EVALUATION OF A FINANCIAL DECISION SUPPORT SYSTEM IN BUSINESS EDUCATION: AN EXPERIMENTAL STUDY

William R. King  
University of Pittsburgh

G. Premkumar  
University of Pittsburgh

K. Ramamurthy  
University of Pittsburgh

Follow this and additional works at: http://aisel.aisnet.org/icis1988

Recommended Citation
http://aisel.aisnet.org/icis1988/29
EVALUATION OF A FINANCIAL DECISION SUPPORT SYSTEM IN BUSINESS EDUCATION: AN EXPERIMENTAL STUDY

William R. King  
G. Premkumar  
K. Ramamurthy  
Joseph M. Katz Graduate School of Business  
University of Pittsburgh

ABSTRACT

This study explores the use of computers in business education and its impact on the performance of students. A controlled experimental study has been carried out to determine the differences in performance of students in a computer assisted instruction (CAI) group with that of a control group. It investigates the relationship between performance and various student-specific characteristics such as aptitude, attitude, sex, domain experience, domain expertise, and system experience. The relationships between attitude towards CAI, fulfillment of expectations, and satisfaction with the system and course have also been explored. The results indicate that CAI has favorable effects on students' performance and that personal attributes have relatively less important roles to play.

1. INTRODUCTION

Computing in higher education has expanded far beyond its traditional uses in hard sciences. No longer limited to scientific and numerically oriented applications, it now supports a much wider variety of tasks for many more people in the university. With the advent of Personal Computers (PC) and the growth of end-user computing (Rockart and Flannery 1983), the role of computers in industry is also undergoing rapid change. Most organizations have started using PCs extensively and require that their managers and supporting staff be conversant with computers to perform their jobs effectively (King and Snitkin 1986; Jones and Lavelli 1986). This has put pressure on business schools to prepare their students, the future managers, to perform well in this new computing environment. As a consequence, during the past few years business schools have invested heavily in hardware and software and started emphasizing greater computer usage in their curriculum.

The use of computers in education and its effect on student performance has been extensively studied in education research and it has been fairly well established that computers improve the performance in primary school education (Edwards et al. 1975). However, there have been only a few studies in higher education (Hebenstreit 1985) and the results have also not been very conclusive (Jamison, Suppes and Wells 1974; Kulik, Kulik and Cohen 1980).

Traditionally computers have been used in business schools for word processing, spreadsheet analysis, statistical analysis, and programming (Frand and McLean 1986), but in recent years decision support systems (DSS) and a variety of courseware have been developed for business school education (Chandler 1984). However, there has been no systematic evaluation of the impact of this software on business education or on performance of students in their coursework or later in their professional work. Most of the courseware development is based on the implicit assumption that computers reduce the computational burden and allow the students to use their mental skills in a more creative manner (Boen 1984). Also, since most business students in their professional career would be using computers extensively in their jobs, this education would provide suitable training to meet their job requirements. It is not clear whether excessive dependence on computers would result in reduced emphasis on teaching the basic principles of the course and turn out students who are only able to generate outputs from computers without sufficiently understanding the underlying principles and assumptions. Hence, it would be interesting to study the role of computers in business education and evaluate their performance.

The objectives of this study are:

a) To explore the use of computers in business education and describe a decision support system that was specifically developed to provide a real-world environment for a finance course.

b) To evaluate the performance of computers in business education and test a set of hypotheses on the relationships between different variables using a quasi-experimental research design.

The paper is organized in four sections: a review of past research on CAI, a description of a decision support system developed for teaching and research, formulation of a set of hypotheses and a research design to test them,
and an analysis of the data and examination of the relationships between different variables.

2. COMPUTERS IN BUSINESS EDUCATION

The dramatic growth of end-user computing (Rockart and Flannery 1983) and the popularity of decision support systems in organizations (Keen and Scott Morton 1978) have necessitated organizations to recruit and train functional personnel, who are conversant with computers (Jones and Lavelli 1986) and business schools are equipping themselves to meet this demand. The low cost of PCs and liberal grants from computer manufacturers have enabled quite a few schools to be self sufficient in hardware. However, it is becoming increasingly realized that mere installation of computers will neither increase the effective utilization of computers nor improve the computer awareness of students (Kling 1986). In most cases they tend to get used as word processors, utilizing only very minimal capabilities of the equipment. Hence, for effective utilization of this equipment, it is essential that installation of hardware be supported by planned programs to:

a) educate the students to use the entire range of facilities available in the computer

b) identify areas for innovative use of computers in coursework

c) motivate the faculty to take active interest in the development of computer integrated coursework

d) develop systems and support structure to provide stability in the operation of the computer center

e) develop evaluation procedures to motivate and monitor the utilization of computers in the coursework

Dempster et al. (1987) identify four stages of evolution of computers in business education. They are a) integration of computers in the curriculum by providing the hardware, software and service, b) development of business education coursework, c) integration of instruction through realistic corporate databases, and d) development of DSS to support various facets of instruction. While most CAI programs have tended to concentrate on the first stage using standard commercial packages to teach programming, accounting, etc., we were interested in studying the last two stages, which are more sophisticated uses of computers in education.

