EFFECTIVENESS OF COMPUTER ASSISTED INSTRUCTION

IN NURSING EDUCATION:

AN INTEGRATIVE REVIEW

by

M. Joan Belfry

B.N., McGill University, 1969

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE (EDUCATION)

in the Faculty

of

Education

© M. Joan Belfry

Simon Fraser University

July, 1984

APPROVAL

Name: M. Joan Belfry

Degree: Master of Science (Education)

Title of Thesis: Effectiveness of Computer Assisted
                               Instruction in Nursing Education:
                               An Integrative Review

Examining Committee

   Chairperson: A. C. Kazepides

_____

                               P. H. Winne
                               Senior Supervisor

_____

                               A. O. Dawson
                               Associate Professor

_____

                               E. Knor
                               Vancouver Community College
                               Vancouver, B. C.
                               External Examiner

                               Date approved   *July 20/84*

PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay

Effectiveness of Computer Assisted Instruction in Nursing Education:

An Integrative Review

Author: _____
   (signature)

   M. Joan Belfry
      (name)

   July 20, 1984
      (date)

ABSTRACT

This study had three parts. First, reports of the use of computers and computer assisted instruction in health care agencies, in general education, and in nursing education were surveyed. These reports summarize current applications of computer technology and identify variables which have been used to measure instructional effectiveness. The survey indicated that despite the rapid growth in use of computers for instruction and the increasing body of literature on the topic, few attempts had been made to integrate the research findings.

Second, recent trends and developments in review methodology were studied to determine approaches for integrating research findings. The advantages and disadvantages of both qualitative and meta-analytic techniques were considered in developing the methodology used in part three of this study.

The third part of the thesis synthesized research findings on computer assisted instruction in nursing education and nursing practice to compare its effectiveness to traditional instruction. Based on part two, selected qualitative and meta-analytic techniques were used to integrate findings from 11 independent research studies of computer assisted learning in nursing. The findings of this synthesis parallel those identified by reviews of research at the

secondary and post-secondary level in disciplines other than nursing. The analysis showed that nursing students using computer assisted instruction generally learned more than those given traditional instruction, particularly when the computer was used as a supplement to other methods. Students who used computers for learning also developed positive attitudes toward computers as an instructional method. In addition, computer assisted learning generally reduced the amount of time students needed to learn the material.

The meta-analysis revealed few relationships between studies' findings and design features of the research, settings of the studies, level of student, subject matter, or the type of computer assisted instruction. Based on these results, recommendations were made regarding application of the findings of this study to nursing education, future research, and future integrative reviews on computer assisted instruction.

TABLE OF CONTENTS

Page

TABLE OF CONTENTS (Continued)

Page

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

The past two decades have been a period of rapid technological change and growth of information. Today there is a greater emphasis on knowledge-based industries than on the production of goods and services. As Molnar (1983) states, "information has become a national commodity and a national resource and has altered the very nature of work" (p. 25). Indeed, it can be said that the United States and Canada have shifted from being predominantly industrialized societies to having information-based economies (Naisbitt, 1984). Coupled with this economic shift has been a rapid growth in electronic technology, particularly that of computers. Because of a need to be able to organize, store, and access the expanding informational base, computer applications are influencing all sectors of society -- business, government, entertainment, schools, homes. Hence, it is becoming increasingly important for people to become familiar with and able to use computers effectively.

Educational institutions have used computers for both administrative and instructional purposes. The early applications of computerized instruction involved little coordination or sharing of information among institutions. The applications were specialized and the computer systems, which were large and expensive, often could be supported only with the help of extra public or private funds. In

many cases, if the funding ceased, the educational application did as well. The advent of microcomputers in the 1970s reduced the cost of computers significantly and fostered growth in their use at all levels of education. Research reports (Jamison, Suppes & Wells, 1974; Kulik, Bangert & Williams, 1983; Kulik, Kulik & Cohen, 1980) indicate that computer assisted learning can be as effective as traditional instruction.

Student outcomes commonly studied in research on the effectiveness of computer assisted learning in general education include student achievement on examinations, student attitudes toward computers, instructional time taken, and the correlation between aptitude and achievement. Reviews of research have indicated that, for elementary students, normal instruction supplemented by computer assisted learning is often more effective than normal instruction alone (Edwards, Norton, Taylor, Weiss & Dusseldorp, 1975; Jamison, Suppes & Wells, 1974; Visonhaler & Bass, 1972). At the post-secondary level, computer assisted learning is as effective as traditional instruction when used as a replacement (Jamison et al., 1974; Kulik, Kulik & Cohen, 1980). For both secondary and post-secondary students, computer assisted instruction has positive effects on student attitudes toward instruction and the subject matter, and reduces the amount of time that students need for learning (Kulik, Bangert & Williams, 1983; Kulik, Kulik & Cohen, 1980).

The use of computers in health education has followed the same

trend as that of general education, but on a smaller scale. The majority of applications of computer assisted learning in nursing have been to teach specific topics and to foster independent decision making with many types of students. The applications have occurred in hospital-based, baccalaureate, graduate, inservice, continuing education, and refresher programs. It is only in the last few years that computer literacy for health professionals has become an issue because of the increasingly varied applications of computers in the health care system and to patient care.

Computers have been used to educate nurses since the late 1960s. For approximately a decade, the majority of published articles about applications of computer assisted learning to nursing education were narrative reports of experiences at a particular school or hospital. More recently, there have been increasing numbers of research studies investigating the effectiveness of computer assisted learning in nursing education. These studies cover a variety of instructional settings, subject areas, and types of students. This makes it difficult to identify trends or form generalizable conclusions about the usefulness of computer assisted learning as an instructional method in this field. This study synthesizes the research findings on computer assisted learning in nursing education to determine its effectiveness as compared to traditional instruction.

Definitions

One major problem in referring to computer applications in education is that many terms are used interchangeably to describe the same activity. Computer assisted learning has been referred to as computer aided instruction, computer based instruction, computer assisted instruction, computer managed instruction, computer simulated instruction, computer based education, automated education, computerized instruction, and a computer controlled teaching device (Conklin, 1981). As well, any one of these terms can be used to refer to instructional activities which involve very different applications of the technology. A reader must scrutinize each report carefully to determine exactly how the computer is being used.

The terms that are fairly widely accepted in the health field and which will be used in this paper are:

1. Computer assisted learning (CAL) - any educational techniques that rely on a computer to facilitate learning. CAL can be subdivided into computer assisted instruction and computer managed instruction.

2. Computer assisted instruction (CAI) - the use of a computer system to perform the teaching function of communicating information to the learner independently of direct interaction between the student and an instructor. Both the content and the instructional

logic are stored in the computer memory.

3. Computer managed instruction (CMI) - an overall system
   for educational management in which detailed student
   records, curriculum data including schedules and
   timetables, and information on available learning
   resources are stored in the computer. The computer is
   then programmed to integrate this information to
   develop unique units of instruction for individual
   students and to facilitate optimal educational resource
   management (Hannah, 1983a).

Overview of the Study

Chapter 2 overviews methods for doing reviews as research studies
and discusses problems associated with doing them. The purpose of
reviewing original research studies on a topic is to synthesize the
findings of the studies so that generalizable trends can be
identified. In approaching this objective, attention is paid to
particular settings, subjects, or conditions which influence or alter
relationships investigated in particular studies. As a result of the
review, recommendations may be made for educational policy, practice,
further research, and future reviews.

The process involved in carrying out a systematic review of
research on a particular topic parallels that for doing original
research except that the data sampled for the review come from primary

studies rather than human subjects. However, one of the major problems associated with doing an integrative review is that review methodology is not outlined clearly. As a result, each reviewer must make several significant decisions. These include deciding if the sample of studies is adequate to be considered representative of the population, developing criteria for determining which subset of studies shall be included in the review, and deciding on an appropriate balance of qualitative and quantitative methods for analyzing the data. In order to ensure that a review is valid and replicable, the reviewer should report on these phases and decisions in detail, pay close attention to details of the primary research studies, use appropriate techniques to analyze data, and attempt to explain variations in the findings across the studies.

Chapter 3 is an overview of the use of computers and computer assisted learning in health care agencies, general education, and nursing education. Computers were first used in the health care system for administrative tasks such as payroll, inventory control, and patient charges. As these applications became more sophisticated, nurses realized the computer's potential for improving nursing practice and the quality of patient care. However, the development of effective applications in this domain progressed slowly because nurses had limited understanding of computers and felt threatened by automation. To date, it has been demonstrated that computers can be used successfully for patient monitoring, charting, quality assessment

and assurance, interdepartmental communication, continuing education, and research. Since the number and variety of applications are increasing rapidly, it is essential that health care professionals become aware of the possible functions and limitations of computers so they can be active participants in planning and implementing systems that meet their practice and educational needs.

Use of computers as an instructional method in general education has been developing slowly. This lag in development compared to society in general has been linked to lack of knowledge about computers and computer assisted learning on the part of teachers, insufficient funding for developing instructional software, and lack of transferability of software between microcomputer systems. Computers are now seen as a method for providing more cost-effective education because they can be used by large numbers of students for independent study either within the institution or through distance education. Teachers then can be released from repetitive teaching and can focus on the learning needs of individual students.

Computer assisted instruction has been used to teach nurses a variety of topics, such as obstetrics, pharmacology, and post-operative nursing care. The programs have been implemented in colleges, universities, and hospitals to provide basic, inservice, and continuing education. While these applications are reported to be effective, no integrative reviews of the research findings have been done to determine generalizable benefits and trends.

Chapter 4 is an integrative review of research on computer assisted instruction in nursing practice and nursing education to determine its effectiveness compared to traditional methods. The questions framing the review were whether students learn more, are able to apply the information to other situations, and take less time to learn when computer assisted instruction is used. Student attitudes toward use of computers for learning and other variables which might influence the effectiveness of instruction were also studied.

Following on the findings concerning integrative reviewing (Chapter 2), the methodology used for this integrative review is discussed in detail. Different sections of the chapter outline the decisions and procedures for identifying the concepts studied, locating and selecting the research studies, summarizing the characteristics and findings, analyzing the findings, and interpreting the results.

Chapter 5 discusses the findings from the review and makes recommendations based on them. The relationship of the findings of this review to existing educational theory and the consistency between findings in general education and nursing education are presented. As well, the implications of the results of the review for nursing education are outlined and recommendations are made for further research and future reviews of computer assisted learning in education.

Chapter 2

Research Review and Integration

The information explosion that has been affecting society in
general over the past few decades has also had an impact on
educational researchers and consumers of educational research. The
number of published studies on many topics has grown so rapidly that
it has become difficult for any one person to read and comprehend the
vast amount of information available. Recently, there has been
increasing interest in and discussion of the use of integrative
reviews of the literature to synthesize the research findings from
primary studies and to identify generalizable trends.

"In the past ten years, scientists and methodologists have
worried about how the findings of research can be synthesized and
organized into coherent patterns" (Glass, McGaw & Smith, 1981, p. 9).
This chapter provides an overview of integrative reviews as a research
method and outlines the problems associated with this type of
research. The first section of the chapter focuses on the purposes
for which reviews have been done and summarizes recent trends and
developments in methodology for reviews. Next, the process involved
in completing a review is discussed. Here, the emphasis is on the
decisions a reviewer must make and the methodological problems which
can arise in each phase of conducting a review. The final section of
the chapter summarizes criteria for determining the validity and

replicability of an integrative review.

Purposes of Reviews

The main purpose for synthesizing research findings on a topic is to generate an orderly and systematic approach to accumulating knowledge and developing theory. As the number of original or primary research studies on a topic grows, periodic review and integration of the findings is needed to summarize the current knowledge base and to make generalizations about the phenomena being studied. This process is generally referred to as producing an integrative review. Beyond helping to define approaches to a broad area of inquiry, integrative reviews can serve three other specific purposes: identifying critical experiments needed for theory development, influencing educational practice, and making recommendations for methodological improvements in research studies.

As Glass (1976a) has noted, educational studies tend to be less cumulative and evolutionary than research in laboratory sciences, so identifying trends can be difficult. Experiments in educational contexts are usually applied research studies which vary widely in their research designs, contexts, subjects, variables, and outcome measures. Synthesizing the findings from heterogeneous studies assists researchers in identifying critical experiments needed to clarify particular aspects or increase understanding of a phenomenon. When they can be designed, critical experiments eliminate the need for

further replications of some experiments and provide clearer guidelines for future research.

Another purpose for completing integrative literature reviews is to influence educational policy and practice. For example, recommendations from a research summary can help administrators make decisions about the advisability of continuing or adopting particular courses of action. Teachers and learners benefit from research summaries when guidelines describe conditions or practices that have been found to make learning more effective.

Many reviews of educational research have been done to point out flaws in previous research and to recommend methodological improvements. The results of these reviews often sound extremely pessimistic because so many criticisms are made of the completed studies. While this type of review is very useful to researchers, others find them of little practical value (Ladas, 1980).

Trends in Review Methodology

Methods for conducting integrative reviews have changed over the years as the volume of research literature increased. During the 1940s, 1950s, and 1960s, reviews generally took the form of a narrative discussion of the findings of a set of primary research studies. This traditional, or qualitative, method was quite satisfactory as long as the number of studies reviewed was small.

By the end of the 1960s, the bodies of research literature were

growing so rapidly that it became exceptionally difficult to integrate studies narratively. Reviewers began to develop strategies for categorizing and classifying results of studies. Frequently, the studies were grouped in tables according to whether outcomes reached statistical significance. These tables, commonly referred to as box scores, recently have been viewed with scepticism because this approach tends to oversimplify the interpretation of positive results and can cause the reviewer to ignore the implications of statistically unreliable, negative, or conflicting findings (Ladas, 1980).

Traditional reviewers have also applied several other methods for summarizing the research findings. Some created lists of all the factors that were shown to have an effect on the phenomena being studied. Others, using the statistics on a given outcome, computed overall averages across the set of studies. And some reviewers opted to do an in depth analysis of only one or two studies that they judged to be typical of all the research on the topic. Each of these methods has been criticized because it fails to cope adequately with the complexities of the research findings (Light & Smith, 1971).

As Glass (1976b) stated in his 1976 American Educational Research Association presidential address, there was a need for an alternative to the casual, narrative integrations of research studies. At this time several researchers had already begun work on methods for improving integrating studies. Although they worked somewhat independently, they all (Cohen, 1977; Edgington, 1972; Glass, 1976a;

Kraemer & Andrews, 1982; Light & Smith, 1971; Rosenthal, 1978; Rosenthal & Rubin, 1979) have recommended methods involving statistical analyses of the data.

Glass (1976a) proposed an approach called meta-analysis, "the statistical analysis of a large collection of analysis results from individual studies" (p. 3). The term meta-analysis is now used to refer to any review that uses quantitative or statistical analyses to integrate findings. Some of the meta-analytic techniques that have been developed for data synthesis include: cluster analysis (Light & Smith, 1971), effect size estimation (Cohen, 1977; Glass, McGaw & Smith, 1981), combining probability values of significance tests (Edgington, 1972), and the blocking technique (Rosenthal, 1978).

During the past few years, reviewers have debated whether qualitative or quantitative reviews are more useful (Cook & Leviton, 1980; Cooper & Rosenthal, 1980; Jackson, 1980; Pillemer & Light, 1980). One positive result of this debate has been that many problems associated with reviews have been identified. One debate that remains unresolved concerns the methodological standards that research studies should meet in order to be used as part of a review. Some feel that experiments with methodological flaws should not be included (Eysenck, 1978), while others argue that exclusion of studies may eliminate valuable data from the review (Glass & Smith, 1978).

Another criticism of qualitative methods is that they are too subjective. Reviewers' biases can influence the interpretation of

results. However, the same problem can affect quantitative reviewers if they do not define their constructs clearly. Cooper and Rosenthal (1980) found that, while qualitative reviewers tend to be more conservative in estimating the magnitude of an effect, both types of data synthesis were prone to problems that could lead to misinterpretation of results.

Perhaps the most serious problem associated with either type of review is the lack of specific guidelines for conducting a synthesis of the data provided by primary studies. One result of this lack is that different reviewers of the same body of literature may reach different conclusions because the studies included or the operational definitions of the constructs vary (Cook & Leviton, 1980). Another is that there are few consistent standards for judging the quality of reviews. Also, Jackson (1980) has pointed out that this lack of consistency could mean that reviewers are not using methods as powerful as could be developed and therefore the accumulation of knowledge is hindered. More stringent methodological guidelines for conducting reviews have now been proposed by Jackson (1980) and Cooper (1982) in order to promote consistency, replicability, and validity of research integration.

Review Methodology

The goals of an integrative review are to summarize the accumulated state of knowledge concerning the relationship(s) of

interest and to highlight the important issues that research to date

has left unresolved (Taveggia, 1974). Jackson (1980) identified six

tasks involved in performing an integrative review. These tasks will

be used as an outline to organize the discussion which follows on how

the nature of research synthesis and the problems associated with it

interfere with achieving the goals set for such reviews. The six

tasks are:

> (1) selecting the questions or hypotheses for the review,
> (2) sampling the research studies that are to be reviewed,
> (3) representing the characteristics of the studies and
> their findings,
> (4) analyzing the findings,
> (5) interpreting the results, and
> (6) reporting the review.
> (Jackson, 1980, p. 441)

Selecting the questions or hypotheses. As with any research

study, the first step in conducting a review is to formulate questions

and hypotheses, and to define the variables to be considered. Jackson

(1980) suggests four sources that should be consulted in order to

accomplish this task: available theory; prior research, including

previous reviews; primary research that is expected to be included in

the review; and one's intuition, insight, and ingenuity.

Variables used in the review must be defined both conceptually

and operationally (Cooper, 1982). The most definitive conclusions

from reviews are reached by those who use broad conceptual definitions

so that a large number of variables can be examined across the primary

studies. Although many of the operational definitions pertaining to variables in primary studies are specified during this first phase, reviewers should be prepared to add other concept-relevant definitions which arise from the literature search. When reviewers use different operational definitions for the same concepts, the conclusions reached from reviews on the same topic may vary considerably.

   Sampling the studies to be reviewed. When conducting a search of the literature for relevant studies, the reviewer must remember that it is virtually impossible to collect all the studies on a topic. And, since the parameters of the universe containing all studies relevant to a topic are not known, reviewers will not be able to determine if the sample obtained is representative of the population (Feldman, 1971). However, the larger the sample of primary studies and the greater the variety of sources consulted, the more representative the sample of primary studies is likely to be.

   Rosenthal (1979) has expressed concern that published articles are a biased sample of completed research because many studies with nonsignificant results remain unpublished. Hunt (1975), in counting tests of significance in published articles, found that, of the 110 tests reported in three studies, only 29% were significant at the .05 level or better. He concluded that "it cannot be said that only significant findings are published" (p. 592). Putting the problem in perspective, Gage (1978) has pointed out that:

No matter how consistent and significant the published results may be, they can never refute the argument that even more nonsignificant or negative findings have gone unpublished ... [the unknowability argument] reduces to futility any effort to accumulate evidence and draw conclusions in the behavioral sciences. (p. 231)

Discrepancies in the conclusions from different reviews can result from reviewers using different sources to retrieve studies, and therefore having different samples (Cooper, 1982). It is recommended that relevant studies be retrieved through on-line computer searches; the use of abstracting sources; tracking citations from articles, bibliographies, and indexes; and informal contact with members of the relevant "invisible college". Reviewers should provide detailed documentation of the search strategies used, including sources, years, and keywords, so that others will be able to judge the adequacy of the search. Precise reporting also facilitates replication or extension of the review.

Once the search is complete, standards must be set for determining which of the obtained studies are appropriate for inclusion in the review. Eysenck (1978) believes that only those studies which meet stringent research standards should be included. Glass (1976a), on the other hand, feels that exclusion of methodologically poor studies discards a large amount of important data. Moreover, he remarks that "it is an empirical question whether relatively poorly designed studies give results significantly at variance with those of the best designed studies" (Glass, 1976a,

p. 4). Yin, Bingham, and Heald (1976) found that only 25% of studies published in sociology and psychology had sound methodological characteristics. If they had limited their investigation to these studies, both the intention and the results of their review would have been altered markedly because of the limited number of variables investigated in this smaller number of primary studies. As a result, they recommend that reviewers should set their own standards for determining which studies should be included.

Pillemer and Light (1980) suggest that all studies included in a review should meet basic standards for research reporting. The researcher must provide summary statistics, perform statistical analyses correctly, and describe outcome measures adequately. However, they do recommend that a reviewer exercise discretion in selecting studies for the review. Some studies which do not meet basic standards may have been designed purposefully to elicit specific information. Since there are no standard assessment techniques for reviews, Yin, Bingham, and Heald (1976) recommend that a case study approach be used to determine what effect discarding poorer quality studies would have on review outcomes. Because reviewers must set their own standards for inclusion of studies, the review report should outline the criteria used to identify and select the primary studies, and should specify the potential shortcomings of the strategy adopted.

Representing characteristics of the primary studies. "The representation of the characteristics of the primary studies is, in effect, the data collection of integrative reviews. The manner in which this is done can substantially affect the results and interpretation of the review" (Jackson, 1980, p. 445).

In his research on integrative reviews, Jackson (1980) found that most reviewers reported some of the findings of the primary studies, but they did not give the direction and magnitude of the results as frequently or consistently. As well, the use of narrative reviews to describe findings related to a particular variable made it difficult to understand the results. He suggests that the best way to represent findings from primary studies is to convert all results to a common metric and report this along with the direction of the relationship. An inferior method would be to describe findings in terms of the statistical significance and direction of the relationship. This approach should be used only if there is insufficient data reported in the primary studies to do anything else.

