (Bluhm, 1987, p. 128)

Operational Definitions

Computer generated paper and pencil test is a test created by a computer program, printed out on paper, completed by pencil in the traditional way, and then scored by the computer upon input of answers to the questions by the students.

On-line test is a test presented on the computer screen, requiring the student to enter his answers to the questions by using the keyboard.

Stratified random test-item generation refers to the process used by computers in the selection of test items during test construction such that unique, parallel tests are created by randomly selecting questions from each objective or each group of objectives.

Measure of Achievement for this study is the score, in percent, that a student attains on tests written to demonstrate mastery of the lab modules, Heating Controls and Motor Magnetics, in the second year electrical apprenticeship program.

Measure of Test Anxiety is the score received on Sarason's Test Anxiety Scale (TAS), a scale composed of 37 true/false items with higher scores indicating greater test anxiety.

Measure of Computer Attitude is the score received upon the completion of a 12-item, 5 point Likert scale, constructed by selecting appropriate items from two validated computer attitude scales. Scores range from 12 to 60, with a higher score indicating a more positive attitude toward computer usage.

QUIZMASTER, authored by J.F.D.Ilott and W.S.Latta, is "a CAT program designed for use on the Apple II series of computers. In addition to administering tests it maintains student records, allows control of student programs, is capable of providing appropriate feedback statements, and enables the teacher to author quizzes" (Latta, 1988, p.8).

CBTS-CML is a computer managed learning program distributed by the Computer Based Training System of Calgary, AB, and designed to operate on VAX minicomputers. It requires large memory capability for holding a substantial number of items for test banking, record keeping, item analysis, student tracking and resource referral.

Summary

This chapter addressed the general topic to be researched, the feasibility of using microcomputers for CML. Since CML has several components, the problem selected for this study focused on the evaluation component; in particular, a comparsion of on-line and off-line testing of adults in a postsecondary vocational program.

The following chapter provides an examination of literature pertaining to the problem and subproblems, in order to develop a background for the study. The third chapter provides a description of the design and setting of the study, and the methodology employed in the data collection process. The fourth chapter presents the results of the analyses of the data collected in the study. Finally, the fifth chapter provides a summary and analysis of the findings of the study, giving a more detailed interpretation of the results and possible theoretical and practical implications, as well as suggestions for further research.

CHAPTER 2 REVIEW OF RELATED LITERATURE

Introduction

This study examined the effects on achievement, test anxiety, computer attitudes, testing time, and potential for cheating, when students are tested on-line by microcomputer as compared to off-line by computer generated paper and pencil tests. To provide a background to this investigation the literature review includes the following subject areas that were relevant to this study: the instructional applications of computers to the learning environment; self-paced mastery learning; testing theory and methods; test anxiety and its measurement; attitudes, attitude change and its measurement; and the potential for cheating.

Development of Computer-Managed Learning

Computer terminology in education

From a modest beginning in the early 1960s computers have been playing an ever increasing role in education, and it appears that this trend will continue. (Ward,1984; Vermette, Orr, & Hall, 1986) All of the early growth of computer-based education (CBE) involved mainframe computers; but with the advent and improvement of the microcomputer, the trend has been to utilize it where possible. (Brzezinski, 1984) The first developments were in the area of computer-assisted instruction (CAI). (Alessi & Trollip, 1985, pp. 47-48) Then with growing interest in individualized instruction in the mid-sixties, the concept of computer-managed instruction (CMI) emerged. (Baker, 1978, p. 4) During approximately the same time period computer-assisted testing (CAT) systems, which focused on the creation of tests from computer stored question pools, evolved also. Another category of CBE appearing in the literature from time to time is computer enriched instruction (CEI). (Bangert-Drowns, Kulik, & Kulik, 1985, p. 61) The Kulik, Kulik, and Shwalb (1986) meta-analytic study of CBE applications to adult and technical training identified three overall types: CAI, CMI, and CEI.

This categorization appears fine on the surface, but the field of CBE is so new that there is no absolute agreement among authors and developers about which categories or acronyms to use. For example, many educators prefer to use computer-assisted learning (CAL) which they feel emphasizes the role of the learner rather than computer-assisted instruction (CAI), where the emphasis appears to be on instruction. On the other hand, some trainers prefer to use computer-assisted training (CAT). (Beech, 1986; Bork, 1985; Gery, 1987) This ambiguity causes confusion, and this led Bork (1985) to state: "the jargon of acronyms should be banned everywhere" (p. 60). Gery (1987), on the other hand, took a more enlightened view providing a table for readers to pick their choice (see Figure 1).

PICK YOUR OWN TERMINOLOGY

A	B	C
Computer	Assisted Aided Managed Based Enhanced Mediated Interactive	Instruction Learning Education Training Teaching Development Study
Select one from each column		

Figure 1

Computer Terminology in Education (Adapted from Gery, 1987)

Developing a definition for CML

The major field of interest for this study was computer-managed learning (CML). Thus an operational definition of the term CML would be helpful in providing a cornerstone as well as a guide to give a sense of direction to the study. Baker (1978) believed that because of the unrestrained growth in its formative years and "the existing diversity in what constitutes CMI, it is probably futile to attempt a precise definition" (p. 12). Instead, he offered some explanations of what CML is not. Bozeman (1979), however, defined it this way: "CMI systems are generally considered management information systems designed to support the management process or functions associated with individualized education" (p. 118). Burke's (1982) definition required that computer programs contain all the required elements; such as, testing, diagnosis, learning prescriptions, thorough record keeping, etc., before it could be classified as CMI. (Bluhm, 1987, p. 8) The definition of preference for this study was a more general definition given by Leiblum (1982), who identified 12 components of a CML system. He then divided these 12 components or functions into three groups (see Figure 2). If a computer program or system contains at least one component from each group it may be classified as a CMI system. Romaniuk and Lee (1984) adopted the same definition in their study for the Alberta Correspondence School regarding the evaluation of the CBTS-CML system, the system that was used as the control in the present study. (pp. 5-8) According to Alessi and Trollip (1985), "CMI programs [may] vary from simply scoring tests and making instructional assignments, through integrating the management of all parts of the instructional process: testing, scoring, assignment making,resource and student scheduling, grading, and reporting" (p. 50).

The roots of CML

To find the real genesis of computer-managed learning and to forecast its chance for long term survival, one must look at the needs and conditions which underpin its creation and evolution. "Educators should realize that a 'true' CML program is rooted in mastery learning and individualization" (Bluhm, 1987, p. 8). Even though the ideas of individualization and behavioral objectives had been around from the early 1900s, it was not until the work

Computer-Managed-Learning Elements

GROUP ONE

1. **Objective Banking.**
The collection and maintenance of a structured set of objectives related to a lesson, task, a course, a discipline of curriculum.
2. **Learning Resources Banking and Library Information Retrieval.**
The collection and maintenance of structured lists of educational facilities/resources/packages and/of a library information retrieval system(just books or printed materials).
3. **Learning Material Banking.**
The collection and maintenance of instructional materials (a learning resource center) limited to those types which can be stored in computer hardware/software of peripheral units.
4. **Item Banking.**
   a. Collection of a structured set of items.
   b. On-line editing of items.
   c. Maintenance of item statistics.
   d. Maintenance of bank usage statistics.

GROUP TWO

5. **Item Generation.**
   a. Generation of items via a framework or macro facility, e.g. random number generation.
   b. Generation of items through randomly selecting/substituting existing parts of the item (stem of answer choices).
   c. Generation of items based on a model/grammar representing the structure of the learning materials.
6. **Test Generation.**
   a. Allocating items for individual study quizzes.
   b. Generating parallel forms of a specific quiz.
   c. Generating standardized final examinations from an item bank.
7. **Assessment.**
   a. Scoring of individual quizzes.
   b. Scoring of final examinations.
8. **Reporting.**
   a. Reporting (usually printed form) of individual study results (tests).
   b. Reporting of individual and group study progress.
9. **Evaluation.**
   a. Providing test analysis.
   b Providing educational product analysis (formative evaluation).
   c. Providing educational process analysis, e.g. information about study time, intervals between attempts, number of attempts, questionnaire processing, etc.

GROUP THREE

10. **Assignment.**
Assigning tasks (e.g. homework) based on objectives or study results.
11. **Counselling.**
Providing individual advice relating to study goals.
12. **Scheduling.**
Creating and maintaining schedules of educational facilities, e.g. of manpower or physical/instructional resources.

Figure 2

Elements of Computer-Managed Learning (Adapted from Leiblum, 1982, P. 131) of Gagne (1965), Glaser (1967), Mager (1962), and others, that some educators started to use instructional design principles which entailed performing a task analysis of a subject matter area to develop a set of behavioral objectives. The course would then be divided into a number of smaller units, each containing relevent objectives; and students, as individuals working at their own pace, would learn these objectives to the point of mastery. Mastery of a unit was usually obtained by attaining some predetermined criterion score.

Attempts at implementation of individualized instruction and mastery learning encountered two main obstacles. First, there was the extensive and frequent amount of testing required, and secondly, the massive increase in record keeping. The ability of the computer to handle these sorts of tasks much more efficiently than teachers led to experimentation and pilot projects, both funded and unfunded, with the result being the development of CML. (Baker, 1978, p. 3-11)

<u>CML by microcomputer</u>

All of the early projects were implemented on mainframe computers; and even though they were relatively successful, from the point of view of teachers and students, it was clear that "CMI on a large scale computer was turning out to be an expensive proposition" (McIsaac & Baker, 1981, p. 41). In addition to cost, remote time-sharing systems reduced user control and the convenience of readily available access at any time. With the advent of the microcomputer, educators like McIsaac and Baker, saw the potential it possessed to overcome some of the problems inherent in the larger systems. Drawing upon their experiences with mainframe CML, they led the way in developing a CML system, MICRO-CMI, for the microcomputer. McIsaac and Baker (1981) reasoned that:

> Given the power of the new microcomputers and the relative high cost and inconvenience of time-sharing, the most reasonable approach to the management of instruction is through a school-based computer designed to store student data, score tests, and provide timely reports to staff, parents and students. (p. 42)

Despite this rekindled interest in CML by the introduction of microcomputers, it has not generated any widespread implementation. (Bluhm, 1987) This type of use for computers in education has its supporters, but it also has its critics. Such controversy only serves to hinder the integration of this form of instruction into the learning environment. Bluhm (1987) quoted several surveys (Anderson & Smith, 1984; Becker, 1984; Simpson & Simcox, 1984; Stevens, 1984) which showed that only a small proportion of school districts throughout the United States were using microcomputers for classroom management tasks. (p. 7) Other studies quoted by Bluhm indicated that in 1985 "computer literacy" and "programming" were considered more importart, with CML having a low district priority.

A study by Dickey and Kherlopian (1987), commissioned by the United States National Institute of Education to evaluate the uses made of computers in mathematics and science in South Carolina, found that computers were mainly used for drill and practice, educational games, and tutorials, with no mention being made of CML at all, either by the researchers or the respondents to the questionnaire. Furthermore, even though approximately 70% of the teachers had access to computers, "approximately half of them did not use them" (Dickey & Kherlopian, 1987, p. 14).

Instructional design principles and CML

Perhaps, because the main delivery method in most elementary and secondary schools is still primarily the traditional, fixed-time, variable-achievement lecture method, there has been no real motivation to embrace a CML system. In contrast to this, in postsecondary vocational education, for the last two decades, the general instructional design principles of self-paced, mastery learning has been promoted. And an analysis of 40 models of instructional design by Andrews and Goodson (1980) "leaves little doubt that there is much similarity in the models which are in use today" (Dick, 1981, p. 29). A recent call by the Journal of Industrial Teacher Education, soliciting articles to bring together the current thinking and expected trends in alternate delivery systems for vocational education, gleaned only five articles. (Finch, 1987, p. 5) It was interesting to note that the articles, submitted by Blank, Finch, Fretwell, Pucel, and Swanson

(1987), were all of the self-paced, competency-based genre. Since "computers have unique capabilities for assisting in such courses", it is expected to be an area in which CML will grow. (Bork, 1986, p. 333)

Even though interest in CML has not developed in regular day school systems, in the field of vocational education there still appears to be a continuing interest. Government funded research and development produced guides which addressed the uses of computers in vocational and technical education, including CML. (Oregon State University, 1984; Zahniser, 1983) In addition, some individuals and institutions developed various components of CML to be delivered by microcomputer. For example, the Lake County Area Vocational Center, Grayslake, Illinois, developed a CML program for use on Apple Computers to manage most of the functions, except testing, in a competency-based vocational education environment. (Parker & Sands, 1985) Another example was the development of a computerized data management system for CBVE at the St. Petersburg Vocational Institute, a system alluded to in the introduction to this study. It was developed to run on Radio Shack TRS-80 and IBM PC microcomputers and could handle most of the functions of CML, except testing.

This continued interest in CML has been kept alive because of the potential advantages it has to offer CBVE. One may recapitulate some of these advantages by paraphrasing Bluhm (1987). Bluhm (1987), relying on studies by Dunkleberger and Heikkinen (1983), Hart (1981-82), Milner (1983), and Riedesel and Clements (1985), indicated that CML had the potential to reduce, significantly, the time spent by teachers in preparing, administering, and scoring tests; keeping records; and filling out required forms, thus freeing instructors to provide more time for helping students.

Testing Theory

Achievement testing concepts

One of the main components of a CML system is the testing component, and as Bork (1986) says "The heart of any mastery based course is the testing program" (p. 334). These two statements highlight the relationship between computer technology and the way it may be used to facilitate self-paced competency-based learning. Tests measuring achievement in courses employing this type of instructional design are normally referred to as criterion referenced tests. Criterion referenced tests differ from norm referenced test in the following way: norm referenced tests will compare students' performance in relation to the performance of the rest of the class members on that test. Criterion referenced tests, on the other hand, compare students' performance on the test against the list of objectives the student must master before continuing with the new material. (Popham, 1981, pp. 26-27)

Gronlund (1982) identified three types of evaluation with respect to determining the progression of students through a course. First, diagnostic tests may be used to determine if students are ready for a course and where they should be placed to begin learning. Secondly, formative evaluation type tests may be used to check students' ongoing progress and provide feedback for corrective action and motivation. Finally, summative tests are used to provide final evaluation of a course, or parts of a course, to determine promotion, certification, etc. (p. 2) The kind of evaluation in which the present study will be engaged is of the summative type, utilizing two different lab modules. Since the tests for these modules are of the summative type, all feedback will be delayed until each test is completed by the student, a view supported by Alessi and Trollip. (1985, p. 246)

On-line testing and off-line testing

Tests may be created onscreen by the computer and questions answered on-line by entering responses at the keyboard. Or they may be printed out on paper to be completed off-line, with the answers often being entered into the computer by the use of cards, or sheets, and optical scanners. Gwinn and Beal (1988) point out that "on-line or interactive testing appears to have received little attention in the past" (p. 239). Instead, the off-line or batch mode appears to be more frequently used. (Dunkleberger & Heikkinen, 1982, pp. 218-225) However, Gwinn and Beal (1988), whose CMI program runs on a mainframe, endorsed on-line testing and indicated that many early drawbacks to on-line testing, such as, cost, reliability and lack of hardware, have largely disappeared. (pp. 239-240) As a result, there should be an increased trend toward on-line computerized testing. Ward (1984) suggested that computer-administered tests, because of their advantage of speed, flexibility and efficiency, may displace paper and pencil tests in some educational settings. (pp. 16-20)

Another variation of computer testing, computerized adaptive testing, uses interactive tests in which the computer not only displays the test items and accepts the students' responses, it "selects items appropriate to the ability of the individual taking the test" based on the individual's previous responses. (Moe & Johnson, 1988, p. 79)

Computerized adaptive testing offers the advantage of reduced test length while maintaining measurement accuracy. (Ward, 1984) The study by Moe and Johnson (1988), employing this type of on-line testing drew favorable support from students, teachers, and administrators, leading the authors to conclude the study "suggests that computerized testing can be a viable alternative to traditional testing in the secondary schools" (p. 85). Millman (1984), however, suggested that the perceived savings in time may not be as great as some researchers suggest. In addition, this type of testing method required much more disk storage space.

Required testing features of CML

A review of literature in the field of testing and computer-based education (Alessi & Trollip, 1985; Dunkleberger & Heikkinen, 1983; Hambleton, 1983 & 1984; Ward, 1984) provided a laundry list of "do's and don'ts" and "pros and cons" of the various techniques used when employing computers to construct and administer tests. Many of these features, such as, randomly generated, unique tests, immediate knowledge of results, test-item analysis, and recording and displaying of student achievement results, are included features of both the CBTS-CML system (used by control group), and the microcomputer CAT program QUIZMASTER (used by the experimental group in this study).

Test Anxiety and its Measurement

State anxiety or trait anxiety?

Although computers hold great potential for administering tests, a few studies have indicated that they generate more test anxiety than paper and pencil tests. Hedl, O'Neil, and Hansen (1973) found that, contrary to their expectations, college students taking an intelligence test on-line by computer experienced a higher level state of anxiety than students taking paper and pencil tests. A study by Llabre, Clements, Fitzhugh, Lancelotta, Mazzagatti, and Quinones (1987) supported this finding. However, their study did not find any systematic relationship between test anxiety and achievement scores. This finding contrasted with Johnson and Johnson (1981), who suggested that higher anxiety levels would yield poorer test scores. Gwinn and Beal (1988), in a longer study and using a much larger sample, found that general test anxiety increased slightly over time. (p. 245) They also found that "test anxiety was significant for predicting computer test scores" but not a significant predictor of the final grade. (p. 247) This would suggest that the type of anxiety being measured is what Spielberger (1966) calls "state" anxiety, which is temporary and susceptible to change, rather than "trait" anxiety, which is permanent. (Cambre & Cook, 1987, p. 15; Gwinn & Beal, 1988, p. 247)

Measuring test anxiety

In non-computer related studies "research since the early 1950s has consistently verified the significant negative relationship between test anxiety and academic achievement for children as well as adults" (Shaha, 1984, p. 869). Sarason and Mandler (1952) were among the first to research this relationship between test anxiety and achievement. In that study they developed a questionnaire to measure state test anxiety. This instrument, the Mandler-Sarason Test Anxiety Questionnaire (TAQ) has become a classic, with many studies using it, or variations of it, to obtain measures of general test anxiety. (Gwinn & Beal, 1988; Llabre et al, 1987; Shaha, 1984) Two other test anxiety scales, the Inventory of Test Anxiety (Osterhouse, 1972) and the Test Anxiety Inventory (Speilberger, 1978) are two other instruments for measuring test anxiety that are frequently used in research studies.

Test anxiety may not necessarily be bad; "test anxiety may have both interfering and facilitating effects upon intellectual performance" (Ruebush, 1960, p. 205). Alpert and Haber (1960) developed a questionnaire, the Achievement Anxiety Test (AAT) with two independent scales; one to measure anxiety which helps with performance during examinations and tests (facilitating test anxiety), and one to measure anxiety which interferes with performance during examinations and tests (debilitating test anxiety). (p. 213) In their study, in addition to determining general test anxiety by Sarason and Mandler's TAQ, Gwinn and Beal (1988) also modified Alpert and Haber's Facilitating/Debilitating Anxiety scale to measure anxiety induced by on-line testing versus paper and pencil testing. They found that at the end of the first term "computer administered tests engendered more debilitating anxiety and paper and pencil tests more facilitating anxiety"; but at the end of the second term when the scale was administered again, the results were reversed. (Gwinn & Beal, 1988, p. 246) Another scale, the Test Anxiety Scale (TAS), based on the Mandler-Sarason Test Anxiety Questionnaire (TAQ), was developed by I. G. Sarason (1958), a different Sarason than the original anxiety test developer. A modified version of this scale, consisting of 37 true/false statements, was expanded from the earlier 21 item scale to improve its sensitivity, and has been frequently used in various studies involving test anxiety. (Sarason, 1978, pp. 198-201)

All of these well known test anxiety scales have their items arranged according to a Likert scale format, a yes/no format, or true/false format; and have undergone much testing to ensure internal consistency. Studies have also been undertaken to determine the validity of various scales by comparing them with each other under test conditions. (Harper, 1974; Thyer and Papsdorf, 1982)

The present study used the Sarason (1978) Test Anxiety Scale (TAS) as the instrument to try to detect changes in student test anxiety.

Computer Attitude and its Measurement

The concept of attitude

According to Fishbein and Ajzen (1975) an attitude is "a learned predisposition to respond in a consistently favorable or unfavorable manner with respect to a given object" (p. 6). Since attitudes are learned it means that they may be modified or changed. One of the most famous theories dealing with attitude change is the one advanced by Festinger (1957). This theory holds that individuals experience discomfort when they hold logically inconsistent cognitions about an object or event. To reduce this cognitive dissonance, they are motivated to change their cognitions or attitudes. One way to bring about change is through exposure to additional information; for example, exposure to computers for testing purposes. However, the direction and degree of attitude change is a function of situational factors, source, medium, form, and content of the information. (Krech, Crutchfield, & Ballachey, 1962, p. 226)

Measuring attitudes

The most common process employed to attempt to measure attitude and attitude change has been an indirect method, which relies on an individual's opinions and behaviours with respect to some concept, process, action, etc. These opinions and behaviours can be an overt indicator of a person's internal attitudes (Katz, 1967), even though the validity of such a process can be open to question at times. The most common method of collecting an individual's opinions regarding some object or process is by means of a scale, usually in the form of a continuum, containing several items measuring the same concept.

Over the years several scales have been developed to measure attitudes and changes in attitudes, with the most popular scale being the one developed by Likert. This five point scale asked subjects to indicate the degree of agreement or disagreement with each of a large representative set of items pertaining to some object or event. All the items then undergo an item analysis to determine the degree to which attitude items discriminate among individuals who differ in their attitudes towards an object. The most discriminating items will be used in the final scale.