It would be very difficult to develop a comprehensive evaluation measure for CAI, given the diverse set of technical and organizational objectives of most institutions. However, central to the introduction of computers are two major objectives that we believe are important:

a) increase the awareness of student and faculty to computing and facilitate greater usage of the computer

b) to provide an information rich environment (i.e., easy access to large volumes of relevant information) to widen the vision and perspective of students and improve their decision making capabilities.

Quite often, the former (relatively narrower) objective of computer literacy is stressed more than the second objective which takes a much broader perspective of information use in education. By providing an information rich, environment students and faculty, with minimal computer skills, can find new and innovative ways to use the information in their coursework. Some of the features of an information rich environment are:

a) On-line access to real-world financial data available from well established and reliable providers such as Dow-Jones, CRSP (1986), Compustat (Standard and Poor 1985), Dun and Bradstreet, etc.

b) Access to market share profitability analysis of companies provided by Profit Impact Marketing Studies, PIMS (Buzzell and Galec 1987)

c) Access to other on-line databases specific to the area.

d) Access to library information and keyword search facilities.

e) Availability of a wide variety of business and scientific software to manipulate and analyze the data.

f) Communication facilities through networking.

Such an information rich environment not only provides facilities for access to large volumes of on-line data, but tools and techniques to analyze and use them for decision making. The fact that the students from business schools are likely to extensively use DSS and interact with corporate databases in their professional work strongly motivates development of an environment where students become conversant with communicating through electronic mail, accessing real world databases, using decision support systems to retrieve, manipulate, and analyze the data for creative decision making as a part of their coursework.

Computers can be used in a variety of ways as an aid to education (Castellan 1986; Balkovich, Lerman and Parmelec 1985; Athey 1983). They may be broadly classified into a) computer assisted tutoring, b) computer simulation, c) computer as a laboratory, d) computer as a tool, e) computerized modelling, and f) computer as a link to the real world. Our interest is in the context of the computer as a link to the real world and we shall describe below a decision support system that was developed with.
this perspective as an instructional system for a graduate level financial management course.

3. FINALLY: A DECISION SUPPORT SYSTEM FOR INSTRUCTION

Finally was developed as a comprehensive on-line DSS for access to real-world stock market and financial data, retrieved from COMPUSTAT and the CRSP (Center for Research on Stock Prices) databases. COMPUSTAT provides data on annual financial accounting information for the firms listed in the Standard and Poor 500 index for the most recent 20 years. The CRSP data from the University of Chicago provides data on price, volume, cash and stock distribution, and other related information on securities for firms traded in New York, Amex, and other leading stock exchanges. Together, they provide an enormous wealth of data that are useful to carry out a variety of analysis such as:

a) Analysis of an industry
b) Analysis of a company
c) Analysis of a project
d) Risk analysis of portfolios
e) Merger and acquisition analysis, etc.

The data available in these two data banks have to be stored in a form that can be easily retrieved and analyzed. A comprehensive decision support system with facilities for retrieving, manipulating, modelling, and analyzing the data would be an ideal tool for some of the analysis mentioned above.

The availability of the data bank and a user friendly environment (Application System) on an IBM 4381 for development of a decision support system motivated the development of a DSS for use in a financial management course. The data available on CRSP and COMPUSTAT were transferred to a relational database and stored in a SQL/DS environment facilitating easy access and manipulation by Application System. The system was developed with an important objective of integrating the DSS into the coursework so that students could learn the basic concepts of cashflow, capital budgeting, risk analysis, and portfolio management by using data from real world companies rather than with textbook problems dealing with fictitious companies, and hypothetical accounting and stock information. Also, the system was intended to serve the needs of all researchers. To provide access to a wide range of users, a menu driven system with a very friendly user interface was developed.

Finally consists of a set of models each concentrating on a particular aspect of financial management. For instance, the cash flow analysis model provides the facility for analysis of cash flow of real world companies, calculation of net present values, development of weighted cost of capital, and forecasting of cash flow for future periods. Similarly, risk analysis is another model provided by Finally which helps the instructor to teach the concepts of risk, the interrelationship between beta values and other factors such as market rate of return and premium on loans, and variation of risks between companies within an industry or between industries. Facilities are provided to the instructor to retrieve the data for any company of their choice and calculate the various risk related factors using a menu driven interface. A portfolio of projects can also be formed to study the impact of different design criteria on the performance of the portfolio. Hence, the model serves the dual purpose of being an aid to class room instruction as well as a tool for analysis and decision making. There are seven different modules in the total system dealing with time value of money, capital budgeting, financial statements, financial ratio analysis, cashflow analysis, risk analysis, and general information.

4. BACKGROUND

Systematic comparison of computer-based and conventional teaching methods has been extensively carried out in the early 1970s in education research. Most research has examined the difference between an experimental group who received part of their instruction through computers and a control group who received their instruction by conventional teaching methods. These studies have generally reported that computer based teaching in a supplementary mode to conventional instruction in elementary schools was more effective.