The reviewer should also attempt to obtain data missing from any report. These actions should be reported in the review (Jackson, 1980). If data is incomplete, Cooper (1982) recommends placing wider confidence intervals around conclusions from reviews that are based on estimated or omitted information.

Analyzing the primary studies. For integrative reviews, data analysis involves combining the results of the primary studies to make inferences about the relationships common to the set of studies. There are few guidelines as to how this analysis should be done and whether statistical procedures should be used. Therefore, each reviewer must make an individual decision regarding the best method for synthesizing the results.

The task of data analysis in integrative reviews is never easy. The reviewer must consider the implications of how the strengths and weaknesses of the primary studies affect the reported results. There can be wide variation in subjects, research methodology, contexts, outcome measures, and statistical analyses. When attempting to account for the effect of these variations, the reviewer must consider the differences in constructs used, the impact of sampling error, and the adequacy of the research designs of the primary studies (Jackson, 1980).

The first step in an analysis is to determine whether the results of the primary studies can be combined. The reviewer must decide if the conceptual and operational definitions are the same or sufficiently similar across the studies. In some cases, the definitions may vary too widely to be considered of the same order, and therefore, data from the outcome measures can not be combined as part of the review. Methodological weaknesses in the primary studies may also be a reason for not including some results in the data

analysis (Jackson, 1980).

The reviewer must pay attention to the hypotheses within each primary study. It is possible that not all of them will be stated in terms of a null hypothesis. When this occurs, the reviewer must ensure that the direction of the results is reported accurately in the review (Jackson, 1980).

Once the reviewer has decided which results can be combined, a method which will facilitate interpretation of the findings must be selected. Qualitative reviews frequently use box score counts. Here, studies are categorized as to whether the findings related to a particular variable are positive and statistically reliable, show no statistically reliable differences, or are negative and statistically reliable. The problem with this method of summarizing results is that the meaning of the lists is difficult to interpret (Ladas, 1980). One must consider whether the number of positive results should be weighed against those that are nonsignificant or negative, or whether the positive results alone can indicate a trend. As well, the statistically reliable results, whether positive or negative, should be studied to determine if the number of them is greater than would occur by chance alone. If they are greater, then further analysis of the primary studies will be necessary to try to determine the cause of these results (Ladas, 1980).

Meta-analytic techniques have several advantages over traditional methods of synthesizing findings from a set of studies. In such

techniques, conventional statistical procedures are used to synthesize the numerous and diverse findings of a set of primary studies. These procedures are advantageous for several reasons. First, they allow reviewers to integrate much larger samples of studies than is possible with qualitative methods. As well, the full power of statistical methods can be used to interpret the results of the synthesis. Third, because statistical procedures are used, the methodology for the integration and the assumptions underlying this methodology are known. Fourth, meta-analytic methods allow reviewers to investigate relationships among studies' designs, subjects, conditions, treatments, and outcomes. This type of analysis is very difficult with qualitative methods. Fifth, data from both "good" and "poor" studies can be included in the analysis provided the studies do not have severe methodological flaws. Finally, a review based on meta-analytic procedures is easier to replicate than one done by traditional methods (Jackson, 1980).

As with any statistical procedure, there are problems and limitations to the use of meta-analytic techniques. One of the unresolved problems is the definition of what constitutes an adequate sample of studies, particularly when multivariate statistical techniques are used for the data analysis. Another problem is that when there is a large number of studies, reviewers must ensure that the characteristics of the studies are coded in a consistent way. When several coders are used, inter-coder reliabilities should be

reported.

Reviewers may also have difficulties combining primary findings when the different measures used in the primary research studies do not produce outcome values which can be considered equivalent. This lack of a common metric is particularly difficult to resolve when the constructs have broad definitions and the outcome measures have not been validated. Lastly, the results of meta-analytic methods are limited in their interpretation as they should not be used to infer causal relationships between variables. The most that can be achieved with these methods is to identify which characteristics of the primary studies may have influenced the results (Jachson, 1980).

When a decision is made to analyze the findings using meta-analytic techniques, the reviewer must be careful to select statistical methods that are appropriate for the data. Each of the methods developed have certain limitations to their use and interpretation. To overcome these problems, it has been suggested that a reviewer use several different techniques for the data analysis (Pillemer & Light, 1980; Rosenthal, 1978).

One of the oldest meta-analytic techniques is cluster analysis (Light & Smith, 1971). The primary focus of this method is to determine if, in fact, the treatment or program subgroups (clusters) in the primary research studies are similar enough to be combined into one large group. Light and Smith suggest that these groups should have similar means, variances, relationships between independent and

dependent variables, subject-by-treatment interactions, and contextual effects. If these subgroups can be combined, then comparisons can be made between different types of programs using much larger sample sizes than were in each separate study, thus increasing the power of the findings.

One of the major benefits of doing a cluster analysis is that differences between treatments called by the same name can be identified. While this technique does not supply the reason for the differences, it does signal that an atypical treatment exists. The reviewer can then do further analyses to try to learn why that treatment is different from the others. One limitation of this technique is that it requires detailed data about the treatment. Carrying out such in-length examinations may be difficult and time consuming if this data is not reported in the primary studies. As well, the only measure of the effectiveness of different treatments is that done at the end of the analysis, after the studies have been combined (Pillemer & Light, 1980).

Another method for pooling data of similar groups so that the effectiveness of treatments can be analyzed involves combining the separate significance tests from two or more independent studies into an overall test of differences. Rosenthal (1978) evaluated nine different procedures which have been used, and summarized the advantages and limitations of each. These techniques involve simple computations, such as adding probabilities, adding z's, adding t's,

and are based on routinely reported test statistics (t, z, or F) or exact probabilities. Combining these separate significance tests increases power by increasing the sample size so that the overall statistical test is more likely to be statistically reliable even when the sample of studies report nonsignificant findings. These techniques are most appropriate when the primary studies can be considered to be an independent and a random sample of the population so that variations among studies can be attributed to chance. As well, the studies should have tested the same directional hypothesis (Pillemer & Light, 1980).

Rosenthal (1978) has recommended use of the blocking technique in order to study interactions between variables. If the means, sample sizes, and standard deviations are available from each study, the data can be combined into an analysis of variance in which the studies are regarded as a blocking variable. Studying the main effect of treatments from the analysis of variance provides an estimate of their effectiveness. However, the main advantage of the blocking technique is that it can help to identify unusual variations in outcomes. When there are substantial treatment-by-study interaction effects, the reviewer should look for factors to explain these results. Rosenthal (1978) suggests examining the residuals closely "for clues as to the nature of the possible moderating variables affecting the operation of the treatment effects" (p. 190). The one major disadvantage of this technique is the amount of work involved, especially when the number

of studies is large.

Probably the most widely known meta-analytic technique is Glass'
(1976a) estimate of effect size. This method reduces the findings of
each study to a common metric that can be interpreted in terms of
areas under the standard normal curve. A common measure of the
experimental effect size is the difference between means on the
outcome variable for experimental and control groups divided by the
within-group standard deviation of the control group. Glass, McGaw,
and Smith (1981) have suggested several other methods for calculating
effect sizes, such as from the t-statistic, when group means or
standard deviations are not reported in the primary study.

The actual interpretation of the effect size can be done in
several ways. First, the average effect size for a particular
variable can be computed across a set of studies. This approach
provides a single value which can be discussed as a percentile or in
terms of standard deviation units of the control group. Another way
to interpret the magnitude of an effect size is to compare it with
typical effect sizes found in other studies of the same variable.
Glass, McGaw, and Smith (1981) stress that the practice of defining
ranges of effect sizes so that they can be referred to as small,
moderate, or large produces errors in interpretation. The real value
of the statistic is that it is standardized and can be used to compare
findings across a set of studies.

One limitation of the effect size statistic is that it is based

on an assumption that experimental and control group scores are normally distributed and that both are random samples of a population. Kraemer and Andrews (1982) have recommended that their nonparametric techniques for effect size calculation be used when these assumptions are violated. However, the use of nonparametric methods is limited by the fact that the calculations require raw data from a study. As well, Glass, McGaw and Smith (1981) found little difference in the result when effect sizes were calculated using parametric and nonparametric methods.

Ultimately, the decision as to whether to use qualitative or meta-analytic techniques for data analysis rests with the reviewer. Cooper and Arkin (1981) feel that many of the criticisms directed at quantitative methods of analysis are unjustified because the problems are caused by the review procedures in general, not by the data analysis techniques. They suggest that the best approach to data analysis may be thoughtful use of a combination of qualitative and quantitative methods. In any case, the reviewer should describe in detail the decisions and procedures used for this phase of the investigation.

Interpreting the results. Interpreting the results of an integrative review can involve relating the findings from the data analysis to existing theory, and making recommendations about how the results of the review relate to existing policies and practices or to

future research. The inferences made in this phase should be stated in the review itself so that others may determine if they are appropriate.

Review conclusions can be useful in identifying generalizable trends that either support or disprove existing theory related to the constructs studied. Also, by trying to find explanations for conflicting results in the primary studies, the reviewer may be able to identify particular cases or variables which influence or alter the phenomena being studied. Light (1979) suggests that reviewers attend to subject-by-treatment interactions and skewed distributions when looking for explanations of varying outcomes. Ladas (1980) recommends a case study approach which involves careful examination of studies which report nonsignificant or negative findings in order to identify reasons for conflicting results. If a particular relationship, such as differences in instruments, participants, testing conditions, or treatments can be identified, the reviewer may suggest that it be investigated through further research.

When appropriate, the reviewer should discuss the implications of the review findings for policy decisions and for practice. How these conclusions are stated is particularly important. Since reviews reach wider audiences, they tend to have a greater effect on knowledge bases and policy decisions than do primary studies (Cooper & Rosenthal, 1980). Applying the findings to actual teaching and learning situations could help readers of the review make practical use of the

review results.

Jackson (1980) found that few suggestions were given by reviewers for future reviews. He recommends that ideas for improved reviews and questions about the method should be included in order to help others who may wish to undertake the task.

Reporting the review. An integrative review is a form of research and, as such, should follow the same stringent guidelines for reports as those outlined for primary research. All phases of the review methodology should be reported in precise detail so that the reader can judge the comprehensiveness of the review and so that replication or extension of the review is facilitated (Cooper, 1982).

Jackson (1980) found that many reviews failed to report adequate information about all phases of the review. When this occurs, the validity of the review can be questioned because it cannot be replicated. As well, the reader has no way of determining whether data about variables and relationships have been omitted that others consider important or that could have increased significance in the future. An incomplete report of a review increases the chance that it will become obsolete quickly (Cooper, 1980).

Summary

This chapter has included a discussion of review methodology and the considerations and decisions associated with each phase. Current

literature on review methodology suggests that a valid and replicable integrative review should include the following: clearly defined conceptual and operational variables, precise sampling procedures and standards for inclusion of studies, consistent coding and clear representation of the characteristics and findings of the studies, a considered combination of quantitative and meta-analytic techniques of data analysis, a discussion of the review findings which attends to variations in results as well as consistencies, and recommendations which relate to practice, further research, and to review methodology itself. The most important aspect of the review, however, is that the reviewer pay fastidious attention to detail in both conducting and reporting the review.

Chapter 3

Computers and Nursing

> The computer is an agent of change. Its uncontrolled
> application may metamorphose our institutions - not in
> beneficial ways that correct existing problems, but in
> ways that we do not intend.
>   (Covvey & McAlister, 1980, p. 7)

In the past two decades, computers have become increasingly available and accessible for instructional purposes. The first section of this chapter reviews changes in society that have related to the proliferation of computers. The implications of these events for education are discussed. Next, an overview of the impact of computers on health care in general and on the nursing profession in particular provides a framework for discussing the current and future uses of computers in nursing. This overview will make it apparent that nursing education programs should be offering instruction in ways that will prepare nurses to function in computerized environments.

Current computer applications in education will be considered in terms of the purposes of the programs offered, the types of computer assisted learning being implemented, and the effectiveness of these programs. The type of information that the education system should be providing to students so that they are comfortable with the technology will also be discussed.

Next, current issues of computers in nursing education will be

compared to those in more general educational settings in terms of the purposes, types, and effectiveness of computer assisted learning programs. A summary of the current research findings on computer assisted instruction in nursing education will provide the basic framework for the integrative review presented in Chapter 4.

Computers and Society

Although the first electronic computer was introduced in the 1940s, the economic shift from an industrialized to an information-based society and the rapid technological developments associated with the changes began in the 1950s. Since that time, the energy crisis, the concern over depleted natural resources, the development of microcomputers and satellites have all helped to foster the growth of a society which relies heavily on technology.

Today, the amount of information available is doubling approximately every 10 years and the largest companies in the United States -- AT&T, IBM, banks, publishers -- are ones whose services almost solely are to process information. Manufacturing computers has become the seventh largest industry, and computer-related jobs are the second fastest growing occupation. Mass production of microcomputers has meant that they are affordable for small businesses, homes, and schools (Merchant & Sullivan, 1982-83). In fact, today's washing machines and ovens often include a computer, in the form of a silicon chip, as a fundamental component.

The rapid introduction of computers into most aspects of individual and societal life has left us with little time to consider the impact of the changes that are taking place and to develop policies to take maximum advantage of them (Ide, 1980). Concern is now being expressed for more cautious implementation of computer technology so that we can study the effects this is having, and so that more people can become knowledgeable and participate actively in the developments. There is a fear that, without an educated populace, both information and computer applications will become controlled by a computer elite (Barton, 1977; Covvey & McAlister, 1980; Miller, 1983).

Molnar (1983) states that computer literacy has become as important as reading literacy. Educational programs should teach students how to make computers work for them rather than focusing on programming skills and the specifics of how computers work. The information explosion also will result in a greater need for continuing education, updating of skills, and job retraining. Educational institutions should be planning strategies and developing programs to meet adults' anticipated needs for further education (Smith & Sage, 1983; Strain, 1983). Computer assisted learning is seen as one solution because this instructional method is claimed to provide individualized, self-paced, effective instruction to large numbers of students.

Computers and Health Care

During the 1950s, health care agencies used computerized systems primarily for administrative tasks such as patient charges, payroll, inventory, and statistics. In the next decade, more health care activities became automated, and nurses recognized the computer's potential to improve nursing practice and the quality of patient care. There has been continual development and expansion of computerized hospital systems since that time. However, their implementation has been slow and successful applications are limited (Ronald, 1979). On the one hand, the systems and programs are often designed by computer specialists who have limited understanding of the needs of the health care agency or its workers. On the other hand, health professionals know little of the uses and capabilities of computers.

Current reports indicate that computers now are being used by nurses in many patient care activities (Edmunds, 1982a; Powell, 1982; Tamarisk, 1982; Yucha & Reigeluth, 1983). One of the first applications in this vein was automated patient monitoring. The computer is connected directly to instruments or other data gathering devices placed on the patient and is used to collect and interpret information on vital functions. The nurse, freed from the role of a technician, then becomes able to devote more time to direct patient care (Badura, 1980; Hannah, 1976; Johnson, Ranzenberger, Herbert, Gardner, & Clemmer, 1980).

When computers are used to record observations about patients'

conditions, it has been found that less time is spent in charting (Hughes, 1980). Computerized recording and integration of data also allow for better assessment, planning, and evaluation of patient care. Quality assessment and assurance efforts are facilitated as both standard and individual care plans can be generated by the computer, and the data stored about the patient care can be accessed easily for nursing audits. As a result of the more systematic planning and recording, patients receive more consistent and higher quality care (Cook, 1982; Edmunds, 1982a).

A part of the nurse's role in hospitals involves coordinating activities and communicating with other departments. When inter-departmental communication is computerized, the speed and accuracy of communication is increased, duplication of work is decreased, more complete records are obtained, and the nurse is relieved of many clerical duties (Cook, 1982; Hannah, 1976).

Computers have also been used as an informational and educational resource for staff. Current information on drugs, treatments, or procedures can be obtained quickly when it is stored as part of the system. As well, inservice education can become more effective and can reach more staff members. Computer assisted learning programs can be used at any time and, unlike lectures or workshops, are always available for review (Pogue, 1982; Tymchyshyn & Helper, 1981).

Some hospitals have also developed computer assisted learning programs for patients, either for specific health problems or for

general health education. Not only are these programs effective in promoting patients' learning, they also free nurses from time spent giving information to patients. This allows nurses to act as a resource and facilitator of learning (Gold & Duncan, 1982; Lyons, Krasnowski, Greenstein, Maloney & Tatarczuk, 1982).

One of the obvious uses of computerized systems is for research in nursing. Because of the storage capabilities of the computer, the selection of samples for study and access to relevant data is facilitated. Staff members can gain a greater appreciation and understanding of research, especially when they are involved as the gatherers of data (Edmunds, 1982a).

One of the administrative uses of computers, staff scheduling, also has had positive effects on the nurse's role. Hospitals have found the task of staff recruitment easier because they are able to predict staffing needs systematically. As well, the computer helps plan for adequate ward coverage and allows for greater flexibility in working hours. The result is increased job satisfaction for employees (Ballantyne, 1979; Duraiswamy, Welton & Reisman, 1981).

Implementation of computerized systems in health care agencies has progressed slowly despite the many reported benefits. The first systems were often designed by people with little knowledge of nursing or of the health care system. These applications sometimes created more work, rather than reducing it, because of the increased volume and complexity of information needed. Many agencies spent time

re-inventing systems because there was limited sharing of experiences among institutions (Powell, 1983). Many nurses know little about computers and have been resistant to automation. They feel threatened by the unknown and think that using computers is complicated and difficult to learn (Edmunds, 1982b; Hannah, 1976; Zielstorff, 1980).

Others suggest that approaches to patient care and the very nature of nursing are in jeopardy because of automated systems (Birckhead, 1978; Farlee, 1978; Zielstorff, 1978). Computers allow nurses the possibility of treating patients as numbers, of using data inaccurately, of developing an overreliance on machines to make decisions, and of violating the patient's right to privacy (Levine, 1980). "As the number of 'menial' tasks is decreased, the nurse may be assigned larger patient loads, to the point where knowledge of patients either as people or 'cases' becomes impossible" (Tamarisk, 1982, p. 48). The quality of patient care may also suffer if nurses spend less time communicating with patients, relying more on standard care plans instead of looking for personalized, innovative approaches to patient problems (Levine, 1980; Tamarisk, 1982). The introduction of computers may help nurses to identify current practices which are unnecessary or are not founded on solid reasoning; but, at the same time, the process of clarifying the nurse's role may result in increased formalization and decreased independence and flexibility in practice (Farlee, 1978; Zielstorff, 1980).

In recent years, the number of computer systems in hospitals has

grown rapidly. This has forced the nursing profession to focus on identifying exactly what nurses need to know in order to function effectively in an automated environment. The first and most obvious task is to help nurses become "computer literate", to increase their knowledge of the functions and limitations of computers. Hardin and Skiba (Skiba, 1983) have developed a framework for computer literacy for nurses based on a modification of the works of the Minnesota Educational Computing Consortium (MECC) and the International Federation for Information Processing.

The MECC Computer Literacy Project developed six categories of objectives which they felt covered the total domain of computer literacy. The first five categories were related to cognitive learning: hardware, programming and algorithms, software and data processing, applications, and impact (careers, personal privacy, ethics of use). The sixth area focused on affective objectives related to attitudes and values toward computers and motivation to use them (Klassen, 1983).

The Fourth Technical Committee of the International Federation of Information Processing has suggested three levels of education for health professionals based on their degree of involvement with computers. The first level involves general education about computer hardware and software, methods of information processing, and examples of computer applications and should be given to all nurses and doctors. The second level, designed for those more involved in design

and implementation of automated systems, includes greater detail about hardware architecture, programming functions and languages, systems analysis techniques, and medical applications. The third level would include intensive training in both computer and health sciences to prepare specialists who would be involved in the actual application of computers to health care systems (Zielstorff, 1980). The computer literacy framework proposed by Hardin and Skiba combines these two models, as well as incorporating both short-term and long-term learning needs of nurses (Skiba, 1983).

Hannah (1983b) and Ronald (1983) have suggested that is should be the responsibility of nursing education programs to teach students about computers so that they are prepared to participate in the development, implementation, and evaluation of computer systems. Hardin and Skiba's framework could serve as a guide for developing these computer literacy programs.

Another effective way to help nurses become "computer literate" is to expose them to the technology. Whenever nurses have worked with hospital-based systems to provide patient care or used computer assisted learning to increase their knowledge, the exposure alone has helped them to become familiar with and less threatened by computers (Butters, Feeg, Harmon & Settle, 1982; Lee, 1983). However, neither of these applications promotes positive attitudes toward computers unless they are perceived to be effective (Badura, 1980; Porter, 1978; Schleutermann, Holzemer & Ferrand, 1983).

Many nurse educators (Conklin, 1983; Pogue, 1982) are recommending increased use of computer assisted instruction in both basic and continuing education of nurses. To date, most of the applications of computer assisted learning have occurred in elementary and secondary education; however, there is increasing evidence that this may be an effective teaching method to use in nursing education as well.

Computers and Education

During the 1960s and 1970s, the majority of computers were found in colleges and universities where they were used for administrative purposes, research, and to teach computer science. There were few institutions using computers for instruction because of the teachers' lack of knowledge of computers and the expense of developing instructional programs. The introduction of microcomputers not only made computers more affordable and accessible, but also created a public demand that education about computers be incorporated at all levels of instruction (Tinker, 1983; Tucker, 1981).