The construction and validation of computer attitude scales

The increased numbers of microcomputers, and their increased utilization in society, have created a need to collect empirical data on the attitudes of instructors, students, parents, and even the general public, toward computers. Because of this need several scales for measuring attitudes toward computers have been developed and validated. (Bannon, Marshall, & Fluegal, 1985; Bear, Richards, & Lancaster, 1987; Dambrot, Watkins-Malik, Silling, Marshall, & Garver, 1985; Loyd & Gressard, 1984; Loyd & Loyd, 1985; Reece & Gable, 1982) Most of the scales employ a Likert scale design to collect the opinions of those being tested, and from that infer an attitude or attitude change. Except for two (Bannon et al., 1985; Bear et al., 1987) the studies did not build, to any significant extent, upon the work of others. In all cases it appeared that the different researchers, working independently of each other, developed their own scales because this was a new field and "there were no appropriate standardized scales available" (Dambrot et al., 1985, p. 84). However, despite this independent development, the ideas and wording of many of the items on the various scales are quite similar.

Some items on all scales tried to assess computer anxiety as a component of general computer attitude. In addition to computer attitude scales, there have also been developed computer anxiety scales, such as the Computer Anxiety Index (CAIN) and the Computer Anxiety Scale (CAS). (Marcoulides, 1988; Simonson, Maurer, Montag-Torardi, & Whitaker, 1987) Upon inspection, it was found that many of the items used in the anxiety scales were similar to the ones used in computer attitude scales. Thus it would appear that researchers see anxiety as one of the main components affecting a person's attitude towards computers.

A review of the literature indicated that early studies were mainly descriptive in nature. Then Reece and Gable (1982) reported the development and validation of a 10 item general attitude scale, called Attitudes Toward Computers, to measure the attitudes of students toward the use of computers. In the preliminary construction of the instrument, with the help of 5 "expert" judges, 30 items were selected to cover the cognitive, behavioral, and affective components of attitudes. After testing, by administration to a sample of 233 seventh and eighth graders, and analysis, it was reduced to a 10 item scale. The internal consistency reliability was found to be .87. A 5-point Likert Scale was used, and items were stated both negatively and positively, with a high score indicating positive attitudes toward computers.

Bannon, et al. (1985) constructed "a scale to determine computer attitudes among students, teachers, educational administrators, and other educators" following a method similar to the one used by Reece and Gable. (p. 679) Their scale consisted of 14 items, each arranged in a 5-point Likert Scale, ranging from strongly agreeing (0) to strongly disagreeing (4). The scale was arranged in two seven-item parts; one part measured positive cognitive computer attitudes and the other measured negative affective computer attitudes. A sample of 2525 students and teachers, male and female at various age levels, were used to test the instrument. An analysis of the results identified two factors, cognitive and affective, and internal consistency estimates of reliability of .922 for the cognitive factor and 0.896 for the affective factor. The results from the study showed the first scale had a low reading suggesting that students and educators had generally positive cognitive attitudes; the second scale had higher average scores indicating a tendency of rejecting negative affective computer attitudes.

Two other researchers, Loyd and Gressard (1984), because of the accelerated introduction of computer based programs into the schools, also recognized the need for a scale to determine computer attitudes of educators and students. The scale they constructed, called the Computer Attitude Scale (CAS), was made up of three subscales - Computer Liking, Computer Confidence, Computer Anxiety. The instrument was composed of 30 items, 10 in each subscale, of the Likert variety. The 30 items were selected from a larger pool of questions by a panel of judges.

The researchers then checked the validity and reliability of the instrument by trying it out on a sample of 155 students from grade 8 through 12. An analysis found that the coefficient alpha reliabilities were quite high (.86 to .95) on each of the subscales as well as on the total score. The item scores were arranged such that a "higher score on any of the subscales or on the total scale indicated a more positive attitude toward using or learning about computers" (Loyd & Gressard, 1984a, p. 503). When correlations were calculated, it was found that high positive correlations (from .64 to .93) existed among all the subscales and the total scale score. From this the researchers suggested that a high total score could be "interpreted to represent a general attitude toward working with computers that reflects liking, confidence, and freedom from anxiety" (Loyd & Gressard, 1984, p. 505).

Another study by Loyd and Gressard (1984) used their newly constructed CAS scale to "examine computer anxiety, computer confidence and computer liking among high school, community college, and senior college students and to explore the possible effects of computer experience, age and sex on these computer attitudes" (Loyd & Gressard, 1984b, p. 68). The size of the sample (354 participants) was much larger this time, and the study, in addition to investigating the research problems posed, served to confirm the validity and reliability of their instrument.

This study concluded that more computer experience correlated positively with more positive computer attitudes; and secondly, there was no significant difference in computer attitudes between sex groups.

Loyd and Loyd (1985) added another component, Computer Usefulness, to the CAS instrument. In addition, they wanted to see if the scale differentiated among persons with different amounts of computer experience. The subjects of this study were teachers, a group that the instrument had not been validated on previously. This study served to check the validity and reliability of the scale with this group, and at the same time revalidate the instrument with the new 10 item computer usefulness subscale added. This brought the number of items for the whole scale to a total of 40. Analysis of the data showed that "all four subscales were related closely enough to support the use of the total score as a measure of general anxiety" (Loyd & Loyd, 1985, p. 908). In addition, the study showed that the scale had the ability to differentiate among teachers with different amounts of computer experience.

Dambrot et al. (1985) constructed a new Computer Attitude Scale (CATT) which consisted of 20 items, 9 positive and 11 negative statements about computers. The statements were rated using a 5-point Likert type scale, with a high score indicating a negative attitude toward computers. The instrument was validated on a large sample before being used in the study, and resulted in an internal reliability of .79. This study reported that there were significant differences in computer attitudes between males and females, with the females holding more negative attitudes toward computers. This was in marked contrast to the Loyd and Gressard (1984) study referred to previously.

Researchers often select instruments developed and validated by others, and modify them to fit conditions of their particular investigations. Sanders (1988), in his study with first-year electrical apprentices at NAIT, selected a combination of appropriate items from two validated scales, the Computer Attitude Scale (CATT) (Dambrot et al., 1985) and the Computer Attitude Scale (CAS) (Loyd & Loyd, 1985, p. 47). The computer attitude scale thus created was then validated by a panel of experts. This modified scale by Sanders was the instrument that the current study used for measuring changes in computer attitude since the type of subjects, content, and setting were practically identical.

Cheating

Gwinn and Beal (1988) state that "cheating is a potential problem in any testing situation" (p. 241). This is no less true for on-line testing than it is for paper and pencil testing. In fact, in the Gwinn and Beal study (1988), students "report that actual knowledge of cheating is about the same on computer tests as on paper and pencil tests" (p. 241). However, they perceived that the chances for cheating were greater using the on-line method. These perceptions may come from a lack of knowledge of the testing security procedures built into the system.

Control of access to a test by the person who is supposed to write that test is probably the first step in attempting to insure security and reduce the chance of cheating in a computer testing program. This can be done in many programs by requiring students to use identification numbers, giving the student access to only specified exams, or setting sign-on time limits and test length times; or access may have some manual controls, like students having to show identification before the appropriate exams are released, or instructors issuing test by floppy disks. (Alessi & Trollip, 1985, p. 250; Qwinn & Beal, 1988, p. 241) Alessi and Trollip (1985), commenting on the use of floppy disks, observed that "although this generally requires more work by the instructor, it also is a method that is successful" (p. 250).

Demonstrative actions that convey to students the impression that cheating is taboo and will not be condoned during the testing process should be displayed by those administering the testing program. In the computerized testing program of the Gwinn and Beal (1988) study, to try to reduce the opportunity for cheating by collaboration during testing, students were randomly assigned to terminals and "test in progress" signs were placed on the top of assigned terminals. In addition, upon signing-on to the system students were given warnings onscreen that talking, looking at others' screens, and the use of notes and books were not permitted. At NAIT, where the present study took place, students were required to write tests in a testing room with glass windows between it and the learning center so that instructors could monitor the testing process. Students were also advised that they were being monitored both visually and auditorily. The creation of large test item banks such that items for exams may be selected from them randomly during the creation of a test, and the student having the knowledge that this is the way the tests are created could possibly be the most potent deterrent to cheating.

Summary

Computer-managed learning (CML) is a tool institutions can use to manage instructional programs, especially ones that are operated in an individualized, mastery learning mode. Despite the generally positive agreement among researchers of CML's ability to manage effectively and reduce the time spent on clerical duties such as test correction and record keeping, there are educators who, from a philosophical standpoint, do not support CML, for the same reasons they do not support mastery learning. With the advent of microcomputers it was predicted that CML would experience a rapid growth. However, this has not occurred, partly because of hardware restrictions and partly because of the traditional structure of the day school system. On the other hand, in post-secondary vocational institutions, where programs are switching to individualized, competency-based/mastery learning instructional designs, the advantages and potential benefits of CML are seen as the tool which can ensure success of this delivery methodology.

One of the key features of CML that makes it so important in determining the success or failure of individualized, competency-based instruction is the testing component. Tests measuring achievement in programs based on mastery learning designs are keyed to a set of behavioural objectives and are normally referred to as criterion referenced tests. The development of testing procedures employing computers has added much flexibility to the various requirements of diagnostic, formative and summative evaluation. For example, testing routines can be created to cause early termination during diagnostic evaluation, immediate feedback during formative evaluation, and delayed feedback during summative evaluation. Recently there has been evidence of a growing interest in exploring the potential of on-line tests as a substitute for the off-line computer generated paper and pencil tests and scan sheet computer scoring methods. This renewed interest has been brought on, mainly, by the development and advances made in microcomputer capabilities. The few relevant studies found, generally express support for on-line testing as a substitute for off-line paper and pencil testing. However, the effects on achievement of being tested on-line by microcomputer, one of the major focuses of this study, has not been investigated to any great depth.

Different testing methodologies give rise to concerns about the generation of excess test anxiety. The findings of the few studies located dealing with test anxiety and on-line testing generally indicate that they create more state test anxiety than paper and pencil tests do. The findings, however, have been not been overwhelming in support of that conclusion. One study, for example, found that tests in courses of longer duration were more facilitating than debilitating. A variety of scales for measuring test anxiety have been developed and validated during the last thirty years, with one of the most popular being the Test Anxiety Scale (TAS) developed by Sarason (1978), the scale used in the current study.

The attitudes that one develops toward some object or event are learned, thus they can be changed through further learning and experiences. The development of computers, especially microcomputers, has led to the creation and validation of a host of computer attitude scales to try to measure the attitudes toward computers of people, or classes of people, exposed to them under a variety of circumstances and environments. These scales measure attitudes indirectly by using peoples' expressed opinions. Since scales for the measurement of attitudes toward computers are relatively new, it appeared the primary focus of most of the researchers dealt with developing appropriate items for their instruments and the validating and checking the reliability of these scales. The present study attempts to detect any change in computer attitudes when using on-line test as compared to off-line tests by using a modified version of two of the validated computer attitude scales described earlier.

Cheating is always a concern during the administration of most forms of testing. The passing on of answers to test questions from one student to another under an individualized system, since all students do not have to write a specified exam at the same time, is particularly disturbing. The suggestion for creating large test item banks from which the computer can draw items randomly to construct unique exams for each student could be the solution to most problems in cheating, no matter what the delivery method used. Even then, Madaule (1988), in a paper presented to the First European Conference on Computers in Education, made the observation that no matter what method of testing we use there must still be huma. supervision.

CHAPTER 3: METHODOLOGY

Setting

At the Northern Alberta Institute of Technology training of first and second year apprentices for the electrical trade is facilitated by a CML program delivered on VAX minicomputers and a group of hardcopy terminals. The CML system being used is the one that was developed at the Southern Alberta Institute of Technology (SAIT) and now commercially available from Computer Based Training Systems (CBTS), Calgary, Alberta.

The CBTS-CML system provides features for administering exams and tracking student progress, and the electrical apprenticeship training division of NAIT has used it to set up test item banks and other program management options for all modules of their programs. Students are issued printed learning modules which state the objectives to be learned and the learning resources needed for their completion. Once given an identification password and granted access, any student can use the hardcopy terminal to print out a unique module exam which can be used as a self-check test, a review, or an assignment. This paper and pencil test can be taken home, and resources or other people may be used to help in its completion. To check for mastery the student then signs on to the computer again and enters answers to the test questions. The CML system corrects responses, states criteria score required, and keeps a record of the results. After failing three attempts, the system will deny continued access to the student, directing him to see the instructor. After several modules are completed, the student must generate a more substantial paper and pencil summative test to see if he has mastered the module information. These tests are written under supervised conditions in a testing room, a part of the learning resources center, and are the tests that are used to evaluate student progress.

In addition to the theory training, the programs have a lab component. The students are divided into groups, to work in the labs. Each lab section has a different instructor and offers a different module. All of the students in one lab work on the same module for approximately two weeks. At that time each student prints a unique, computer generated paper and pencil test, which is written under supervised conditions, to test lab theory. Each group then exchanges labs to complete the module that the other group had just finished. Two such modules of the lab component of the second year electrical apprenticeship program were used for this study.

Facility

This study took place at Plaza One of the Northern Alberta Institute of Technology where the first and second year electrical apprenticeship training programs are located. The CML room houses approximately 10 hardcopy terminals to give students access to the various tests stored in the CBTS-CML system. There are a large number of study carrels and several video carrels. There is an instructor facilitating/command area for helping students as well as carrying out administrative functions. This area is connected to the CBTS-CML system by video terminals, and the instructors control and monitor students' progress through them. In one corner of the CML room there is another room, capable of holding approximately 30 people, for testing purposes. This room has glass windows, and it is where students write their supervised exams. It was in this room that five Apple II type computers with dual disk drives were placed so that students would be writing the tests on-line under the same conditions as when writing the paper and pencil exams. The two practical labs in which practicum was conducted were separate rooms, located down the hall from the CML room.

Subjects

The sample used for this investigation was 38 second year electrical apprenticeship students who returned to class at the Northern Alberta Institute of Technology during the months of May and June, 1989, to complete the second portion of their in-class training. The students comprising this sample were part of the population of second year electrical apprentices from Central and

Northern Alberta and were selected according to normal procedures used by the Apprenticeship and Trade Certification Division of Alberta Career Development and Employment. A comparison between the sample and the population with respect to gender, age, and education level was made to determine if the sample was representative of the larger population. (Refer to appendix A for details of this comparison.) The sample comprised 37 males and one female. The sample was divided into three classes and all students in all classes participated in the study. Two of the students withdrew from the course; the data collected from them were eliminated from the study. The data collected from another student who participated in most of the study but who refused to do the on-line test by microcomputer was included. This provided a total sample of 36 subjects for most portions of the statistical analyses.

Design of the Study

Assignment of Groups

The 38 students in the sample were randomly assigned to three groups: an experimental group (13 students), a primary control group (13 students), and a secondary control group (12 students). The experimental group was then randomly divided into two subgroups and each half assigned to either the Heating Controls lab module class or the Motor Magnetics lab module class. Likewise the primary control group was also randomly divided into two subgroups and each half assigned to one or the other of the lab classes. The third group, the secondary control group, was assigned a Heating Controls module lab class as a whole group. The experimental group completed their first lab exams on-line by microcomputer; both control groups completed their first lab exams by computer generated paper and pencil tests (see Figure 3). In other words, for the first trial each class was divided approximately equally into experimental and primary control groups. The experimental group did on-line testing and the control group did computer-generated paper and pencil testing. This controlled for instructor variable. For the second lab module exams (second trial) the primary control group became the experimental group, taking the examination on-line by microcomputer; and the experimental group acted as the primary control group, taking the examination by computer generated paper and pencil tests. The secondary control group completed the exam on their second lab module, again by the paper and pencil method.

General Design

Since it was not possible to select a true random sample from the target population this study is considered to be of the quasi-experimental type. The study employed a variety of quasi-experimental designs. (Ary, Jacobs, & Razavich, 1985) The main part of the study, to investigate differences in achievement using on-line tests by microcomputer as compared to computer generated off-line

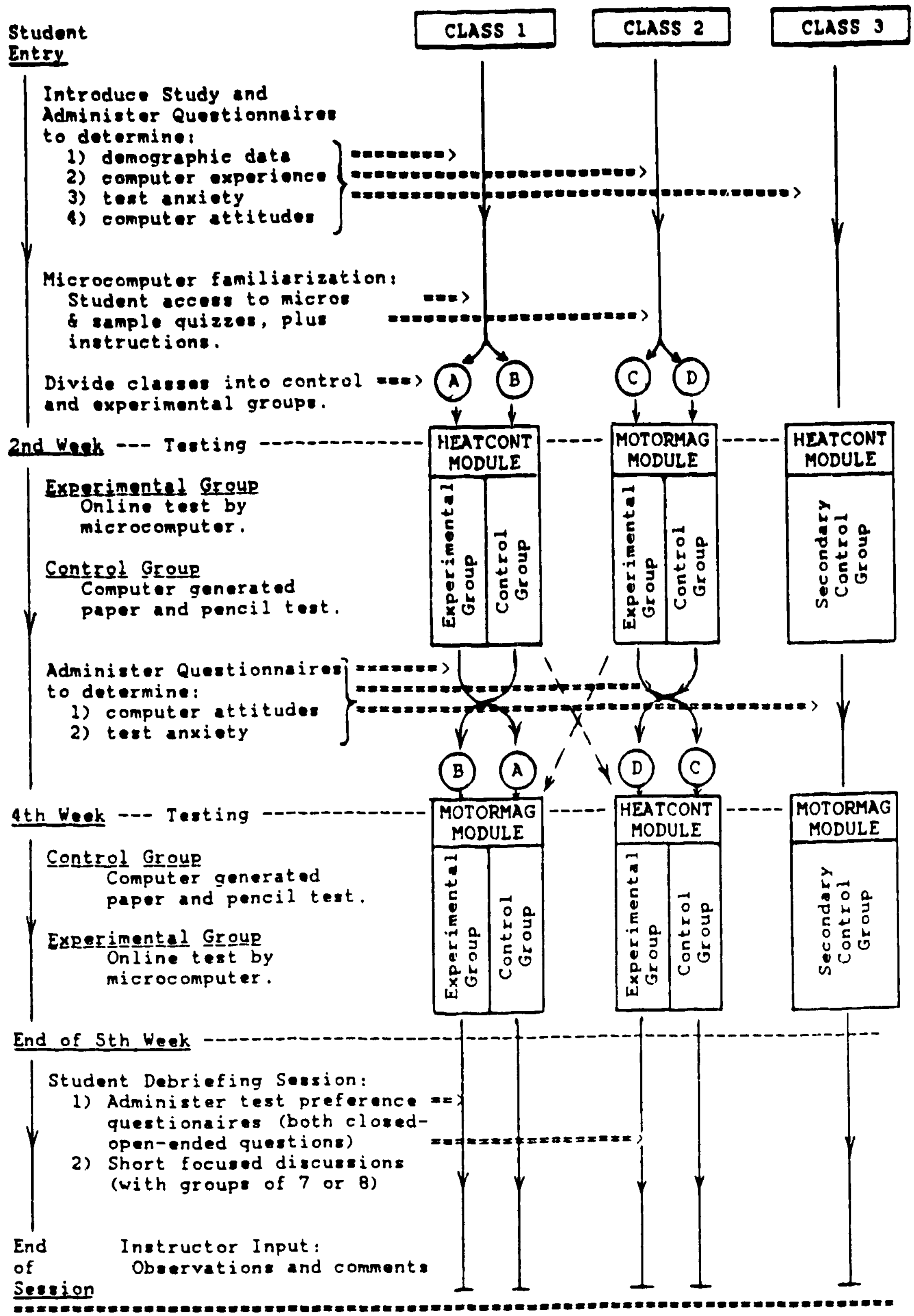

Figure 3

Study Design Model

paper and pencil tests, was a posttest-only randomized control group design. The independent variable was the testing method, on-line tests by microcomputer or computer generated paper and pencil tests, and the dependent variables were achievement scores and testing times. Since the posttest-only design is considered a weak one, the study set up a hybrid of the counterbalanced design by interchanging the subject matter that two classes studied, and at the same time the experimental group and the primary control group were interchanged for the second testing period. This permitted combining the results for each module to yield larger experimental and control groups in order to add more credence to the findings. Refer to Figure 3. In addition, a randomized control group pretest-posttest design was used to determine what effect the independent variable, on-line exams by microcomputer, had on the dependent variables, test anxiety and computer attitudes.

Experimental Procedure

On the morning of May 2, 1989, the day second year electrical apprentices began training, they were as a group given a brief introduction to the study as well as a committment to data confidentiality. In addition, they were informed that participation was voluntary but was encouraged. Questionnaires were administered, with an introductory cover letter attached, to: (1) gather demographic and computer experience data, and (2) determine pretest computer attitudes and test anxiety. On this first day students were randomly assigned to the three groups.

Two microcomputers were made available immediately so students could become familiar with them through the use of sample quizzes and personal instruction in order to counterbalance student experience on the hardcopy terminals. It was found that the rapid pace of the electrical program prevented students from finding enough time to thoroughly familiarize themselves with the microcomputer. Thus to compensate for this, in addition to the initial familiarization through personal instruction, students were also given another short briefing immediately prior to taking the supervised exam since "this is when anxiety may be heightened or reduced" (Alessi & Trollop, 1985, p. 252). The same explanation was given to each subject, and it was assumed that this briefing would negate any disadvantage that might exist by the student taking the test on-line as compared to off-line by paper and pencil.

At approximately the end of two weeks, each student then completed the achievement test by one of the two methods, off-line paper and pencil test using the hardcopy terminals, or on-line by microcomputer, depending on which group they were assigned to as described above. Since students were working at their own pace, they could take the exams early if they completed the lab module early. There was, however, a deadline date, May 16, 1989, by which the tests for the first lab modules had to be written. More than half of the students did not write until the deadline day. Times required to complete the tests by each method were recorded to the nearest minute. Test times were recorded manually for the microcomputer system and automatically on the CBTS-CML system. Immediately following the completion of the test, each student from all three groups was readministered posttest questionnaires for computer attitudes and test anxiety to assess any pretest-posttest changes in attitude or anxiety.