However, the results have been inconclusive in the case of CAI in higher education (Jamison, Suppes and Wells 1974), PLATO (Programmed Logic for Automatic Teaching Operators) and TICIT (Time Shared Interactive Computer Controlled Information Television) are two large systems where extensive development and evaluation of CAI in colleges have been conducted (Alderman 1978). With inconclusive results, the earlier enthusiasm of expecting the same magnitude of effects in favor of CAI as experienced in elementary schools faded away. However, Kulik, Kulik and Cohen (1980) carried out a meta-analysis of past research, a process of teasing out generalizations from past research (Glass 1976), and found that the computer group out-performed the control group and that this advantage translated to an increase from the fiftieth to the sixty-sixth percentile in final examinations in a variety of courses. In 37 of the 39 studies analyzed, they found the performance of students in the computer group to be better than the control group, signifying a clear preference for computers.

Clark (1983) argues that, while most analyses have shown positive learning effects for the newer media over more conventional treatments, there has been considerable potential for confounding in the reviewed research. For example, Clark argues that the difference between the two groups drops from 0.5 standard deviations to 0.13
standard deviations if the instructor effect is removed by having the same instructor for both the groups. Another source of confounding was suggested to be novelty, as evidenced by a decrease in the differences between the media and conventional treatments with lapse of time (Clark and Salomon 1986). Clark and Snow (1975) and Kulik, Kulik and Cohen (1980) also report significant reduction in effect as the time duration for treatment increased, indicating that novelty may play a part in the reduction of effect. Clark (1983) claims that there is compelling evidence that the larger effects have been due to systematic but uncontrolled differences in content, novelty, and/or teaching method between the two groups but not necessarily to CAI per se. Hence, Clark and Salomon (1986) suggest that, though such overall studies may be useful for understanding the role of media in education, it is necessary to understand the cognitive aspects of the media attributes (Salomon 1979) for better understanding of the process of learning.

The outcomes that have been typically studied have been:

a) Student achievement
b) Correlation between aptitude and achievement
c) Course completion or attrition rate
d) Student attitudes towards the course
e) Instructional time
f) Time taken for completion of the experiment
g) Student satisfaction with instructional media and with the course.

There are other important factors that influence the outcome. They are:

a) **Course level**: The course could be at an introductory or higher level. The fit between the CAI system and course level is critical. For instance, a CAI tutor may be an excellent mode for an introductory level course, while a more focussed DSS may be required for a higher level course.

b) **Quality of the system**: There has been extensive research in information systems linking satisfaction to the quality of the system (Ives, Olson and Baroudi 1983; Srinivasan 1985). CAI mode of instruction will be fruitful only if the software quality is good and leads to satisfaction. Various factors, such as user friendliness, reliability, response time, and versatility, influence the quality of the system.

c) **Fit between the system and the coursework**: In situations where CAI is used as a supplement to traditional mode of instruction, it is important that the two processes are integrated in a manner that provides synergies for learning (Criswell and Swezey 1984; Nelson 1985).

d) **Structurability of the course**: Bok (1986) found structurability of the course to be an important cri-
teria for introduction of CAI. Use of computers can relieve the students/users of routine activities and enable development of creative and intellectual problem solving skills (Adler 1986; Athey 1983).

e) **System experience**: Prior experience with computers tends to alleviate the initial fears and problems in getting familiarized with the system and this to a large extent improves the performance of students (Lucas 1978).

f) **Duration of treatment**: Kulik, Kulik and Cohen (1980) found duration of interaction with the computer system in the course to be a significant dimension influencing the outcome. In some studies the computer was used for the full duration of the course while in others it was used only for a few classes. It can be expected that the impact of CAI will be influenced by the length of exposure (Clark 1983).

g) **Instructor Characteristics**: Though computers have been found to reduce the time for instruction (Kulik, Kulik and Cohen 1980), sometimes the instructors may perceive them as an intrusion into their course-work and may not fully approve of its use. Also, the instructors’ style should match the students’ learning style, environmental demands, and system constraints (Gregore 1979). Simon and Boyer (1974) delineated four facets of teaching behavior: cognitive, socio-emotional, substantive, and communication. It is clear that the socio-emotional and communication facets are dependent on the instructor characteristics and play a major part in the learning process (Dunkin and Barnes 1986). Hence, it is not surprising that instructor effect has been observed to have the maximum explanatory power on student achievement (Kulik, Kulik and Cohen 1980).

h) **Personal attributes**: Various other personal attributes such as sex (Gattiker 1987), domain expertise, previous domain experience, aptitude (Dunkin and Barnes 1986; Gage 1979), and cognitive style of individual (Clark and Salomon 1986) influence the effectiveness of CAI.

**Summary**

Although there are inconsistent results on the effects of CAI, it is generally accepted that CAI improves students' performance. There are a few intervening variables, which can significantly affect the outcomes, that need to be controlled in experiments. It has been found that instructor effect is one of the most significant and once it is controlled the difference in performance between the CAI and non-CAI group considerably diminishes. It has also been found that correlation between aptitude and achievement is higher in the non-CAI group compared to the CAI group. There is no conclusive evidence re-
5. RESEARCH DESIGN

Evaluation of effectiveness of CAI is a complex task, requiring complete understanding of the context of the use of CAI. Most earlier studies have evaluated only the tangible outcomes, but there are other intangible or less direct outcomes, such as heightened confidence in using the computers, improved quality of course material, faculty development, and providing groundwork for future innovation (Kulik, Kulik and Cohen 1980). Hence, any decision on a research design is dependent on the objectives for introduction of CAI, purpose of the research study, and the variables proposed to be investigated in the study. Isaac (1977) provides an excellent overview of the learning outcomes and the possible methods of their evaluation. Nelson (1985) has provided a good survey of the strengths and weaknesses of each one of the methods.