In Canada, computer applications in education have followed the same trends as those in the United States. Computer assisted learning systems are being implemented in an attempt to provide more cost-effective instruction that can be adapted to individual learning styles and needs. As well as having courses to familiarize students with computers and to develop computer skills, educational

institutions are using computer assisted learning as an adjunct to traditional classroom instruction. However, because education is a provincial responsibility and there are no national standards for computers in schools, each jurisdiction has been left to make its own decisions about which computer applications are desirable. This approach has limited the interchangeability of systems and software between provinces. It has also led to a lack of Canadian content in computer programs because of heavy reliance on software developed outside the country (Science Council of Canada, 1981).

In British Columbia, the current emphasis is on the introduction of microcomputers. By 1982, 89% of secondary schools in the province offered an elective course in computer science and literacy, and the rest had plans to implement the course in 1983-84 (Wuhrer, 1983). Computers were also being used to teach mathematics, language arts, business education, and science (Jones, Porter & Rubis, 1983).

Computer assisted learning may take a variety of forms, but all may be categorized as either adjunct or primary computer assisted learning (Chambers & Sprecher, 1983). Adjunct CAL occurs when the computer is used as a supplement to other forms of instruction. An example of adjunct CAL is a drill and practice program. If computer assisted learning replaces other forms of instruction, it is called primary CAL. In this case, the program is a self-contained unit which presents new information to the student, as in a tutorial. While the distinctions between these two types of instruction have not been

emphasized in the literature, each may have definite implications for the effectiveness of student learning. According to Chambers and Sprecher (1983), the bulk of research studies to date have concerned the use of adjunct CAL. These studies indicate that this method of computer assisted instruction is as effective or better than traditional classroom approaches for improving student achievement, improving student attitudes toward the use of computers, and for reducing learning time. Research studies involving primary CAL tend to achieve results similar to those for adjunct CAL only when there is sufficient human interaction accompanying its use. Without this teacher involvement, completion rates for courses were considerably less than those given traditional instruction and student attitudes toward the computer assisted instruction were not positive (Chambers & Sprecher, 1983).

The most common way to categorize computer assisted learning is according to the purpose for which the computer is being used. Most frequently, the programs are referred to as drill and practice, tutorial, simulation, or utility programs.

Drill and practice programs are designed to reinforce previously learned materials. The computer generates questions or problems and gives the student immediate feedback on the accuracy of each response. In some programs, successive problems are selected based on responses to previous ones, thus structuring the experience to match the individual student's learning needs. Drill and practice programs are

particularly effective in increasing retention and transfer of learning (Strain, 1983).

Tutorial programs are used to present new material to the student. Usually, information is presented in small segments and then the student is tested for comprehension of the material. The student's response is evaluated and appropriate feedback is given immediately. For incorrect responses, the student may be given prompts, hints, or additional information to help understand the concept. Although this form of instruction is similar to programmed instruction texts, the CAL program can be more flexible and adaptive to student's learning needs if the selection of content in successive frames is based on responses given to previous questions rather than presenting only one means for progressing through the material.

Simulations are programs developed to help students learn decision making skills or to manipulate variables in order to achieve some preset goal. The student is presented with a "real life" situation requiring a solution. As each decision is made, the student can observe the effect of the action on the situation. Simulations are particularly useful in conditions where, in real life, equipment costs are high, the dangers or costs of errors are great, or when opportunities for practice do not occur frequently enough to develop necessary skills (Strain, 1983). Computer games are an adaptation of simulations, but they may not portray real life situations. Often, they include an element of competing against the computer.

Utility programs are those used for some administrative function such as record keeping or evaluation of learning. When these programs are used to evaluate learning, they may be solely a computer-generated test, or they may be similar to drill and practice programs or simulations. One advantage of using utility programs for testing is that the students receive immediate feedback on the accuracy of their responses and a summary of how well they have achieved the instructional objectives. Using computers for testing and evaluation can relieve teachers of repetitive marking and, at the same time, provide them with summaries of individual students' learning needs.

On the whole, computer assisted learning has been found to have advantages for both students and teachers (Strain, 1983). Well-developed programs can have flexible sequencing which is responsive to the learning needs of individual students. Those who understand the material or who learn it quickly will spend little time on the material, while those who encounter difficulties can be provided with additional information or further practice until they master the content. In either case, the student receives more immediate feedback and reinforcement of learning than can usually be provided by a teacher in the classroom. As well, computer simulations can provide students with opportunities to have experiences and practice decision making in situations that would not be available otherwise in the classroom or laboratory. In situations where there is a shortage or rapid turnover of teachers, computer assisted learning programs allow

for more consistency and continuity of instruction for students.

Research studies on the effectiveness of computer assisted learning have reported on such student outcomes as achievement on examinations, attitudes toward the subject matter and computers, and instructional time taken. Elementary students received higher examination scores when normal instruction was supplemented by computerized drill and practice programs than when only classroom instruction was used (Visonhaler & Bass, 1972). At the post-secondary level, primary computer assisted learning has been found to be as effective as traditional instruction (Jamison, Suppes & Wells, 1974; Kulik, Kulik & Cohen, 1980). For both secondary and post-secondary students, computer assisted learning has positive effects on student attitudes toward the subject matter and computers, and reduces the amount of time that students need to learn the material (Kulik, Bangert & Williams, 1983; Kulik, Kulik, & Cohen, 1980).

For teachers, use of computer assisted learning can help them to make more effective use of their instructional time. Drill and practice programs, particularly those that generate problems randomly, can free the teacher from time-consuming and tedious tasks such as test construction, grading, and record keeping. Then, they can be available to help students on an individual basis, acting as a tutor and consultant. Computers can also be used to reach larger numbers of students without increasing the workload of the teacher appreciably.

Applications of computer assisted learning can be found

throughout the educational system. Molnar (1983) recommends that computers be available at all levels so that people can understand their potential and learn to use them effectively. Colleges and universities, in particular, will need to be responsive to their adult students who have had little opportunity to learn about computers, especially if the current trends toward increased independent study and distance education are implemented through use of computer assisted learning.

Computers and Nursing Education

The use of computers in nursing education has developed slowly, lagging behind that of general education. The first report of computer assisted learning appeared in the 1960s (Bitzer, 1966), but it was not until the late 1970s that discussion of the use and advantages of these programs began to appear more consistently in the literature (Cheung, 1979; Franz, 1976; Levine & Wiener, 1975; Meadows, 1977; Porter 1978). In the same period, nursing education in Canada and the United States experienced a major transition. Nursing programs moved from a hospital-based three-year apprenticeship system to the more academically-oriented two-year college setting. The transition in educational orientation and preoccupation with developing curricula which were appropriate for the new setting may have contributed to lack of interest in instructional innovations.

Faculty in nursing education experienced the same fears and

doubts as teachers in the general education system. Their lack of knowledge of computers and their fear that the technology would take over legitimate teaching functions created resistance to experimentation with this developing instructional method (Ackerman, 1982; Frantz, 1976). Lack of funds for courseware development and, therefore, limited use of computer assisted learning meant that there were few research studies which reported on the relative benefits of this instructional method as compared to traditional techniques.

In the 1970s, however, several factors contributed to the search for and implementation of alternate methods of nursing instruction. The continuing shortage of nurses in the health care system created a demand for increased production of qualified nurses. At the same time, the cost of the low teacher-student ratios for clinical instruction was questioned as it appeared that the expense was greater than the educational system could support. Although nursing faculty were expected to maintain proficiency in the roles of educator, practitioner, and administrator, they had limited time in which to accomplish any one of these (Ackerman, 1982).

As the cost of implementing computer technology decreased and the demands on faculty increased, alternative instructional methods were considered. It soon became apparent that computer assisted learning could be used not only to introduce students to new concepts, but also to provide them with practice in decision making and the application of knowledge without hazard to actual patients (Mirin, 1981). The

development of computerized simulations could relieve stress on the already overcrowded areas in hospitals used for student education (Frantz, 1976). In addition, the research on computer assisted learning in general education lent support to its increasing application in nursing.

Computer assisted learning has been used at all levels of nursing education -- diploma, baccalaureate, graduate, and refresher -- as well as for hospital orientation and continuing education. Computers have been applied to everything from management of the instructional environment (grades, storage of data, scheduling of rooms, student placements) to initial instruction, practice with concepts, patient simulations, and testing of knowledge. At the same time, research on nursing and methods of nursing education has been facilitated by the introduction of computers (Meadows, 1977).

Individual applications of computer assisted learning in nursing education have covered a wide variety of topics, such as: psychiatric nursing (Kamp & Burnside, 1974), midwifery (Naber, 1975), epidemiology (Donebedian, 1976), pharmacology (Timpke & Janney, 1981), immobility (Hannah, 1978), post-operative nursing care (Collart, 1973; Conklin, 1983; Kirchhoff & Holzemer, 1979), decision-making (Brennan, 1981; Taylor, 1980), psychomotor skills (Larson, 1981), medical-surgical nursing (Hoffer, Mathewson, Herbert, Loughrey & Barnett, 1975; Huckabay, Anderson, Holm & Lee, 1979; Sweeney, O'Malley & Freeman, 1982; Valish & Boyd, 1975), computer literacy and awareness (Hannah,

1983b; Ronald, 1983).

The majority of the published reports of computer assisted learning in nursing education are narrative discussions of the development and implementation of a program. In most cases, the application was judged to be effective, although these conclusions usually were based on subjective reports of students and teachers rather than on objectively gathered and analyzed data.

The growing body of systematic research indicates that applications of computer assisted learning in nursing have similar outcomes to those found at the college and university level in general education. In undergraduate nursing education, both primary and adjunct computer assisted learning are claimed to be as effective as traditional instruction in terms of student achievement on examinations (Bitzer & Boudreaux, 1969; Conklin, 1983; Hannah, 1978; Kirchhoff & Holzemer, 1979; Larson, 1982). Substantial reductions in the amount of time students spent learning the material have been reported in several studies (Bitzer & Boudreaux, 1969; Hannah, 1978; Larson, 1982). Students who have been exposed to computer assisted instruction developed positive attitudes toward use of computers as an instructional method (Conklin, 1983; Kirchhoff & Holzemer, 1979; Morin, 1983; Schleutermann, Holzemer & Farrand, 1983).

As the number of research studies grows, however, it becomes increasingly difficult to identify whether these results are limited to the situation and students on which the studies were conducted, or

if these are widely generalizable to computer assisted learning in nursing education. The studies have been conducted in a variety of settings, using differing levels of students, and covering many subject areas. As well, many different outcome measures have been used to reach the reported conclusions. To date, there has been no systematic review of the research findings to draw conclusions and identify generalizable trends regarding the effectiveness of computer assisted learning in nursing education. The next chapter reports an integrative review conducted to determine if computer assisted instruction in nursing is as effective as traditional methods in terms of student learning, retention and application of learning, attitudes, and time taken to learn. Other variables which could influence the effectiveness of instruction, such as cognitive style, are also considered. Discussion of the results of the review and recommendations for nursing education, research, and future reviews are presented in Chapter 5.

Chapter 4

Methodology and Results

The literature reviewed in Chapter 2 indicated that several specific methods for reviewing and synthesizing research outcomes could be identified. Although these methods diverge, it is agreed upon consistently that all phases of the review methodology should be reported in detail. This allows the reader to judge the comprehensiveness and facilitates replication or extension of the review.

This report will follow Jackson's (1980) guidelines for integrative reviews. The first section of this chapter presents the questions to be addressed and defines the variables involved. Next, the procedures used in locating studies and determining which to include in the review are outlined. A summary of the characteristics and findings of the selected studies precedes the analysis of data from the primary studies. Finally, the results of the data synthesis, based on both qualitative and quantitative measures, are presented. Chapter 5, which follows, discusses the conclusions and recommendations resulting from this integrative review.

Selecting the Questions and Variables

The literature about computer assisted learning in general, reviewed in Chapter 3, indicated that effectiveness could be defined

in several ways. Measures of effectiveness which have been used by researchers include student achievement on examinations; ability to transfer knowledge to other situations; student attitudes toward the subject matter, computers, and computer assisted learning; the amount of time students needed to learn; and cost of production of the materials (Edwards, Norton, Taylor, Weiss & Dusseldorp, 1975; Jamison, Suppes & Wells, 1974; Kulik, Bangert & Williams, 1983; Kulik, Kulik & Cohen, 1980; Visonhaler & Bass, 1972). Synthesizing these approaches, the present investigation of computer assisted learning in nursing addresses four questions:

1. Do nursing students using computer assisted instruction learn more information than those given traditional instruction?

2. Are nursing students who learn through computer assisted instruction able to retain or apply this information in other situations (e.g. clinical practice or simulations) better than those given traditional instruction?

3. Do nursing students who learn through computer assisted instruction take less time to learn the same material than those given traditional instruction?

4. What other factors, such as learning styles, attitude, aptitude, age, cost, influence the effectiveness of

computer assisted learning?

For this study, the following operational definitions were created:

1. Computer assisted learning (CAL). Any educational techniques that rely on a computer to facilitate learning. CAL can be subdivided into computer assisted instruction and computer managed instruction (Hannah, 1983a). Both of these terms have been defined in Chapter 1.

2. Traditional instruction. Any educational techniques that do not use a computer to facilitate learning. These techniques may include lecture, discussion, slides, films, readings, printed learning modules, or any combination of these.

3. Achievement. Acquisition of new knowledge or concepts by students measured by written examinations given at the completion of instruction (immediate) or several weeks later (retention).

4. Application. Application of knowledge from instruction which is measured by simulated conditions (e.g. written examinations, programs on a computer) or by observation of performance in clinical situations.

Locating and Selecting the Studies

The first step was to locate and collect primary research studies that investigated the effectiveness of computer assisted learning in nursing education. Searches were made of the following computerized data bases: (a) MEDLINE, 1966 to December 1983, a multidisciplinary index of publications in the health field; (b) ERIC, 1966 to December 1983, a data base of educational materials from the Educational Resources Information Centre, consisting of two files: Research in Education and Current Index to Journals in Education; and (c) NTIS, 1966 to December 1983, the index of reports on government-sponsored research, development, and engineering of the National Technical Information Service. As well, manual searches were conducted of the following indexes: Comprehensive Dissertation Abstracts (1975 to March 1984); and the Cumulative Index to Nursing and Allied Health Literature (1966 to April 1984). The keywords used for these data searches were autoinstructonal aids, automation, competency based education, computer assisted instruction, computer assisted learning, computer oriented programs, computers, Dial Access Information Systems, educational technology, health education, instructional development, lifelong learning, medical education, nursing, nursing education, printed materials, programmed instruction, program evaluation, research, and teaching methods. The actual keywords used for each search were subsets of this list matched to the particular data base being accessed.

Citations in articles located through the computer and manual searches provided a second source of research studies for the review. As well, bibliographies that have been prepared on computer applications in nursing were used. These bibliographies were obtained from the Registered Nurses' Association of British Columbia (a pamphlet file of bibliographies on various topics, one of which was related to computers and computer assisted instruction) and from Hannah and Conklin (1983).

As a result of the various literature searches, 27 studies or reports of studies were obtained. At this point, standards were set for determining which primary research studies would be included in the integrative review. First, the study had to be conducted in an educational setting rather than in an experimental laboratory. It was decided that only studies from educational settings would provide information on the effectiveness of computer assisted instruction as it occurs in practical applications. In this review, all of the studies obtained took place within an educational context. Second, the application being researched had to involve computer assisted learning by nurses or nursing students. In this case, the literature search had eliminated all irrelevant studies. Third, the design of the research study had to include experimental and comparison groups and had to report some findings based on objective outcome measures (scores on achievement tests, numerical data on the relationship between variables studied, estimates of time the students spent

learning, development or operating costs of computer assisted instruction). Seventeen of the 27 studies obtained did not meet the third standard. Therefore, the sample was reduced to 10. The 17 studies that were eliminated were narrative discussions of the effectiveness of computer assisted instruction. Although many of them indicated that their conclusions were based on objective data, no such data were presented in the materials.

In order to increase the sample size, attempts were made by telephone to contact the senior author of each of the studies that had been eliminated in order to obtain the data needed for including the study in this review. Only two further studies which met the criteria were obtained.

In some cases, the same study was reported in several independently published papers. When this occurred, the report that contained the most complete data on the study findings was selected for the review. This was done to ensure that each primary study was counted only once in the analysis. One study (Bitzer & Bitzer, 1973) was eliminated by this process. Therefore, the sample for this integrative review contained 11 studies.

Two articles (Schleutermann et al., 1983; Valish & Boyd, 1975) described more than one experiment on the same group of students and reported the results from each of these separately. Since the same students were used, the studies could be viewed as having a repeated measures research design rather than testing independent treatments.

Usually, the unit of analysis for integrative reviews is the study rather than the treatment (Glass, McGaw & Smith, 1981; Jackson, 1980). As a result, the findings in each study related to student achievement were pooled by calculating the average of the means for the experimental and comparison groups. If this procedure had not been followed, each of these studies would have been counted more than once in the review.

Characteristics of the Primary Studies

Representing the characteristics of the primary studies involves developing categories to describe both the features of the primary studies and their findings. This process of quantifying the data from the research studies allows particular factors which may have influenced the outcomes to be identified.

Describing the features of the primary studies provides the reviewer with data that can help explain particular review findings. For the purposes of this review, a coding sheet was created to quantify the information about the 11 studies selected (see Appendix A). Table 1 summarizes the studies in terms of the 19 categories developed to describe features of publication, aspects of the application of computer assisted learning, the setting of the study, the research design and methods, and the outcome measures. All studies were coded by the reviewer.

Table 1

Categories for Describing Studies and
Number of Studies in Each Category

Coding Category	Number of Studies
Publication date: 1966-69	2
1970-74	0
1975-79	5
1980-84	4
Publication form: Journal	9
Book	1
Unpublished	1
Country of origin: Canada	2
USA	9
Computer use: Tutorial	9
Simulation	2
Implementation: Substitute	9
Supplement	2
Subject: Medicine	7
Surgery	2
Psychiatry	1
Obstetrics	1
Approximate time on CAI: Less than 3 hr.	5
3 - 6 hr.	1
More than 6 hr.	1
Not stated	4
Software and course development: Local	8
Others	1
Not stated	2
Study setting: Hospital leading to RN	2
College leading to RN	0
University - baccalaureate	4
- graduate	2
Hospital inservice education	3

Table 1 (Continued)

Coding Category	Number of Studies
Length of study: 0 – 4 weeks	1
4 – 8 weeks	2
More than 8 weeks	6
Not stated	2
Experimental design: Pretest – posttest	7
Posttest only	3
Other	1
Sample: Volunteer	2
Convenience	9
Sample size: 1 – 25	3
26 – 50	5
51 – 75	0
76 – 100	2
More than 100	1
Assignment to groups: Random	7
Matched	3
Uncontrolled	1
Historic effects: Same time period	9
Different time period	2
Instructor effects: Same instructor	6
Different instructor	0
Not stated	5
Pretest: Knowledge	8
Ability	4
Cognitive style	2
Attitude	2
Demographic	4
Application/transfer	2
Outcome: Knowledge/learning	10
Retention	1
Application	2
Time	7
Attitude	9
Cost	2

Table 1 (Continued)

Coding Category		Number of Studies
Achievement measure:	Instructor developed	10
	Commercial	0
	Not measured	1
:	Based on lesson	5
	Not based on lesson	5
	Not measured	1

Publication features. The information on publication dates of
the studies was collected so that the developmental trends for
research conducted on computer assisted learning in nursing education
could be identified. The form of publication and country in which the
research was conducted provide data useful in describing the field of
research. Since no theses or dissertations were obtained for the
review, this descriptor was not used.

Application of computer assisted learning. Several variables
describe the application of computer assisted instruction. The first
of these covered the type of computer assisted learning used, i.e.
drill and practice, tutorial, simulation, and utility. Each of these
applications was described in Chapter 3. None of the 11 studies used
drill and practice or utility programs.

Another variable considered was whether the computer was used as
a supplement to or a substitute for traditional instruction. In
studies in which the computer served as a substitute, it replaced all
other forms of instruction such as lectures, readings, and independent
learning modules. In studies where the computer was used as a
supplement, it did not replace other forms of instruction but was used
to provide additional instruction to students in the experimental
group.

Next, studies were grouped according to the subject area of the
instruction -- medicine, surgery, psychiatry, obstetrics, or

pediatrics. None of the studies used instruction related to
pediatrics.

The data on the approximate duration of the computer assisted
learning unit is based on information given in each study. This data
was provided either as the average time that students took to complete
the unit or as the maximum length of time that it took to use the
computer assisted instruction. Because the groupings within this
category were fairly broad, interpretation of results reported
subsequently are correspondingly less precise.

The last variable considered was where the software for the
instructional unit had been produced. If it was developed by the
author(s), either alone or in conjunction with others, the software
was considered to be locally produced. When the instructional unit
was obtained from another institution, the software was categorized as
being produced by others.

Study setting. Another variable used to describe the primary
studies was the setting in which the research took place. Information
was categorized according to the type of institution and the level of
student.

Research design. Seven variables were used to describe the
research design and methods. First, the studies were categorized
according to the duration of the treatment. While some studies took

only a few weeks, others were conducted over several semesters.

Next, the research design was considered. Studies were coded according to Campbell and Stanley's (1963) classifications. Only two categories were necessary as studies had either a pretest-posttest or a posttest only design.