After the supervised examinations were finished for the first lab session, the lab classes composed of students from the experimental and primary control were assigned the alternate modules for the second two week session; that is, those who took the Heating Controls module during the first lab session were assigned to the Motor Magnetics module, and those who took Motor Magnetics now took Heating Controls. In addition, during this second session, the students who formed the primary control group during the first trial now became the experimental group, and the former experimental group became the primary control group. The secondary control group who were in a class by themselves, were assigned the Motor Magnetics lab module during this second session.

At the end of two weeks, approximately, to determine successful completion of the alternate modules, students again wrote supervised tests; on-line by microcomputer if they were part of the experimental group, but by computer generated paper and pencil test on the CBTS-CML system if part of either control group. Again the times required to complete tests by each method were recorded.

At approximately one week after the second achievement testing session, the part of the sample comprising the experimental and primary control groups were administered a short questionnaire containing both closed- and open-ended questions to elicit responses with regard to test preference, strengths, weaknesses, potential for cheating,and so on, of the two methods of delivery. Then debriefing sessions, taking approximately 30 minutes each, were held with the paricipants by breaking them into three small groups of 7 or 8 for short, focused discussions. Also, a short two-question questionnaire was administered to the secondary control group to determine their feelings about not being required/permitted to use the microcomputers for writing the exams.

When the study was conceived and planned, it was hoped that the instructors would be more involved with it. However, because of the instructional and program development demands on them, there was less opportunity for participation than anticipated. To help counterbalance this so as to extend and broaden the range of input from instructors, a short familiarization session and demonstration of the on-line testing program used in the study was held on June 20, 1989. The instructors were then given a short questionnaire containing open-ended questions, primarily, to gather observations and comments on various aspects of computer testing and the study.

Instruments

Demographic data and computer experience questionnaire: A short questionnaire was administered to the sample to determine the following : (1) demographic data regarding age, education, and sex; and (2) level and kind of computer experience. This questionnaire was attached to the pretest computer attitude and test anxiety questionnaires, and administered to all subjects at the same time. See Questionnaire #1 in Appendix B.

Achievement instruments: The instruments used to measure achievement were criterion-referenced module tests consisting of a combination of 40 multiple-choice and true/false questions, selected by a stratified, random format from computer test item banks, in such a way that the test for each subject was unique. The control tests were administered by paper hardcopy terminals on a VAX computer using the CBTS-CML system (see Appendix I for brief overview); the tests for the experimental group were administered using Apple II family computers utilizing a modified version of the software program QUIZMASTER (see Appendix H for overview). The questions in the CBTS- CML test item banks that apply to the Heating Controls and Motor Magnetics modules of the second year electrical apprenticeship program used in this study were entered onto question disks in QUIZMASTER to provide equivalent tests to those generated by the CBTS-CML system. Modifications were made to the QUIZMASTER program to give the student as much flexibility as possible when taking the test. For example, the student may browse through questions and answer later, change answers after any selection has been made, and review complete exam and change any answers he wishes before indicating that he wants the test scored. Several pilot trial runs were made on the transferred test item banks to ensure that the program performed properly during the testing process. In addition, the test items were proofread and answers debugged to minimize the chance of something going wrong when under actual testing conditions.

Because of the limited number of times a test item was drawn from the test pools due to the stratified, random selection process and the small size of the groups, it was difficult to compute test reliability coefficients by the usual Spearman-Brown split half procedure or the Kuder-Richardson formula 20 procedure. As a result a simpler method, the Kuder-Richardson 21 formula, was used. It requires data only on the number of items in the test, the variance of the scores, and the mean of the scores in order to calculate a reliability coefficient for the test. However, it does require the assumption that all items in each of the unit tests a were of equal difficulty.

Test anxiety instrument: Harper (1984) pointed out that "two instruments have been used to a significant degree in research on test anxiety . . . the Mandler-Sarason Test Anxiety Questionnaire and the Achievement Anxiety Test (AAT) developed by Alpert and Haber" (p. 961). Another instrument that has also gained wide acceptance for measuring test anxiety is I. G. Sarason's Test Anxiety Scale (TAS). A review of the literature indicated that these instruments, or some adaptation of them, are quite often used in many research studies involving test anxiety, and have been validated and cross validated with other scales. The scale used for the present study to try to detect differences in student test anxiety when using on-line test as compared to off-line paper and pencil tests was the Sarason (1978) Test Anxiety Scale (TAS), which used a scale consisting of 37 true/false items. The scale was pilot tested with another group of second year electrical apprentices, a group similar to the study sample, prior to being used in this study. See Questionnaires #1 and #2 in Appendix B.

Computer attitude scale: Several recently developed and validated scales for measuring attitudes toward computers and computer based learning were discussed in the literature review. These opinionnaires, usually of the Likert Scale type, must sometimes be modified because their statements do not apply to the population or the particular circumstances of the study. Sanders (1988) designed attitude instruments to measure changes in computer attitudes of first-year electrical apprentices at NAIT "by combining relevant items from two validated scales; the Computer Attitude Scale (CATT) (Dambrot et al., 1985), and the Computer Attitude Scale (CAS) (Loyd & Loyd, 1985)" (p. 47). The resulting 12-item scale, "designed to measure the students attitudes toward the use of computers in education", was pilot tested and administered in this study in a pretest-posttest design to detect perceptible changes in computer attitudes due to exposure to the experimental treatment, on-line computer testing. (Sanders, 1988, p. 47) See Questionnaires #1 and #2 in Appendix B.

Concluding questionnaire: This wrap-up questionnaire, administered about a week after the completion of the second module achievement tests, was constructed of both closed- and open-ended questions to elicit opinions from the groups regarding on-line microcomputer administered tests versus computer generated paper and pencil tests. This questionnaire asked questions to explore test preference, strengths, weaknesses, cheating, and perception of time required. See Questionnaire #3 in Appendix B. This questionnaire was completed by experimental and primary control groups but not by the secondary control group.

Focused discussion question form: A short list of questions was prepared to insure that all groups commented on the same areas of interest in the small group discussions, in order to obtain relevant, consistent data. Another person accompanied the interviewer during the discussion for the purpose of note taking. See a copy of this form in Appendix B. Experimental and primary control groups were involved in the discussion but not the secondary control group.

Secondary control group questionnaire: This was a questionnaire to elicit responses from the secondary control group about how they felt not being allowed to participate in the microcomputer testing experience. See questionnaire #4 in appendix B.

Instructor questionnaire: This questionnaire was constructed to elicit from instructors involved with the project observations and comments on various aspects of the study; such as, acceptance or rejection by student of on-line testing, strengths and weaknesses of such an implementation, perceptions of test security, etc. See Questionnaire #5 in Appendix B.

Additional Sources of Data

Data on the age, gender, education of the population of second year electrical apprentices of Alberta, as well as first year laboratory and theory marks for the present sample, were supplied courtesy of the Apprenticeship and Trade Certification Branch, Alberta Career Development and Employment, in order to check if the sample were representative of the population. In order to safeguard individual rights neither set of data included names.

Chapter Summary

Thirty-six second year electrical apprentices at the Northern Alberta Institute of Technology participated in a study to compare the effects of taking summative exams on-line by microcomputer as compared to off-line by computer generated paper and pencil exams. Part of this quasi-experimental study was based on a posttest-only randomized control group design, with the independent variable being the testing method and the dependent variables being achievement and testing time. The other part of the study was based on a randomized control group pretest-posttest design with the treatment again being the testing method and the dependent variables being test anxiety and computer attitudes. Random assignment was used to divide the sample into three groups: an experimental group, a primary control group, and a secondary control group. The experimental and primary control groups exchanged roles in the second half of the study to provide a type of counterbalanced design.

The next chapter will present results of the analyses of the data that was collected by the experimental procedure described in this chapter.

CHAPTER 4: RESULTS

Introduction

This chapter presents an analyses of the data in order to answer the research questions presented in chapter one. The data were analyzed using t-tests and ANOVAs (analysis of variance) primarily; and the results of these statistical tests and concommitant conclusions, as well as subsidary analyses, are presented here.

Hypothesis 1: Achievement

There will be no significant difference in achievement between students who are tested on-line by microcomputer and students who are tested off-line by computer generated and scored paper and pencil tests.

After the completion of the module exams for both trials, t-tests were performed to determine if there were significant differences in means of achievement scores between: (1) the two experimental groups completing tests on the Heating Controls module, (2) the two control groups taking tests on the Heating Controls module, (3) the two experimental groups taking tests on the Motor Magnetics module, and (4) the two control groups taking tests on the Motor Magnetics module. The results of the t-tests revealed that there were no significant differences in achievement between groups taking the tests either as experimental groups or control groups (Table 1).

Table 1

Comparison of Achievement Scores of Subgroups by use of t-t[illegible]s

Trial	Treatment (Groups)	N	Mean (SD)	Range	t	D.F.	P
Heating Controls							
1	On-line (A)	6	84.83 (8.91)	70-93	0.65	9	.534
2	On-line (D)	5	87.60 (3.65)	83-90			
1	P & P (B)	6	73.33 (6.44)	64-81	1.06	11	.380
2	P & P (C)	7	78.29 (9.76)	64-90			
Motor Magnetics							
1	On-line (C)	7	77.14 (9.69)	68-95	0.12	11	.910
2	On-line (B)	6	76.50(10.43)	60-88			
1	P & P (D)	5	83.40 (7.89)	74-95	0.15	9	.886
2	P & P (A)	6	82.50(11.54)	64-93			

*Groups (A,B,C,D)--see study design model, p. 46

Since there were no significant differences in achievement between groups at the .05 level of probability, the experimental groups and control groups for each module were combined to provide larger groups (Table 2) on which to perform oneway ANOVAs to ascertain if the differences in achievement between the experimental groups (on-line tests by microcomputer) and the primary and secondary control groups (computer generated and scored paper and pencil tests) were significant. The results of these findings are summarized in Tables 3 and 4.

Table 2

Combined Achievement Scores for Experimental and Control Groups

Group	N	Mean (SD)	Range	KR (21)
Heating Controls				
E (On-line)	11	86.09 (6.86)	70 - 93	.92
C1 (P & P)	14	75.14 (8.73)	62 - 90	.92
C2 (P & P)	11	84.36 (8.38)	67 - 98	.94
Total	36	81.31 (9.34)	62 - 98	
Motor Magnetics				
E (On-line)	13	76.85 (9.61)	60 - 95	.94
C1 (P & P)	11	82.91 (9.58)	64 - 95	.96
C2 (P & P)	11	83.36 (7.92)	69 - 98	.93
Total	35	80.80 (9.36)	60 - 98	

E - Experimental Group
C1 - Primary Control Group (Alternates as experimental group)
C2 - Secondary Control Group (Acts as a control group only)

The results of the ANOVA on the achievement scores in the Heating Controls module yielded an F-Ratio of 6.75, giving a level of significance of .0035 (Table 3). Thus there are significant differences ($p < .05$) between at least two of the mean scores of the three groups. The Scheffe posthoc test, at the .05 confidence level, indicated that the mean achievement score of the primary control group (paper and pencil tests) was significantly lower than the mean score of the experimental group

Table 3

Oneway Analysis of Variance of the Achievement Scores for the Three Groups (Heating Controls module)

Sources of variation	Sum of squares	D.F.	Mean squares	F Ratio	F Prob.
Between Groups	886.74	2	443.24	6.76	.0035
Within Groups	2165.17	33	65.61		

(on-line tests) and the mean score of the secondary control group. On the other hand, the Scheffe test indicated no significant difference between the mean scores of the experimental group and the secondary control group.

The ANOVA that was performed on the achievement scores for the Motor Magnetics module produced an F-Ratio of 1.96, yielding a level of signficance of .1577 (Table 4). It is apparent from the results that there are no significant differences between the scores of the three groups at the $p < .05$ level. Hence a Scheffe test to compare the means was not calculated.

Table 4

Oneway Analysis of Variance of the Achievement Scores for the Three Groups (Motor Magnetics module)

Sources of variation	Sum of squares	D.F.	Mean squares	F Ratio	F Prob.
Between Groups	324.45	2	162.23	1.96	.1577
Within Groups	2651.15	32	82.85		

However, inspection of the means of the groups for Motor Magnetics (Table 2) indicated that the experimental group (on-line test) mean is lower (although not significantly so) than that of the other two groups. In addition, inspection revealed that this group of subjects with lower mean scores by on-line testing were the same group with the lower mean scores by paper and pencil tests during the first trial when they acted as the primary control group. This apparent anomaly raised the question about whether the random assignment procedures used at the beginning of the study did indeed create three equal groups.

To investigate the possibility that the groups may not be equal, the subjects' first year lab and theory marks were procured, and oneway ANOVAs were performed on the achievement scores (Table 5).

Table 5

Oneway Analysis of Variance of First Year Laboratory and Theory Achievement Scores for the Three Groups

Sources of variation	Sum of squares	D.F.	Mean squares	F Ratio	F Prob.
Laboratory					
Between Groups	625.87	2	312.94	5.84	.0069
Within Groups	1714.69	32	53.58		
Theory					
Between Groups	301.33	2	150.66	2.72	.0809
Within Groups	1770.56	32	55.33		

The results of the oneway ANOVAs confirmed that there was a significant difference p = .05 between groups on the first year lab achievement scores, and applying the Scheffe test (p = .05) revealed that the difference was significant between group one and group two. Although the oneway ANOVA for the theory marks indicated there were no significant differences between the groups at the .05 level, an inspection of the probability (p = .0809) indicated that it is approaching the .05 level. Applying the Scheffe test at p = .10, a fairly "coarse screen", showed that there was a significant difference between groups one and two. An inspection of the means of the first year lab and theory achievement scores showed that the mean achievement scores of group one consistently exceeded the mean scores of group two (Table 6). Figure 4 displays graphically this consistent difference. Thus despite the random assignment of subjects, the groups do not appear to be equal.

To adjust for this uncontrolled variable (biased sampling), an ANCOVA (analysis of covariance) on the achievement data was performed using both the first year lab achievement scores and the first year theory marks as covariates, thus statistically taking into account and adjusting initial differences in the data (controlling the uncontrolled variable). The results of these analyses (Table 7) yielded probability levels of .088 and .086 for the Heating Controls module and the Motor Magnetic module

Table 6

Mean Achievement Scores in First Year Lab and Theory of the Three Groups

Group	N	Mean (SD)	Range
Laboratory			
Group One	11	85.55 (7.37)	68 - 96
Group Two	13	75.46 (6.78)	64 - 90
Group Three	11	78.45 (7.88)	67 - 88
Total	35	79.57 (8.30)	64 - 96
Theory			
Group One	11	83.91 (6.35)	71 - 92
Group Two	13	77.08 (7.65)	63 - 92
Group Three	11	78.45 (8.15)	63 - 92
Total	35	79.65 (7.80)	63 - 92

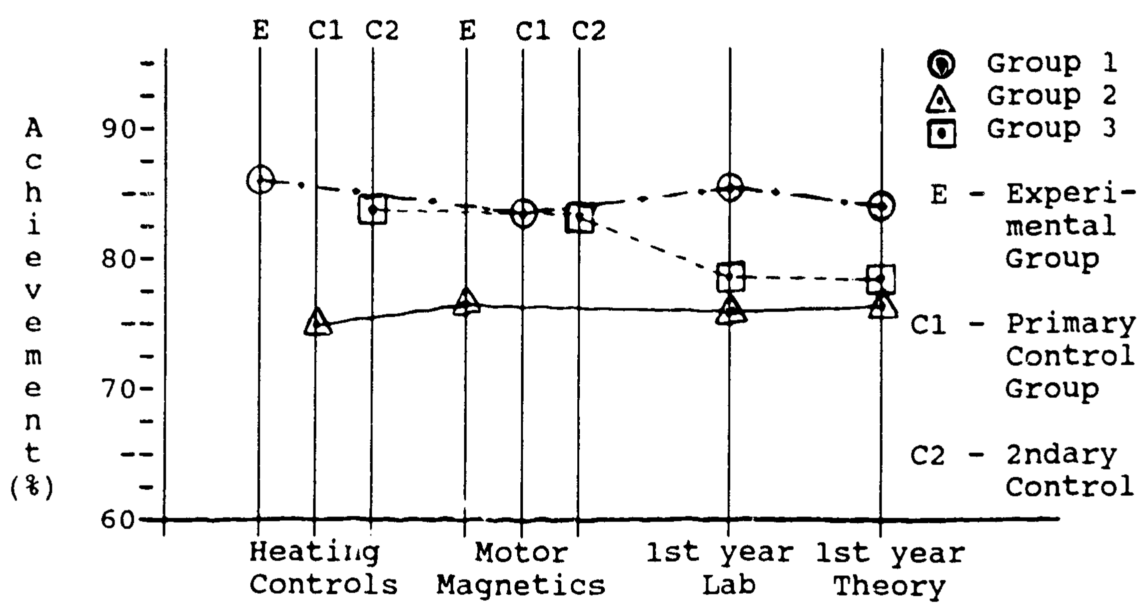

Figure 4

Comparison of Achievement Means (This Study and First Year)

Table 7

<u>Analysis of Covariance of the Achievement Scores by Group (Covariates--First year Lab and Theory Marks</u>

Sources of variation	Sum of squares	D.F.	Mean squares	F Ratio	F Prob.
<u>Heating Controls</u>					
Covariates	604.93	2	302.67	5.72	.008
1st year lab	49.77	1	49.77	.94	.340
1st year theory	115.92	1	115.92	2.19	.149
Main Effects	279.22	2	139.61	2.64	.088
Subgroups	279.22	2	139.61	2.64	.088
Explained	884.15	4	221.038	4.18	.009
Residual	1533.29	29	52.87		
Total	2417.44	33	73.26		
<u>Motor Magnetics</u>					
Covariates	1507.13	2	753.50	18.41	.000
1st year lab	38.35	1	38.35	.94	.341
1st year theory	457.91	1	457.91	11.18	.002
Main Effects	218.94	2	109.47	2.68	.086
Subgroup	218.94	2	109.47	2.68	.086
Explained	1726.06	4	431.51	10.54	.000
Residual	1186.91	29	40.92		
Total	2912.97	33	88.27		

respectively, thus indicating that there are no significant differences at $p < .05$ among the achievement score means of the three groups in either the Heating Controls module or the Motor Magnetic Controls module.

Based on the results of the analysis of covariance and the rationale for using it, the study failed to reject the first null hypothesis and thus no significant difference in achievement was found between the groups taking tests on-line by microcomputer and those taking off-line computer generated paper and pencil tests.

Hypothesis 2: Test Anxiety

There will be no significant difference in test anxiety levels between students who are tested on-line by microcomputer and students who are tested off-line by computer generated paper and pencil tests.

Summaries of the means, variances, and ranges of the test anxiety scores for each of the three groups, under both pretest and posttest conditions, are presented in Table 8. Inspection of the means in the table, as well as the graphical presentation (Figure 5), indicate that only slight differences exist between the three groups E, C1, and C2, both on the pretest and posttest. In addition, the means of the test anxiety scores for the posttest were only slightly greater than the pretest means. However, the pretest-posttest difference was smaller for the on-line computer testing group (E) than for the primary control group (C1), as indicated by the intersected lines on the graph (Figure 5).

To determine if these differences (between groups and from pretest to posttest) were large enough to be significant, a two by three analysis oı variance with repeated measures (ANOVAR) was performed on the test

Table 8

Comparison of Test Anxiety Scores by Group before and after Treatment

Group	Pretest Mean (SD)	Pretest Range	Posttest Mean (SD)	Posttest Range
E	15.23 (7.06)	7 - 33	15.85 (7.34)	5 - 31
C1	14.67 (9.23)	3 - 22	16.25 (9.73)	2 - 30
C	14.46 (8.66)	3 - 32	15.27 (7.63)	6 - 28

* Maximum score possible=37

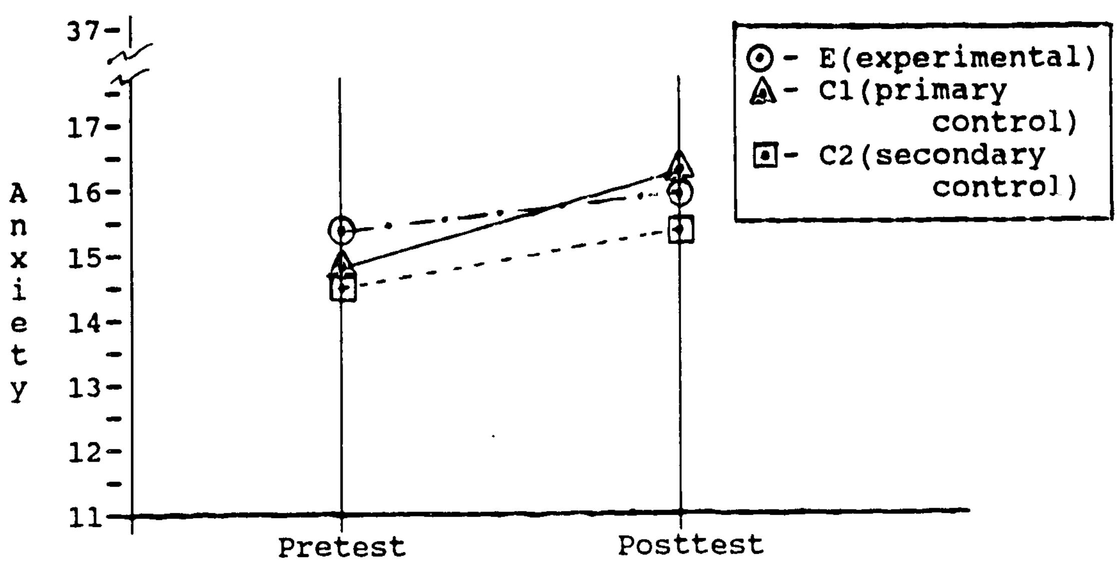

Figure 5

Test Anxiety Mean Scores by Group and Pretest-Posttest

anxiety data. The results of this analysis are summarized in Table 9. The F-ratios for the "between groups", "within groups", and "interaction effects" were quite small, and thus confirm that there are no significant differences ($p < .05$) in the test anxiety levels either between the

Table 9

Two by Three Analysis of Variance with Repeated Measures of Test Anxiety by Group before and after Treatment

Sources of variation	Sum of squares	D.F	Mean squares	F Ratio	F Prob.
Between Subjects					
Groups (A)	6.49	2	3.25	0.025	.975
Error within	4270.02	33	129.40		
Within Subjects					
Pre-Post Measures (B)	18.12	1	18.12	2.09	.158
Interaction AB	3.14	2	1.56	0.18	.836
Error within	285.81	33	8.66		

three groups or from the pretest to the posttest for each group.