Since the objectives of this study were to measure the effectiveness of CAI in a business education context, as well as link it with various psychological and demographic variables, a quasi-experimental research design (Campbell and Stanley 1966) augmented by a survey instrument to measure the psychological and demographic variables was deemed to be most suitable. This type of design is well in line with the general design guidelines that have been extensively used in educational research, particularly in the context of evaluation of CAI (Kulik, Kulik and Cohen 1980).

The research design consists of using a group of students taking a graduate level course in financial management for evaluating the CAI mode of instruction that was specifically developed for that course. The students were randomly assigned to two groups, CAI (experimental) and non-CAI (control) group, and their performance evaluated based on a case study that required utilizing the principles and theories taught in the course. The experimental group used the CAI system to analyze the case study while the control group had to analyze the case manually with the assistance of a calculator.

Experimental Instrument: Since the students were enrolled in a finance course of the MBA program, a fairly complex case incorporating the concepts of risk analysis, cost of capital, and capital budgeting was generated in close coordination with the concerned finance faculty. The case study involved not only calculation of various financial values, but also required articulation of the assumptions on various risk factors and a detailed analyses before arriving at a final decision. This tested the students' grasp of the course material and their ability to use it for analysis. The students were evaluated for both accuracy and quality of decision making. To maintain a fair degree of compatibility of task complexity between control and treatment group and avoid the potential trap of ending up with evaluating the computer capability per se, additional data and information were provided to the control group to minimize the computational burden. The students in the CAI group were well conversant with the system and had used it as a part of the course for various other course assignments.

Psychological and demographic variables were measured using two sets of questionnaires, one administered at the beginning of the term and the other at the end of the term. The first questionnaire measured the attitude of students to CAI, their prior experience with computers, and various demographic variables. The second questionnaire measured the satisfaction of the students with the CAI system as well as with the mode of instruction and included items eliciting behavioral manifestations.

Experimental Error Control: In view of the conflicting results that have been reported in the past (Clark and Salomon 1986), it is extremely important to avoid contamination and confounding from extraneous variables by achieving control on variability. The three types of variances that need to be controlled are:

a) Experimental variance which should be maximized by designing the experimental conditions to be as different as possible. In this study, there were two treatment conditions: traditional instruction as the control group and traditional instruction supplemented with CAI as the treatment group.

b) Extraneous variances that are to be controlled through any one of the following: blocking, randomizing, treating the extraneous variables themselves as independent variables, matching the subject on one or more variables, or statistical method of analysis of covariance.

c) Error variance, to be minimized through careful control of measurement conditions and using reliable and valid measures (Kerlinger 1973).

As discussed in the earlier section, there are a number of intervening factors that have led to inconsistent and contradictory results in past research studies. Control for variances that could otherwise lead to conflicting results have been ensured through the following precautions.

a) Kulik, Kulik and Cohen (1980) found that instructor effect was the most significant determinant of performance and that if it was controlled there was very little significant difference in performance between the CAI and non-CAI group. In this study, both groups were taught by the same instructor.

b) The variability caused by student specific factors has been controlled through randomized assignment of students to the two treatment groups, by making use
of hashing and random number generation techniques.

c) The third aspect of variance control, minimization of error variance, has been ensured through the use of appropriately validated and robust measuring instruments.

d) An additional source of variance could occur due to bias in the evaluation of performance. A set of benchmarks for evaluation was established to ensure objectivity in evaluation of performance. Two independent evaluations of the performances were carried out and inter-rater reliability ensured before an average of the two scores was used as a performance measure.

e) The historical effect (Kulik, Kulik and Cohen 1980) of difference in outcome due to the groups undergoing the treatment in different semesters was controlled by subjecting both groups to the treatment at the same time.

Research Hypotheses

Multi-dimensional criteria have been recommended for evaluating effectiveness of information systems (King and Epstein 1983). In this study, we have evaluated three dimensions of effectiveness:

a) Student performance
b) Student attitude
c) Student behavior intentions

Performance: This was assessed in terms of two dimensions: decision making quality and accuracy. Although Jamison (1974) found it hard to show an improvement in performance for the CAI group, subsequent studies have generally shown that CAI mode of education results in improved performance (Kulik, Kulik and Cohen 1980). In this study we were interested in knowing if there is an improvement in performance due to CAI in a business education context where such studies have been quite rare. Also, the relationship between the demographic variables and performance were explored. Seven hypotheses, stated in a null form, are listed below.

Hypothesis 1a: Accuracy of performance will not be different between the CAI and non-CAI group.

Hypothesis 1b: Decision making quality will not be different between the CAI and non-CAI group.

Domain Expertise: Domain expertise has been observed to play a major role in the decision making process (Simon 1979). Given the importance of this variable it is expected that it will be a major factor influencing decision performance. It is also expected that domain expertise will play a more important role in the non-CAI rather than in the CAI group as the DSS would have already captured some of the domain expertise in its models.