Third, each study was categorized as to how the participants were obtained. In some cases, volunteers from one or more classes were used; in others, the participants were all members of a particular class or group, i.e. a convenience sample. Each study was then coded according to the total number of participants obtained.

Data was also coded as to whether students were assigned randomly to experimental and comparison groups or whether the groups were created by matching participants in these groups on some variable such as age or ability. As well, each report was classified according to whether the two groups were involved in the study concurrently. In some cases, the class forming the experimental group followed in the semester after the comparison group.

Another variable related to the research methodology was whether the same instructor taught both the experimental and comparison groups. It was considered important to code information on this variable, a control for instructor effects, because two previous reviews of research (Kulik, Kulik & Cohen, 1980; Kulik, Kulik, & Cohen, 1979) had reported that this was the only study characteristic that had a statistically significant correlation with study outcomes.

However, none of the 11 studies reported that different instructors were used for each group. Thus this variable could not be considered in the data analysis.

For those studies using a pretest-posttest design, the measures used for the pretest were examined. Pretest instruments were coded as to whether they were used to collect data on students' prior knowledge of the subject matter, ability or aptitude, cognitive style, attitudes to computer assisted instruction, ability to transfer the subject matter to other situations, and demographic information such as age, previous experience, and educational background.

Outcome measures. Two variables were related to the outcomes measured in the reports. The outcomes studied were categorized according to whether they gauged student learning, retention of knowledge, transfer or application of knowledge to other situations, time needed for student learning, or student attitudes. For both the pretest and outcome measures, the number of entries in Table 1 exceeds the number of studies because several reports collected information on more than one variable.

Studies were also coded to describe who had developed the instrument used to measure student achievement. In all cases, the instructor of the course had developed the test. No commercial achievement tests were used.

Of the studies that measured student achievement as an outcome,

all of them described how the test was developed and what information was tested. As a result, it was possible to code these outcome measures according to whether they were based specifically on information in the instructional unit or whether they tested the topic in general.

Methodology for the Data Analysis

This section describes the methodology used for the meta-analytic and qualitative analysis of the studies' findings on the effectiveness of computer assisted instruction versus traditional instruction. Basically, the data analysis was conducted using estimates of effect size for objective outcomes combined with Ladas' (1980) microanalytic approach to explaining statistically unreliable or negative findings.

The effect size (ES), defined as the difference between the means of the experimental and control groups divided by the standard deviation of the control group (Glass, McGaw & Smith, 1981), was used to quantify the outcomes of each study. For studies that reported means and standard deviations of both groups, ES was calculated from these measures. For less fully reported studies, ES was calculated from the t-statistic using procedures described by Glass, McGaw, and Smith (1981). The effect sizes then were examined in relation to study characteristics to determine if there were any factors that may have influenced the studies' findings.

Previous reviews on the effectiveness of computer assisted

instruction versus other methods (Kulik, Bangert & Williams, 1983; Kulik, Kulik & Cohen, 1980) have correlated effect sizes for achievement with characteristics of studies to determine if any one finding could be explained by differences in the studies themselves. These reviews classified outcomes on student achievement according to the direction and statistical reliability of the differences between computer assisted and other forms of instruction. The following 4-point scale was used for the classification: 1 = difference favoured traditional instruction and was statistically reliable; 2 = difference favoured traditional instruction but was not statistically reliable, 3 = difference favoured computer assisted instruction but was not statistically reliable, and 4 = difference favoured computer assisted instruction and was statistically reliable (Kulik, Bangert & Williams, 1983).

For the current review, correlations between effect size statistics or ratings on the 4-point scale and characteristics of studies had limited usefulness. First, since the number of primary studies in the review was small, correlations between features of studies and student achievement would be quite tenuous. Also, it was recognized that one study with a particularly deviant effect size could produce spurious findings. Second, many categories describing the studies yield only nominal data. Measures of association based on such data are difficult to interpret, especially under conditions of small sample size. Therefore, it was decided not to attempt

correlations between the findings and the study characteristics in this review. The 4-point scale was used, however, to determine if there was a statistically reliable correlation between achievement effect sizes and the statistical reliability of the studies' results.

For most outcomes measured in the studies, a qualitative approach was taken for data synthesis. The reviewer felt that using statistical methods would be both difficult and misleading because of the number of subjective judgments that would be necessary to quantify the results. This decision was made because of the limited reports on many outcomes and the diversity of methods used for measuring them. As well, the studies often contained insufficient data to use any of the other meta-analytic methods discussed in Chapter 2.

The qualitative method selected to examine studies' findings in detail was Ladas' (1980) microanalytic approach. Particular emphasis is placed on examining studies with statistically unreliable or negative findings to see if aspects of their research design or methodology account for these findings. Study aspects such as the difficulty of the treatment, failure to assign participants to groups randomly, demographic characteristics of the students, length of the materials used, explanations given to students, and reliability and validity of the outcome measure may all contribute to obtaining statistically unreliable results.

<u>Results</u>

This section describes the results of the meta-analytic and qualitative analyses of the effectiveness of computer assisted instruction versus traditional instruction. The findings are discussed under the following topics: immediate achievement, retention of knowledge, application of knowledge, instructional time, student attitudes, and other variables influencing effectiveness of instruction. Effect sizes were calculated whenever there was sufficient data in the primary study. The majority of the effect sizes were related to immediate achievement, although it was possible to calculate a few related to retention, application, and attitude. For the data analysis of this review, findings reported in the primary studies were considered statistically reliable at the .05 level. This probability level was selected because it is a commonly accepted level at which to reject the null hypothesis. As well, many studies did not report data unless it reached this level of significance. On occasion, one or more studies reported similar results that might indicate a trend but not all were statistically reliable at the .05 level. In these cases, any results which were reported at the .10 level or less were also considered in the analysis.

<u>Immediate achievement</u>. Ten of the 11 studies measured students' achievement on examinations given immediately after the instruction. Table 2 summarizes the results according to the statistical

Table 2

Statistical Reliability of Study Findings
for Student Achievement on Examinations

	$p < .05$	$p > .05$	Total
Favour CAL	4	4	8
Favour traditional	0	2	2
Total	4	6	

reliability of the findings.

Four of the ten studies reported statistically reliable differences (p < .05) between computer assisted instruction and traditional methods. All of these studies favoured computer assisted instruction. As well, four other studies reported that the average score for those taking computer assisted instruction was higher than that for traditional instruction although these results were not statistically reliable. In the other two studies, students taking conventional instruction had higher scores, on the average, than the experimental group but these results were not statistically reliable. Overall, the box score results favour computer assisted instruction over other methods.

Several other factors related to the combined results were examined in order to describe the effectiveness of computer assisted instruction more precisely. Means scores for the experimental and comparison group of each study were converted to a percentage of maximum scores. One study (Hannah, 1978) contained insufficient data for this calculation and so the sample was reduced to nine. Across the nine studies, the average examination score for the experimental (CAL) group was 67.63% while that for the comparison group was 62.77%. For each study, the difference between the experimental group's mean (as a percentage of maximum) and the comparison group's mean (as a percentage of maximum) was computed. Pooling across the nine studies, a t test for correlated samples revealed a statistically reliable

difference in group means favouring students exposed to computer assisted instruction (t = 2.65, df = 8, p < .05).

Table 3 summarizes the findings and reports effect sizes for immediate achievement by students according to author, year of publication, type of student, the use made of computer assisted instruction, the treatment received by the comparison group, the statistical reliability of the results, and the effect size. The average effect size was .48 with a standard deviation of .66. This average effect size implies that, in a typical class, performance of students receiving computer assisted instruction was raised by .48 standard deviations. Referring to areas of the standard normal curve, one can see that an average student experiencing computer assisted instruction performed at the 68th percentile on their examinations in a distribution where the typical student in the comparison group performed at the 50th percentile.

Because of the small number of studies in the sample, it was important to determine how well different measures of effects agreed. The correlation between the effect sizes and scores on the 4-point scale reflecting direction and reliability of results was .61. This correlation indicates a statistically reliable agreement between the two measures at the .10 probability level. After converting means to percentages of maximum, the correlation between effect size and the difference between means of the experimental and comparison groups was .92 (p < .01). This indicates that the difference in group means is a

Table 3

Effect Size for Achievement with Computer Assisted Instruction Versus Other Forms of Instruction

First Author	Year	Student Level	CAI Use	Comparison Group	Statistics			Effect size[1]
Bitzer	1966	Diploma	Subst.	Lecture	t=1.9	df=12	p=.09	1.02
Boettcher	1981	Baccal.	Subst.	Printed Instruction	t=-.70	df=81	p≐.70*	-.08
Conklin	1983	Baccal.	Suppl.	Modules & reading	t=3.73	df=23	p<.01*	1.95
Hannah	1978	Baccal.	Subst.	Lecture	t=.017		p=.01	.01
Hoffer	1975	Inserv.	Subst.	Lecture	t=.495	df=32	p≐.70*	.18
Huckabay	1979	Grad.	Suppl.	Lecture & reading	t=2.256	df=29	p<.05*	.67
Kirchhoff	1979	Bacc.	Subst.	No treat.	t=4.043	df=98	p<.01	.85
Pogue	1984	Inserv.	Subst.	Lecture	t=1.675	df=25	p<.20*	.52
Schleutermann	1983	Grad.	Subst.	Printed Instruction	t=.0025	df=8	p≐.00*	.00
Valish	1975	Inserv.	Subst.	No treat.	t=-1.57	df=122	p≐.08*	-.32

M=.48 s=.66

Md=.35

*Computed by the reviewer.
[1]A positive ES indicates that the computer assisted group outperformed the comparison group. A negative ES indicates the opposite.

much better predictor of effect size for student achievement than the direction and statistical reliability of results.

Although the average effect size for the 10 studies was not large, these measures ranged in size from very positive (an increase in achievement scores of 1.95 standard deviations for the experimental group) to slightly negative (a decrease in achievement scores of .32 standard deviations, see Table 3). To determine if these findings on immediate achievement could be explained by particular study features, average effect sizes for studies were calculated for each sub-category of the 19 study characteristics coded. These average effect sizes are reported in Table 4.

The initial coding of study characteristics did not include data on the treatment received by the comparison group. As part of the data analysis, however, the comparison groups were categorized according to the treatment they were given: lecture, printed instruction, and no treatment. Average effect sizes for immediate achievement were calculated for each of these categories. These results have been summarized in Figure 1.

Figure 1 illustrates the difficulty one has in interpreting means of effect sizes, especially when the number of studies in the sample is small. For example, the mean of .63 for the comparison group treatment of printed instruction is influenced by the large effect size obtained in one study. If that one study were eliminated, one would conclude that there is very little difference in achievement

Table 4

Average Effect Size for Student Achievement
According to Study Characteristics

Coding Category	Number of Studies[1]	Mean ES
Publication date: 1966-69	1	1.02
1970-74	0	
1975-79	5	.28
1980-84	4	.60
Publication form: Journal	8	.53
Book	1	.01
Unpublished	1	.52
Country of origin: Canada	2	.98
USA	8	.36
Computer use: Tutorial	8	.52
Simulation	2	.34
Implementation: Substitute	8	.27
Supplement	2	1.31
Subject: Medicine	7	.30
Surgery	2	1.40
Psychiatry	1	-.08
Approximate time on CAI: 0 - 3 hr.	5	.93
3 - 6 hr.	1	.52
Not stated	4	-.10
Software and course development: Local	8	.56
Others	1	-.32
Not stated	1	.67
Study setting: Hospital leading to RN	1	1.02
College leading to RN	0	
University - baccalaureate	4	.68
- graduate	2	.34
Hospital inservice education	3	.13

Table 4 (Continued)

Coding Category	Number of Studies[1]	Mean ES
Length of study: 0 - 4 weeks	1	1.02
4 - 8 weeks	2	.82
More than 8 weeks	6	.27
Not stated	1	.52
Experimental design: Pretest - posttest	7	.61
Posttest only	3	.18
Sample: Volunteer	2	.09
Convenience	8	.58
Sample size: 1 - 25	2	.51
26 - 50	5	.66
51 - 75	0	
76 - 100	2	.39
More than 100	1	-.32
Assignment to groups: Random	7	.32
Matched	2	.84
Uncontrolled	1	.85
Historic effects: Same time period	8	.41
Different time period	2	.76
Instructor effects: Same instructor	5	.87
Not stated	5	.09
Achievement measure: Instructor developed	10	.48
: Based on lesson	5	.65
Not based on lesson	5	.31

[1]The sample size for student achievement was 10 studies because one study did not report on this outcome.

between students using computer assisted and printed instruction as
the mean effect size is reduced to -.04. However, as the figure
stands currently, one must conclude that the comparison group
treatment had little influence on the effect sizes obtained.

Effect
Size

2.0
1.8
1.6
1.4
1.2
1.0
.8
.6
.4
.2
0.0
-.2
-.4
-.6

Lecture   Printed   No
               Instruction   Treatment

Number of studies: 5   3   2
   Effect size mean: .48   .63   .27

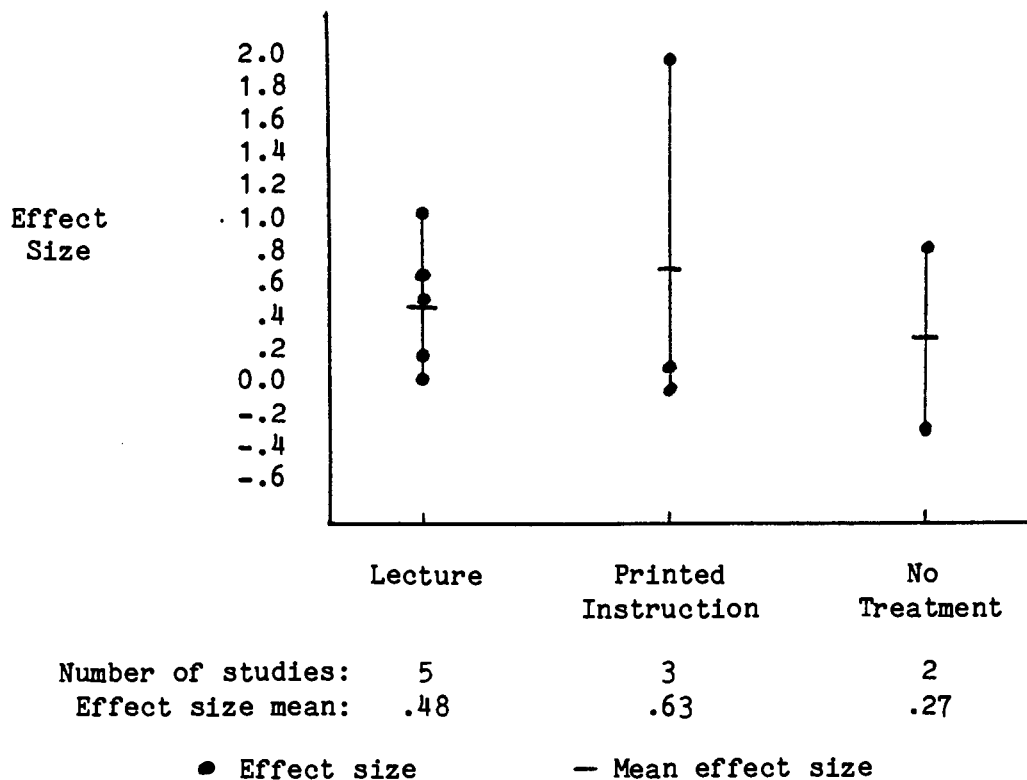● Effect size   — Mean effect size

FIGURE 1. Effect sizes for student achievement according to
comparison group treatment.

In order to see if there were other explanations for the
variation in findings, the study characteristics were considered.
Several of the features could not be interpreted in terms of effect
size because of the small number of studies in one or more categories.

These categories included the form of publication, the length of the
computer assisted instruction and of the study, software development,
historic and instructor effects, and selection of participants. As
well, mean effect sizes were influenced greatly by a single aberrant
study in the categories of country of origin of publication, the
subject area of the computer assisted instruction, and assignment to
groups. The findings related to effect size means and the other study
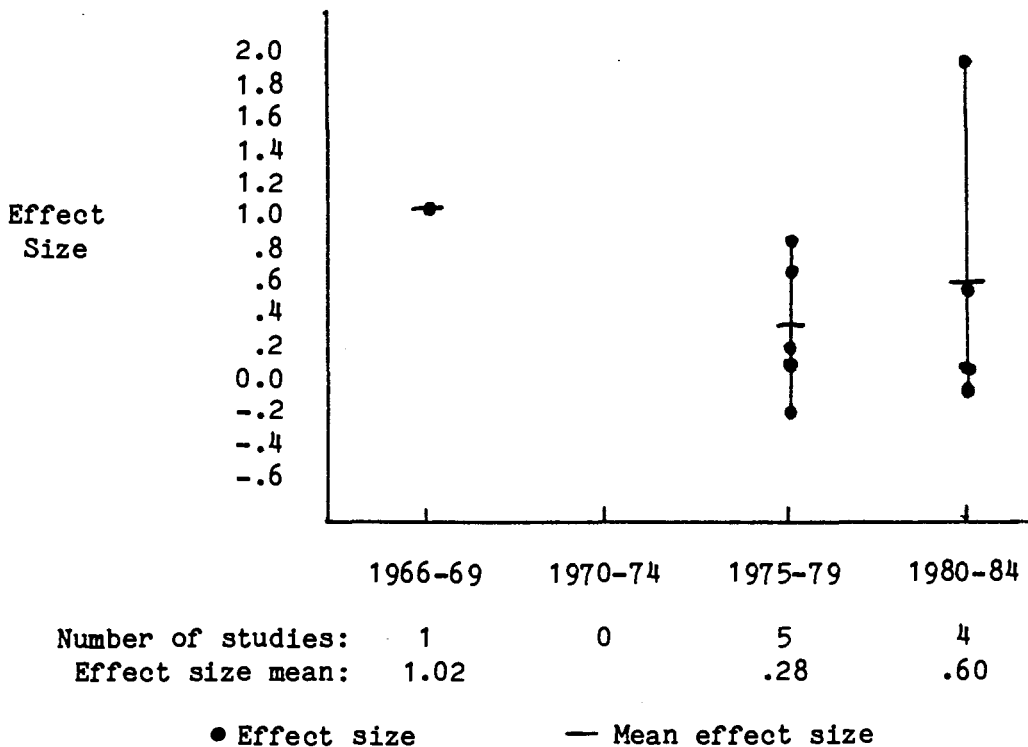characteristics follow.

FIGURE 2. Effect sizes for student achievement by year of publication
of study.

Figure 2 indicates that the year of publication did not influence

the effect sizes achieved. In fact, the mean effect sizes would be almost identical if the single study reported in 1966-69 and the one outlying study from the 1980-84 category were not considered.

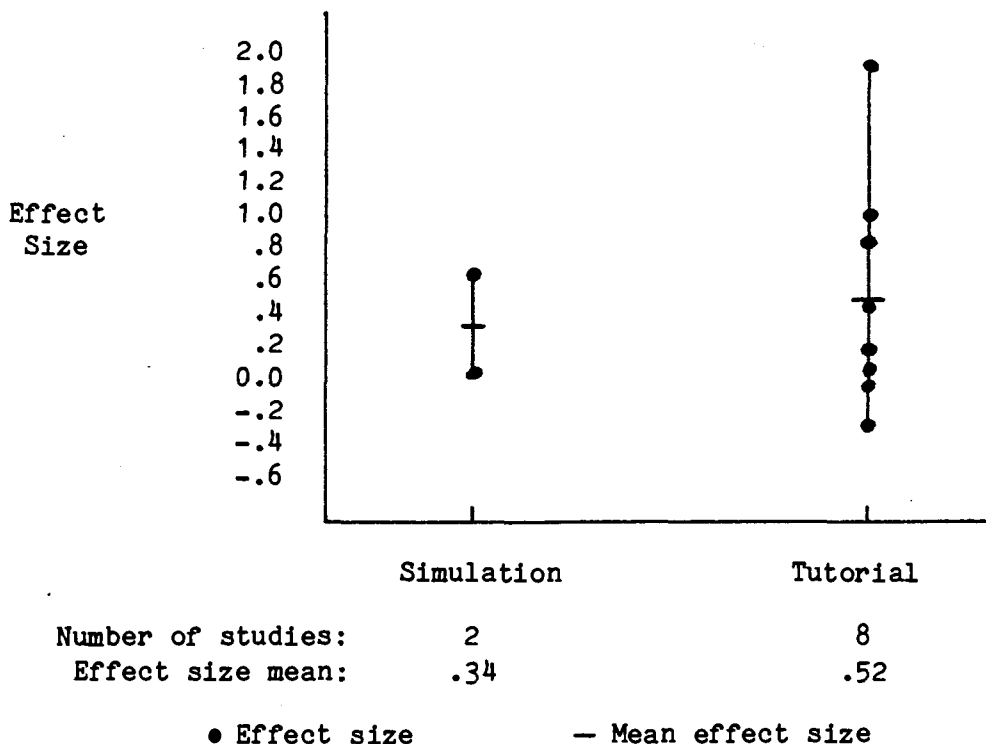|	Simulation	Tutorial
Number of studies:	2	8
Effect size mean:	.34	.52

● Effect size   — Mean effect size

FIGURE 3. Effect sizes for student achievement according to type of computer assisted instruction.