Analysis performed for the whole sample revealed significant coorelations between achievement scores for the Heating Control module and the test anxiety scores both on the pretest ($r = -.3174$, $p = .030$) and posttest ($r = -.3923$, $p = .009$). On the Motor Magnetics module negative coorelations between achievement and test anxiety scores were also recorded; however, they were not large enough to be significant at $p < .05$.

The findings of this study thus failed to reject the null hypothesis since it appears that test anxiety levels were similar for students taking tests on-line by microcomputer and off-line by computer generated paper and pencil tests.

Hypothesis 3: Attitude

There will be no significant difference in attitude toward the role of computers and CML in training between students who are tested on-line by microcomputer and students who are tested off-line by computer generated paper and pencil tests. Table 10 presents a summary of the means, variances, and ranges of the computer attitude scale scores for each of the three groups, under both pretest and posttest conditions. An inspection of the means reveal that computer attitudes were approximately equal across all three groups and remained relatively unchanged between pretests and posttests. The high variance in control group (C1) is attributable to one score. The on-line computer test group (E) was the only group to show an increase from pretest to posttest, however miniscule it may be (Figure 6).

Table 10

Comparison of Computer Attitude Scores by Group before and after Treatment

Group	Pretest		Posttest	
	Mean (SD)	Range	Mean (SD)	Range
E	45.40 (6.68)	31 - 56	47.47 (5.10)	41 - 57
C1	46.83 (5.74)	39 - 55	46.13 (10.87)	19 - 56
C	46.27 (8.28)	25 - 53	45.37 (8.31)	27 - 56

*Minimum score = 12 Maximum score 60

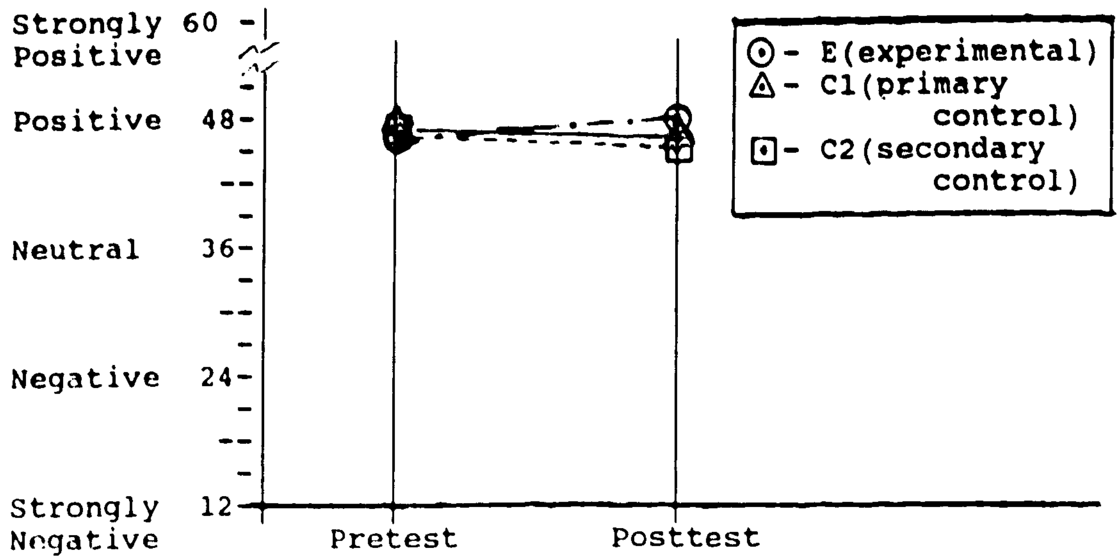

Figure 6

<u>Computer Attitude Scores by Group and Pretest-Posttest</u>

A two by three ANOVAR was performed on the computer attitude scores to see if the differences across groups and from pretests to posttests were large enough to be statistically significant at p = .05 level. The results of this analysis are summarized in Table 11 and confirm that there were no significant differences in the means, either between the experimental and control groups or from the pretest to the posttest for each group.

This study failed to reject the third null hypothesis since students taking tests on-line by microcomputer were not found to have attitudes and attitude changes different from those taking tests off-line by computer generated paper and pencil tests.

Table 11

Two by Three Analysis of Variance with Repeated Measures of Computer Attitudes by Group before and after Treatment

Sources of variation	Sum of squares	D.F	Mean squares	F Ratio	F Prob.
Between Subjects					
Groups (A)	6.44	2	3.22	0.034	.967
s-within	3134.38	33	94.98		
Within Subjects					
Pre-Post Measures (B)	0.42	1	0.42	0.019	0.890
Interaction AB	32.99	2	16.49	0.760	0.476
BS-within	716.25	33	21.71		

Hypothesis 4 : Test Times

There will be no significant difference in test completion time between students who are tested on-line by microcomputer and students who are tested off-line by computer generated paper and pencil tests.

In order to ascertain if the testing times for the experimental and control groups for both trials could be combined, t-tests were performed on the means. The results of the t-tests, summarized in Table 12, reveal that there were no significant differences at $p < .05$ level in completion times between groups taking the same module tests

Table 12

Comparison of Testing Times of Subgroups by use of t-tests

Trial	Treatment (Groups)	N	Mean (SD) (Minutes)	Range	t	D.F.	P
Heating Controls							
1	On-line (A)	6	28.17 (8.89)	18-39	0.65	9	.533
2	On-line (D)	5	31.4 (7.34)	24-40			
1	P & P (B)	6	34.83 (6.65)	27-43	1.70	11	.117
2	P & P (C)	7	45.14(13.45)	18-62			
Motor Magnetics							
1	On-line (C)	7	48.00(12.70)	30-65	0.39	11	.701
2	On-line (B)	6	45.33(11.47)	24-55			
1	P & P (D)	5	50.20(18.67)	30-71	1.18	9	.270
2	P & P (A)	6	39.33(11.83)	21-58			

*Groups (A,B,C,D)--see study design model, p. 46

either as experimental or control groups. Thus the experimental groups and control groups for each module were combined to provide larger groups (Table 13) on which to perform oneway ANOVAs to see if the differences in test times between the experimental groups (on-line tests by microcomputer) and the primary and secondary control groups (computer generated and scored paper and pencil tests) were significant. The results of these findings are summarized in Tables 14 and 15.

Table 13

Combined Testing Times for Experimental and Control Groups

Group	N	Mean (SD) (Minutes)	Range
Heating Controls			
E (On-line)	11	29.64 (7.99)	18 - 40
C1 (P & P)	13	40.38 (11.72)	18 - 62
C2 (P & P)	11	36.82 (13.29)	20 - 66
Total	35	35.88 (11.83)	18 - 66
Motor Magnetics			
E (On-line)	13	46.77 (11.72)	24 - 65
C1 (P & P)	11	44.27 (15.54)	21 - 71
C2 (P & P)	11	43.45 (11.46)	27 - 58
Total	35	44.94 (12.66)	21 - 71

E - Experimental Group
C1 - Primary Control Group (Alternates as experimental group)
C2 - Secondary Control Group (Acts as a control group only)

Table 14

Oneway Analysis of Variance of the Testing Times for the Three Groups (Heating Controls module)

Sources of variation	Sum of squares	D.F.	Mean squares	F Ratio	F Prob.
Between Groups	702.28	2	351.14	2.77	.078
Within Groups	4053.25	32	126.66		

Table 15

Oneway Analysis of Variance of the Testing Times for the Three Groups (Motor Magnetics module)

Sources of variation	Sum of squares	D.F.	Mean squares	F Ratio	F Prob.
Between Groups	77.67	2	36.33	0.22	.807
Within Groups	5377.22	32	168.04		

Since the ANOVAs indicate that there were no significant differences among the means at the .05 level of confidence, the study failed to reject the fourth null hypothesis. Thus no significant difference was found in the time required to complete the tests by either method.

Hypothesis 5 : Cheating

There will be no perceived increase in attempts at cheating, or increased opportunities for cheating, when using on-line microcomputer tests as compared to using computer generated and scored off-line paper and pencil tests.

The data gathered to support or reject such an hypothesis are primarily of a qualitative nature, derived from observations and perceptions of students, instructors, and the investigator.

The term cheating, with respect to taking exams, is very general and there are several sides to it. The two aspects that usually were addressed by respondents were

(1) the deliberate attempts at cheating while testing was in progress, and (2) the more subtle outside sharing and passing on of test answers. A majority of students and instructors commented on either one or both of these facets. The gathering of hard data on such a question is difficult and elusive at best, and how reliable the data that one gathers from the perceptions of others on such a topic compounds, even more, the chance of drawing erroneous conclusions. Despite this, the responses received from both students and instructors on questionnaires and through discussion appeared to be forthright and does provide at least a partial answer to the research question.

From observing the students in an unobtrusive manner as they wrote exams by each method, the researcher detected no attempts at cheating either when students wrote by paper and pencil tests or on-line by microcomputer.

The responses by students to the question, "Under which testing method do you feel it would be easier to cheat?", (see Appendix C) were as follows: 34.8% said it would be easier to cheat by paper and pencil tests, 4.3% said it was easier to cheat by microcomputer, and 60.9% said it was about equally as easy under either method. This question and other related aspects of cheating were explored further in focused discussion sessions (see Appendix F). The general consensus of the student groups was that very little cheating was occurring; and they perceived that the random generation of unique tests for each student, and a secure, supervised testing area, were the main factors for this.

Input from instructors regarding cheating was elicited by the following question: "Which system do you feel provides greater opportunities for cheating? Why?" Of the 10 respondents 50% felt that the opportunity for cheating was equal for each system; 40% felt paper and pencil tests offered more opportunity for cheating. Table 16 compares students' and instructors' perception of ease of cheating using the two systems. The views of both groups were quite similar, with the highest proportion in each group considering both testing methods about equal with respect

Table 16

Comparison of Student and Instructor Views on Ease of Cheating by each Method

Group	N	P & P Tests	On-line Tests	About Equal	No Response
Students	23	8 (35%)	1 (4%)	14 (60%)	
Instructors	10	4 (40%)		5 (50%)	1 (10%)

to ease of cheating. The on-line testing method was perceived as the system by which students were less likely to be able to cheat. This may be a biased view since the length of exposure to on-line testing by instructors and students was limited. However, with such a perceived difference between on-line testing and computer generated paper and pencil tests, one should be able to say, at the very least, that there are no increased attempts at cheating when using on-line tests. To obtain a variety of instructor responses for the reasons why they selected the option they did see Appendix E.

Subsidary Analysis

Thus far the results of the study have been extrapolated from the collected data in a quantitative manner using a variety of statistical analysis techniques. However, qualitative data of a more humanistic nature are also important when trying to assess various learning environments and the impact of different methodologies and technologies on them. Data of this kind was obtained for this study by eliciting the opinions and feelings of students and instructors by means of questionnaires and group discussion.

Responses by students to questions (see Appendix C) regarding test preference, test completion time, and ease of use, when comparing the two testing methods were mixed with neither method standing out as a method of choice (Table 17). However, if the test items were to contain mathematical calculations the testing method of preference shifted highly in favour (83%) of the paper and pencil method. Preference for paper and pencil tests when calculations were involved also surfaced in the focused group discussions and in instructor comments (see Appendices E and F).

Table 17

Comparison of testing methods by students

	N	P & P Test	On-line Test	Equal
Testing Method Preference	24	9(37.5%)	6(25%)	9(37.5%)
Fastest Testing Times	24	9(37.5%)	6(25%)	9(37.5%)
Easiest to use	24	9(37.5%)	7(29.2%)	8(33.3%)
Testing Method/with Math	24	19(82.6%)	2 (8.7%)	2 (8.7%)

A question regarding the inconvenience of using the systems indicated that the use of floppy disks may be more convenient than the hardcopy terminal (37.5% to 20.8%). However, the majority feel that both testing procedures had the same level of inconvenience (41%). The inconvenience of entering answers at the hardcopy terminal, after the paper and pencil tests were completed, appeared to be the main irritant for students, according to the responses made to open-ended questions on questionnaires and during focused discussion.

Students, in response to questions regarding necessary features and requirements while taking a test, indicate quite strongly (87% of respondents) that knowledge of results (marking of the questions) on summative exams should be delayed until the exam is completed and reviewed, as opposed to obtaining immediate feedback after each question. Also the ability to be able to change and review answers at any time during the testing process was seen by students and instructors as an important feature in on-line testing programs, and suggestions were made indicating that they must go even farther in emulating paper and pencil test flexibility. Nine out of ten instructors rated as very important a feature for testing programs to insure that test information is not lost after students start writing exams. In other words, if the system goes down, the proportion of the test the student has answered should not be lost. Another question of interest for the study was the students' and instructors' perceptions of the advantages and disadvantages of each testing method. Table 18 presents a summary of the most frequently stated responses by students with respect to the advantages and disadvantages of writing tests by each method (see Appendix

Table 18

Summary of Advantages and Disadvantages of on-line and off-line testing as stated by students.

Testing Method	Advantages/Disadvantages
Computer generated paper and pencil tests	
Advantages	**Disadvantages**
1. It is easy to preview whole test, and use other parts of tests for clues to other questions.	1. A lot of paper to contend with.
2. Able to underline key words.	2. Easy to miss questions or to misread questions because of quality of print.

3. No inconvenience switching between screen and paper on desk.

4. All questions are available all the time.

5. Especially advantageous for computation type questions.

6. Hardcopy useful for instructor-student consultation.

7. No advantages.

3. Takes too long to enter answers at terminal and get your results. If you make a mistake on entry you cannot correct your response.

4. Too much noise from terminals.

5. If there is a power outage the whole system goes down.

6. No disadvantages.

On-line Testing by Microcomputer

Advantages

1. Questions are clear and easy to read.

2. No paper required, correction is immediate when exam is finished.

3. One question at a time presentation makes it easy to concentrate.

4. Cannot miss a question.

5. No advantages.

Disadvantages

1. Cannot see question once it is skipped. Only one question at a time is permitted on screen.

2. Found it hard on the eyes.

3. Time taken to review answers too long.

4. Inconvenient to look between desk and screen when doing calculations.

C). The instructors' responses appear in Appendix E and are integrated with the discussion in Chapter 5.

Summary

The results of the statistical analysis of the present study suggest that student achievement levels, testing times, changes in attitudes toward computers, increases in test anxiety levels, and attempts at cheating do not differ significantly whether students complete tests on-line by microcomputer or off-line by computer generated and scored paper and pencil tests. Qualitative data collected on the testing methods support certain aspects of the statistical findings; however, it also provided advantages and disadvantages of the testing methods and suggestions that could make the on-line tests by microcomputer a realistic contender in CML. The next chapter contains a more detailed discussion of these results.

CHAPTER 5: SUMMARY, DISCUSSION, AND IMPLICATIONS

This study investigated the feasibility of using microcomputers to perform the testing function in a computer-managed learning environment. Specifically, the primary interest was in exploring the possibility of using on-line microcomputer based testing as a substitute for computer generated and scored paper and pencil tests in self-paced, competency-based, postsecondary vocational programs. To do this on-line tests parallel to existing computer generated paper and pencil tests were prepared and placed directly into the second year electrical apprenticeship program. A substantial part of the study compared empirical data gathered in both testing modes. In addition, qualitative data was collected from instructors and students and compared.

Summary of the Results

Several factors should be kept in mind as one attempts to draw conclusions from the findings presented in this study. First the size of the total sample was small, thus the experimental and control groups were small making it difficult to extrapolate beyond the present setting. The counterbalanced design of the study revealed that the random assignment of students to experimental and control groups did not create equal groups so covariates from another source were obtained to adjust the data. This does not mean that the information gained from the study is not useful; but that caution has to be used in interpreting the results.

Another factor the reader should bear in mind is that the control groups for the study were not using paper and pencil tests in the traditional manner; they were handling the hardcopy terminals to retrieve and score tests daily. For example, only one of the students in the sample did not complete his first year electrical apprenticeship training using the CBTS-CML system. This made the groups much more familiar with computer hardware and terminology, and one should be aware of it when the findings of this study are compared with the findings of others. Appendix G summarizes the variation of computer experience, ranging from simply playing computer games to taking formal computer courses, that the groups had. An inspection of this appendix indicates that there were not a wide range of differences in computer experiences between the groups.

Analyses of the data failed to reject the four null hypotheses that tested the first four research questions of the study. The fifth hypothesis for research question five was tested in a less rigorous manner, and rejected on the basis of analysis of qualitative, descriptive data collected during the experiment. It appeared that students achieved equally well whether writing tests on-line by microcomputer or off-line by computer generated and scored paper and pencil tests, and showed no significant differences in the length of time required to complete the tests. Similarly, there were no significant increases in test anxiety levels because of the testing methods used. In addition, no significant changes in attitudes toward computers were recorded in either group. Moreover, according to respondents' perceptions, attempts at cheating should be no different using on-line tests or computer generated paper and pencil tests.

Discussion and Implications of Findings

Achievement

Statistical analyses of the achievement data in the present study suggest that no significant differences exist between groups taking tests on-line by micrcomputer and off-line by paper and pencil tests. This finding generally supports the conclusions reached by other studies which have compared the on-line and off-line test methods, even though most of the studies were primarily a mix of various standardized achievement, aptitude, and personality tests and inventories. (Vansickle and Kapes, 1988) The present study, on the other hand, dealt with non-standardized achievement tests, of the type typically used for evaluation in a classroom setting.

Despite the lack of significant differences in achievement between the two methods, there appears to be a predominant perception (see Appendices C, E, and F) that the paper and pencil test has a significant advantage over the on-line test because of the freedom it allows students for viewing the whole test so that scanning back and forth through test questions can provide hints and answers for other questions in the test. According to the test scores this perceived advantage may be more apparent than real. Frame by frame question presentation, as is the usual on-line pattern, may offer equivalent corresponding advantages since the results were not significantly different. Vansickle and Kapes (1988) suggested that when only one item at a time is presented on the monitor it "could focus the examinee's attention on the task better" (p. 10).

Test Anxiety

Hedl et al. (1973) and Llabre et al. (1987) found that on-line computer tests generated more test anxiety than paper and pencil tests. The present study, on the other hand, found no significant difference in test anxiety scores between the groups tested on-line by microcomputer and those tested off-line by computer generated paper and pencil tests, either pretest or posttest. What caused this difference in study results is open to speculation. Whether familiarity through extensive student use of computer hardcopy terminals to generate and score paper and pencil tests impacted on the outcome of the experiment in a way that the traditional paper and pencil tests did not, was not investigated.

Hedl et al. (1973) reported that remarks by students "revealed a number of procedural variables in the administration of the computer test were important determiners of the observed affective reactions, and not the computer per se" (p. 221). They found that clear instructions, terminal familiarity, and increased student control are important in lowering anxiety levels. The present study tried to eliminate these variables to reduce their effect on the test anxiety levels. To do this, familiarization practice sessions were conducted covering the operation of the microcomputer and testing programs prior to testing. In addition, QUIZMASTER, the testing program used for this study, had many options that permitted placement of control in the hands of the students. Such options as the ability to skip questions, change answers, and review questions, all contribute to a sense of student control and freedom of choice while taking the tests. Student responses to question 5 (Appendix F) confirmed the level of importance placed on these features by students while writing tests on-line.

Hedl et al. also found that the lack of ability of the program to cease testing procedures when the student has failed the criteria score (or alternatively lack of a feature to delay knowledge of results) significantly increased test anxiety. QUIZMASTER, which can give either immediate or delayed knowledge of results, was set to delay the mark until the student finished the exam. When queried about student preference regarding this point on Questionnaire 3 (Appendix B), 87% of the students responded that they would prefer that the knowledge of test results be withheld until the test is completed, and the primary reason for this was the effect it would have on the rest of the exam because of anxiety.

These features, along with the increased exposure to computers, may have helped in producing the finding of this study that there was no significant difference in test anxiety when taking tests on-line by microcomputer and off-line by computer generated and scored paper and pencil tests. Gwinn and Beal (1988) found that general test anxiety increased slightly over time. The present study was of relatively short duration and whether the same pattern would have materialized is not known. No attempt was made to determine the relationship of test anxiety to curriculum content studied.

It appears that there is a certain amount of test anxiety no matter which testing method is used. Investigations of the relationship between academic achievement and test anxiety have consistently found a significant negative correlation between the two variables. (Shaha, 1984, p. 869) The present study supports this general finding. When the sample was taken as a whole, significant negative relationships were found between achievement scores for the Heating Controls module and the test anxiety scores on both the pretest and posttest. On the Motor Magnetics module a negative correlation between achievement and test anxiety was also found. An inspection of the test anxiety scores indicated the range (from a low score of 2 to a high score of 33 out of a possible 37) was quite large. However, this was caused by just a few subjects; most scores were concentrated in a much narrower range. These few outliers were present on both the pretest and posttest. If these students suffer from "state" test anxiety, then a longer exposure time should reduce their anxiety. On the other hand, if these scores are an indication of "trait" test anxiety, then it would be difficult to reduce it. No attempt was made to identify the students with outlier scores, either on the pretest or posttest; and no attempt was made to investigate the causes for such variation.

Attitude Towards Computers

One is never certain if the responses made by a student to items on an attitude scale actually reflect the true feelings of that person. However, it is the only type of instrument available for measuring computer attitudes; and it is hoped (and one must assume that is the case) that after being given reasonable assurance of confidentiality, participants will respond to all scale items in a frank, unbiased manner.