Hypothesis 2a: Prior domain expertise will not influence performance accuracy.

Hypothesis 2b: Prior domain expertise will not influence decision quality.

Domain Experience: Previous work experience may influence the performance as it is fair to assume that familiarity with the domain area increases knowledge of the area and thereby improves performance (Sanders and Courtney 1985).

Hypothesis 3a: Past domain experience will not influence performance accuracy.

Hypothesis 3b: Past domain experience will not influence decision quality.

Sex: Mixed results have been reported on the effects of sex over performance. Some studies have reported that women are less computer literate than their male counterparts even after attending a computer course (Johnson, Johnson and Stanne 1986), whereas Anderson (1987) found that females were better than males in analyzing algorithms. Gattiker (1987) found difference in performance among the two sexes. The following hypotheses are tentatively proposed.

Hypothesis 4a: The sex of students will not influence performance accuracy.

Hypothesis 4b: The sex of students will not influence decision making quality.

Attitude: The linkage between prior beliefs and expectations leading to attitudes, which in turn influence the person's intentions to perform an action (e.g., use the system) is a subset of the fairly well accepted and validated Fishbein's model (Fishbein and Azjen 1972). It is also well established in IS literature that satisfaction with the system is related to the attitude of the individual (Lucas 1978). Hence it is hypothesized that:

Hypothesis 5: Satisfaction with CAI mode of instruction is not related to the attitude the students hold towards CAI.

Hypothesis 6: Satisfaction with CAI system is not related to the attitude the students hold towards CAI.
Behavior: Fishbein's behavioral model discussing how situational variables operate through social and psychological variables to influence behavior is widely known (Liang 1986). Behavior is influenced by expectations, prior beliefs, and by attitudes. The model also suggests a bi-directional relationship between belief and attitude. Behavioral manifestations in turn are closely linked to satisfaction. We therefore propose that:

Hypothesis 7: Satisfaction with CAI has no impact on the students' future action plans.

Operationalization of the Constructs

The experimental model has many new constructs that have not been operationalized before; hence, considerable care needs to be taken in the measurement phase to ensure that the instruments measure the underlying constructs. Most studies in IS research as well as in educational research, in the context of CAI, lay very little emphasis on the measurement aspects, leading to inconclusive and contradictory results. Extreme care was taken to operationalize the constructs based on available research literature and, wherever possible, available instruments were adapted to suit the study.

Performance: Some of the major factors considered in evaluating the performance of DSS are the time taken to arrive at a decision, number of alternatives considered, confidence in the decision, and quality of decision. In this study it was thought fit not to consider "time" taken to solve the case as an important factor because, even a priori, it would be decidedly in favor of the CAI group (comparison of computer versus calculator).

If performance is measured in terms of an objective test, as in the past media research, then total marks could be an useful indicator of outcome. Since our experiment was a case study which involved both mathematical computation as well as analysis of the data, we have two dimensions of performance: accuracy and decision making quality. Accuracy is determined by the precision of the final solution values while decision making quality is evaluated by the alternatives considered, assumptions made, reasoning process adopted, etc.

Attitude: The importance of value perceptions and attitude towards information systems has been reported to be an important determinant of effectiveness (King and Rodriguez 1978) and has also been found to play a major role in the effective use of CAI (Clark and Salomon 1987). Some of the major advantages of CAI according to Boen (1983) are:

a) Greater time flexibility
b) Quick feedback leading to greater motivation and enthusiasm to learn

c) Better maintainability of the quality and reliability of the coursework irrespective of the variation in instructors' abilities

d) Flexibility to pace the speed of learning by choosing different modules

e) Ability to experiment and thereby learn creatively

f) Reduction in routine computational activities thereby freeing the mind for more inferential analysis and better problem solving.

Attitudes were measured with respect to these attributes which were suitably reworded to elicit their agreement/disagreement on a five point Likert type scale. Attitude was theorized to be a composite measure of these attributes. Also, past difficulties, experience, and confidence in interacting with computers were measured using single item statements.

Experience with computers: The experience with computers was measured using seven items measuring familiarity with some of the standard software packages extensively used in business schools (Ffrand and McLean 1985). This instrument was developed for a previous study and was found to be a reliable measure of experience with computers in the context of business education.

Satisfaction: Satisfaction as a construct has been extensively researched in different disciplines. In the context of CAI, we were interested in measuring two dimensions of satisfaction: satisfaction with the course and satisfaction with the system. Satisfaction with the CAI system is very similar to measuring satisfaction with an information system and various validated instruments are available for measurement (Bailey and Pearson 1983; Ives, Olson and Baroudi 1983). The Ives, Olson and Baroudi instrument, which has been extensively tested for various psychometric properties, was adapted to suit an educational CAI environment by deleting some items that specifically dealt with DP services which were absent in this context. We used a fourteen item scale to measure satisfaction with the system. Although there are well formulated instruments for evaluation of a course, they deal with the total course including content, schedule, and instructor capabilities, etc., whereas we were more interested in evaluation of the course in the CAI context. Hence we have used a single item measure of satisfaction with the course.