Figure 3 depicts the effect sizes for student achievement according to the type of the computer assisted instruction -- tutorial or simulation. As can be seen, the type of computer assisted instruction has very little influence on the mean effect sizes of the studies.

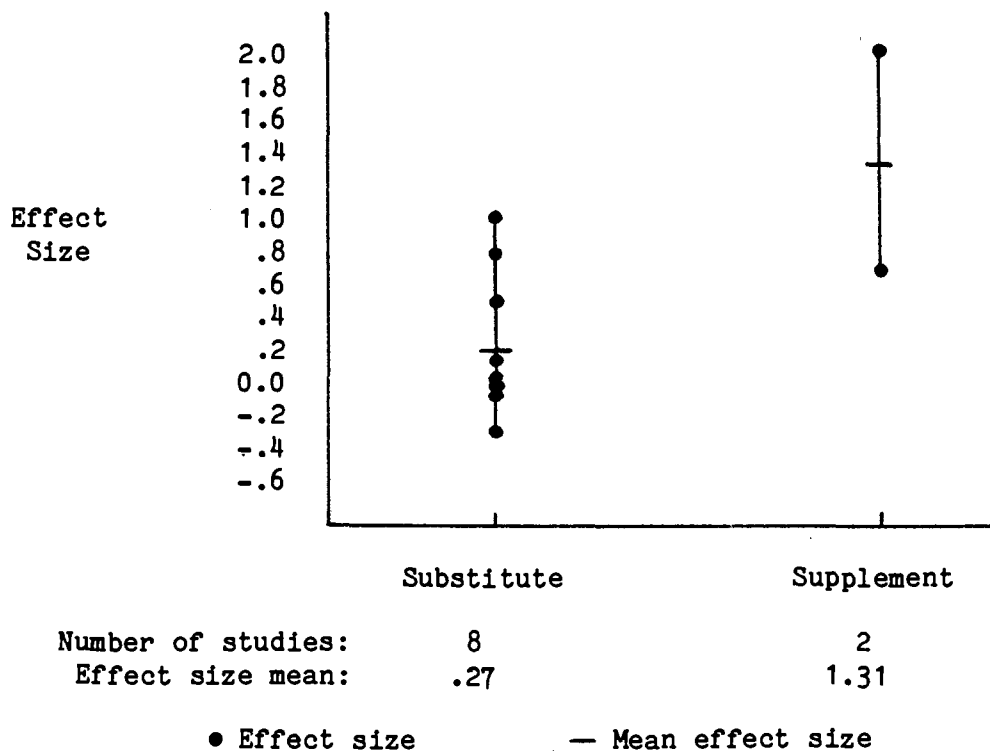|		
	Substitute	Supplement
Number of studies:	8	2
Effect size mean:	.27	1.31

● Effect size   — Mean effect size

FIGURE 4. Effect sizes for student achievement according to use made of computer assisted instruction.

Figure 4 summarizes the relationship of mean effect sizes for student achievement and the use made of the computer assisted instruction. Both the means and the ranges of the effect sizes indicate that student achievement is greater when computer assisted instruction is used as a supplement to rather than a substitute for traditional instruction.

One other aspect that could influence the effect size means on student achievement was the type of student. The results for this

category are shown in Figure 5. Once again, the effect size means would be almost identical if the single study from the hospital group and the one outlying study in the baccalaureate group were not considered.
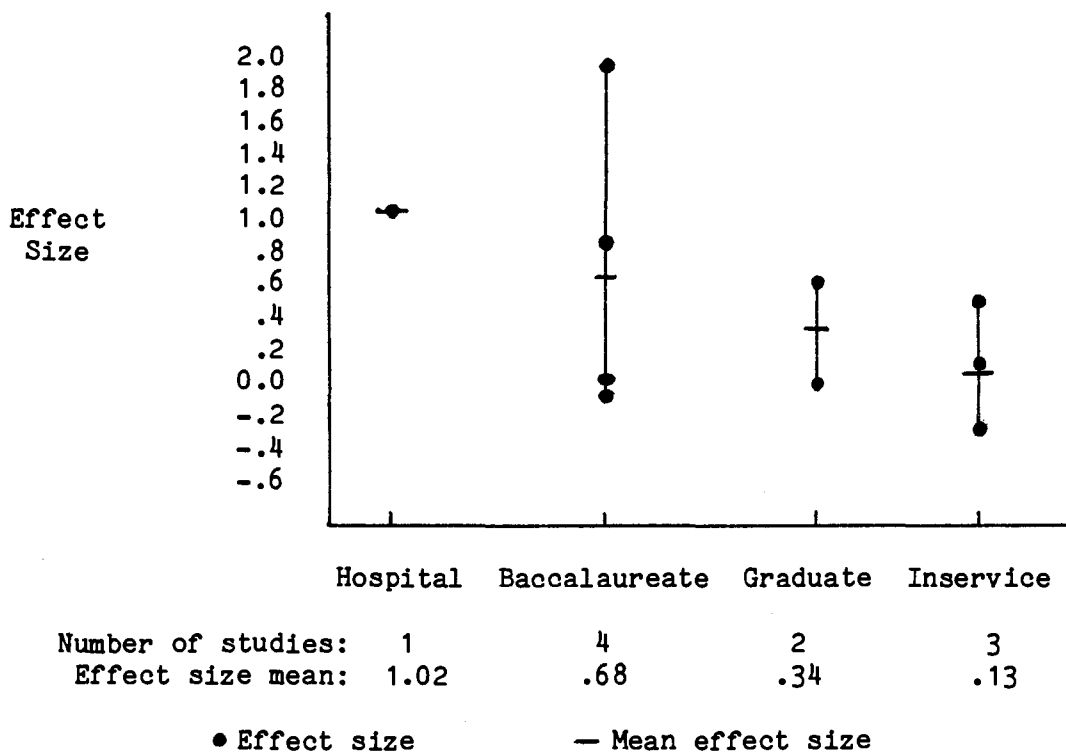

FIGURE 5. Effect sizes for student achievement according to type of student.

In considering the research design, the length of time over which the study took place produced some interesting results which are summarized in Figure 6. The mean effect sizes for the studies that were over eight weeks long were fairly consistent as compared to the mean effect sizes of the studies that were less than eight weeks long.

These shorter studies produced both the largest and the most negative
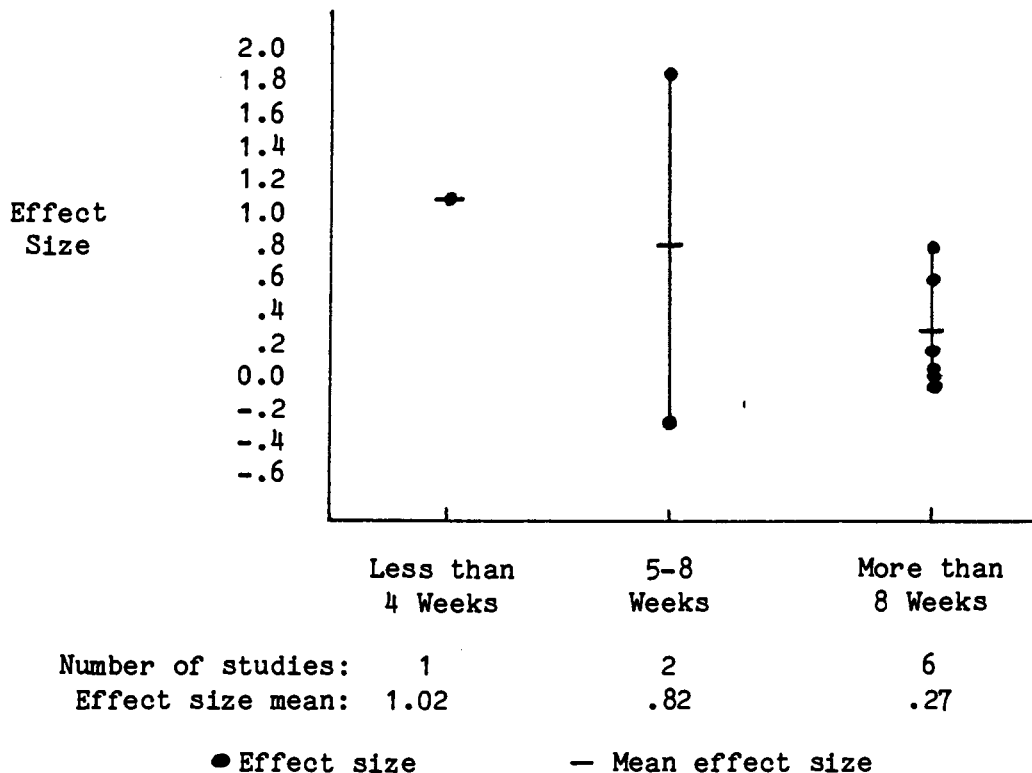effect sizes.

FIGURE 6. Effect sizes for student achievement according to length of
study.

Another feature that influenced effect sizes on student
achievement was whether the study had a pretest-posttest or posttest
only design. Figure 7 summarizes these results. The mean effect size
produced from studies with a pretest-posttest design is somewhat
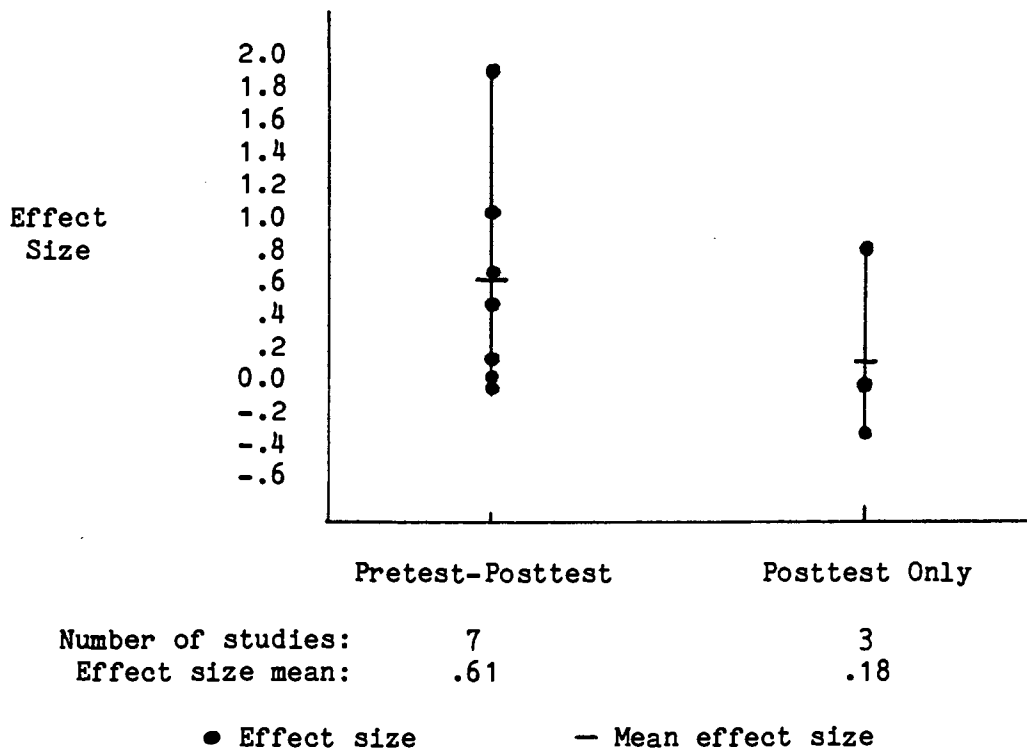larger than that from a posttest only design.

FIGURE 7. Effect sizes for student achievement according to research design.

One aspect influencing whether a study will produce statistically reliable results is the size of the sample. In cases where the sample size is small, one needs a much larger difference between means to obtain a statistically reliable result. Figure 8 summarizes the results of the effect sizes for student achievement according to sample size. As can be seen, the mean effect sizes do not appear to be. influenced by the number of participants in the study.

Effect
Size

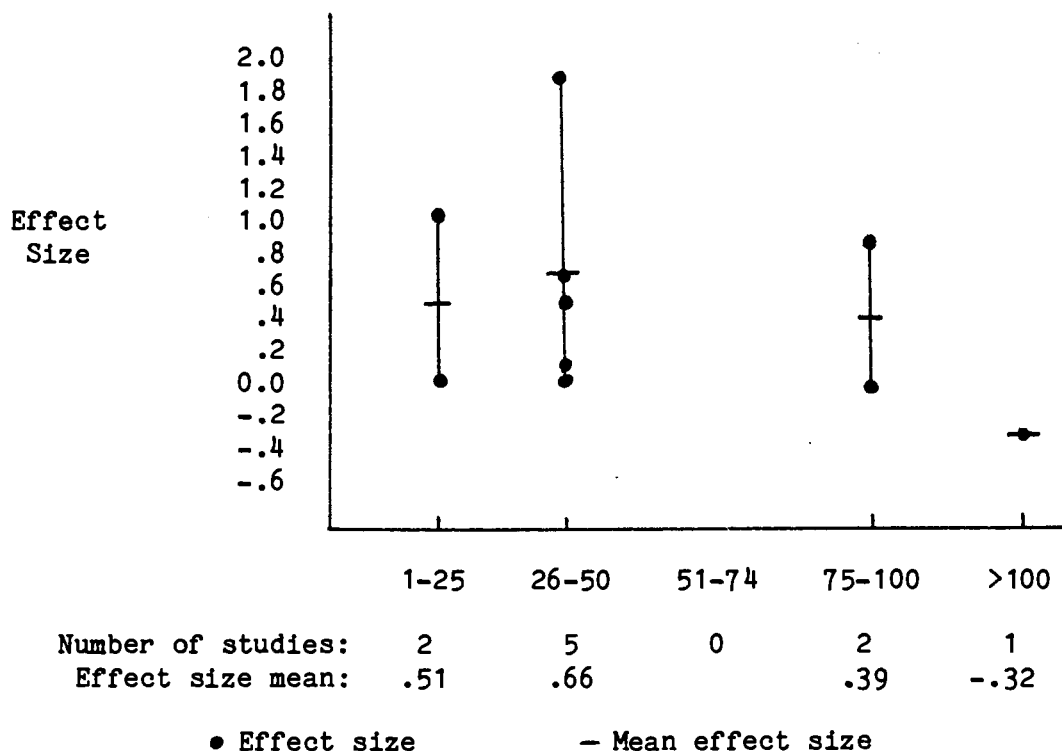|	1-25	26-50	51-74	75-100	>100
Number of studies:	2	5	0	2	1
Effect size mean:	.51	.66		.39	-.32

● Effect size   — Mean effect size

FIGURE 8. Effect sizes for student achievement according to size of
sample.

The final characteristic of studies considered in terms of its

relationship to effect sizes for student achievement was the outcome

measure itself. The studies had been coded according to whether the

achievement measure was based only on material contained in the

instruction or whether it was on the general subject area. Figure 9

summarizes the relation between the content of the test and the effect

sizes for student achievement. Mean effect sizes were larger when the

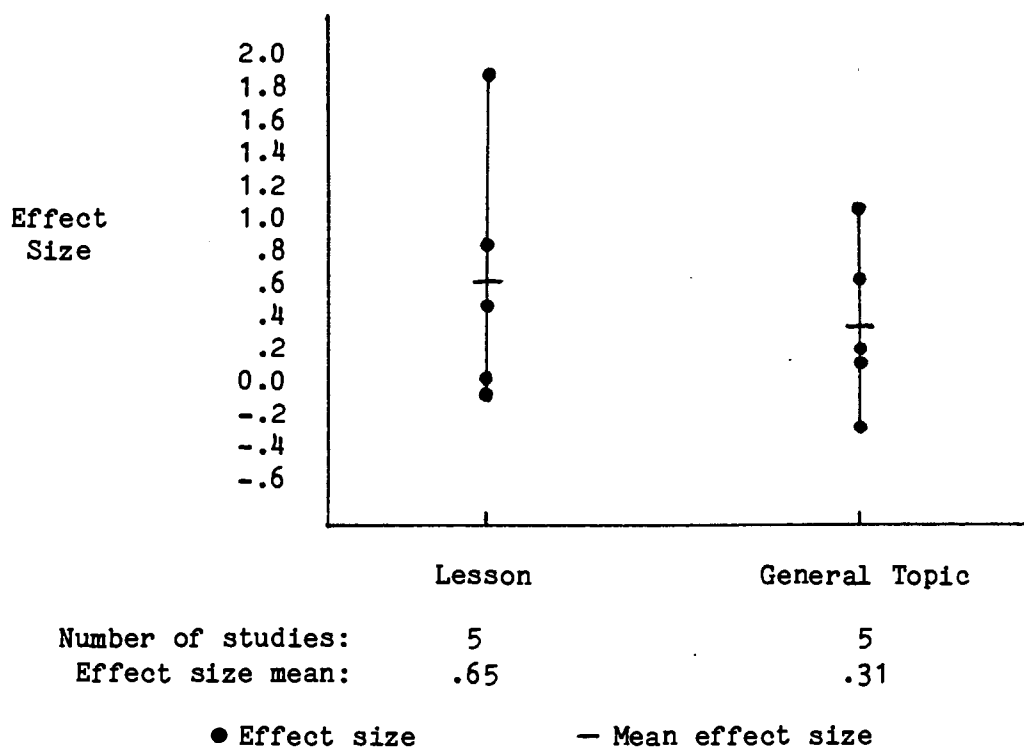test was based only on material in the lesson.

FIGURE 9. Effect sizes for student achievement according to content of the test.

Other aspects related to the achievement measure that could influence effect sizes are the objectivity, reliability, and validity of the test. All of the studies used an objective test to measure student achievement. All of the tests used multiple choice questions and two studies also used true/false and matching items. As well, all of the tests were deemed to have face validity as they were developed either in consultation with the instructor or by a panel of experts.

Only three studies (Conklin, 1983; Huckabay et al., 1979; Kirchhoff & Holzemer, 1979) reported estimates of the test

reliability. In all three cases, the authors felt that the test reliabilities were acceptable. The studies that reported test reliability were examined in relation to studies that did not report this statistic to determine if there were differences in effect sizes on student achievement. The findings are presented in Figure 10.
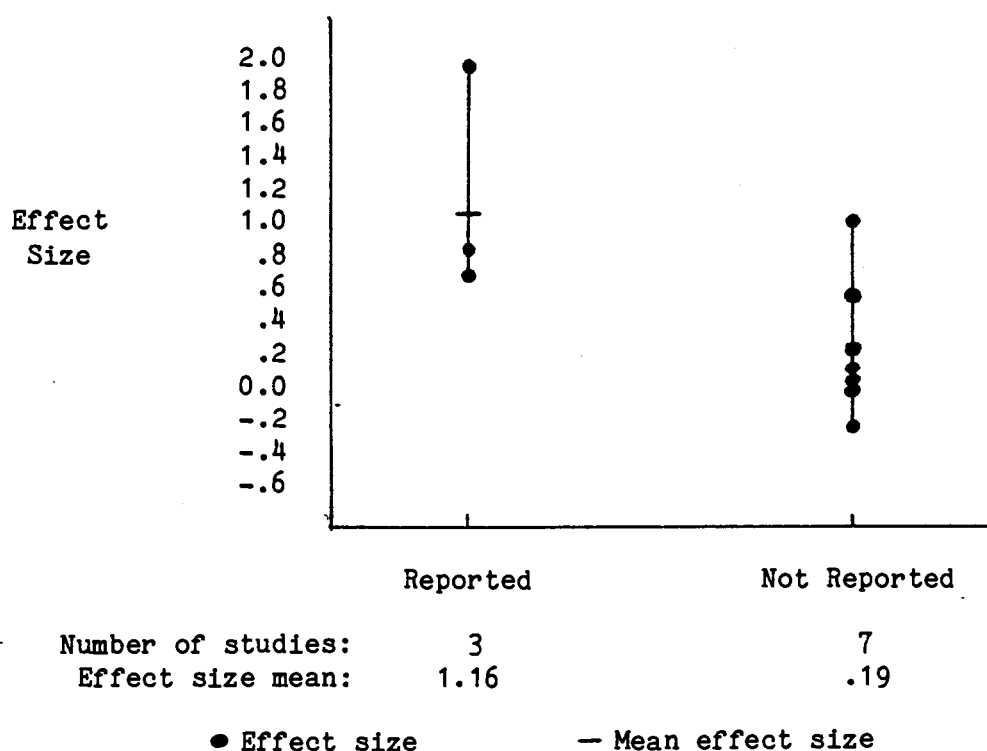

FIGURE 10. Effect sizes for student achievement according to whether test reliability was reported.

Larger effect sizes favouring the students who used computer assisted instruction were obtained from the three studies that used measures having demonstrated reliability. As well, two of these three studies (Conklin, 1983; Kirchhoff & Holzemer, 1979) used psychometric test

development procedures to maximize the measurement of differences between individuals. These two studies have some of the largest effect sizes for student achievement as well as reporting statistically reliable differences between treatment groups that favour students using computer assisted instruction.

In summary, study characteristics which appeared to be related to obtaining larger effect sizes for student achievement were use of the computer as a supplement to instruction, having a pretest-posttest research design, and use of reliable tests to measure outcomes.

Studies having small or negative effect sizes for immediate achievement. The studies that had effect sizes of less than .01 were examined in more detail to see if explanations could be found for the results obtained. Valish and Boyd (1975) explained that their findings, which favoured the comparison (no treatment) group slightly (ES = -.32), were obtained because the unit of computer assisted instruction was on information which both the experimental and comparison groups already knew.