Dambrot et al. (1985) cite a number of studies which indicated that general attitudes toward computers in the early 1980s were negative, with females exhibiting more negativism than males. Three reasons advanced for this negativism toward computers were as follows:

1) functional problems, such as limited usefulness, limited availibility, downtime, computer incompatibility, poor documentation, and frustration of new and potential users; 2) inadequate training during initial contact overwhelmed new users because of the volume of information to be assimilated; and 3) a general resistance to change and a fear of technology. This negativism seems to have been replaced by more positive computer attitudes in recent years. Sanders (1988) found that the first year electrical apprentices expressed positive attitudes toward computers, even on the pretest. (p.78) Sanders' study is of interest in evaluating the findings of the present research because the same attitude scale was administered to the same type of student population. The present study supported these findings with the attitude scores of the three groups on the pretest being quite positive, falling within 4% of the scores obtained by the Sanders study.

There was no significant increase or decrease between pretest and posttest mean attitude scores across all three groups, thus indicating that attitude toward computers did not change because of the treatment. Since the attitude toward computers, according to the pretest, was already quite positive it would be harder to make gains in the positive direction. The only increase from pretest to posttest was by the on-line testing group, but the increase was so small it was statistically insignificant.

A few negative comments were voiced during the focused discussion students indicating some negativism toward both methods of computer testing and a preference for the lecture delivery method and traditional paper and pencil test evaluation. However, when the groups were polled to see how many felt that a return to the older system was preferable to the self-paced competency-based CML system, only 3 out of 21 were in favor. The general consensus was that although improvements were needed to the CML system, the overwhelming majority preferred it to returning to traditional instruction and testing modes.

Test Times

Differences in the time taken to write tests for each module, either on-line by microcomputer or off-line by computer generated paper and pencil tests, were not statistically significant. There was, however, quite a large range in test completion time for each of the modules by both testing methods. This variation was expected and is generally quite normal under most testing conditions. Only one quarter (25%) of the students thought that the on-line test required less time to complete, 37.5% thought testing was faster by the paper and pencil method, and 37.5% could see no difference in test time for either method. This perception is slightly at odds with the actual test time measurements because, although it was shown there was no significant difference ($p < .05$) in test times between the two methods, an inspection of the means for the Heating Controls module revealed that the on-line test had a mean completion time of 29.64 minutes whereas the primary control group had a time of 40.38 minutes. (The Motor Magnetics module had practically equal test times by each method.)

Data collected during the focused discussion (Appendix F) revealed that the students felt that the review feature on the on-line testing program was time consuming. They said the review feature was essential, but it should be restructured to be more time efficient. For example, if only one question needed to be changed part way through the test, to make that change the review function required that a person cycle through the test from question one to the point where the change was required. This kind of "dead" cycling time may give the perception of taking longer to complete a test than actually is the case. This line of discussion led to several suggestions for modifying the review feature to improve its efficiency. One recommendation was to give the program the ability to scroll back and forth between questions and make changes at any time the test was in progress. Another suggestion was to build a feature to tag questions and be able to jump back to them when needed. Yet another suggestion was to give the program a feature to permit jumping to a specific test question by just typing in its number. Some of these suggestions were also mentioned by instructors (Appendix E).

With respect to the off-line testing method some students expressed the view (Appendix C) that congestion sometimes occurred when trying to access the hardcopy terminals to enter answers to test questions.

The times required to complete tests were recorded automatically by the CBTS-CML system. In addition, CBTS could set maximum test limits to which students had to adher or they would be locked out. These time limits helped provide some measure of automatic control with respect to motivating students to make efficient use of their time during the testing process, and at the same time eliminate one more task an instructor would have to perform. The on-line microcomputer program, on the other hand, had no such feature, and testing times were recorded manually. However, this is a feature that should be able to be built into a microcomputer testing program today without much difficulty.

Cheating

Madaule (1988) stated that an on-line testing system "must offer as much security as traditional exams, written with paper" (p. 663). If this statement is accepted as the baseline for any summative evaluation used to judge progress or decide promotion, then we must also accept that there must be instructor supervision. Madaule (1988) shared this view and elaborated, "this means that in no case could the test be taken in a self-service modality:

actual human supervision is required, exactly as usual" (p.669). The electrical program at NAIT has a special enclosed room for supervising tests. Similarly, if one were completing summative tests on microcomputer, supervision would be required. Responses from instructors (Appendix E) confirm the belief that adequate supervision is the main way to prevent cheating. At other times when students are in the formative evaluation mode they need not be supervised.

Students perceived that attempts at cheating would be no more prevalent when on-line tests were being written than when paper and pencil tests were being written. In fact, they believed it would be easier to cheat on the paper and pencil variety (Appendix F). Students believed that one of the biggest deterrents to cheating, not only during the writing of tests but also outside by the sharing of answers, was the use of large test item banks from which unique tests are generated each time. Instructors seem to agree with these statements, in fact many of the instructors' responses closely parallel the responses of the students. Compare responses in Appendices C, E, and F.

Thus it appears that, apart from such features as the use of passwords, denying access to certain modules, lockouts, etc., the two main techniques which have the most influence in reducing cheating are 1) the building of large test item banks capable of creating a large number of parallel unique tests, and 2) an efficient test supervision system.

The number of microcomputers used in the present study was small. If one were to realistically use on-line microcomputer testing to operate a CML program, many more machines would be required. Under these conditions more attention would have to be paid to the physical arrangement of the work stations so as to reduce the chance of cheating. In addition, overall security of a system is probably lessened when floppy disks are involved because of the potential for the outright stealing or copying of disks. Thus, in this respect, a floppy disk based system may require increased vigilence on the part of instructors with respect to hardware and software security.

Two Issues

Two separate yet closely connected issues regarding on-line testing kept arising throughout the study. a) The first one concerned the feasibility, from an evaluation standpoint, of writing tests on-line by microcomputer as compared to off-line by the paper and pencil method; and it was the principal focus of this study. b) The other issue that kept surfacing was the economic, logistical, and practical feasibility of providing a full CML system on-line by microcomputer.

a) With respect to the first issue, the statistical data indicates that it is feasible to use on-line testing methods without negatively impacting on student achievement and test time. However, because of the shortness of the study and because tests were mainly of the recognition/ objective type there may be other factors that could impinge upon this finding. Two such concerns seem to stand out as one looks at the responses to questions on student and instructor questionnaires and focused discussions (see Appendices C, E, and F). The first concern was the perception by both students and instructors that answering questions involving mathematical calculations, formulas, and problem solving would be hampered more when tests were taken on-line than by paper and pencil methods. The second was the big advantage of paper and pencil tests for overviewing, reviewing, and comparing answers.

One cannot deny the existence of these advantages. All that can be done is to see what changes could make the on-line test "more" equivalent. For questions involving calculations, the major recommendation by students (see Appendix F) was that, although more inconvenient than paper and pencil tests, a suitable solution may be a bigger work station to provide a space to work with paper and pencil before entering the answers on the screen.

With respect to the advantage of being able to view the whole paper and pencil test, student suggestions (see Appendix F) were mainly related to making the test as much like the paper and pencil tests as possible. They viewed the features on the QUIZMASTER program which permit reviewing and changing of answers as a big step in the right direction. They made further suggestions to speed up the process for moving from one test question to another. Another suggestion was to use a split screen to display two questions on the screen at once. Others suggested that printed output tests, when needed for consultation, were all that was required. Thus, apart from these concerns, it appears that from an evaluation standpoint on-line tests are feasible. A large majority, of the instructors, recognizing the importance and benefit of a computer-managed learning system to the success of a self-paced mastery learning environment said that they would consider trying on-line testing by microcomputer, if they did not have access to the CBTS-CML system.

b) With respect to the second issue, the economic, logistical, and practical feasibility of on-line testing, the answer is not clear. The limited size and comprehensiveness of the study was not sufficient to extrapolate any conclusions from the data collected. However, from personal observations and responses from instructors, a few guarded remarks are presented.

One of the possible problems with on-line tests is the amount of time per student that it takes to write exams. One machine is tied up by a student for the full length of time it takes to write the test. If students are writing both formative self-check module tests as well as summative tests by this method, it would require many work stations because of the demands of self-paced instruction. In contrast to this, the hardcopy terminals can quickly print paper and pencil tests so that each terminal can handle several students. But even then, these terminals became backlogged when a lot of students are accessing them. Thus with a full blown system working with microcomputers the machine to students ratio is a question that would have to be resolved. Maybe the microcomputer would have a much better chance in an independent class with a limited number of students rather than in a total integrated, centralized setup where the demands for machines can fluctuate from light to extremely heavy.

A microcomputer network may do better. But the question about whether a microcomputer network can handle on-line testing efficiently enough to prevent undue student delays in a self-paced program is something that needs to be tested before one can answer definitively. Possibly some hybrid system consisting of on-line testing for formative evaluation to take advantage of immediate feedback, and off-line tests scored by optical scanner for the summative evaluation might be a possibility worth considering.

Microcomputer Program Trade-offs

The selection of microcomputer software to setup a CML program in an institution or classroom depends, not only upon the features one required, but also upon the features that can be done without if the particular program selected does not have them. It is not usually a question of all or none, but one of compromise. For example, one might wish to have a forty column screen to avoid eye fatigue. QUIZMASTER had this screen format, and an overwhelming number of students preferred it to the printed test from the hardcopy terminal, just the reverse of the findings of most studies just a few years ago. QUIZMASTER uses floppy disks; other programs are available which use hard drives. Some programs, like QUIZMASTER, offer a variety of features that give students control over the testing process; others give little control, offering only immediate feedback which produces a lot of anxiety. Yet, at the same time, the program with the limited features may have a large amount of memory for test banking; whereas the more flexible program may only be able to offer tests of limited size. All things being equal, one should opt for a program that will give maximum flexibility to the student in the test situation and has the ability to print out hardcopies of any student's test when required. (See Appendix F)

The successful implementation of any testing program is contingent on several factors. First, it must be educationally sound so that it does not jeopardize the students' rate of progress or produce an unduly high stressful testing environment. A high proportion of students and staff should see the testing program as an acceptable method of evaluation to avoid any undermining from taking place before the methodology has proven itself. It must be affordable, in the short term, and economical in the long term. In addition, the logistics dealing with changes in staff load and the ability to handle efficiently the required numbers of students cannot be overlooked.

Suggestions for Further Research

Because of the limited scope of this study with respect to the number of modules tested, the number of microcomputers available was sufficient. However, further research is needed to find an acceptable machine to student ratio when testing for a complete self-paced competency-based course is set up on microcomputer such that machine shortages will not impact negatively on student progress rates. An investigation could also be made to see if hard drive systems can be an effective medium for the delivery of on-line testing in self-paced environments with a view to finding out the maximum number of students that a typical network could handle without undue delays that would impede student progress.

Another area of research that may be useful to pursue, according to the concerns expressed by students and instructors, is to investigate differences in achievement levels on tests involving mathematical calculations, formulas, and problem solving; and to experiment with alternative methods to overcome any differences.

This study tried to make the tests on-line by micro-computer as equivalent as possible to paper and pencil tests. The study, therefore, did not explore the potential capabilities, like graphics, animation, adaptive testing, etc., that microcomputers may be able to bring to the testing environment. Opportunities for further research in this area are many. With the technological advances in computers, the prospect for on-line testing may just be beyond the horizon. As a result educators should keep an open mind and explore the potential benefits that could accrue from it.

REFERENCES

Alessi, S. M., & Trollip, S. R. (1985). Computer-based instruction: Methods and development. Englewood Cliffs, NJ: Prentice-Hall.

Alpert, R., & Haber, R. M. (1960). Anxiety in academic achievement situations. Journal of Abnormal and Social Psychology, 61(2), 207-215.

Ary, D., Jacobs, L. C., & Razavieh, A. (1985). Introduction to research (3rd ed.). New York: CBS College Publishing.

Baker, F. B. (1978). Computer managed instruction: Theory and practice. Englewood Cliffs, NJ: Educational Technology Publications.

Bangert-Drowns, R. L., Kulik, J. A., & Kulik, C. C. (1985). Effectiveness of computer-based education in secondary schools. Journal of Computer-Based Instruction, 12(3), 59-68.

Bannon, S. H., Marshall, J. C., & Fluegal, S. (1985). Cognitive and affective computer scales: A validity study. Educational and Psychological Measurement, 45(3), 679-681.

Bear, G. G., Richards, H. C., & Lancaster, P. (1987). Attitudes toward computers: Validation of a computer attitudes scale. Journal of Educational Computing Research, 3(2), 207-218.

Beech, G. (Ed.). (1986). Interactive learning on the IBM-PC. Winslow, England: Sigma Press.

Blank, W. E. (1987). A statewide system for competency-based instruction. Journal of Industrial Teacher Eucation, 24(4), 36-45.

Blank, W. E. (1985). Computerized Data Management System for Competency-Based Vocational Education. (Project Report). St. Petersburg Vocational Institute, St. Petersburg, FL. (ERIC Document Reproduction Service No. ED 268 228)

Bluhm, H. P. (1987). Administrative uses of computers in the schools, Englewood Cliffs, NJ: Prentice-Hall.

Bluhm, H. P. (1987). Computer-managed instruction: A useful tool for educators? Educational Technology, 27(1), 7-12.

Bork, A. (1986). Mastery learning with computers. Proceedings of 1986 National Education Computing Conference, (pp. 333-336).

Bork, A. (1985). Personal computers for education. New York: Harper & Row.

Bozeman, W. C. (1985). Computers and computing education: An introduction. Scottsdale, AZ: Gorsuch Scarisbrick Publishers.

Bozeman, W. C. (1979). Computer-managed instruction: State of the art. AEDS Journal, 12(2), 117-137.

Brzezinski, E. J. (1984). Microcomputers and Testing: Where Are We and How Did We Get There? Educational Measurement: Issues and Practice, 3(2), 7-10.

Cambre, M. A., & Cook, D. L. (1987). Measurement and remediation of computer anxiety. Journal of Educational Technology, 27(11), 15-20.

Crocker, L., & Schmitt, A. (1987). Improving multiple-choice test performance for examinees with different levels of test anxiety. Journal of Experimental Education, 55(4), 201-205.

Dambrot, F, H., Watkins-Malik, M. A., Silling, S. M., Marshall, R. S., & Garver, J. O. (1985). Correlates of sex differences in attitudes toward and involvement with computers. Journal of Vocational Behavior, 27(1), 71-86.

Dick, W. (1981). Instructional design models: Future trends and issues. Journal of Educational Technology, 21(7), p. 29-32.

Dickey, E., & Kherlopian R. (1987). A survey of teachers of mathematics, science, and computers on the uses of computers in grades 5-9 classrooms. Journal of Educational Technology, 27(6), 10-14.

Dunkleberger, G. E., & Heikkinen, H. (1983). Computer-made tests set mastery learning free. Curriculum Review, 22(5).

Dunkleberger, G. E., & Heikkinen, H. (1982). A review of Computer-generated repeatable testing. AEDS Journal, 15(4), 218-225.

Finch, C. R. (1987). Instructional systems development. Journal of Industrial Teacher Education, 24(4), 18-2 Finch, C. R. (1987). Alternate delivery systems for vocational instruction. Journal of Industrial Teacher Education, 24(4), 5-6.

Fishbein, M., & Ajzen, I. (1975). Belief, attitude, intention, and behavior. Reading, MA: Addison-Wesley Publishing company.

Fretwell, D. H. (1987). Challenges to implementing competency based vocational programs in developing countries. Journals of Industrial Teacher Education, 24(4), 47-51.

Gagne, R. M. (1965). The Conditions of Learning (First Edition). New York: Holt, Rinehart, and Winston.

Gery, G. (1987). Making CBT happen. Boston, MA: Weingarten Publications.

Gronlund, N. E. (1982). Constructing Achievement Tests. Englewood Cliffs, N.J.: Prentice-Hall.

Gwinn, J. F., & Beal, L. F. (1988). On-line computer testing: Implementation and endorsement. Journal of Educational Technology Systems, 16(3), 239-251.

Hambleton, R. K. (1984). Using microcomputers to develop tests. Educational Measurement: Issues and Practice, 3(2), 10-14.

Hambleton, R. K., Anderson, G. E., & Murray, L. N. (1983, September). Applications of microcomputers to classroom testing. New Directions for Testing and Measurement, (19).

Harper, F. B. W. (1974). The comparative validity of the Mandler-Sarason test anxiety questionnaire and the achievement anxiety test. Education and Psychological Measurement, 34(4), 961-966.

Hedl, J. J., O'Neil, H. F., & Hansen, D. N. (1973). Affective reactions toward computer-based intellignece testing. Journal of Consulting and Clinical Psychology, 40(2), 217-222.

Jelden, D. L. (1987). CMI test item presentation/feedback and its effect on final examination performance: Staff study. Journal of Educational Technology Systems, 16(2), 99-109.

Johnson, J. H. & Johmson, K. N. (1981). Psychological considerations related to the development of computerized testing stations. Behavior Research Methods and Instrumentation, 13, 421-424.

Katz, D. (1967). The functional approach to the study of attitudes. In M. Fishbein, (Ed.). Readings in Attitude Theory and Measurement. New York: John Wiley & Sons, Inc.

Krech, D., Crutchfield, R. S., & Ballachey, E. L. (1962). Individual in society: A textbook of social psychology. New York: McGraw-Hill.

Kulik, C. C., Kulik, J. A., & Shwalb, B. J. (1986). The effectiveness of computer-based adult education: a meta-analysis. Journal of Educational Computing Research, 2(2), 235-249.

Kulik, J. A., Kulik, C. C., & Bangert-Drowns, R. L. (1985). Effectiveness of computer-based education in elementary schools. Computers and Human Behavior, 1, 59-74.

Kulik, J. A., Kulik, C. C., & Cohen, P. A. (1980). Effectiveness of computer-based college teaching: a meta-analysis of findings. Review of Educational Research, 50(4), 525-544.

Latta, W. S. (1988). Evaluation of Quizmaster. Unpublished master's thesis, University of Alberta, Edmonton, AB.

Leiblum, M. D. Computer-managed instruction: An explanation and overview. AEDS Journal, 1982, 15(3), 70-72

Llabre, M. M., Clements, N. E., Fitzhugh, K. B., Lancelotta, G., Mazzagatti, R. D., & Quinones, N. (1987). The effect of computer-administered testing on test anxiety and performance. Journal of Educational Computing Research, 3(4), p. 429-433.

Loyd, B H. & Loyd, D. E. (1985). The reliability and validity of an instrument for the assessment of computer attitudes. Educational and Psychological Measurement, 45(4), 903-909.

Loyd, B. H., & Gressard, C. (1984). Reliability and factorial validity of computer attitude scales. Educational and Psychological Measurement, 44(2), 501-505.

Loyd, B. H., & Gressard, C. (1984). The effects of sex, age, and computer experience on computer attitudes. AEDS Journal, 18(2), 67-77.

Madaule, F. (1988). Correcting examinations by computer. In F. Lovis & E. D. Tagg (Ed.), Computers in Education. (pp. 663-668). Amsterdam, The Netherlands: Elsevier Science Publishers.

Marcoulides, G. A. (1988). The relationship between computer anxiety and computer achievement. Educational Computing Research, 4(1), 151-157.

Mager, R. F. (1962). Preparing Insturctional Objectives, Palo Alto, CA: Fearon.

Mandler, G., & Sarason, S. B. (1952). A study of anxiety and learning. Journal of Abnormal and Social Pshchology, 47(2), 166-173.

McIsaac, D. N., & Baker, F. B. (1981). Computer-managed instruction system implementation on a microcomputer. Educational Technology, 21(10), 40-46.

Millman, J. (1984). Using microcomputers to administer tests: An alternate point of view. Educational Measurement: Issues and Practice, 3(2), 20-21.

Moe, K. C., & Johnson, M. F. (1988). Participants' reactions to computerized testing. Educational Computing Research, 4(1), 79-86.

Oregon State University. (1984). Computer skills in vocational education: A methodology for computer skills improvement in vocational secondary programs. Eugene, OR: Author. (ERIC Document Reproduction Service No. ED 259 168)

Osterhouse, R. A. (1972). Desensitization and study skills training as treatment for two types of test anxious students. Journal of Counseling Psychology, 19, 301-307.

Parker, R., & Sands, J. (1985). Instructor Guide for the Lake County Vocational Center Computer-Managed Instructional System for Competancy Based Vocational Education. (ERIC Document Reproduction Service No. ED 260 224)

Popham, W. James (1981) . Modern Educational Measurement. Englewood Cliffs, N.J.: Prentice-Hall.

Pucel, D. J. (1987). The performance-based instructional design system. Journal of Industrial Teacher Education, 24(4), 27-35

Roblyer, M. D., Castine, W. H., & King, F, J. (1988). Assessing the impact of computer-based instruction. Computers in the Schools, 5(3-4), 11-149

Reece, M. J., & Gable, R. K. (1982). The development and validation of a measure of general attitudes toward computers. Educational and Psychological Measurement, 42(4), 913-916.

Romaniuk, E. W., & Lee, D. (1984). Evaluation of the Alberta Correspondence School computer-managed learning project (Research Report). Edmonton, AB: ECR Educational Courseware and Research Inc.

Ruebush, B. K. (1960). Interfering and facilitating effects of test anxiety. Journal of Abnormal and Social Psychology, 60(2), 205-212.

Sanders, J. (1988). The effects of pair students for work on a graphically oriented CAI simulation. Unpublished master's thesis, University of Alberta, Edmonton, AB.

Sarason, I. G. (1978). The test anxiety scale: Concept and research. In C. D. Spielberger & I. G. Sarason (Eds.), Stress and Anxiety (Vol. 5). New York: Hemisphere/ Wiley.

Shaha, S. H. (1984). Matching-tests: reduced anxiety and increased test effectiveness. Journal of Educational and Psychological Measurement, 44(4), p. 869-881.

Simonson, M. R., Maurer, M., Montag-Torardi, M., & Whitaker, M. (1987). Development of a standardized test of computer literacy and a computer anxiety index. Journal of Educational Computing Research, 3(2), 231-247.

Spielberger, C. D., Gonzales, H. P., Taylor, C. J., Algaze B., & Anton, W. D. (1978). Examination stress and test anxiety. In C. D. Spielberger & I. G. Sarason (Eds.), Stress and Anxiety (Vol. 5), New York: Hemisphere/ Wiley.