Behavioral intention: It measures the behavioral manifestations in terms of recommending the course to others and taking more CAI courses in the future. It is expected that satisfied users would desire more CAI courses as well as recommend them to others. Behavioral intention was measured on a five point Likert type scale varying from strong agreement to strong disagreement.
Demographic variables: Sex, domain experience, and domain expertise were all measured on dichotomous categorical scales. Aptitude was measured using GMAT scores, which have been widely adopted as a surrogate in past educational research (Kulik, Kulik and Cohen 1980).

6. DATA ANALYSIS AND RESULTS

Testing of Psychometric Properties of the Instrument

Since some of the instruments used for measuring the construct have not been previously tested and some instruments were modified, it was necessary to ensure content validity, construct validity and reliability of the instruments. Content validity implies that all aspects of the attribute being measured are considered by the instrument, i.e., the measurement is complete and sound. The exhaustive literature review process through which the lists of items measuring the constructs were identified lends credibility to ensure face validity. These items were further refined through expert opinions.

Construct validity consists of two major validity concepts: convergent validity and discriminant validity. Convergent validity measures the degree to which multiple attempts measuring the same concept through maximally different methods are in agreement. In our study convergent validity was examined by measuring the correlation of each individual item with the aggregate less that item (Ives, Olson and Baroudi 1983). In all the cases, the correlation is significant at a level of less than 0.00, indicating that the items hang together to form a single construct. Discriminant validity is the degree to which a concept differs from other concepts and it is usually established through factor analysis. Factor analysis identifies the items which measure the same construct and those items that do not load to any factor significantly and need to be dropped (Green 1978). Factor analysis of the data for "attitude" revealed that, except for one item, all items were loading onto the same factor. Hence, this item was considered an outlier and removed from the list of items measuring this construct.

Reliability. Some of the measures of reliability are test-retest reliability, inter-item reliability, and internal consistency. Internal consistency was assessed using Cronbach’s Alpha (Cronbach 1951), and was found to be 0.7952 for the attitude measure and 0.909 for the satisfaction measure. These values indicate that the measures exhibit sufficient internal consistency and are therefore reliable.

Randomization. To test for randomized allocation of students to the two groups, we examined the aptitudes of the students before they entered the course through a common measure (GMAT score). T-test of the difference in mean GMAT scores between the two groups, as displayed in Table 1, reveals insignificance of the differences (p = 0.000) confirming random assignment of students.

<table>
<thead>
<tr>
<th>Table 1a. T-test for Randomization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>Aptitude (GMAT)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 1b. Test for Inter-Rater Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance</td>
</tr>
<tr>
<td>-------------</td>
</tr>
<tr>
<td>Accuracy</td>
</tr>
<tr>
<td>Decision Making</td>
</tr>
</tbody>
</table>

Discussion of the Results

Table 2 provides a summary of the results examining each hypothesis. Detailed discussion of each result follows.

<table>
<thead>
<tr>
<th>Hypoth. Number</th>
<th>Table Ref.</th>
<th>Independent Variable</th>
<th>Dependent Variable</th>
<th>Findings &amp; Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>H-1a</td>
<td>3</td>
<td>CAI versus TRADITIONAL EDUCATION</td>
<td>ACCURACY</td>
<td>CAI group superior</td>
</tr>
<tr>
<td>H-1b</td>
<td></td>
<td></td>
<td>DEC QUALITY</td>
<td>CAI group superior</td>
</tr>
<tr>
<td>MULTIPLE REGR. ANALYSIS</td>
<td>7</td>
<td>Attitude Expertise Sex Dom. Exp. System Exp.</td>
<td>ACCURACY</td>
<td>No variable important in CAI group for any performance dimension. Aptitude and domain experience important for non-CAI group.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hypoth. Number</th>
<th>Table Ref.</th>
<th>Independent Variable</th>
<th>Dependent Variable</th>
<th>Within Groups Non-CAI group</th>
<th>CAI group</th>
<th>Pooled group</th>
</tr>
</thead>
<tbody>
<tr>
<td>H-2A</td>
<td>4</td>
<td>EXPERTISE</td>
<td>ACCURACY</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>H-2B</td>
<td></td>
<td></td>
<td>DEC QUALITY</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>H-3A</td>
<td>5</td>
<td>DOMAIN EXPERIENCE</td>
<td>ACCURACY</td>
<td>+</td>
<td>+</td>
<td>ns</td>
</tr>
<tr>
<td>H-3B</td>
<td></td>
<td></td>
<td>DEC QUALITY</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>H-4A</td>
<td>6</td>
<td>SEX</td>
<td>ACCURACY</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>H-4B</td>
<td></td>
<td></td>
<td>DEC QUALITY</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>H-5</td>
<td>8</td>
<td>ATTITUDE</td>
<td>CAI PROCESS SATISFACTION</td>
<td>n/a</td>
<td></td>
<td>ns</td>
</tr>
<tr>
<td>H-6</td>
<td>8</td>
<td>ATTITUDE</td>
<td>CAI SYSTEM SATISFACTION</td>
<td>n/a</td>
<td></td>
<td>ns</td>
</tr>
<tr>
<td>H-7</td>
<td>8</td>
<td>PROCESS &amp; SYSTEM SATISFACTION</td>
<td>BEHAVIORAL ACTIONS &amp; PLANS</td>
<td>n/a</td>
<td></td>
<td>+</td>
</tr>
</tbody>
</table>

ns - Not statistically significant n/a - Not applicable + - Positive influence - Negative influence
Performance

As discussed earlier, the two constructs of performance are labelled "accuracy" and "decision making quality." The results of the T-test, evaluating the difference in performance between the control group and treatment groups are shown in Table 3.