Schleutermann et al.'s (1983) study, in which the effect size was zero, had several problems. First, they reported that the unit of computer assisted instruction was programmed poorly and students became frustrated using it. Second, the total sample size (12) was small. In such cases, it is difficult to obtain statistically reliable differences between groups. They recognized this problem

and, accordingly, did not perform any statistical tests on the group means. Third, one group of students achieved higher scores consistently no matter which method of instruction they received. The research design had two groups of students complete three different situations -- one using computer assisted instruction, one using printed instruction, and one where the student chose one of these two methods. For the first simulation, one group was taught using computer assisted instruction while the other group used printed instruction. The two groups changed methods of instruction for the second simulation. In the first unit, the group using the computer assisted instruction had the higher test scores. In the second, the group using the printed simulations achieved more. If the results for this one group are pooled over these two simulations and compared to the mean scores of the other group, there is considerable difference between the group means. This finding could indicate that the two groups were not similar in ability despite being assigned randomly. Schleutermann et al. did collect data on the students' grade point average and scores on the Graduate Record Examination but did not use this data to check for differences in ability between the two groups. Probably the authors did not do this test due to the small sample size; however, they do not comment on this in the study. In any case, the validity of the study's results should be questioned.

Both Boettcher et al. (1981) and Schleutermann et al. (1983) used printed materials for the comparison group treatment. Boettcher et

al.'s (1981) study was designed specifically to test whether students learned more in the cognitive categories of knowledge and application (Bloom, 1956) using computer assisted or printed instruction. They developed one lesson to teach the knowledge base and another lesson to teach application. Once developed, the lessons were then produced as computer assisted and printed instruction. By developing the materials in this way, they ensured that the instructional logic was parallel in both methods. For example, in the lesson which taught the knowledge component, both the printed instruction and the tutorial method used for the computer assisted instruction first presented some information to the student and then questioned them about their understanding. The only difference in progress through the two lessons was that the students using the computer assisted instruction had to answer questions correctly in order to proceed, while those using the printed instruction could continue without answering the on textual questions. The same approach was used to develop the simulations which taught application of knowledge.

Schleutermann et al. (1983) used a similar approach to develop their computer assisted and printed simulations. The simulation itself was prepared first, complete with all decision options, and then it was produced in both computer assisted and latent image format. Again, the instructional logic was parallel in both methods. A student would progress through the unit in exactly the same way no matter which form of instruction was used. The effect sizes for

student achievement obtained from both studies (-.08 in Boettcher et al. and .00 in Schleutermann et al.) have probably been affected by this approach to preparing the instructional materials. Indeed, if the units of instruction have similar content and parallel instruction approaches, one would not expect to find differences in achievement between computer assisted and printed instruction.

In looking at the characteristics of the studies as coded, there are a few similarities among these three studies. All of them used computer assisted instruction as a substitute for traditional methods. As well, two of the three used printed instruction for the comparison group treatment. The other study (Valish & Boyd, 1975) had an untreated comparison group. None of the three studies reported the reliability of the test used to measure achievement, nor did they mention use of psychometric methods for test development. Finally, two of the three studies were longer than eight weeks in length. Other studies with low effect sizes for student achievement also fell in these categories.

<u>Summary of results for immediate achievement</u>. On the basis of these findings for student achievement, one can conclude that nursing students using computer assisted instruction achieve better scores on examinations than those using other forms of instruction. The most positive benefit to learning occurs when the computer is used as a supplement to instruction.

Retention of knowledge. Only one of the eleven studies (Boettcher, Alderson & Saccucci, 1981) tested students' retention. The experimental (computer) and comparison (printed instruction) groups were tested 11 weeks after exposure to the learning material for both their retention of knowledge and ability to apply the knowledge to new situations. Both groups had scores on the two retention tests that were approximately five percent lower than scores on the initial examinations. The effect size for the retention of knowledge examination was .06. There was no statistically reliable difference between the two groups on this measure, but the drop in scores on the delayed posttest was statistically reliable (p < .045). This result could be interpreted to mean that computer assisted instruction is no more effective than printed instruction in helping students retain information.

Application of knowledge. Two studies examined whether students who learn through computer assisted instruction could apply this knowledge to other situations better than those having traditional instruction. One study (Boettcher, Alderson & Saccucci, 1981) found no statistically reliable differences between the two groups (computer assisted versus printed instruction) for either their application of knowledge immediately following the treatment (ES = -.03) or when tested 11 weeks later (ES = .03). Both groups showed statistically reliable gains (p < .0005) in their ability to apply the

knowledge after the initial instruction, and a statistically reliable drop (p < .003) in application scores on a posttest given 11 weeks later. These results indicate that both computer assisted and printed instruction are effective in helping students learn to apply information to new situations. As was discussed in immediate achievement, it appears that the lack of reliable differences between the two groups in this study can be attributed to the fact that the two units of instruction were very similar in content and instructional approach.

The second study (Huckabay, Anderson, Holm & Lee, 1979) also concluded that there were no statistically reliable differences between the experimental (computer) and comparison (lecture) groups in their ability to apply knowledge to case studies, although their interpretation of the results is debatable (p = .09, ES = .67). As well, the students in the experimental group had statistically reliable increases (p < .01) over pretest scores in their ability to transfer the knowledge learned through instruction to case studies. In this case, the computer assisted instruction was used as a supplement to traditional instruction. Therefore, it can be concluded that computer assisted instruction was more effective in helping students learn to apply knowledge than lectures alone.

Instructional time. Seven studies reported on the instructional time taken by students using computer assisted instruction. In two of

these reports (Conklin, 1983; Huckabay, Anderson, Holm & Lee, 1979), the computer was used to supplement traditional instruction. In both studies the results were statistically reliable (p < .05), indicating that students in the experimental group, while taking longer to learn the material, achieved higher examination scores than those in the comparison group. In these two studies, the computer assisted instruction took approximately 2 - 2.5 hours and resulted in an average effect size for student achievement of 1.31.

Of the five studies that reported the length of the computer assisted instruction when it was used as a substitute for traditional methods, only three gave comparative data on the time students took to learn. In two studies (Bitzer, 1966; Bitzer & Boudreaux, 1969) the learning time for the experimental group was reduced. Bitzer (1966) reports a reduction in time of 25%, while Bitzer and Boudreaux (1969) report that learning time was reduced by a minimum of 40%. In the third study (Pogue, 1984), students spent 37% longer learning with computer assisted instruction. Pogue speculated that both the novelty of the instruction and the fact that students may have completed the optional review exercises could have contributed to the longer time taken to learn with computer assisted instruction. Bitzer (1966) found a statistically reliable difference (p < .09) indicating that students learned more using computer assisted instruction. The other two studies found no difference in achievement between the two groups.

Kirchhoff and Holzemer (1979) studied the effectiveness of

computer assisted instruction using a modified posttest only design in which the control group received no treatment. Students spent 1 - 2 hours in computer assisted instruction and achieved higher examination scores, on the average, than did the comparison group (ES = .85). These results were statistically reliable at the .01 level. However, in a study by Valish and Boyd (1975), there were no reliable differences in achievement between the experimental and an untreated comparison group. Their study involved registered nurses in a hospital setting and the computer assisted instruction covered information that all nurses could have learned either in their basic programs or from practical experience in the hospital.

In summary, students using computer assisted instruction as a substitute for traditional methods learn as much or more than those given conventional instruction and may do so in less time. When the computer is used as a supplement to instruction or when no other form of instruction is given, the additional investment in time for students produces greater learning.

Student attitudes. Nine of the 11 studies discussed student attitudes toward computer assisted instruction. Of the five studies that collected objective data that could be used to measure differences between experimental and comparison groups, only two reported comparisons of pretest and posttest scores. In these two studies (Conklin, 1983; Huckabay, Anderson, Holm & Lee, 1979) no

statistically reliable differences were found between the experimental and comparison groups in their overall attitude to computer assisted instruction before or after the treatment. The effect size on the attitude measure in Conklin's study was .57, while that in Huckabay's et al.'s was .40. Conklin found that students given computer assisted instruction were more positive toward this method of instruction after having used it (p < .05).

Two studies reported data on whether students preferred computer assisted instruction or the comparison group treatment. Boettcher, Alderson, and Saccucci (1981) report a statistically reliable result (p < .001) indicating that students who used computer assisted instruction had more positive attitudes toward their assigned mode of instruction than the comparison group (printed instruction) did to theirs. Huckabay et al. (1979) collected data on whether students preferred their assigned method of instruction, computer assisted instruction or lectures, over the other method. They found a reliable difference (p < .09) indicating that more students in the experimental group than in the comparison group preferred their assigned method of instruction. Also, they found that 72% of the participants preferred computer assisted instruction to lectures. Table 5 presents a summary of these findings.

Kirchhoff and Holzemer (1979) had students rate their attitudes toward a unit of computer assisted instruction covering postoperative nursing care. Their Adjective Rating Scale had five factors:

Table 5

Students' Preference for Instruction
and Assigned Method of Instruction

Preferred Method

		CAL	Lecture
Assigned Method	CAL	12	3
	Lecture	9	5

(Huckabay et al., 1979)

emotional appeal, interest value, perceived difficulty, perceived dullness, and practical value. Correlations between scores on these five factors and other variables are reported in the study. A correlation of $\pm$ .17 was statistically reliable at the .05 level and one of $\pm$ .24 at the .01 level. They found that the less experience students had with surgical patients, the more the computer assisted instruction was seen as having emotional appeal ($r = -.22$) and practical value ($r = -.24$). Those students who were having clinical experience in surgical nursing at the time they were given the computer assisted instruction were more likely to perceive the method as having emotional appeal ($r = .29$), having interest value ($r = .24$), being not difficult ($r = .40$), and having practical value ($r = .28$). In terms of learning, students tended to learn more the less they perceived computer assisted instruction as being difficult ($r = .18$) or dull ($r = .39$).

Most of the five attitudinal factors had positive, statistically reliable correlations with each other. The more positively students rated computer assisted instruction for practical value, the more likely they were to perceive it as having emotional appeal ($r = .62$), having interest value ($r = .66$), and being not difficult ($r = .45$). In addition, the less difficult the students perceived the method, the more that it had emotional appeal ($r = .53$), had interest value ($r = .38$), and was considered not dull ($r = .40$). Students who perceived computer assisted instruction as having emotional appeal also saw it

as having interest value (r = .70).

Six studies reported information on student attitudes toward computer assisted instruction that had been obtained from question-naires. When asked if they preferred computer assisted instruction over other methods, 50% preferred CAI over any other method (Bitzer & Boudreaux, 1969), 77% preferred it over lectures (Bitzer, 1966; Huckabay et al., 1979; Pogue, 1984), 93% liked it better than reading (Pogue, 1984), and there was no preference over printed instruction (Schleutermann et al., 1983). In no study did students as a whole prefer another method to computer assisted instruction; however, there were some students in each study who had a personal preference for other methods (Bitzer, 1966; Bitzer & Boudreaux, 1969; Huckabay et al., 1979; Pogue, 1984).

Generally, students liked computer assisted instruction because they were actively involved in learning (Bitzer, 1966; Pogue, 1984; Valish & Boyd, 1975), got immediate feedback on their answers (Bitzer, 1966; Schleutermann et al., 1983), could experiment without harm to patients (Bitzer, 1966; Bitzer & Boudreaux, 1969), found the experience enjoyable (Bitzer & Boudreaux, 1969; Kirchhoff et al., 1979; Pogue, 1984; Schleutermann et al., 1983), could work at their own pace (Bitzer, 1966), and found the method time efficient (Schleutermann et al., 1983). Students did not feel that they needed typing skills (Bitzer & Boudreaux, 1969; Pogue, 1984) or that adjusting to the mechanics of working the computer interfered with

learning (Bitzer & Boudreaux, 1969). In the study that reported the most negative responses to computer assisted instruction (Schleutermann et al., 1983), the students experienced problems accessing the computer and there were errors in the program itself.

In summary, it appears that students generally have favourable attitudes toward computer assisted instruction and prefer it over other methods of instruction once they have had experience with it. As well, students are more likely to have positive attitudes toward computer assisted instruction when the information is new and is seen to have practical value, when they are having concurrent clinical experience related to the topic of instruction, and when the information is not perceived as difficult or dull.

Other variables. Other variables investigated in the studies were cognitive style, ability, type of clinical experience, demographic information, time taken for development of the instruction, and cost of development. The findings related to each of these variables are presented in this section.

Two studies (Bitzer, 1966; Kirchhoff & Holzemer, 1979) reported findings on the relationship of cognitive style to student learning with computer assisted instruction. Bitzer (1966) used Kagan's (no citation in Bitzer, 1966) test for cognitive style to categorize students according to whether they used associative, analytical, or inferential methods for learning. During the computer assisted

instruction, students who had an inferential style were more likely to solve problems by conducting experiments (r = .72), checking results with selected conditions (r = .92, p < .05), and then answering the question. Of the three cognitive styles, students favouring inferential methods had the highest correlation with posttest scores, although this relationship was not statistically reliable.

Kirchhoff and Holzemer (1979) used Kolb, Rubin, and McIntyre's (cited in Kirchhoff & Holzemer, 1979) learning style inventory to classify students into one of four learning modes -- concrete experience, reflective observation, abstract conceptualization, and active experimentation. They found that students experiencing computer assisted instruction who used active experimentation tended to get higher posttest scores, although this relationship was not statistically reliable. One could conclude that learning style does not affect the amount of learning accomplished using computer assisted instruction.

Four studies (Bitzer, 1966; Boettcher et al., 1981; Kirchhoff et al., 1979; Schleutermann et al., 1983) collected pretest data on student ability. Grade point averages and scores on the Graduate Record Examination were used to measure this variable. None of the studies reported any statistically reliable relationship between ability and posttest scores for either experimental or comparison groups.

Two studies (Conklin, 1983; Kirchhoff & Holzemer, 1979) mentioned

that students were having concurrent clinical experience in areas relevant to the computer assisted instruction. Only Kirchhoff and Holzemer (1979) reported findings as to whether clinical experience had an effect on outcomes. They found that students who were having clinical experience in an area where they could apply the information learned in the computer assisted instruction tended to have higher posttest scores ($r = .25$, $p < .01$) than those placed in nonrelated clinical areas.

Four studies (Boettcher et al., 1981; Huckabay et al., 1979; Schleutermann et al., 1983; Valish & Boyd, 1975) collected demographic information on the participants. Depending on the study, information was collected on age, sex, educational background, previous experience in nursing, marital status, current work area, and previous experience with computer assisted instruction. In all but one, Huckabay et al.'s (1979) study, none of these variables correlated significantly with any other. Huckabay et al. found that age had an inverse correlation to pretest scores of knowledge of the subject ($r = -.33$, $p < .05$) and to posttest scores on ability to transfer information to other situations ($r = -.41$, $p < .01$). These correlations were based on information from all participants and, therefore, are not related directly to the use of computer assisted instruction.

Four studies (Bitzer, 1966; Hannah, 1978; Kirchhoff & Holzemer, 1979; Schleutermann et al., 1983) reported information on the amount of time to prepare the instructional materials. However, each chose a

different method for reporting this information. Kirchhoff and Holzemer (1979) stated that it took two months to convert an existing assignment into a 1 - 2 hour unit of computer assisted instruction. Bitzer (1966) took approximately five weeks to prepare two hours of material suitable for computer assisted instruction. Hannah (1978) stated that it took approximately 11 weeks to prepare the computer assisted instruction used in her study, but she does not state how long it took students to complete the unit. Schleutermann et al. (1983) reported that it took them 170 hours to prepare the computer assisted instruction versus 188 hours for printed instruction. They also stated that their preparation time reduced with each successive unit as they became more efficient.

Two studies (Bitzer & Boudreaux, 1969; Schleutermann et al., 1983) reported costs related to computer assisted instruction. Schleutermann et al. (1983) found that it cost $11,277.83 to develop and run three units of computer assisted instruction for twelve students over nine months. The control group treatment -- printed instruction using a latent image format -- cost $8,609.18 to develop and use. No conclusions were made about whether they felt these costs were reasonable or whether the additional $2,668.65 spent on computer assisted instruction was justified. They do mention, however, that the development cost of these materials reduces substantially with each successive use, whereas other forms of instruction, such as lecture, have fairly consistent costs.

Bitzer and Boudreaux (1969) did not report development costs but focused on the expense of the hardware needed for computer assisted instruction. They expressed this as a cost per student hour at a terminal and amortized the costs over five years. In comparing costs for PLATO III and PLATO IV, they reported that the former system cost between $1.90 and $2.90 per student hour while the latter should cost between 31 and 68 cents per student hour. Much of this reduced cost seems to be related to the cost of terminals for the different computer systems. The PLATO III student console cost approximately $5000.00 while the PLATO IV one was anticipated to cost $1500.00.

One final observation related to variables which might affect the study outcomes was that all 11 studies used computer assisted instruction as opposed to computer managed instruction. Therefore, the type of computer assisted learning, CAI versus CMI, was not a relevant variable.

This concludes the description of the synthesis of the findings of the 11 research studies included in the integrative review. Chapter 5 discusses the implications of these findings in relation to theory, nursing education, and previous research. As well, recommendations are made for future research and integrative reviews.

Chapter 5

Discussion and Recommendations

"...in looking to what has gone before, one can often
gain insight on possible roads to take to the future."
(Townsend, 1983, p. 341)

The purpose of this study was to integrate the findings of

research studies on computer assisted learning in nursing education to

determine its effectiveness as compared to traditional methods. Prior

to conducting the review, the literature on methods for doing reviews

and the problems associated with them was summarized. As well, the

use of computers in health care agencies, general education, and

nursing education was examined to determine current trends and

applications of computer technology in education. This chapter

discusses the findings from the integrative review in terms of the

implications for theory, practice, and research, and makes

recommendations for future applications and reviews.

The first section of this chapter summarizes the findings on the

effectiveness of computer assisted learning in nursing education

according to the following topics: immediate achievement, retention

of knowledge, application of knowledge, instructional time, student

attitudes, and other variables influencing effectiveness of

instruction. Each topic is discussed in terms of the consistency of

the findings from the current review with those of previous reviews

and research. Next, the findings are related to current theories of

learning. Then, the implications of the findings for nursing education and further research are considered. Finally, recommendations are made for further integrative reviews on computer assisted instruction.

Discussion of the Findings

Immediate achievement. The results of this meta-analytic and qualitative review showed that nursing students using computer assisted instruction learned more initially than those given traditional instruction, particularly when the computer was used as a supplement to other methods. In the typical implementation, computer assisted learning raised examination scores by approximately five percentage points, or about .48 standard deviations. Thus, when the distribution of scores on examinations testing immediate achievement is based on the average student who was given traditional instruction receiving a score at the 50th percentile, the typical nursing student using computer assisted instruction scored at the 68th percentile.

These results are consistent with previous reviews on computer assisted instruction in general education. Earlier meta-analytic and box-score reviews (Edwards et al., 1975; Hartley, 1977; Jamison et al., 1974; Visonhaler & Bass, 1972) were unanimous in stating that, in elementary education, computer assisted instruction used as a supplement was more effect than traditional methods alone. At the

secondary (Kulik, Bangert & Williams, 1983) and post-secondary levels (Kulik, Kulik, & Cohen, 1980), smaller effect sizes for achievement have been obtained than those in elementary education; however, these reviewers have still supported the effectiveness of computer assisted instruction over other methods.

In this review, several of the characteristics of the primary studies were found to be associated with the effect sizes achieved. The most obvious difference related to how the computer assisted instruction was used. When used as a supplement to other methods, computer assisted instruction had an average effect size of 1.31 for immediate achievement. This means that the average student who used computer assisted instruction as an adjunct to other instruction scored at the 90th percentile on examinations. When computer assisted instruction was used as a replacement for other methods, the mean effect size was .27. These findings parallel those reported by Edwards et al. (1975) for elementary students and by Jamison et al. (1974) for secondary and post-secondary students. In fact, both Thomas (1979) and Hallworth and Brebner (cited in Foreman, 1983) found that when computer assisted instruction was used as a supplement, it generally produced greater achievement regardless of the teaching strategy, computer system, testing method, or level of education.

One question raised in the research on computer assisted instruction is whether it is more effective than other, less expensive methods used to augment traditional instruction. Suppes and

Morningstar (cited in Jamison et al., 1974) found that classroom drill

and practice can accomplish the same results as computer assisted

instruction except that the classroom approach takes twice as long.

Dick and Latta (cited in Visonhaler & Bass, 1972) found that eighth

grade students had greater achievement with programmed instruction

than they did with computer assisted instruction. On the other hand,

Hartley's (1977) review reports that although effect sizes produced by

computer assisted instruction were not as large as those produced by

tutoring, they were larger than effects produced by programmed

instruction or the use of individual learning packets.

In the current review, the effects for student achievement always

favoured computer assisted instruction over lecture; however, if the

outlying study is eliminated, there were virtually no measurable

advantages when computer assisted instruction was used as a

replacement for printed instruction. One explanation for this latter

finding relates to the similarity in formatting that can exist with

both computer assisted and printed instruction. In this review, the

two studies that used printed instruction as the comparison group

treatment (Boettcher et al., 1981; Schleutermann et al., 1983) used

similar approaches to prepare the instructional materials. The actual

lesson or simulation was developed first and then produced as computer

assisted and printed instruction. By preparing the materials in this

way, the authors ensured that both the content and the format for

delivery in the two forms of instruction were as comparable as

possible. As a result, it is not surprising that the effect sizes for student achievement comparing computer assisted and printed instruction are close to zero as the units of instruction were so similar in content and in instructional approach.