Spielberger, C. D. (1966). Theory and research on anxiety. In C. D. Spielberger (Ed.), Anxiety and Behavior. New York: Academic Press.

Suzuki, W. N., & Finch, C. R. (1987). Alternate delivery systems: critique and comment. Journal of Industrial Teacher Education, 24(4), 52-55.

Swanson, R. A. (1987). Training technology system: A method of identifying and solving training problems in industry and business. Journal of Industrial Teacher Education, 24(4), 7-17.

Telem, M. (1982). CMI and the MIS--An integration needed. AEDS Journal, 9(2), 48-55.

Thyer, B. A., & Papsdorf, J. D. (1982). Discriminant and concurrent validity of two commonly used measures of test anxiety. Journal of Educational and Psychological Measurement, 42(4), 1197-1205.

Vermette, S. M., Orr, R. R., & Hall, M. H. (1986). Attitudes of Elementary School students and teachers toward computers in education. Journal of Educational Technology, 26(1), 41-47.

Ward, W. C. (1984). Using microcomputers to administer tests. Educational Measurement: Issues and Practice, 3(2), 16-20.

Zahniser, G., Long, J. P., & Nasman, L. O. (1983). Microcomputers in Voc Ed: A Decision Guide. (Research & Development Series No. 239A). Columbus, OH: The National Center for Research in Vocational Education, Ohio State University.

APPENDIX A

COMPARSION OF SAMPLE AND POPULATION

Table 19

<u>Comparison of Sample to Population by Education, Gender, and Age</u>

Demographic Catagories	Provincial Population (N=500)	Sample (N=36)
<u>EDUCATION LEVEL</u>		
< 11	9.1 %	11.1 %
11	14.3 %	8.3 %
> 11	76.6 %	80.6 %
Mode	Grade 12	Grade 12
<u>GENDER</u>		
Male	98.6 %	97.2 %
Female	1.4 %	2.8 %
<u>AGE</u>		
< 25	47.2 %	50.0 %
25 - 29	31.8 %	36.1 %
> 29	21.0 %	13.9 %
Mean (yr.)	26.15	25.52
Range (yr.)	19 - 51	20 - 37
Mode (yr.)	[illegible]	24

Note: Using the categories listed it appears that the sample is approximately representative of the population, the second year electrical apprentices of Alberta.

APPENDIX B

DATA COLLECTION INSTRUMENTS

INTRODUCTION

May 1, 1989

Dear Second-year Electrical Apprenticeship Students:

My name is Roland Rumbolt, and I am a graduate student in Industrial and Vocational Education at the University of Alberta. This study is being undertaken in partial fulfillment of my course requirements. NAIT, NAIT instructional staff, and the Alberta Apprenticeship Board approve of the study and are interested in the results. The purpose of the study is to investigate the potential for using microcomputers in the delivery of computer-managed learning (CML) programs.

Your participation in this study would be greatly appreciated and is important to the success of the research. The research addresses the feasibility of taking exams on-line on the microcomputer as compared to off-line by the paper and pencil method using the hardcopy terminals, and explores student attitudes and opinions regarding these methods. During the study some students, selected at random, take one exam by microcomputer; the remainder of the students will be tested in the usual manner. In addition, each student will fill out three questionnaires over the course of the 8 week block of training.

The information collected will be used in a statistical survey and will be kept strictly confidential. To ensure confidentiality all questionnaires and interview materials will be kept in a secure location and will be available only to the researcher and his thesis committee, none of whom have any connection with your course, or are in a position to evaluate your performance as a student. Your anonymity will be strictly ensured and your involvement is voluntary. To ensure anonymity I.D. numbers will be used to match your completed questionnaires, interviews, and tests. Then once the data is compiled, responses will be analyzed in summary form and will not be identifiable with any particular person.

I appreciate your cooperation and thank you in advance for your time.

Sincerely yours,

R. Rumbolt

Roland Rumbolt

COMPUTER-MANAGED LEARNING SURVEY

This questionnaire was designed to find out how second-year electrical apprentices feel about computers and the different forms of computerized testing.

Your participation in this survey would be greatly appreciated and is considered important. The information collected will be used in a statistical study and will be kept strictly confidential. Your anonymity is guaranteed and your involvement is voluntary. Again, I would like to thank you for your cooperation.

Please ignore the numbers in the right margin.

My student I.D. number is : ☐☐☐☐☐

Do not write in this space

1 2 3 4 5

PART 1: BACKGROUND INFORMATION

Please complete the following statements either by filling in the blank spaces or by circling the appropriate response. 1/6

Demographics

1. My year of birth is: 1 9 ☐☐ 8-9

2. My sex is : 1. Female 2. Male 11

3. The highest school grade I completed was : 7 8 9 10 11 12 13-14

4. I completed this grade by : 15
   1. Regular Day School
   2. Adult Upgrading
   3. Correspondence School
   4. Grade Equivalent Diploma (GED)

5. In courses other than electrical apprenticeship, the number of years of training beyond grade 12 that I have completed is: 16
   1. Zero years
   2. One year
   3. Two years
   4. Three years
   5. Four years

Computer Experience

6. Did you take your first-year apprenticeship training using the computer-managed learning system at NAIT ?	1. No	2. Yes	21
7. If the answer to Question #6 is "No", were computers used to administer the tests and program at the institution where you received your apprenticeship training ?	1. No	2. Yes	22
8. Do you presently own a microcomputer ?	1. No	2. Yes	23
9. Have you used microcomputers to play games ?	1. No	2. Yes	24
10. Have you ever used microcomputers before ?	1. No	2. Yes	25
11. Have you had to use computers at your worksite ?	1. No	2. Yes	26
12. Have you ever taken any formal courses, for example, basic programming, wordprocessing, spreadsheeting, etc., on microcomputers at high school or some other institution ?	1. No	2. Yes	27

PART 2: ATTITUDE SURVEY

Please circle the number which expresses the extent to which you agree with each statement.

Scale to use: STRONGLY AGREE = 1, AGREE = 2, UNSURE = 3, DISAGREE = 4, STRONGLY DISAGREE = 5

(Your answers are neither right nor wrong. Please read each question carefully, but do not spend too much time on any one question.)

	STRONGLY AGREE	AGREE	UNSURE	DISAGREE	STRONGLY DISAGREE	
1. A computer could make learning more fun for me.	1	2	3	4	5	30
2. Given a little time and training, anybody could learn to use computers.	1	2	3	4	5	31
3. Using a computer could be enjoyable.	1	2	3	4	5	32
4. Computers are so complicated I would rather not use one for learning.	1	2	3	4	5	33
5. Even though computers are valuable and necessary, I still have a fear of them.	1	2	3	4	5	34
6. I feel very negative about computers in general.	1	2	3	4	5	35
7. Computers do not scare me at all.	1	2	3	4	5	36
8. Working with a computer would make me very nervous.	1	2	3	4	5	37
9. I am sure I could do work with computers.	1	2	3	4	5	38
10. I'm not the type to do well with computers.	1	2	3	4	5	39
11. Once I start to work with the computer, I would find it hard to stop.	1	2	3	4	5	40
12. I will do as little work with the computer as possible.	1	2	3	4	5	41

OMISSION OF COPYRIGHTED MATERIAL

The following pages have been removed from this thesis:

pp. 113 - 114.

These pages contained copies of I. G. Sarason's Test Anxiety Scale (TAS), which was used during the study to measure pre and post test anxiety.

The source of this instrument was:

Sarason, I. G. (1978). Test anxiety scale: Concept and research. In C. D. Spielberger and I. G. Sarson (Eds.), Stress and Anxiety (Vol. 5). Washington: Hemisphere/Wiley.

COMPUTER-MANAGED LEARNING SURVEY

Questionnaire 2

This questionnaire was designed to find out how second-year electrical apprentices feel about computers and the different forms of computerized testing.

Your participation in this survey would be greatly appreciated and is considered important. The information collected will be used in a statistical study and will be kept strictly confidential. Your anonymity is guaranteed and your involvement is voluntary. Again, I would like to thank you for your cooperation.

My student I.D. number is : |__|__|__|__|__|

Do not write in this space

1 2 3 4 5

PART 2: ATTITUDE SURVEY

Please circle the number which expresses the extent to which you agree with each statement.

Scale to use: STRONGLY AGREE = 1, AGREE = 2, UNSURE = 3, DISAGREE = 4, STRONGLY DISAGREE = 5

(Your answers are neither right nor wrong. Please read each question carefully, but do not spend too much time on any one question.)

2/6

Statement	STRONGLY AGREE	AGREE	UNSURE	DISAGREE	STRONGLY DISAGREE	
1. A computer could make learning more fun for me.	1	2	3	4	5	30
2. Given a little time and training, anybody could learn to use computers.	1	2	3	4	5	31
3. Using a computer could be enjoyable.	1	2	3	4	5	32
4. Computers are so complicated I would rather not use one for learning.	1	2	3	4	5	33
5. Even though computers are valuable and necessary, I still have a fear of them.	1	2	3	4	5	34
6. I feel very negative about computers in general.	1	2	3	4	5	35
7. Computers do not scare me at all.	1	2	3	4	5	36
8. Working with a computer would make me very nervous.	1	2	3	4	5	37
9. I am sure I could do work with computers.	1	2	3	4	5	38
10. I'm not the type to do well with computers.	1	2	3	4	5	39
11. Once I start to work with the computer, I would find it hard to stop.	1	2	3	4	5	40
12. I will do as little work with the computer as possible.	1	2	3	4	5	41

OMISSION OF COPYRIGHTED MATERIAL

The following pages have been removed from this thesis:

pp. 116 - 117.

These pages contained copies of I. G. Sarason's Test Anxiety Scale (TAS), which was used during the study to measure pre and post test anxiety.

The source of this instrument was:

Sarason, I. G. (1978). Test anxiety scale: Concept and research. In C. D. Spielberger and I. G. Sarson (Eds.), Stress and Anxiety (Vol. 5). Washington: Hemisphere/Wiley.

COMPUTER MANAGED LEARNING SURVEY

Questionnaire 3

This questionnaire, the last one in this study, was designed to find out how second year electrical apprentices feel about computers and the different forms of computerized testing.

Your participation in this survey has been greatly appreciated and is considered important. The information collected will be used in a statistical study and will be kept strictly confidential. Your anonymity is guaranteed and your involvement is voluntary. Again, I would like to thank you for your cooperation.

My student I.D. number is : :_:_:_:_:_:

Do not write in this space

1 2 3 4 5

Please complete the following statements by filling in the blank spaces or by checking the appropriate response. A space is provided at the end for any extra comments or suggestions, if you have any.

3/6

1. Which one of the two testing methods did you prefer? — 8
   1. Computer generated paper and pencil test. 1. :_:
   2. On-screen by microcomputer. 2. :_:
   3. About equal preference for each. 3. :_:

2. Which method do you feel required less time to complete and correct an exam? — 9
   1. Computer generated paper and pencil exams. 1. :_:
   2. On-screen by microcomputer. 2. :_:
   3. About equal time for each. 3. :_:

3. Which testing method did you find easiest to use? — 10
   1. Computer generated paper and pencil exams. 1. :_:
   2. On-screen by microcomputer. 2. :_:
   3. About as equally easy by each method. 3. :_:

4. Under which testing method do you feel it would be easier to cheat? — 11
   1. Computer generated paper and pencil exams. 1. :_:
   2. On screen by microcomputer. 2. :_:
   3. About equally easy under each method. 3. :_:

5. How would you compare the convenience of <u>handling</u> floppy disks in order to take a test as compared to using the hard copy terminal to print and score a paper and pencil exam? — 12
   1. Floppies are more convenient to use than the hard copy terminals. 1. :_:
   2. Floppies are less convenient to use than the hard copy terminals. 2. :_:
   3. Both are about the same level of inconvenience. 3. :_:

6. Which test marking method would you prefer? 13

   1. Have immediate marking of each question as it is completed. 1. :__:

   2. Have the marking delayed until all questions are completed and reviewed. 2. :__:

7. Which of the two testing methods do you think you would prefer if the test questions involved calculation problems? 14

   1. Computer generated paper and pencil test. 1. :__:

   2. On-screen by microcomputer. 2. :__:

   3. About equal preference for each. 3. :__:

   4. Not sure. 4. :__:

8. What do you see as the biggest advantage or like best about taking tests by the usual computer generated paper and pencil exam?

   ____
   ____

9. What do you see as the biggest disadvantage or like least about taking tests by the usual computer generated paper and pencil exam?

   ____
   ____

10. What do you see as the biggest advantage or like best about taking tests on-sceen by microcomputer?

   ____
   ____

11. What do you see as the biggest disadvantage or like least about taking tests on-screen by microcomputer?

   ____
   ____

<u>Comments and suggestions</u>: (Use back of questionnaire also, if necessary.)

____
____
____
____
____
____
____
____
____
____
____
____
____

COMPUTER-MANAGED LEARNING SURVEY

Questionnaire 4

This questionnaire was designed to find out how second-year electrical apprentices feel about computers and the different forms of computerized testing.

Your participation in this survey would be greatly appreciated and is considered important. The information collected will be used in a statistical study and will be kept strictly confidential. Your anonymity is guaranteed and your involvement is voluntary. Again, I would like to thank you for your cooperation.

My student I.D. number is : |_|_|_|_|_|

Do not write in this space

1 2 3 4 5

1. How did you feel when you learned that you did not have to take exams on-line on the microcomputer? 27

   (1) pleased |_|

   (2) disappointed |_|

   (3) neither pleased nor disappointed |_|

2. Why? ____

____

____

____

COMPUTER MANAGED LEARNING SURVEY

Questionnaire 5

Please complete the following questionnaire and seal it in the accompanying envelope. Individuals need not identify themselves. Your input is valuable and is considered important to the success of the study. Any and all responses will be much appreciated. Thank you.

1. What kinds of feedback did you receive (or overhear) from the students regarding their feelings about on-line testing, its acceptance, rejection, problems with it, etc.?

2. From your perspective as an instructor,
   a) What do you see as the major advantages of a paper & pencil testing system over an on-line testing system?

   b) What do you see as the major advantages of on-line testing (by micro or otherwise)?

3. a) What do you see as some of the weaknesses/disadvantages/places where improvements could be made in a computer generated paper and pencil test system such as you now have?

   b) What do you see as some of the weaknesses/disadvantages of testing students on-line by microcomputer?

4. What improvements could be made to an on-line system that would make it more acceptable?

5. If you did not have access to the CBTS-CML system but were faced with an open entry-open exit mastery learning program, do you feel that on-line testing by microcomputer may be a worth trying?

6. a) Which system do you feel provides greater opportunities for cheating?

on-line by micro |__| paper & pencil exam |__| about equal |__|

b) Why? ____
____
____
____

7. How important are the following features in any testing program?

	very important	somewhat important	low importance
a) Provision to insure that test information will not be lost after student starts writing exam.	____	____	____
b) Provision to change and review answers to questions as many times as required within testing time limit.	____	____	____
c) Provision for student to obtain feedback on answers to all questions requested.	____	____	____

8. Please list what you consider the order of importance of the statements in question #7 to be.
____

FOCUSED DISCUSSISON QUESTIONS

The students in the experimental and the main control group will be assembled in groups of 7 or 8 and will be asked, as a group, to respond to the following questions. A consistent presentation format will be followed with each group, and all students will be encouraged to add to the discussion.

Question #1.

a) Do you think that on-line microcomputer testing was a suitable substitute or viable (realistic) alternative for the computer produced paper and pencil tests?

b) For 'Yes' answers ==============> Why?

c) For 'No' answers ===============> Why not?

Question #2.

(If the responses to question #1 are primarily negative, then this question will be asked.)
What are some of the modifications that could be made to make on-line testing a more suitable alternative?

Question #3.

Do you think the amount of practice time on microcomputer before taking the exam on-line was sufficient?

Question #4.

a) How important to you are printed paper copies of supervised exams for the purpose of reviewing them with instructors.

b) When using on-line testing, do you think there would be any need to pass out, temporarily for review purposes, a printed paper copy of the exam if each incorrect test answer had an explanation on-screen of why the answer chosen was incorrect and an explanation of the correct answer?

Question #5.

a) Do the features: the ability to skip questions and have them recycled at the end, the ability to change answers freely, and the ability to review all questions and change answers b fore the exam is marked, cause the on-line test to closely simulate a paper and pencil exam?

b) What other features might improve it even more?

Question #6.

One of the questions on the questionnaire asked about the ease of cheating.

a) What are some of the subtle ways that a person could cheat or beat the system, when being tested:

i) on-line by microcomputer?

ii) computer generated paper and pencil test?

b) (Ask this question if any methods of cheating are suggested) What changes could be made to eliminate, or reduce them?

Question #7.

a) How would you react to being tested on-line if the test questions were problems in which calculations were required?

b) (Ask this question if the answer to (a) is negative.) Do you have any suggestions that might help provide a solution to this problem?

Question #8.

What would be your reaction if all of your supervised exams were on-line tests using floppy disks and microcomputer as compared to:

i) computer generated off-line paper and pencil exams.

ii) regular paper and pencil exams.

APPENDIX C

STUDENT RESPONSES TO QUESTIONNAIRE #3

Responses to Student Questionnaire #3

Part of the data obtained from the students by this questionnaire was collected by open-ended questions. The responses to these questions are summarized below.

Question 8: What do you see as the biggest advantage or like best about taking tests by the usual computer generated paper and pencil exam?

Responses:

1. You have pencil (in your hand) for calculations.
2. It makes it easy to look over all the questions and to preview whole exam.
3. You can set your own pace.
4. It is convenient for looking back at questions.
5. You don't have to look up at the screen.
6. All questions are available at a glance.
7. None.
8. It provides proof of calculations.
9. I am used to paper and pencil tests.
10. You can underline key words.
11. Scanning is important and easy for making appropriate corrections.
12. It is easier to review.
13. It provides a hardcopy of the exam.
14. You can compare your answers.

15. Nice to be able to look at questions and write right on the paper.
16. Only an advantage for calculations.

Question 9: What do you see as the biggest disadvantage or like least about taking tests by the usual computer generated paper and pencil exam?

Responses:

1. A lot of paper to contend with.
2. There is more writing involved.
3. The computer does not allow for human error.
4. It takes too long to punch in answers and get the results.
5. It is easy to skip (miss) questions.
6. No disadvantages.
7. If you are mentally tired you could misread a question.
8. You can't check your answers.
9. Small wording may cause misinterpretation of questions.
10. Sometimes you enter the wrong answers and don't realize it.
11. Greater chance for errors.
12. Takes too long to mark.
13. Too much noise from computer.
14. Allows for cheating at the entry terminal.
15. Having to punch in answers on the hardcopy

terminal after finishing exam.

Question 10: <u>What do you see as the biggest advantage or like best about taking tests on-screen by microcomputer?</u>

Responses:

1. Questions are easy to read.
2. There is less writing involved.
3. Immediate feedback at end of exam.
4. It marks test for you and provides correct answers.
5. No paper involved and it is corrected faster.
6. The question is there and there is no chance to miss anything.
7. There are no particular advantages.
8. You can't miss a question. The clarity of questions on-screen.
9. There are no advantages.
10. Did not like on-screen testing.
11. It has potential.
12. The answers that you enter are easily seen.
13. One question at a time makes it easier to concentrate.
14. You are provided with the answers when you finish the test.
15. You only have to key-in the answer once.
16. It is harder to cheat.

Question 11: <u>What do you see as the biggest disadvantage or like least about taking tests on-screen by microcomputer?</u>

Responses:
1. You can't see the question once it is skipped.
2. The use of floppy disks.
3. The screen was hard on the eyes.
4. Punching in the answers.
5. There were no part marks given.
6. It was difficult looking up at the screen while doing calculations.
7. The chance of having a floppy disk.
8. You can't see all the questions at the same time.
9. Having to make calculations.
10. No disadvantages
11. It felt uncomfortable and you couldn't get a feel for the questions.
12. You can't scan, thus causing a mental block.
13. The time taken to review answers is too long.
14. There is no room to calculate and you can't review answers, until questions are completed, if an answer to a previous question comes to mind.
15. It reviews all answers instead of going one at a time.

16. You can't write on it.

17. You had to wait too long for test marking.

Other Comments and Suggestions

Responses:

1. With paper and pencil tests you can get hints for an answer later in the test.

2. Microcomputer is better.

3. No difference.

4. If you can see all the questions at the same time you can gain some knowledge pertaining to the present problem.

5. Bigger tables needed for working on when using microcomputer.

6. For calculataions and formulas, circuits diagram could be provided on-screen into which values could be substituted so that predictions may be made.

7. Both the same.

8. Good if whole course were done on microcomputer but I didn't feel prepared.

9. Prefer microcomputer during course but paper and pencil for exams.

10. Easier to review.

11. Needs to be able to review one question at a time.

12. Preferred to use paper and pencil.

13. Less time to correct but longer to do the exam.

14. Printer attached to microcomputer so test questions could be printed out as you do test on screen.

15. Feared for low mark -- keep your ideas in your own school.

16. When test is worth 33.3 % of mark different computers should not be used.

APPENDIX D

STUDENT RESPONSES TO QUESTIONNAIRE #4

Responses to Student Questionnaire #4

The responses to this short questionnaire were elicited from the secondary control group to try to obtain some perception of how they felt about being left out of the main part of the study.

Question 1: How did you feel when you learned that you did not have to take exams on-line on the microcomputer?

Responses:	1.	pleased	(5 out of 12)	42%
	2.	disappointed	(1 out of 12)	8%
	3.	neither pleased nor disappointed	(6 out of 12)	50%

Question 2: Why?

Responses:
1. Didn't want to use the microcomputer because I was too involved in other school work.
2. Don't like computers.
3. Prefer test paper.
4. Wanted to use the microcomputer.
5. Need to work answers out on paper.
6. Could skip questions and go back.
7. No difference.

APPENDIX E

INSTRUCTOR RESPONSES TO QUESTIONNAIRE #5

Instructor Responses to Questionnaire #5

When interpreting these responses the reader should keep in mind that some of the instructors, because of work demands, did not have a lot of exposure to and involvement with the on-line testing project; thus part of the responses are based on professional judgement and other experiences related to computer testing.