Table 3.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Group</th>
<th>Size</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>T</th>
<th>Sig. Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>Non-CAI</td>
<td>21</td>
<td>27.80</td>
<td>12.75</td>
<td>-5.42</td>
<td>0.000</td>
</tr>
<tr>
<td>Decision</td>
<td>CAI</td>
<td>21</td>
<td>26.78</td>
<td>15.79</td>
<td>-3.57</td>
<td>0.001</td>
</tr>
</tbody>
</table>

As evident, the performance of the CAI group is superior to the non-CAI group on both the dimensions of performance (p < 0.000, and p < 0.001). Hence both hypotheses H1a and H1b are rejected. This leads us to believe that CAI has been useful in not only improving the final objective performance (accuracy) in the experiment, but also in better understanding the principles of the course and applying them in the case study (decision making quality).

Having confirmed that the CAI group performs better than the non-CAI group, further subcategorization of the control and treatment groups was performed based on their domain expertise. The results of the T-test examining the difference in mean performance dimensions between the two subgroups in each group is shown in Table 4.

Table 4.

<table>
<thead>
<tr>
<th>T-test - Domain Expertise</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>----------------------------</td>
</tr>
<tr>
<td>Accuracy</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>Decision</td>
</tr>
<tr>
<td>2</td>
</tr>
</tbody>
</table>

The results indicate that domain expertise does not appear to influence either of the two dimensions of performance. After a lapse of three weeks, to eliminate other potential bias, the members of the control group were also asked to perform an identical case using the system. Fourteen of the original 21 students in the non-CAI groups participated in the experiment. The entire data set was pooled and an analysis of the influence of prior domain expertise (represented by their undergraduate specialization) was made. The results are shown in Table 3. Contrary to normal expectations, it can be observed that the differences in the two dimensions of performance between the two groups are not statistically different. There is insufficient evidence to reject hypotheses H2a and H2b, which suggests that domain expertise does not influence performance.

Students were separated into two subgroups based on domain experience, whether they previously had one or more years of finance/accounting experience or none. The difference in the two dimensions of performance between the two groups have been assessed using T-test for all the three combinations as discussed above for domain expertise.

Table 5.

<table>
<thead>
<tr>
<th>T-test - Domain Experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>----------------------------</td>
</tr>
<tr>
<td>Accuracy</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>Decision</td>
</tr>
<tr>
<td>2</td>
</tr>
</tbody>
</table>

The results, shown in Table 5 reveal partial support for hypotheses H4a and H4b. Higher domain experience appears to have some effect on the accuracy dimension for both the CAI (p < 0.053) and the non-CAI (p < 0.035) groups, but not for the pooled group(p < 0.581). However, there is no influence on decision quality for any of the groups. These findings do not conform to past research or prior expectations. It is quite likely that the nature of past work experience may not have matched the type of problem (risk analysis, cost of capital, and investment decision) examined in this study.
Gattiker (1987) found in his study that sex was a major determinant of performance. Hence, we were interested in evaluating the impact of sex on performance. Table 6 shows the results of the T-test of the difference in the performance dimensions between the two sexes for the three combinations. The results are not significant enough to reject the hypotheses H4a and H4b, leading us to infer that sex is not a major criterion in influencing performance.

Table 6.

<table>
<thead>
<tr>
<th>T-test - Sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size Mean</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>Subgroup:</td>
</tr>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

Having evaluated separately the link between performance and a few of the independent variables, it was decided to use a multiple regression model to evaluate the joint influence of the various student-specific attributes on the performance dimensions. Since the tests for the control group and experimental group were held under different test conditions, it was decided to run separate multiple regression models for each of the groups. The independent variables considered were aptitude of the student (GMAT scores), domain experience (WEXP), domain expertise (UNGRAD), SEX, and system experience (SYSEX). In the case of the non-CAI group, system experience was not used in the model as it was irrelevant to the context.

In the case of the CAI group, it is observed that none of the predictor variables referred to above have entered the model for either dimension of performance except domain expertise where decision making quality is the dependent variable. However, the explanatory power of this model ($R^2 = 4.6\%$) is so low that its practical worth is questionable. In the case of the non-CAI group, however, we observe that aptitude and domain experience appear in the model (Table 7).

These results support Gage's (1978) findings of a higher correlation between aptitude and achievement in the non-CAI group relative to the CAI group. It seems to indicate that CAI is helpful in not only relieving the users of the computational burden, but more importantly in minimizing the dependence of the quality of decision performance on the personal characteristics of the users. Having established (through hypotheses H1a and H1b) that the CAI group performed significantly better than the non-CAI group and that none of the student specific independent variables have contributed significantly to the performance, it can be concluded that the CAI mode was the main determinant of performance in this study.