In this review, the mean effect size obtained favouring computer assisted instruction over lectures may have been affected by other factors. All of the tests used to measure the differences in student achievement between these two methods were based on the general subject area rather than on the specific information contained within the unit of instruction. If tests of general knowledge do not measure true effect sizes as well as achievement tests based specifically on the unit of instruction, then the effect sizes obtained for differences in student achievement between computer assisted instruction and lectures may be underestimated. Further research would be necessary to determine if this hypothesis is true.

One trend that was noted but could not be considered reliable given the small number of studies was that effect sizes tended to decrease with the amount of nursing education and practice students had received. Computer assisted instruction tended to be more effective for basic education of nurses (hospital training and baccalaureate levels) than it was for graduate nurses using it for continued education. This finding would need further study to determine if it could be considered reliable.

Three characteristics related to the studies' research design and

methodology were associated with the effect sizes obtained. One of these characteristics was the length of the study. Studies of shorter duration tended to produce larger effect sizes for student achievement. Similar findings have been reported in reviews of computer assisted instruction (Kulik et al., 1983) and of tutoring methods (Cohen et al., 1982). One interpretation for this finding is that studies which take place over a shorter time span are better controlled than those that take longer. Therefore, shorter studies are more likely to measure the true effects of instruction because there is less time for other variables, such as experience, to influence student achievement. If studies occur over a period of weeks and use one or two classes from the same institution for the participants, it would be possible for them to share their experiences with one another and to teach each other the content. If this occurred, there would be very little observed difference in achievement between the experimental and comparison groups, and the true effect of computer assisted instruction would be underestimated. This interpretation must be considered speculative as the studies did not provide data on whether sharing of information occurred among the participants. However, it is interesting to note that, in this review, the most extreme effect sizes, both positive and negative, were associated with studies that were less than eight weeks long.

Another study characteristic that was associated with the effect sizes obtained for student achievement was the research design.

Studies with a pretest-posttest design produced larger effect sizes in favour of computer assisted instruction than those with a posttest only design. One possible explanation for this finding is that pretesting the participants increased their awareness of being part of a research study and, as a result, they put more effort than they normally would have into the task. Again, this interpretation must be considered speculative as the studies did not provide data to support this explanation.

One study characteristic that has not been mentioned before as influencing effect sizes was the reported reliability of the outcome measure. In the current review, those studies that reported testing the outcome measure for reliability achieved a higher average effect size (ES = 1.16) than those that did not report this statistic (ES = .19). One could interpret this finding as indicating that those studies that developed reliable measures and reported this were more likely to be measuring true effects of computer assisted instruction on student achievement. In the absence of data from the studies that did not report test development procedures, this finding must be regarded as speculative.

Other reviews of computer assisted instruction have noted that publication features (year of publication, form of publication) (Hartley, 1977; Kulik et al., 1983; Smith, 1980) and controlling for instructor effects (Cohen et al., 1982; Kulik, et al., 1983; Kulik, et al., 1980) are related to effect sizes for achievement. None of these

characteristics was found to be related to effect sizes in this review. The small sample size prevented any meaningful investigation of these variables.

Retention of knowledge. Only one study (Boettcher et al., 1981) reported measuring retention of knowledge. In this case, there were no differences in achievement on retention examinations between nursing students who had used computer assisted instruction or printed instruction. This finding is similar to those reported in other reviews (Edwards et al., 1974; Kulik et al., 1983). As well, both Edwards et al. (1974) and Splittgerber (cited in Foreman, 1983) suggest that retention rates may actually be lower for computer assisted instruction than for other methods. Further research is needed to determine the effect of computer assisted instruction on retention of knowledge.

Application of knowledge. Two studies examined whether nursing students who learned through computer assisted instruction could apply this knowledge to new situations better than those given traditional instruction. In one study (Boettcher et al., 1981), there was no reliable difference in ability to apply information between the groups using computer assisted instruction and printed instruction. However, both groups had statistically reliable gains in their ability to apply information to new situations. This finding indicates that computer

assisted instruction and printed instruction are equally effective for helping nursing students learn to apply knowledge. As has been discussed, the lack of difference between the groups is probably related to the fact that both units of instruction were extremely similar in content and instructional approach. The authors state that they may not have made good use of the capabilities of the computer as they wanted their materials to be as parallel as possible for the purposes of the research.

Huckabay et al. (1979) found that nursing students who used computer assisted instruction as a supplement to lectures had a statistically reliable increase in their posttest scores for application over pretest scores, while those given lectures alone did not. While the differences in achievement between the two groups were not statistically reliable, the results from this study indicate that nursing students do benefit from using computer assisted instruction to learn how to apply information to new situations. The findings from both of these studies must be considered tentative as no other reviews or research studies reported on ability to apply or transfer knowledge.

Instructional time. In cases where computer assisted instruction was used as a supplement to other methods (Conklin, 1983; Huckabay et al., 1979), nursing students spent longer learning the material but achieved higher examination scores ($p < .05$) than did the comparison

group.

The review findings are less conclusive for the three studies where computer assisted instruction replaced other forms of instruction. In two of the three studies, instructional time was reduced for the students using computer assisted instruction. One of these studies (Bitzer, 1966) reported that this was accompanied by statistically reliable differences in examination scores that favoured the group using computer assisted instruction, while the other study (Bitzer & Boudreaux, 1969) found no differences in achievement. In the third study (Pogue, 1984), nurses using computer assisted instruction took longer to learn the material and there were no reported differences in achievement. These three studies suggest that nursing students using computer assisted instruction as a substitute for other methods learn as much or more than those given conventional instruction and may do so in less time.

Although the number of primary studies which reported on this aspect was small, this finding parallels previous reviews. Edwards et al. (1975) and Jamison et al. (1974) report that, on the average, it took students less instructional time to learn with computer assisted instruction than with other methods. Also, both Kulik et al. (1983) and Kulik et al. (1980) found substantial reductions in instructional time for students using computer assisted instruction but their results are based on reports from only two or three primary studies.

Student attitudes. Generally, the findings from this review indicate that all nursing students have favourable attitudes toward computer assisted instruction, although no statistically reliable differences in attitude were found between experimental and comparison groups in the two studies that measured this. However, nursing students given computer assisted instruction were more positive toward this method once they had used it. Several previous reviews (Hallworth & Brebner, cited in Foreman, 1983; Kulik et al., 1983; Kulik et al., 1980) also report that the general reaction of students to computer assisted instruction is positive at all levels of education.

Kirchhoff & Holzemer (1979) correlated their attitude measure with other variables. They found that nursing students were more likely to have positive attitudes toward computer assisted instruction when the information they were learning was new and was seen to have practical value, when they were having concurrent clinical experience related to the topic of instruction, and when the instruction was not perceived as difficult or dull. No other studies or reviews are known that have reported on these aspects. It would be valuable for future research to determine if these conditions, when met, result in increased achievement.

Other variables. The other variables considered which could influence the effectiveness of computer assisted instruction in

nursing education were cognitive style, ability, type of clinical experience, demographic information, time taken for development of the instruction, and costs. No statistically reliable differences in achievement between the experimental and comparison groups were found that could be attributed to cognitive style or ability. One study (Kirchhoff & Holzemer, 1979) reported on the effect of concurrent clinical experience in a related area. In this case, nursing students using computer assisted instruction who were placed in an area where they could apply the information learned had higher posttest scores than those in an unrelated area.

None of the four studies that collected demographic information on the participants found any statistically reliable differences between groups that could be related to the effectiveness of computer assisted instruction.

The last two variables consider the effectiveness of computer assisted instruction in relation to faculty and administration. On the average, it took approximately two months to prepare two hours of computer assisted instruction. All four studies recognized that initial development of instructional materials is expensive, but that the costs are reduced with successive use and should be considered in relation to the number of students taught. As well, the cost-effectiveness of computer assisted instruction, which can be used repeatedly over long periods of time, should be compared to instructional costs for other methods, such as lectures. The actual

costs reported for development and maintenance of computer assisted instruction were meaningless as no comparative data was provided which encompassed both aspects.

In their review, Chambers and Sprecher (1983) state that the average time for courseware development is 100 hours for one hour of computer assisted instruction. Reports of costs for development and maintenance vary widely because of the variation in factors considered. Norris (cited in Foreman, 1983) has pointed out that costs for traditional instruction have been increasing at the rate of 13% a year (1975-78) while those for computer assisted instruction were decreasing at 5% a year. Braun (cited in Foreman, 1983) found that student attendance increased when computer assisted instruction was used and this resulted in a tax cost saving to the public.

If the cost estimates of computer assisted instruction are adjusted to include the hidden costs of hardware, software, maintenance, and support, then it may be more expensive than traditional instruction (Foreman, 1983). On the other hand, Kearsley (cited in Foreman, 1983) has stated that cost estimates are based on the assumption that computer assisted instruction is as effective as traditional methods. Considering that students achieve as much or more with computer assisted instruction and in less time; that some types of computer assisted instruction, e.g. simulations, can not be presented by traditional methods; and that students have positive attitudes toward this form of instruction; cost estimates for computer

assisted instruction versus traditional methods should actually be interpreted as giving computer assisted instruction the edge.

Implications for Theory Development

It was not a purpose of this integrative review to be able to relate the results to theories of learning. The majority of the results in this review were obtained from data in the primary studies based on objective measures of student achievement, instructional time, and attitudes. Since the actual process involved in achieving these outcomes was not considered in the primary studies, any statements relating these results to theories of learning would have to be based on inference alone.

One theory which should be considered, however, relates to the effects of instructional technology on school learning. Kulik (1981) developed a model which suggests that computer assisted instruction would be more effective at lower levels of education because younger students need the stimulation and guidance of a highly reactive teaching medium. Post-secondary students profit more from solving problems themselves. Hence, a highly interactive medium, such as computer assisted instruction, may actually interfere with learning. For example, drill and practice and tutorial programs usually require the student to progress through the lesson following a sequence set by the programmer. The question and answer format pre-supposes that questions have one right answer. While this approach may be effective

for teaching younger students and for lower levels of cognitive learning (knowledge and comprehension), it "allows little or no practice in challenging what others take to be the 'facts', of assembling data into theories and testing them, or solving novel problems in novel ways" (Garson, 1983, p. 123). Kulik, Bangert, and Williams (1983) state that findings from recent reviews of computer assisted instruction support Kulik's (1981) model. The effect size for achievement by elementary students was .41 (Hartley, 1977), by secondary students was .32 (Kulik et al., 1983), and by post-secondary students was .25 (Kulik et al., 1980).

The findings from the current review do not support this trend. Two explanations are possible. First, Hartley's (1977) study involved only computer assisted instruction when used as a supplement, whereas the other two reviews (Kulik et al., 1983; Kulik et al., 1980) included studies where the instruction had been a supplement to or a substitute for traditional methods. In reviews that include both applications or computer assisted instruction, one could anticipate lower effect sizes. Since both of the latter reviews and the current review included both types of computer assisted instruction, the findings could be used as evidence against Kulik's (1981) statement that the effectiveness is a function of instructional level.

Current evidence suggests that the effectiveness of computer assisted instruction is more likely to be a function of the way it is used. For example, most review findings are quite consistent in

reporting that computer assisted instruction is most effective when it is used as a supplement to traditional methods. The one problem that has not been addressed with any consistency is determining why computer assisted instruction is more effective for student learning. In nursing education, very little research has been done to study whether some types of computer assisted instruction are more effective than others or whether such things as the degree of control the student has over the learning experience actually make a difference in terms of achievement. It would seem that factors such as these are more likely to be related to the effectiveness of computer assisted instruction than the level of instruction.

The second explanation of why the findings from this review do not support Kulik's (1981) model is that other variables have influenced the results. In particular, the sample size is small and may not be representative of the population. Also, the majority of studies had been published in journals and, in previous reviews (Kulik, 1981; Kulik et al., 1983; Smith, 1980), this factor has been shown to be related to positive effect sizes for student achievement. These effect sizes may overestimate the true effects of instruction if it is true that journals are more likely to accept studies with statistically reliable results for publication.

Another variable which may have influenced the results of this review is the participants themselves. The review was restricted to primary studies in nursing education. It is possible that nursing

students differ in their approaches to or motivation for learning in comparison to other post-secondary students. However, further research would be needed to determine if this is the case.

Limitations of the Study

The major factor limiting the generalizability of the findings of this integrative review is the small number of studies included in the sample. While the body of research on computer assisted instruction in nursing education is known to be limited, it is highly likely that the sample obtained for this review is not representative of the population. First, most of the reports were obtained from published sources. The large number of narrative reports that were eliminated by the sampling procedures suggests that the actual number of primary studies which could have been included in the review is much larger than that achieved. These studies could not be obtained due to limitations of time.

Second, the effect sizes for student achievement had a skewed distribution. With a sample of studies that is representative of the population, the distribution of the effect sizes might approximate a normal distribution. However, the consistency of the findings with those from previous reviews of computer assisted instruction does indicate that the findings from this study are similar to what might be anticipated with a larger sample of studies.

One further limitation to this integrative review is that all

studies were coded by one person. The results of the coding could include biases of the reviewer. This would reduce the replicability of the review and make the results questionable. Having several people code the same studies and then checking for inter-coder reliability would increase the validity of the coding.

In light of these issues, the results of this review on computer assisted instruction in nursing education should be considered as preliminary. Further reviews of the same topic are needed to establish whether the findings of this review are generalizable.

Implications for Nursing Education

The results of previous reviews on the effectiveness of computer assisted instruction versus other methods, in combination with the findings from the current review, indicate that computer assisted instruction is an effective method for teaching nursing students. For cognitive learning, computer assisted instruction is most effective for student achievement when it used as a supplement to other forms of instruction. Moreover, students may learn the material in less time. The technology has the potential to adapt instruction to individual learning needs and to allow students to learn at their own pace and at a time convenient to them. The positive attitudes that students develop toward this form of instruction after having used it may increase motivation to learn and, as a result, the amount of learning that actually takes place. This review only considered outcomes that

have been used frequently to measure effectiveness of instruction. No attempt was made to identify subtle or unique outcomes, such as confidence with computers, which may also result from the use of computer assisted instruction.

Boettcher et al. (1981) suggested the importance of computer assisted instruction for nursing education is related to how the technology is used rather than the fact that it is used. For example, the results obtained by Boettcher et al. (1981) and Schleutermann (1983) suggest that printed instruction can be as effective as computer assisted instruction when the units cover the same content and use comparable approaches to instruction. However, in these two studies, the capabilities of the computer beyond those of printed instruction were ignored for the purposes of the research. One might anticipate that computer assisted instruction would be more effective than printed instruction if use were made of the computer's capabilities for complex branching within sequences of instruction so that the unit is adapted to the individual student's learning needs rather than following a set format for all students. As well, once a program is established for a simulation, the computer is capable of providing students with a great variety of experiences as the data for case studies can be changed each time the program is run. With printed instruction, the student is confined to reviewing the same case study over and over again. The variety of experiences possible with computer assisted instruction may help students become more

effective problem solvers and may result in improved ability to apply knowledge to new situations. Certainly, more research would be needed to establish whether the quality and variety of instruction possible with computer assisted instruction is more effective than other forms such as printed instruction.

As well, there is an indication that computer assisted instruction is more effective when the computer is used in combination with other media, such as audio or visual aids. The two studies which obtained the largest effect sizes in favour of achievement with computer assisted instruction (Bitzer, 1966; Conklin, 1983) had augmented their learning units with films or slides. This finding must be considered preliminary as there is no other evidence to support this result.

One point that was made quite clear in Schleutermann et al.'s study (1983) was that the unit of computer assisted instruction must be free of major errors in programming. In their study, students became frustrated with the inefficiencies and errors in the program. They suggest that these difficulties may have interfered with learning and, therefore, reduced the potential effectiveness of the unit of computer assisted instruction.

Valish and Boyd's study (1975) emphasizes the problems associated with using software that has been developed by other institutions. In their study, the posttest scores for achievement actually favoured the untreated comparison group at a significance level that could be

considered statistically reliable (p ≐ .08). Their conclusion was that the information in the unit of computer assisted instruction could have been learned in basic nursing programs or through clinical experiences and, therefore, the computer assisted instruction was not appropriate for their study. One lesson that can be learned from their experience is that software obtained from other institutions should be reviewed prior to use to ensure that its content is consistent with learning objectives and with the learning needs of the students.

One unresolved issue is the cost-effectiveness of computer assisted instruction in nursing education. Hoffer et al.'s study (1975) involved nurses at a remote site who accessed the instructional unit via telephone lines. Results of the study indicated that these nurses learned as much as those taught by traditional methods. Therefore, it seems that computer assisted instruction could be used effectively for distance education. Using computer assisted instruction in this way for continuing education of nurses could result in greater access to information about recent developments that affect professional practice by larger numbers of nurses. The result would be savings in cost to both the institution providing the instruction and the nurses accessing it.

This integrative review reached no conclusions related to the cost-effectiveness of computer assisted instruction for nursing education. Perhaps the most important consideration at this time is

whether the information can be taught to students as effectively and more economically by other methods of instruction. There is some indication that printed instruction may be as effective as computer assisted instruction for cognitive learning. Further research is needed in this area before any definitive conclusions can be reached.

Implications for Further Research

The limited amount of primary research on the effectiveness of computer assisted instruction in nursing education suggests that much more research needs to be done. Further work is needed to determine if the positive attitudes that nursing students develop toward computer assisted instruction do, in fact, affect the amount of information learned.

While it appears that computer assisted instruction can be more effective than other forms of instruction, we still do not know what makes it effective or if one type of computer assisted instruction is more effective for learning than another. In order to make the most appropriate use of computer assisted instruction in nursing education, further research is needed to compare the effectiveness of the different types of computer assisted instruction for any one level of learning, whether it be cognitive, psychomotor, or affective. Here, the focus of the research should be on studying the process of instruction rather than just the outcome. As well, research is needed to determine whether there are certain learning styles or ways of

problem solving that are more effective with certain types of computer assisted instruction. One question that could be asked is whether those students who have a similar learning style to that of the person who programmed the unit of instruction are the ones for whom the instruction is most effective.

As the results of this integrative review indicate, further research is needed to determine if computer assisted instruction in nursing education is as effective as other methods in promoting retention of learning. The results to date have been inconclusive. As well, further research is needed on whether application of knowledge to new situations is facilitated by using computer assisted instruction. The research studies included in this review used either printed or computerized simulations to teach application of knowledge and then tested the results using written examinations. The results have been favourable, indicating that both of these methods of instruction were effective in helping students transfer knowledge to written simulations. However, it is equally important to study whether nursing students can transfer information learned through computer assisted instruction to clinical situations as well.

Some of the findings from the review indicate that further research is needed to determine other variables which influence the effectiveness of computer assisted instruction. For example, do nursing students learn more when computer assisted instruction is combined with concurrent clinical experience related to the topic of

instruction? Do nursing students learn more when the computer assisted instruction is augmented with audio and visual aids? Do nursing students learn more when they perceive the information they are learning through the computer assisted instruction to have practical value? Is computer assisted instruction more effective for some levels of nursing education than for others?

In terms of the research itself, several recommendations can be made as a result of this review. First, researchers should attempt to develop reliable and valid tools to measure outcomes to increase the likelihood that the effects measured are true effects. However, the most important aspect to consider when developing these achievement tests is to ensure that the test actually measures the construct being studied. For example, in this review, a few studies used psychometric test development procedures to create examinations which were considered to be reliable measures of the content taught. These procedures, in fact, produce tests designed to measure differences between individuals and, therefore, may not necessarily be measuring the amount of information learned through instruction. In order to measure student achievement more accurately, especially for purposes of research on the comparative effectiveness of different methods of instruction, researchers should consider using edumetric procedures (Carver, 1974) as tests developed in this way will measure gains in knowledge rather than differences between individuals.

Second, it seems that short, well controlled studies would

provide more meaningful results when comparing the effectiveness of one method of instruction to another. Finally, researchers should be encouraged to provide more detailed descriptions of the methods and statistics used for the research so that replication of studies is facilitated and future reviews of research can be completed.

Implications for Future Integrative Reviews

The growing number of integrative reviews appears to be having some benefits in synthesizing findings of primary studies and in interpreting results of subsequent reviews. The results of the current review, while not generalizable, could be interpreted with more confidence because of the availability of previous review findings on computer assisted instruction.

The problems related to review methodology have been discussed in the literature. However, some questions remain unresolved. As Kulik et al. (1980) have pointed out, further work needs to be done to determine whether the level of sampling for reviews should be the studies themselves or whether study results should be weighted according to numbers of participants, numbers of variables studied, or numbers of classes used. In terms of the current review where only a limited number of primary research studies were available, the reviewer questioned if the results from the narrative reports could have been included in the review to add more confidence to the findings. It is not known how this could be done or whether this

would have altered the results.

One aspect of meta-analytic reviews that has impact on the results is the manner of coding studies. In the current review, a coding form was created based on the reviewer's perception of what might be important to study. As the review progressed, other variables which appeared to influence the review results were identified. It seems, therefore, that more thought should be given to the coding of data and coding forms should be based on an analytical framework of whatever is being measured. This framework should be reported as part of the review.

At the present time, findings on particular characteristics of the primary studies are reported sporadically because there is no consistency to the coding process. More standardized coding forms should be developed so that findings from reviews are reported in a more consistent manner, thus allowing results to be compared across reviews. For example, characteristics of the primary studies related to publication features, study settings, and research design and methodology could be coded in a standardized format for any review. As well, previous reviews on a particular topic can provide guidelines for coding study characteristics related to the particular construct being studied and the outcome measures used.