Question 1: <u>What kinds of feedback did you receive (or overhear) from the students regarding their feelings about on-line testing, its acceptance, rejection, problems with it, etc.</u>?

Responses:
1. None (not directly involved).
2. I did not solicit comments, the only thing I heard was negative. They seemed to be more used to the paper and pencil tests.
3. Did not hea_ any comments for or against.
4. Some students stated it was more difficult than using the printed exam. On the printed exam they could look at all questions at once, therefore, answering or reading questions often helped him answer them.
5. Most had no complaints. A few were upset because they got a bad mark.
6. Very little negative feedback. Only one student seemed uncomfortable with the

on-line testing.

7. Some of the students thought it would be nice to have a hardcopy to review with an instructor after the exam.

8. Some felt it was optional and different and didn't care for it. We didn't hear too many compliments.

9. Some did not like it because it was new and unfamiliar. Most liked it and the quickness that they received their mark. Some felt that cheating by discussing tests with those who had already written them was very common.

Question 2: From your perspective as an instructor,

a) What do you see as the major advantages of a paper and pencil testing system over an on-line testing system?

Responses: 1. Time to examine all questions (overview) and to be able to look back at specific questions.

2. The student's freedom to move back and forth between items.

3. Able to concentrate on question - found computer screen distracting - room on desk or table for code book - calculation - paper on desk for diagram drawing to assist thinking process.

4. When the student wants to review his exam and ask the instructor questions paper and pencil

testing is much easier.

5. For mathematical problems all the work can be shown and errors spotted.
6. The student has a hardcopy for review.
7. Can view and review answers easily - can go back and see what mistakes were made after the fact.
8. The instructor has a better knowledge of the students and can more easily diagnose problem areas - somewhat more personal contact with the students.

b) What do you see as the major advantages of on-line testing (by micro or otherwise)?

Responses:
1. Micro low initial cost; otherwise - lower paper cost.
2. Question is displayed on screen when entering answers - opportunity to compare answer to question.
3. No paper, therefore less expensive.
4. Tests are marked immediately and recorded.
5. The traditional room and time, could be replaced if a more flexible system that is better at accommodating the students.
6. The student receives immediate feedback in regards to exam results.
7. No paper is needed unless it is required by the student - 40 character lines are very

legible and convenient - statistics and analysis.

8. No subjectivity by instructor in grading and the student receives feedback very quickly while the exam is still fresh in his mind.

Question 3: a) What do you see as some of the weaknesses/disadvantages/places where improvements could be made in a computer generated paper and pencil test system such as you now have?

Responses:

1. More questions and better reference to objectives and modules for questions wrong.
2. Graphics would be nice in some situations.
3. Tendency to put answer in wrong order - eg. question prompt should spell out question not just #2. a.
4. Some questions can be worth 13% of total test. This system does not allow for partial marks.
5. It is difficult to supervise and prevent the possibility of students sharing information.
6. The lack of questions in some areas. To create a larger number of unique exams more questions must be developed.
7. Graphics with questions.
8. A better controlled exam environment such as audio supervision of the exam room to prevent conversation, and separate terminals

for supervised or recorded exams, also do not look at enough of students' work/diagrams.

b) <u>What do you see as some of the weaknesses/disadvantages of testing students on-line by microcomputer</u>?

Responses: 1. Emotion, no overview of tests, not able to look back at questions.

2. Unable to scan the entire test with ease.

3. Screen is distracting - should be able to move to a desk to work out problems.

4. The inconvenience of trying to review questions with the students.

5. Hardware costs.

6. The students must be totally dependent on the computer, and a power outage could create problems.

7. Students spend more time on terminals so more terminals needed.

8. The lack of "seeing" the students actual methods of problem solving. I have seen many students arrive at a correct answer even though their reasoning was faulty. Very good for fast feedback and testing of straight "recall" material.

Question 4: <u>What improvements could be made to an on-line system that would make it more acceptable</u>?

Responses: 1. Scrolling back and forth.

2. Print questions on paper - allow students to leave terminal to problem solve - return to terminal to enter answers.

3. The students could stand a longer "warm-up" time. (Introduction to new microcomputers)

4. Diagnostics for questions missed.

5. Some method of "hardcopy" retention of exam material so student could see what he has done as a complete package.

Question 5: <u>If you did not have access to the CBTS-CML system but were faced with an open entry-open exit mastery learning program, do you feel that on-line testing by microcomputer may be worth trying</u>?

Responses: 1. Yes.

2. Yes.

3. Yes.

4. If CBTS was not available I feel that the micro testing would be excellent. The same amount of disagreement between students and the micros could have been stopped, if they had been exposed to it exclusively.

5. Yes.

6. **Yes.**
7. **I believe the use of microcomputers for a mastery learning program is feasible and economical.**
8. **Certainly, but I would want more software.**
9. **Very definitely I feel the open entry/exit system for mastery learning is the only way to go, whatever system is used for placing/evaluating students.**

Question 6: a) Which system do you feel provides greater opportunities for cheating?

on-line by micro 0/9 (0 %)

paper and pencil exam 4/9 (44.4 %)

about equal 5/9 (55.6 %)

b) Why?

Responses:

1. Gathering of students while entering answers.
2. It appears to depend on the degree of supervision rather than the medium of delivery.
3. Easier to have an extra piece of paper with additional information hidden.
4. The paper and pencil exam is identical for each student, whereas the micro could generate different questions.[Traditional?]
5. The student can only cheat if other students are near the terminal - solution provide more

computer space.

6. Students will cheat to some degree if they are not supervised closely.

7. If you are going to cheat you will do it on any system, also discussing the exam with those who have already written it is all too common.

Question 7: <u>How important are the following features in any testing program</u>?

a) Provision to insure that test information will not be lost after student starts writing exam.

very important	<u>9/10</u>
somewhat important	<u>1/10</u>
low importance	<u>0/10</u>

b) Provision to change and review answers to questions as many times as required within testing time limit.

very important	<u>9/10</u>
somewhat important	<u>1/10</u>
low importance	<u>0/10</u>

c) Provision for student to obtain feedback on answers to all questions requested.

very important	<u>8/10</u>
somewhat important	<u>2/10</u>
low importance	<u>0/10</u>

Question 8: <u>Please list what you consider the order of importance of the statements in Question #7 to be</u>.

Responses: 1. a, c, b

2. a, b, c (but all very important).

3. b, a, c

4. b, a, c

5. a, b, c

6. b, c, a

7. c, a, b

8. c, b, a

APPENDIX F

RESPONSES DURING FOCUSED DISCUSSIONS

Responses During Focused Discussion

Question #1.

a) Do you think that on-line microcomputer testing was a suitable substitute or viable (realistic) alternative for the computer produced paper and pencil tests?

b) For 'Yes' answers ========================> Why?

c) For 'No' answers ==========================> Why not?

Responses:

In the discussion session it was difficult to obtain answers focused on only one question. As a result the responses to one question quite often partially answered another question.

About three quarters of the students said that on-line testing could be a viable alternative, but upon giving the answer most of the "yes" replies had conditions attached of the type:

- - Yes, but keep it to theory and lab type work, not mathematical calculations.

- - Yes, but random selection of questions [by students when writing test] would be better.

- - Yes, but some training needed beforehand.

- - Yes, but forward and reverse review feature needed.

- - Yes, but random or block access needed.

- - Yes, but should be echoed to a printer.

- - Yes, but one should be able to place any two questions on the screen at the same time.

For those who said "No", the main reasons given were:

- - Paper and pencil tests allow you to see any part of the test whenever you wish for review and help with other questions.

- - It would be difficult to handle problems dealing with formulas and calculations.

Question #2.

(If the responses to question #1 are primarily negative, then this question will be asked.) What are some of the modifications that could be made to make on-line testing a more suitable alternative?

Responses:

This question was basically answered in question number one. When it was asked specifically, the responses given to question number one were reiterated.

Question #3.

Do you think the amount of practice time on microcomputer before taking the exam on-line was sufficient?

Responses:

This question was somewhat irrelevant because the original plan of students having approximately one hour practice time for familiarization did not materialize. Rather it turned into a 15 to 20 minute familiarization at the beginning of the study and then another short review of the program functions before students took the test. Most felt that they were familiar enough, after that length of time to handle the program and they knew that someone was near if they ran into trouble.

Question #4.

a) How important to you are printed paper copies of supervised exams for the purpose of reviewing them with instructors?

b) When using on-line testing, do you think there would be any need to pass out, temporarily for review purposes, a printed paper copy of the exam if each incorrect answer had an explanation on-screen of why the answer chosen was incorrect and an explanation of the correct answer?

Responses:

The responses were mixed. The general concensus seemed to be that a hard copy is important for review with instructors, especially for the tests involving calculations. Approximately half felt that if the explanations of incorrect responses were provided onscreen, there would be no need of a hardcopy, except on any occasion it is requested by the student. Many felt that onscreen explanations would help reduce instructor consultation time.

Question #5.

a) Do the features: the ability to skip questions and have them recycled at the end, the ability to change answers freely, and the ability to review all questions and change answers before the exam is marked, cause the on-line test to closely simulate a paper and pencil exam?

b) What other features might improve it even more?

Responses:

The more freedom of movement that a student had to move from point to point in the test, the more it simulates a paper and pencil test. They saw as "musts" the features listed here and made further suggestions to give students more freedom of control. Some comments were:

- - Good to be able to skip and go back - a must.

- - Review - yes, definitely.

- - These features were good, without them things would be ridiculous.

Other suggestions that would make test better:

- - Tagging questions is important.

- - Display two or more questions at one time.

- - Select any question for review by typing its number.

- - Review button would be good.

- - Scroll backward and forward anytime during test.

- - Something to tell you where a certain question is located - something like a rollodex located at one side of the screen.

Question #6.

One of the questions on the questionnaire asked about the ease of cheating.

a) What are some of the subtle ways that a person could cheat or beat the system, when being tested:

i) on-line by microcomputer?

ii) computer generated paper and pencil test?

b) (Asked this question if any methods of cheating are suggested). What changes could be made to eliminate, or reduce them?

Responses:

Students generally talked around the questions. Most felt that little cheating is taking place. They seemed to think that it would be easier to cheat with paper and pencil tests by students exchanging papers and discussing answers when entering responses at hardcopy terminals. The suggested method for cheating with on-line testing were looking over someone's shoulder, copying disks, or deliberate sabotage like unplugging the microcomputer or scratching the disks. But most agreed that because of the random generation of questions for tests, cheating is difficult for either method. Some specific responses include:

- - Question choice makes it difficult to cheat.

- - Cheating would be at a minimum.

- - Cheating is short changing ourselves - our job is at stake.

- - Smuggle in paper with answers on it.

- - Looking over someone's shoulder or you may look at the screen at someone's answer.

- - Pencil and paper makes it easy to pass paper to the next fellow.

- - You must trust your own rather than someone else's answers.

- - Neither is easy to cheat on because of random selection of test questions.

- - Unplug machine or scratch disks.

Question #7.

a) How would you react to being tested on-line if the test questions were problems in which calculations were required?

b) (Ask this question if the answer to (a) is negative). Do you have any suggestions that might help provide a solution to this problem?

Responses:

A large majority expressed doubts about using on-line testing for tests containing problems involving calculations. Most felt that it would be "no good unless given paper and bigger testing station." Worksheets could be handed into instructor. Some felt that looking back and forth from screen to worksheets would be irritating and entry mistakes could be made.

Question #8.

What would be your reaction if all of your supervised exams were on-line tests using floppy disks and microcomputer as compared to:

i) computer generated off-line paper and pencil exams?

ii) regular paper and pencil exams?

Responses:

About three quarters of the students expressed positive reactions to writing tests using floppy disks providing the tests dealt with subject matter like explanation information like theory and code, but if calculations were involved they would prefer the paper and pencil type.

The students expressed the thought that if feedback were provided onscreen for all incorrect answers, it could serve as a very suitable substitute for paper and pencil tests. Some liked the speed at which the mark was returned. Another reason for expressing a positive reaction toward microcomputers was the idea that if one machine gives trouble it is independent of the functioning of the rest, whereas when the big system fails all the terminals go down.

A few people expressed a preference for the traditional instructor marked tests. They felt that this way one could obtain part marks for what they get correct.

Added Questions Arising out of Discussion:

Question #9

Would you prefer to return to the traditional method of instruction and testing (that is lecture and group testing?)

Responses:

There were a couple of students who would prefer learning by the old system, but the overwhelming majority said that they would not like to go back to that method, even though there are a few bugs in the new self-paced system.

Question #10

In the past, the use of microcomputers used to bring complaints of eye fatigue and eye strain. How do you compare the testing on-line with the paper and pencil test?

Responses:

Only a couple of students said they found the onscreen print worse than the paper print; the majority said that the onscreen print was superior to the paper print (draft dot matrix). A few said it was about the same.

APPENDIX G

COMPUTER EXPERIENCE SURVEY (From Questionnaire 1)

Table 20

Computer Experience Survey (From Questionnaire 1)

Totals	E N=11	C1 N=14	C2 N=11
1. Did you take your first-year apprenticeship training using the computer-managed learning system at NAIT?	Y=91% N=9%	Y=93% N=7%	Y=91% N=9%
2. If the answer to Question #1 is "No", were computers used to administer the tests and program at the institution where you received your apprenticeship training?	Y=100% N=0%	Y=0% N=100%	Y=100% N=0%
2. Do you presently own a microcomputer?	Y=9% N=91%	Y=7% N=93%	Y=9% N=91%
3. Have you used micro-computers to play games?	Y=73% N=26%	Y=57% N=43%	Y=55% N=45%
4. Have you ever used micro-computers before?	Y=82% N=18%	Y=57% N=43%	Y=55% N=45%
5. Have you had to use computers at your worksite?	Y=9% N=91%	Y=14% N=86%	Y=0% N=100%
6. Have you ever taken any formal courses, for example, basic programming, word-processing, spreadsheeting, etc., on microcomputers at high school or some other institution?	Y=45% N=55%	Y=43% N=57%	Y=37.5% N=62.5%

* E -- Experimental Group
C1 -- Primary Control Group
C2 -- Secondary Control Group

APPENDIX H

OVERVIEW OF QUIZMASTER

QUIZMASTER Program Description

This overview of QUIZMASTER is a reprint of the description given by Latta (1988), one of the authors of the QUIZMASTER program. (pp. 122-135) This description applies to the standard version. The QUIZMASTER program used in this study was modified so that test length was extended to forty questions with the test item pool capable of holding 160 questions, and a review option was added so that answers to questions could be reviewed and changed prior to scoring of the tests.

* * * * * * *

Introduction

QUIZMASTER was written to solve some of the testing problems encountered by teachers who, by the nature of their courses, must individualize instruction for their students. A good example of a course requiring individualization is the multiple activity Industrial Arts program. In this program it is normal to find each student reponsible for a different learning activity. The use of traditional paper and pencil testing methods often leads to problems in test security and relatively complex test management and record keeping systems. All too frequently these activities occupy valuable time which could be better spent on instruction and laboratory supervision.

QUIZMASTER provides an alternative to paper and pencil testing methods. If the teacher wishes, test administration can be initiated, and carried out, by the student at a time that the student finds convenient. Each time the program is used a different form of the test is generated.

The QUIZMASTER Unit

With QUIZMASTER, each test is called a Unit. The QUIZMASTER Unit consists of a structured test item pool and a number of special information screens. The way that the program assembles a test and presents it to a student is described in the following paragraphs.

The Test Item Pool

When assembling a test, QUIZMASTER draws its questions from a structured item pool. The structure makes it possible to have, for any given Unit, a number of different tests which will have approximately the same level of difficulty and to cover the same range of learning objectives. To accomplish this, the test item pool is structured on the basis of the Topic. Each Topic represents a different learning objective. A Unit can contain as many as twenty Topics and each Topic can contain four questions. The test item pool can contain four times twenty or eighty questions. The item pool structure for a Unit is illustrated in Figure 7.

When QUIZMASTER assembles a test it does so by selecting one question at random from each Topic. A QUIZMASTER Unit can contain a minimum of one Topic and a maximum of twenty Topics. A test assembled by QUIZMASTER, then, could be as small as one question or be as large as twenty questions.

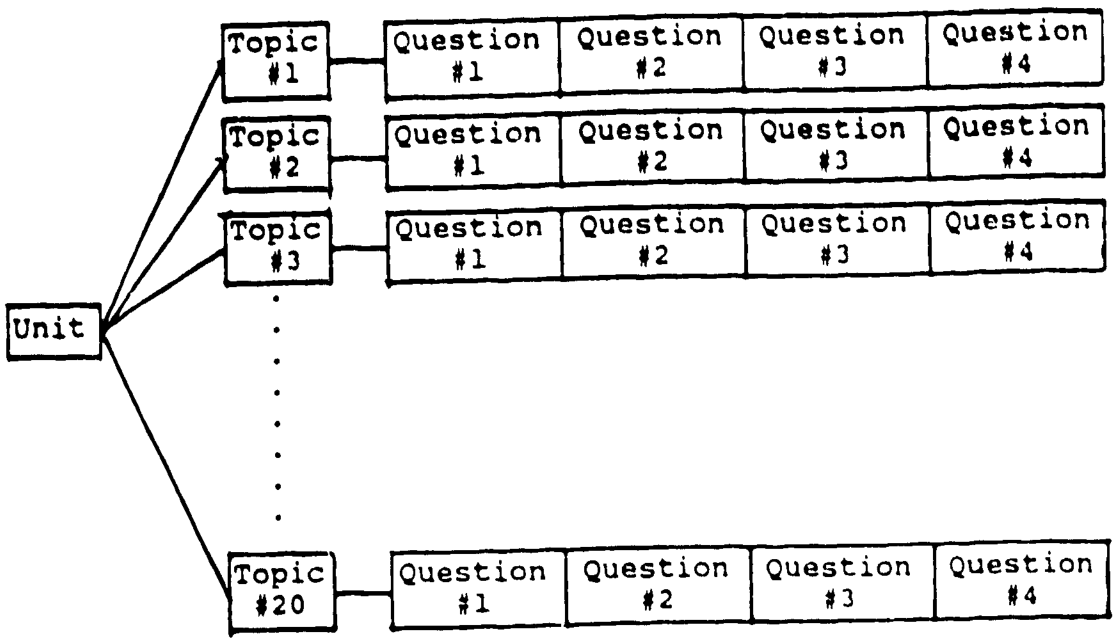

Figure 7.

Item Pool Structure

Using a structured selection process has the advantage that the teacher knows that every exam presented to the students will cover all of the intended objectives. The random selection process within Topics reduces the likelihood that two students will be given the same test thereby improving the level of test security. The intention is that within any Topic the four questions should all test the same objective and the questions should all have about the same level of difficulty.

Special Information Screens

One of the more important functions of QUIZMASTER is to facilitate the learning process by providing additional information to the student. This information is provided on special screens that may be presented before, during, and after a test. In all cases the use of these screens is optional. If the teacher decides not to use any or all of the special information screens, the student would not be aware of anything missing from the test.

Pretest Information. QUIZMASTER uses two screens to present pretest information to the students: The Unit Objective screen and The Student Introduction screen. The intent of the Unit Objective screen is to make the student aware of the test objectives prior to taking a test. The Student Introduction is intended to provide any special instructions for the student. These two screens are available to the student as a help function prior to attempting a Unit. They are presented again when the student attempts a Unit.

Resource Lists. QUIZMASTER has provision for two types of resource lists: The Unit Resource List and the Topic Resource List. These lists provide a means of directing the student to information for further study. The Topic Resource List is presented only when the student answers a question incorrectly. Since a Topic is intended to contain four equivalent questions testing the same objective there is only provision for one resource list per

Topic. The Unit Resource List is displayed after a Unit is completed and only in the event that the student's score on the Unit was less than the preset mastery level.

Feedback. One of the important features of QUIZMASTER is the program's ability to provide the student with a high le of Feedback. As part of the entry procedure for every Question the teacher is given the opportunity to enter Mastery Feedback and Non-mastery Feedback that would be unique to that question. Mastery Feedback is presented in the event that a question is answered correctly and Non-mastery Feedback is displayed when the student's response is incorrect. Mastery and Non-mastery Feedback screens are also available at the Unit level. The appropriate Unit feedback screen is presented after a Unit is completed.

Accessing the QUIZMASTER Unit

QUIZMASTER Units are stored on separate data disks. Each data disk will hold two QUIZMASTER Units and the corresponding response frequency data used to perform an item analysis for each Unit. Access to the Unit's contents and response frequency data is controlled by programs which exist on separate Teacher and Student Disks.

Item Analysis Capability

The random selection feature of assembling a test raises questions regarding statistical reliability and test difficulty. It is desirable that the four questions in a Topic be as closely matched as possible in so far as the difficulty level is concerned. QUIZMASTER facilitates this by accumulating data on each question as it is used. The program will calculate a difficulty index and display response frequencies to allow the teacher to assess the items in any Topic. The teacher may then make changes to the test items as needed.

Student Records

QUIZMASTER's record keeping system allows the teacher to monitor student progress and review the tests as they were written by the students. The program also maintains an ongoing item analysis which permits the teacher to monitor the quality of the test items. QUIZMASTER uses a different Student Disk for each class to be managed. The class list on each student disk will accommodate forty-five students and will maintain marks for as many as thirty different Units.

QUIZMASTER's record keeping system includes a Grade Book and a system for recreating tests. This information is available on the screen or it can be printed on the computer system printer. Printout formats are available that are suitable for teacher records or individual progress reports.

The Student Grade Book

Two marks may be available for each Unit. The first mark is the mark the student earned on the first attempt at the Unit. The second mark is the mark the student earned on the last attempt of the Unit. If a student made only one attempt at a Unit the Grade Book would contain only one mark. If two or more attempts are made the second mark would be the mark earned on the last attempt.

Several report formats are available to the teacher. A class summary is available for each Unit. The Teacher can also request a mark summary for any student. The program is capable of printing out the entire Grade Book or just a final mark summary.

The Grade Book has full editing facilities which allow the teacher to enter or change any student mark.