Another aspect which should be considered is topics for future integrative reviews on computer assisted instruction. To date, the focus has been on assessing the effectiveness of computer assisted

instruction for a particular type or level of student. The study
characteristics considered in these reviews are fairly concrete and
easily measured features such as the year of publication of the study,
the topic of instruction, or the research design. One direction for
the future would be to shift the focus for reviews to the nature of
the computer assisted instruction itself. Computers have the
capability to store and synthesize large amounts of data and can be
programmed to provide very individualized instruction to students.
There are programs developed with very complex branching capabilities
that can adapt instruction continually to a particular student based
on information stored about learning needs, ability, cognitive style,
and previous responses to questions (Strain, 1983). When there is
sufficient research on this type of adaptive instruction compared to
the more linear drill and practice programs, it should be possible to
investigate whether the logic and branching capabilities of the
computer are what makes instruction effective for students.

As well, future reviews could focus on what types of computer
assisted instruction are most effective for different levels of
learning. Emphasis should also be placed on trying to identify
exactly what makes computer assisted instruction more or less
effective in any situation. Because of the limited research in the
field, integrative reviews on these topics would need to be combined
across levels of education in order to obtain sufficient numbers of
studies to reach generalizable conclusions.

The current reviewer recommends that this review be replicated within the next few years. By that time, the number of primary studies on the effectiveness of computer assisted instruction in nursing education should be large enough for more conclusive results to be obtained.

APPENDIX A

CODING CATEGORIES FOR CHARACTERISTICS OF PRIMARY STUDIES

Coding Category

Publication date: 1966-69
                    1970-74
                    1975-79
                    1980-84

Publication form: Journal
                    Book
                    Thesis/Dissertation
                    Unpublished

Country of origin: Canada
                     USA

Computer use: Drill and practice
                Tutorial
                Simulation
                Utility

Implementation: Substitute
                  Supplement

Subject: Medicine
           Surgery
           Psychiatry
           Obstetrics
           Pediatrics

Approximate time on CAI: Less than 3 hr.
                                  3 - 6 hr.
                           More than 6 hr.
                                  Not stated

Software and course development: Local
                                   Others
                                   Not stated

Study setting: Hospital leading to RN
                 College leading to RN
                 University - baccalaureate
                             - graduate
                 Hospital inservice education

APPENDIX A (Continued)

Coding Category

Length of study: 0 - 4 weeks
                             4 - 8 weeks
                    More than 8 weeks
                             Not stated

Experimental design: Pretest-posttest
                       Posttest only
                       Other

Sample: Volunteer
          Convenience

Sample Size: 1 - 25
                     26 - 50
                     51 - 75
                     76 - 100
             More than 100

Assignment to groups: Random
                        Matched
                        Uncontrolled

Historic effects: Same time period
                    Different time period

Instructor effects: Same instructor
                      Different instructor
                      Not stated

Pretest: Knowledge
           Ability
           Cognitive style
           Attitude
           Demographic
           Application/transfer

Outcome: Knowledge/learning
           Retention
           Application
           Time
           Attitude
           Cost

APPENDIX A (Continued)

Coding Category

Achievement measure: Instructor developed
                       Commercial
                       Not measured

                  :   Based on lesson
                       Not based on lesson
                       Not measured

BIBLIOGRAPHY

Ackerman, W. B. (1982). Technology and nursing education: A scenario for 1990. Journal of Advanced Nursing, 7, 59-68.

Badura, F. K. (1980). Nurse acceptance of a computerized arrhythmia monitoring system. Heart and Lung, 9(6), 1044-1048.

Ballantyne, D. J. (1979). A computerized scheduling system with centralized staffing. Journal of Nursing Administration, 9(3), 38-45.

Barton, R. (1977). The next ten years. In J. Feldman & C. Mosmann (Eds.), Instructional computing in the university - The second decade (pp. 13-18). San Francisco: San Francisco Press, Inc.

Birckhead, L. M. (1978). Nursing and the technetronic age. Journal of Nursing Administration, 8(2), 16-19.

Bitzer, M. (1966). Clinical nursing instruction via the PLATO simulated laboratory. Nursing Research, 15(2), 144-150.

Bitzer, M. D., & Bitzer, D. L. (1973). Teaching nursing by computer: An evaluative study. Computers in Biology and Medicine, 3(3), 187-204.

Bitzer, M. D., & Boudreaux, M. C. (1969). Using a computer to teach nursing. Nursing Forum, 8(3), 234-254.

Bloom, B. S. (Ed.) (1956). Taxonomy of educational objectives, handbook 1: Cognitive domain. New York: David McKay Company, Inc.

Boettcher, E. G., Alderson, S. F., & Saccucci, M. S. (1981). A comparison of the effects of computer-assisted instruction versus printed instruction on student learning in the cognitive categories of knowledge and application. Journal of Computer-Based Instruction, 8(1), 13-17.

Brennan, P. F. (1981). Establishment of a computer assisted instruction program to teach managerial decision making. In H. G. Hefferman (Ed.), Proceedings: Symposium on computer applications in medical care. Los Angeles: IEEE Computer Society.

Butters, S., Feeg, V., Harmon, K., & Settle, A. (1982). Computerized patient care data: An educational program for nurses. Nurse Educator, 7(2), 11-16.

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research. Chicago: Rand McNally College Publishing Company.

Carver, R. P. (1974). Two dimensions of tests - Psychometric and edumetric. American Psychologist, 29(7), 512-518.

Chambers, J. A., & Sprecher, J. W. (1983). Computer assisted instruction: Current trends and critical issues. In D. O. Harper & J. H. Stewart (Eds.), Run: Computer education (pp. 107-119). Monterey, California: Brooks/Cole Publishing Company.

Cheung, P. (1979). Put test questions on a computer. Nursing Mirror, 149(18), 26-28.

Cohen, J. (1977). Statistical power analysis for the behavioral sciences (rev. ed.). New York: Academic Press.

Cohen, P. A., Kulik, J. A., & Kulik, C.-L. (1982). Educational outcomes of tutoring: A meta-analysis of findings. American Educational Research Journal, 29, 415-428.

Collart, M. E. (1973). Computer-assisted instruction and the teaching-learning process. Nursing Outlook, 21(8), 527-532.

Conklin, D. N. (1981). A study of computer assisted instruction in nursing education. Unpublished Masters thesis, University of Calgary, Calgary.

Conklin, D. N. (1983). A study of computer-assisted instruction in nursing education. Journal of Computer-Based Instruction, 9(3), 98-107.

Cook, M. (1982). Using computers to enhance professional practice. Nursing Times, 78(37), 1542-1544.

Cook, T. D., & Leviton, L. C. (1980). Reviewing the literature: A comparison of traditional methods with meta-analysis. Journal of Personality, 48(4), 449-472.

Cooper, H. M. (1982). Scientific guidelines for conducting integrative reviews. Review of Educational Research, 52(2), 291-302.

Cooper, H. M., & Arkin, R. M. (1981). On quantitative reviewing. Journal of Personality, 49(2), 225-230.

Cooper, H. M., & Rosenthal, R. (1980). Statistical versus traditional procedures for summarizing research findings. Psychological Bulletin, 87(3), 442-449.

Covvey, H. D., & McAlister, N. H. (1980). Computer consciousness. Surviving the automated 80s. Reading, Massachusetts: Addison-Wesley Publishing Company.

Donabedian, D. (1976). Computer-taught epidemiology. Nursing Outlook, 24(12), 749-751.

Duraiswamy, N., Welton, R., & Reisman, A. (1981). Using computer simulation to predict ICU staffing. Journal of Nursing Administration, 11(2), 39-44.

Edgington, E. S. (1972). An additive method for combining probability values from independent experiments. Journal of Psychology, 80, 351-363.

Edmunds, L. (1982a). Teaching nurses to use computers. Nurse Educator, 7(5), 32-38.

Edmunds, L. (1982b). Computer-assisted nursing care. American Journal of Nursing, 82(7), 1076-1079.

Edwards, J., Norton, S., Taylor, S., Weiss, M., & Dusseldorp, R. (1975). How effective is CAI? A review of the research. Educational Leadership, 33(2), 147-153.

Eysenck, H. J. (1978). An exercise in mega-silliness. American Psychologist, 33(5), 517.

Farlee, C. (1978). The computer as a focus of organizational change in the hospital. Journal of Nursing Administration, 8(2), 20-26.

Feldman, K. A. (1971). Using the work of others: Some observations on reviewing and integrating. Sociology of Education, 44(1), 86-102.

Foreman, D. (1983). Search of the literature. In D. O. Harper & J. H. Stewart (Eds.), Run: Computer education (pp. 128-140). Monterey, California: Brooks/Cole Publishing Company.

Frantz, R. F. (1976). Computers in nursing education. Image, 8(2), 23-26.

Gage, N. L. (1978). The yield of research on teaching. Phi Delta Kappan, 60(3), 229-235.

Garson, J. W. (1983). The case against multiple choice. In D. O. Harper & J. H. Stewart (Eds.), Run: Computer education (pp. 120-126). Monterey, California: Brooks/Cole Publishing Company.

Glass, G. V. (1976a). Primary, secondary, and meta-analysis of research. Educational Researcher, 5(10), 3-8.

Glass, G. V. (1976b). Primary, secondary, and meta-analysis of research. Presidential address to the Annual Meeting of the American Educational Research Association, San Francisco, April 21, 1976.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). Meta analysis in social research. Beverly Hills, California: Sage Publicatons.

Glass, G. C., & Smith, M. L. (1978). Reply to Eysenck. American Psychologist, 33(5), 517-519.

Gold, R. S., & Duncan, D. F. (1980). Computers and health education. The Journal of School Health, November, 503-505.

Hannah, K. J. (1976). The computer and nursing practice. Nursing Outlook, 24(9), 555-558.

Hannah, K. J. (1978). Overview of computer assisted learning in nursing education at the University of Calgary. In M. C. Stainton, J. C. Kerr, & J. L. Moore (Eds.), Perspectives: Nursing education, practice and research (pp. 43-56). Calgary: Faculty of Nursing, University of Calgary.

Hannah, K. J. (1983a). Computer assisted learning in nursing education: A macroscopic analysis. In M. Scholes, Y. Bryant, & B. Barber (Eds.), The impact of computers on nursing. An international review (pp. 280-287). Amsterdam: Elsevier Science Publishers B. V.

Hannah, K. J. (1983b). The place of medical informatics in a baccalaureate nursing curriculum. In J.-C. Pages, A. H. Levy, F. Gremy, & J. Anderson (Eds.), Meeting the challenge: Informatics and medical education (pp. 67-77). Amsterdam: Elsevier Science Publishers B. V.

Hannah, K. J., & Conklin, D. (1983). Bibliography. Computers in nursing. Unpublished document, Faculty of Nursing, University of Calgary, Calgary.

Hartley, S. S. (1977). Meta-analysis of the effects of individually paced instruction in mathematics (Doctoral dissertation, University of Colorado, 1977). Dissertation Abstracts, 38, 4003.

Hoffer, E. P., Mathewson, H. O., Loughrey, A., & Barnett, G. O. (1975). Use of computer-aided instruction in graduate nursing education: A controlled trial. Journal of Emergency Nursing, 1(2), 27-29.

Huckabay, L. M., Anderson, N., Holm, D. M., & Lee, J. (1979). Cognitive, affective, and transfer of learning consequences of computer-assisted instruction. Nursing Research, 28(4), 228-233.

Hughes, S. J. (1980). Installing a computer-based patient information system. Journal of Nursing Administration, 10(5), 7-10.

Hunt, K. (1975). Do we really need more replications? Psychological Reports, 36, 587-593.

Ide, T. R. (1980). Canada in the age of microelectronics. In The computerized society: Implications for Canada 1979-1980 (The Walter L. Gordon lecture series, Vol. 4, pp. 9-38). Toronto: Omnigraphics Inc.

Jackson, G. B. (1980). Methods of integrative reviews. Review of Educational Research, 50(3), 438-460.

Jamison, D., Suppes, P., & Wells, S. (1974). The effectiveness of alternative instructional media: A survey. Review of Educational Research, 44(1), 1-67.

Johnson, D. S., Ranzenberger, J., Herbert, R. D., Gardner, R. M., & Clemmer, T. P. (1980). A computerized alert program for acutely ill patients. Journal of Nursing Administration, 10(6), 26-35.

Jones, R. J. D., Porter, D., & Rubis, R. (1983). A Survey of the Use of Microcomputers in B. C. Schools. Unpublished manuscript, Faculty of Education, Simon Fraser University, Burnaby.

Kamp, M., & Burnside, I. M. (1974). Computer-assisted learning in graduate psychiatric nursing. The Journal of Nursing Education, 13, 18-25.

Kelly, E. F., Pascarella, E. T., Terenzini, P., & Chapman, P. (1976). The development and use of the Adjective Rating Scale. A measure of attitude toward courses and programs. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, 1976.

Kirchhoff, K. T., & Holzemer, W. L. (1979). Student learning and a computer-assisted instructional program. Journal of Nursing Education, 18(3), 22-30.

Klassen, D. (1983). Computer literacy. In D. O. Harper & J. H. Stewart (Eds.), Run: Computer education (pp. 34-37). Monterey, California: Brooks/Cole Publishing Company.

Kraemer, H. C., & Andrews, G. (1982). A nonparametric technique for meta-analysis effect size calculation. Psychological Bulletin, 91(2), 404-412.

Kulik, J. A. (1981). Integrating findings from different levels of instruction. Paper presented at the American Educational Research Association, Los Angeles, April 1981. (ERIC Document Reproduction Service No. ED 208 040).

Kulik, J. A., Bangert, R. L., & Williams, G. W. (1983). Effects of computer-based training on secondary school students. Journal of Educational Psychology, 75(1), 19-26.

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

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

Ladas, H. (1980). Summarizing research: A case study. Review of Educational Research, 50(4), 597-624.

Larson, D. E. (1982). The use of computer simulation to teach calculation and regulation of intravenous flow rates to nursing students. Paper presented to the Annual Conference of the Association for the Development of Computer Based Instructional Systems, Vancouver, B. C., June 1982.

Lee, E. R. (1983). The computer's place in nursing education. In Perspectives in nursing - 1983-1985 (pp.212-223). Presentations at the 16th National League for Nursing Biennial Convention. New York: National League for Nursing.

Levine, D., & Wiener, E. (1975). Let the computer teach it. American Journal of Nursing, 75(8), 1300-1302.

Levine, M. E. (1980). The ethics of computer technology in health care. Nursing Forum, 19(2), 193-198.

Light, R. J. (1979). Capitalizing on variation: How conflicting research findings can be helpful for policy. Educational Researcher, 8(9), 7-12.

Light, R. J., & Smith, P. V. (1971). Accumulating evidence: Resolving contradictions among different research studies. Harvard Educational Review, 41(4), 429-471.

Lyons, C., Krasnowski, J., Greenstein, A., Maloney, D., & Tatarczuk, J. (1982). Interactive computerized patient education. Heart and Lung, 11(4), 340-341.

Meadows, L. S. (1977). Nursing education in crisis: A computer alternative. Journal of Nursing Education, 16(5), 13-21.

Merchant, R., & Sullivan, C. (1982-83). Microcomputers for everyone. Community College Review, 10(3), 8-11.

Miller, G. A. (1983). Computers in education: A non-Orwellian view. In D. O. Harper & J. H. Stewart (Eds.), Run: Computer education (pp. 17-20). Monterey, California: Brooks/Cole Publishing Company.

Mirin, S. (1981). The computer's place in nursing education. Nursing and Health Care, 11(9), 500-506.

Molnar, A. R. (1983). The next great crisis in American education: Computer literacy. In D. O. Harper & J. H. Stewart (Eds.), Run: Computer education (pp. 24-29). Monterey, California: Brooks/Cole Publishing Company.

Morin, P. J. (1983). A comparative study of learning styles and attitudes toward computer assisted instruction of nursing students. Paper presented at the Annual Conference of the Association for the Development of Computer Based Instructional Systems, Denver, Colorado, May 1983.

Naber, S. (1975). Computerized nurse-midwifery management: Its usefulness as a learning-teaching tool. Journal of Nurse Midwifery, 20(3), 26-28.

Naisbitt, J. (1984). _Megatrends. Ten new directions transforming our lives_. New York: Warner Books.

Pillemer, D. B., & Light, R. J. (1980). Synthesizing outcomes: How to use research evidence from many studies. _Harvard Educational Review, 50_(2), 176-195.

Pogue, L. M. (1982). Computer assisted instruction in the continuing education process. _Topics in Clinical Nursing, 4_, 41-50.

Pogue, L. M. (1984). _Computer assisted instruction study_. Unpublished research report, Department of Nursing, Medical College of Georgia, Augusta.

Porter, S. F. (1978). Application of computer-assisted instruction to continuing education in nursing. Review of the literature. _The Journal of Continuing Education in Nursing, 9_(6), 5-9.

Powell, J. A. (1983). The computer's place in health care. In _Perspectives in Nursing - 1983-1985_ (pp. 203-211). Presentations at the 16th National League for Nursing Biennial Convention. New York: National League for Nursing.

Ronald, J. A. (1983). Guidelines for computer literacy curriculums in a school of nursing. _Journal of the New York State Nurses' Association, 14_(1), 12-18.

Ronald, J. S. (1979). Computers and undergraduate nursing education: A report on an experimental introductory course. _Journal of Nursing Education, 18_(9), 4-9.

Rosenthal, R. (1978). Combining results of independent studies. _Psychological Bulletin, 85_(1), 185-193.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. _Psychological Bulletin, 86_(3), 638-641.

Rosenthal, R., & Rubin, D. B. (1979). Comparing significance levels of independent studies. _Psychological Bulletin, 86_(5), 1165-1168.

Schleutermann, J. A., Holzemer, W. L., & Farrand, L. L. (1983). An evaluation of paper-and-pencil and computer-assisted simulations. _Journal of Nursing Education, 22_(8), 315-323.

Science Council of Canada (1981). Policy issues in computer aided learning. (Proceedings P81/6, Catalogue No. SS24-17/1981E). Proceedings of a workshop sponsored by the Science Council of Canada Committee on Computers and Communication. Ottawa: Science Council of Canada.

Skiba, D. J. (1983). Computer literacy: The challenge of the '80s. Journal of the New York State Nurses' Association, 14(1), 6-11.

Smith, D. J., & Sage, M. W. (1983). Computer literacy and the education/training interface. Computers and Education, 7(4), 227-234.

Smith, M. L. (1980). Publication bias and meta-analysis. Evaluation in Education: An International Review, 4(1), 22-24.

Strain, A. R. (1983). Aspects of computer based instruction. (Department of Supply and Services, Contract Number 36100 - 1 - 0460). Ottawa: Communications Research Centre.

Sweeney, M. A., O'Malley, M., & Freeman, E. (1982). Development of a computer simulation to evaluate the clinical performance of nursing students. Journal of Nursing Education, 21(9), 28-36.

Tamarisk, N. K. (1982). The computer as a clinical tool. Nursing Management, 13(8), 46-49.

Taveggia, T. (1974). Resolving research controversy through empirical cumulation. Sociological Methods and Research, 2, 395-407.

Taylor, A. P. (1980). Clinical simulations in nursing. Nursing Times, 76(28), 1217-1218.

Thomas, D. B. (1979). The effectiveness of computer-assisted instruction in secondary schools. Association for Educational Data Systems Journal, 12(3), 103-116.

Timpke, J., & Janney, C. P. (1981). Teaching drug doses by computer. Nursing Outlook, 29, 376-377.

Tinker, R. L. (1983). Guest editorial. In D. O. Harper & J. H. Stewart (Eds.), Run: Computer education (pp. 209-211). Monterey, California: Brooks/Cole Pubishing Company.

Townsend, I. (1983). The second coming. Resurrection or Reservation? In M. Scholes, Y. Bryant, & B. Barber (Eds.), _The impact of computers on nursing. An international review_ (pp. 334-346). Amsterdam: Elsevier Science Publishers B. V.

Tucker, R. N. (1981). Microcomputers in education. _Educational Media International_, _3_, 2-3.

Tymchyshyn, P., & Helper, J. (1981). PLATO goes to the hospital. _Nursing Careers_, March-April, 16 & 19.

Valish, A. U., & Boyd, N. J. (1975). The role of computer assisted instruction in continuing education of registered nurses: An experimental study. _The Journal of Continuing Education in Nursing_, _6_(1), 13-32.

Visonhaler, J. F., & Bass, R. K. (1972). A summary of ten major studies on CAI drill and practice. _Educational Technology_, _12_(7), 29-32.

Wuhrer, B. (1983). _The impact of micro computers on computer education in British Columbia_. Unpublished research report, British Columbia Institute of Technology, Burnaby.

Yin, R. K., Bingham, E., & Heald, K. A. (1976). The difference that quality makes - the case of literature reviews. _Sociological Methods and Research_, _5_(2), 139-156.

Yucha, C., & Reigeluth, C. M. (1983). The use of computers in nursing education, practice, and administration. _Computers and Education_, _7_(4), 223-226.

Zielstorff, R. D. (1978). Nurses can affect computer systems. _Journal of Nursing Administration_, _8_(3), 49-53.

Zielstorff, R. D. (Ed.). (1980). _Computers in nursing_. Rockville, Maryland: Aspen Systems Corporation.