Re-creating Tests

The random aspect of test assembly makes it impossible to know in advance which questions will be presented to a student. Although this relieves some of the problems of test security it presents other difficulties in diagnosing learning problems. To overcome these difficulties, QUIZMASTER allows the teacher to re-create any test, as it was presented to the student, and to view the student's responses to the objective questions. It is only possible to know whether a student's response was right or wrong for completion type questions.

Controlling the Testing Situation

QUIZMASTER allows the teacher to customize the testing situation in order to suit most conditions. The teacher can control; student access to any Unit, the mastery level for each Unit, the feedback patterns, the resource list display, and testing mode.

Unit Access

In order to control class progress the teacher may deny or allow class access to any Unit at any time.

Unit Mastery Level

The teacher may set the mastery level for each Unit. The allowable limits are 1% to 100%.

Feedback Patterns

Changing the feedback mode changes the timing of the Question Mastery and Non-Mastery Feedback. In the immediate mode the appropriate Question feedback statement will be presented immediately after the student enters a response to a question. The deferred mode allows the student to complete the test uninterrupted. In the deferred mode the feedback statements are only presented during the test review.

The test review is offered only when QUIZMASTER is operating in the test mode. Once the test is completed the student is taken through a review. The student will see each question in turn together with the response that was entered. The program will indicate whether the student's answer was right or wrong. The correct answer is not given. If Mastery or Non-Mastery feedback statements exist then the appropriate feedback statement is displayed. The Topic Resource list is presented only when the student's response to a question is incorrect.

Testing Mode

The teacher may place any Unit into the "test mode" or the "pretest mode". This change alters several factors in the testing situation. In the test mode the student may decline to answer a question and return to it later in the test. In the pretest mode if the student declines to answer a question it is not presented a second time. Regardless of the feedback pattern settings the feedback statements are never displayed in the pretest mode. Similarly the test review is never offered in the pretest mode.

Taking a Test

Figure 8 illustrates how the various parts fit together when a student elects to take a test. It is important to remember that there are a number of possible variations, depending on the selections made for such things as the test mode, and feedback patterns.

When a student elects to take a test the first screens to be displayed are the Unit Objective and Student Introduction. If these screens were left blank QUIZMASTER will move directly to the first question.

The question administration sequence can have two different patterns depending on the feedback mode selected. These patterns are illustrated in Figure 9 and Figure 10. If the immediate feedback mode is selected then feedback statements are presented imm iiately after the student enters a response to each question. In addition the

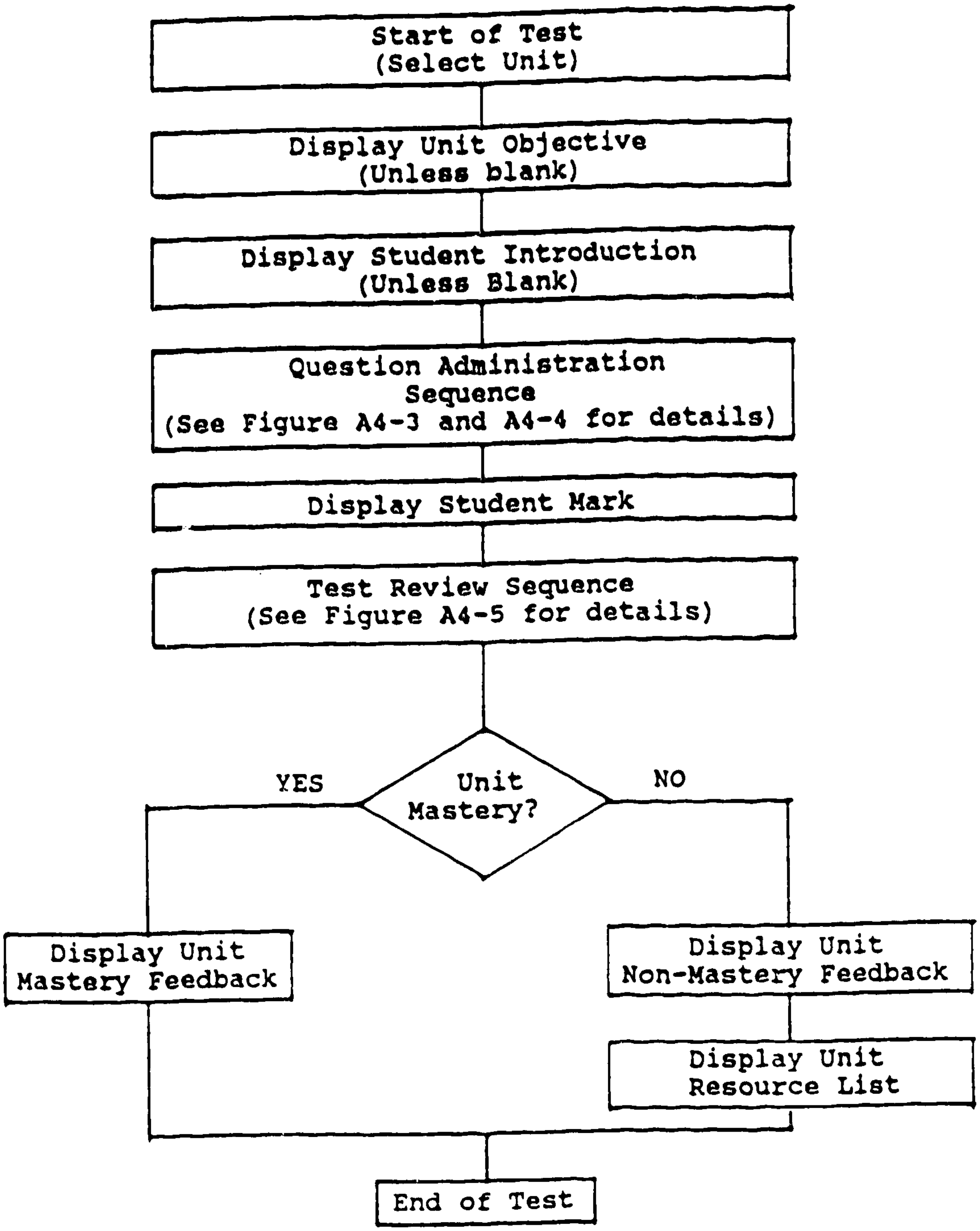

Figure 8.

<u>Typical test sequence</u>

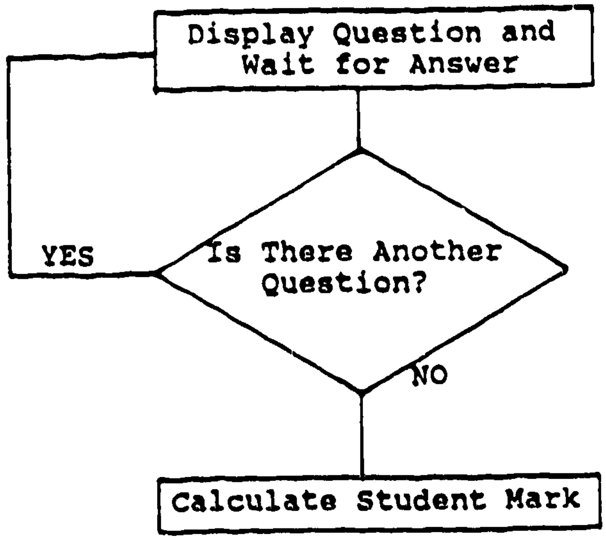

Figure 9.

Question Administration Sequence - Deferred Feedback Mode

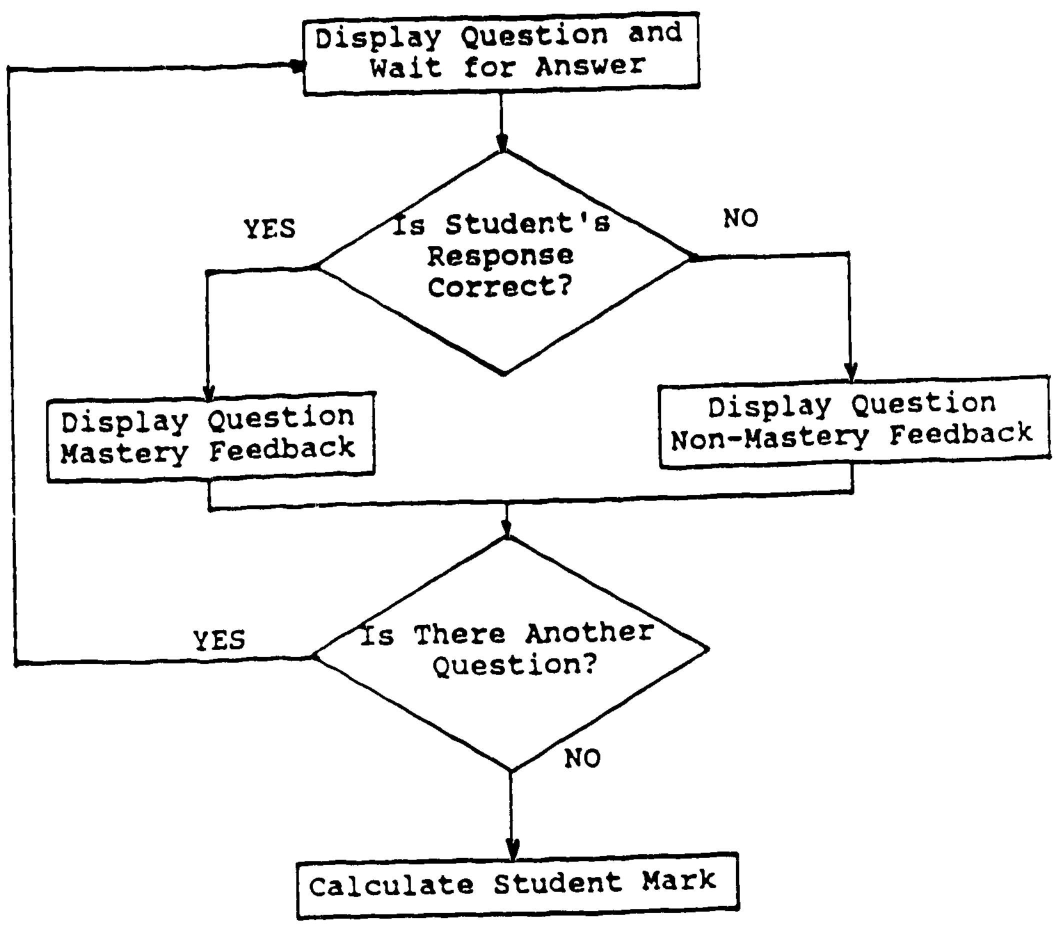

Figure 10.

Question Administration Sequence - Immediate Feedback Mode program supplies a message telling the student whether the response was correct or not. The correct answer is not given. If the feedback mode selection is "deferred" the student completes the test uninterrupted.

Upon completion of the test the student's mark is calculated and displayed. The student is then taken through the test review sequence (Figure 11).

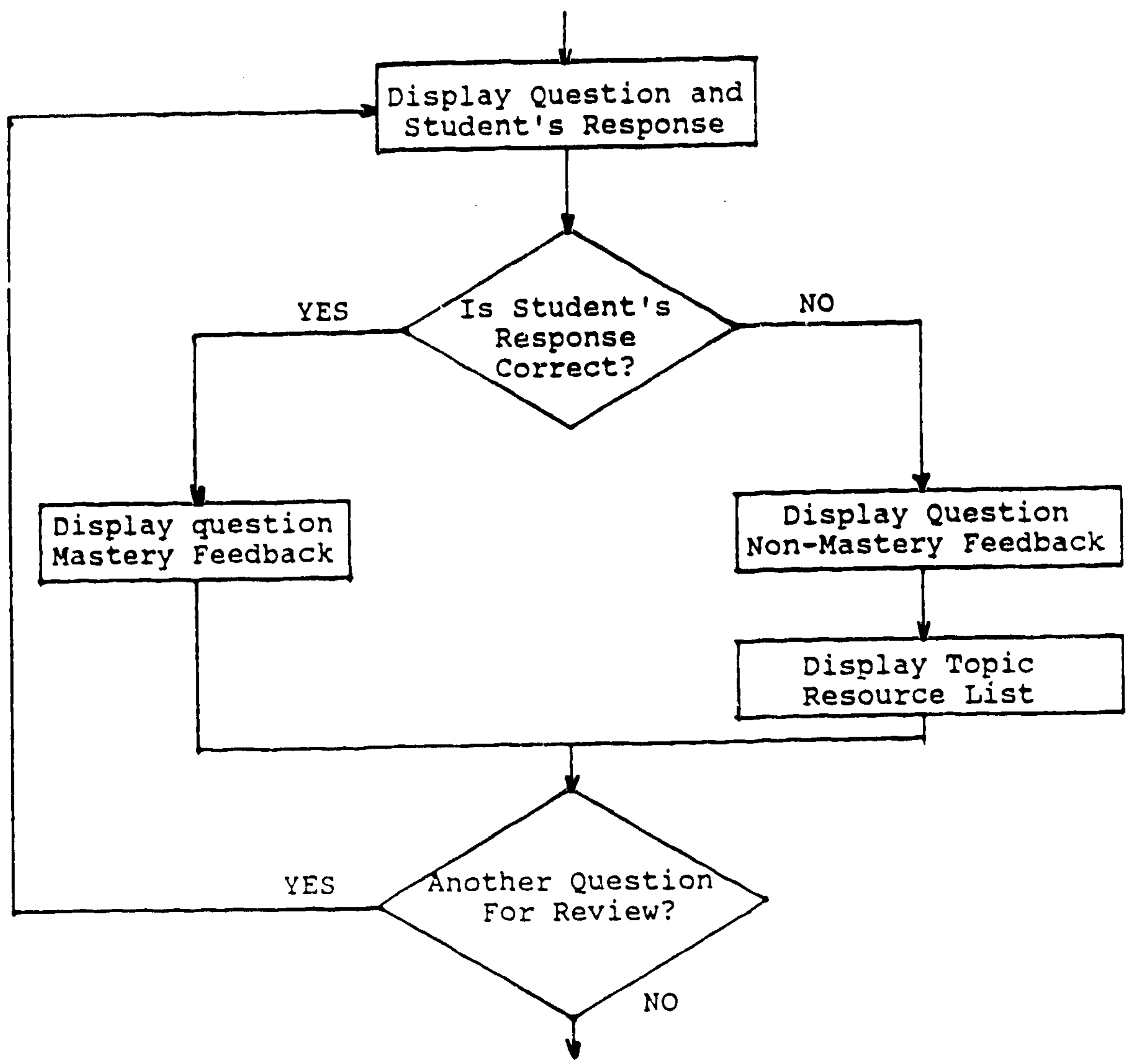

Figure 11.

Test Review Sequence

Once the test review sequence is completed the appropriate Unit Mastery or Non-Mastery feedback is displayed. The Unit Resource List is only displayed in the event that the student fails to achieve the preset mastery level. In the event that the teacher leaves the Unit Mastery and Non-Mastery screens blank QUIZMASTER will supply an appropriate message.

APPENDIX I

OVERVIEW OF CBTS-CML SYSTEM

CBTS-CML OVERVIEW

Introduction

The following discussion of CBTS-CML is not meant to be exhaustive in describing all of the system's features and options, but to provide an overview of some of the more popular and most frequently used ones. Readers wishing a more detailed explanation of the CBTS-CML system should refer to the Training Reference Manual for CBTS Computer-Managed Learning, published by Learning Resources Centre, Northern Alberta Institute of Technology 11762-106 Street, Edmonton, AB, T5G 2R1.

CBTS-CML is a computer-managed learning software program packaged and distributed by Computer Based Training Systems (CBTS), Calgary, AB. The basic hardware requirements for the delivery of this CML program are a VAX minicomputer by Digital Equipment Corporation (DEC) and remote CRT terminals, or the much more commonly used hardcopy printer terminals.

CBTS-CML systems provide a variety of features that contribute to the various components of computer-managed learning. Some of these features include course mapping and modularization, course prescription, course management, test item banking, test generation and marking, record keeping, analyses of individual and group performance, and courseware monitoring.

User Access

CBTS software is designed in menu format, which groups similar information options together. A user simply responds to the menus that are presented by answering questions and making choices. The program has a MAIN MENU, and from it the user can access nine primary menus that allow different users to perform a variety of functions, depending on the level of access granted. For example, the systems manager will have access to all parts of the program, whereas classroom instructors will have access to fewer functions, and students will have access to even fewer functions still. Normally, the department CML manager uses the USER/PRIVELEGES MENU to grant access to those other users. These users, in turn, may be granted the power to assign access to other users. For example, instructors may be granted access so they can add and edit i' m test banks, change test marks, etc.; instructors may then grant access to students so they can take exams, check course objectives, and access course maps, etc.

Testing Function

A major component of the CBTS-CML system is the testing function. Test questions are stored in the computer memory within a testbank. Each testbank can be subdivided into a maximum of 99 modules, each containing a maximum of 9 objectives. Each objective can contain up to 99 questions. Thus, the number of questions that may be stored in any uniquely named testbank is approximately, 99x9x99=88200 questions.

In addition to being associated with specific objectives, questions may be constructed as true/false, multiple choice, completion, problem solving, and essay types. The problem solving type question uses an algorithm to generate unique numbers for each problem issued. The questions may also be assigned different lock categories, different difficulty levels, and different taxonomic categories.

To enter questions into the test item bank, the person must be aware of the code specifications for each type of question and be able to access the TESTBANK MENU (Figure 12).

TEST BANK MENU

1. EXIT	- Return ot MAIN Menu
2. CREATE	- Create a new testbank
3. EDIT	- Edit an existing testbank
4. CHECK	- Check testbank for errors
5. MATRIX	- Difficulty & cognitive level matrix
6. CHARACTERISTICS	- Question characteristics listing
7. DOCUMENT	- Access to curriculum bank documents
8. EXAM	- Draw exams from testbanks
9. GUIDES	- Display Module Study Guide Menu

Figure 12

CLTS Testbank Menu Options

Using this menu an instructor or other user may CREATE a new testbank and then use the EDIT option to add, change, and delete questions and other testbank content. In addition, questions or groups of questions can be moved from one place to another within the testbank or to other testbanks. The CHECK option allows one to check for coding errors to make sure the testbank is completely error free before being used by students. Test questions may be constructed at different cognitive levels (Bloom's Taxonomy) and degrees of difficulty to improve the testbank quality. The MATRIX command shows how many questions from each cognitive level are in each degree of difficulty. The DOCUMENT option provides a list of information describing the testbank. It gives the number of modules, module statements, objectives, objective statements, questions, and the number of questions of each type. The EXAM option is normally used to draw sample exams, or for courses when only the instructor accesses the CML and a few exams are needed. The way students are permitted to access the testbank is through a course map so that access can be monitored and controlled.

Student Options

A look at the options in the STUDENT MENU in Figure 13 clearly shows the features available to the student under this system. The STUDENT MENU provides the student with the list of system functions that he can access. For example, he may use the CHALLENGE option to write exams ahead of schedule, if the instructor approves and permits it. Answers may be entered to a challenge test immediately,

STUDENT MENU

1. EXIT	7. MESSAGE
2. CHALLENGE	8. OBJECTIVES
3. PROCEED	9. PASSWORD
4. GUIDE	10. REVIEW
5. HISTORY	11. STATUS
6. MAP	

Figure 13

CBTS Student Menu Options

or the student may wait until all the questions have been printed. In the latter case, the computer automatically logs the person off CML to complete exam. Students are given the option to review and change answers before the test is marked. The option PROCEED is the command used to draw test from the question pool. Another student option, GUIDE, can prescribe a list of learning material that should be studied to complete a particular module. This study GUIDE option may or may not be used by some programs. The student HISTORY option gives a student a summary of his performance in a subject. Another option, MAP, gives a student the sequence in which he is expected to complete the course elements; and it is often used to see elements remaining to be completed. The OBJECTIVE option gives the student a list of objectives that one is expected to know before the module is completed. The PASSWORD option allows students to set or change a confidential password for security purposes. The REVIEW option allows one to do practice exams on material already passed, in preparation for a final course exam. There is also a MESSAGE option which permits students to send messages to instructors and vice versa. Finally, the STATUS option shows which module a student is expected to complete next in the subjects of a course.

Student Progress and Item Analysis

ANALYSIS MENU

1. EXIT - Return to main menu
2. PROGRESS - Student progress histogram
3. MODULES - Module completion history
4. OBJECTIVES - Question stats by objective
5. QUESTION - Question stats by question
6. GRADES - Student grade reports

Figure 14.

CBTS-SML Student Progress and Item Analysis Menu

The ANALYSIS MENU provides the instructor with several options to allow instructors to determine student progress. (Figure 14). A student PROGRESS histogram presents the instructor with information on all students registered in a CML class. The MODULE option provides a variety of data at the module level, such as, total completion time, average time, number of attempts, number of instructor passed students, and average score for each attempt. This option is helpful in helping instructors identify difficult and time consuming modules. It can provide a report of all students or just individual ones. The OBJECTIVES statistics option reports how often questions for an objective were answered correctly, and the QUESTIONS statistics option reports the number of times a question has been issued and the percentage of times that the question was answered correctly. In addition, in the case of multiple choice questions, this option can indicate the total number of times the correct response, as well as the distractors, were chosen. Finally, the GRADES option permits instructors to generate marks for each student on the average class mark for each subject. This option allows the weight for each element of the subject to be changed for individual students or for the entire class and can print individual student progress reports, complete with a space for individual instructor comments.

Other Noteworthy Features

There are a group of options that the instructor can access in CBTS-CML that lets him give students a conditional pass, a pass , or a fail in a subject. The SCORE option permits the instructor to change the marks for any student on a specific entry. There is an option that causes the system to lock-out any student until he sees his instructor. Another option gives the instructor control over when supervised exams are to be released. Students may be locked out if they do not finish an exam in the required time, but instructors can cancel it.

The CBTS CML system can provide starting and completion times and dates and the total number of attempts made at the test. The scores are reported for the first four attempts and the last attempt and a record is kept on which attempt a pass was obtained.