Table 7.

Multiple Regression Analysis

<table>
<thead>
<tr>
<th>Group - CAI</th>
<th>Sample Size = 21</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Variable</td>
<td>Independent Variable</td>
</tr>
<tr>
<td>Accuracy</td>
<td>None of the variables entered the equation.</td>
</tr>
<tr>
<td>Decision</td>
<td></td>
</tr>
<tr>
<td>Making</td>
<td></td>
</tr>
<tr>
<td>Quality</td>
<td>Constant 49.884</td>
</tr>
<tr>
<td>Group - Non-CAI</td>
<td>Sample Size = 21</td>
</tr>
<tr>
<td>Dependent Variable</td>
<td>Independent Variable</td>
</tr>
<tr>
<td>GMAT</td>
<td>0.1039</td>
</tr>
<tr>
<td>Accuracy</td>
<td>WEXP</td>
</tr>
<tr>
<td>Decision</td>
<td>GMAT</td>
</tr>
<tr>
<td>Making</td>
<td>Quality</td>
</tr>
<tr>
<td>Group - Pooled</td>
<td>Sample Size = 35</td>
</tr>
<tr>
<td>Dependent Variable</td>
<td>Independent Variable</td>
</tr>
<tr>
<td>GMAT</td>
<td>0.097</td>
</tr>
<tr>
<td>Accuracy</td>
<td>WEXP</td>
</tr>
<tr>
<td>Decision</td>
<td>GMAT</td>
</tr>
<tr>
<td>Making</td>
<td>Quality</td>
</tr>
</tbody>
</table>

Attitude

We have labelled satisfaction with the CAI system as "AGGSAT" (aggregate of 14 items measuring the different attributes of satisfaction), and satisfaction with the CAI process as "PSAT". 86 percent of the students in the sample expressed above average satisfaction with the product and about 73 percent with the process. Table 8 presents the Pearson product-moment correlation analyses between attitude (BATT) and the two dimensions of satisfaction.
The correlation between attitude and satisfaction with the process (PSAT) is only 0.1859 (p < 0.135), and between attitude and satisfaction with the CAI system is -0.220 (p < 0.093). Contrary to normal expectations, they are insignificant; therefore, these hypotheses cannot be rejected. However, more of a serious concern is the negative correlation between attitude and satisfaction with the system, particularly when as large a proportion as 86 percent expressed above average satisfaction. Further analysis of the table provides additional insights. The extent of fulfillment of expectations (PEXPEC) is significantly correlated (r = 0.345, p < 0.018; r = 0.519, p < 0.00) with both the dimension of satisfaction and with perceived usefulness of computer supported education (CAREER) (r = 0.4016, p < 0.007). Fulfillment of expectations is, however, not significantly correlated (r = 0.2133, p < 0.102) with prior expectations (BAT3) which may be due to the presence of students who may not have had their expectations fulfilled. Those who had lower expectations probably responded much more favorably to satisfaction with the system (positive disconfirmation) than those who had higher expectations (negative disconfirmation). This seems to confirm the findings in IS research that implementation failures are caused due to unrealistic and unfulfilled expectations of the users (Lucas 1978; Ginzberg 1981).

### Behavior

Ninety-three percent of the students indicated that they wanted to recommend this course to others; 86 percent indicated that they would like to register for more computer supported courses. From Table 8 we see that while "BEHAVIOR" is highly correlated with satisfaction with the computer supported education process (r = 0.39, p = 0.009), the relationship with system satisfaction is very weak (r = 0.027, p = 0.437). It may be that, with growing computer literacy in the last decade, computers or decision support systems by themselves may not be perceived as novel enough to drastically influence behavior. Rather, satisfactory and innovative use of such computer support in the form of the CAI education process has contributed significantly to favorable behavior. Hence this hypothesis is partially rejected. It can also be observed that correlation between perceived usefulness to career (CAREER) and behavioral manifestations (BEHAVIOR) is highly significant (r = 0.5347, p < 0.000). This once again confirms the fact that, though their expectations have not been adequately fulfilled, students perceive computer supported education to be useful. Besides their fairly high satisfaction with the process as well as with the system, students exhibit favorable behavioral manifestations.

### 7. CONCLUSION AND FUTURE DIRECTIONS

This study examined the degree to which computers could add synergy to the traditional mode of instruction. While the role of computers in education has been subject to considerable study in education literature, the results have been inconclusive. Through a careful experimental design, controlling for the effects of extraneous influences,
the results of this study demonstrate favorable support for CAI. The superior performance of the CAI group appears to be emerging from the characteristics of the CAI process and product. The relationship between satisfaction and prior expectations has also been examined and the importance of nurturing realistic expectations highlighted.

In terms of contributions, this is one among the very few studies on CAI in business education which has evaluated performance dimensions, after controlling the confounding effects of extraneous variables, through a well designed experimental study. Constructs such as attitude, system experience, and satisfaction have been operationalized. A multi-dimensional performance measure has been used which provides greater validity. The study has highlighted the fact that, contrary to normal expectations, individual characteristics played a less significant role in contributing to performance.

There are limitations to this study as well. To minimize the effects of confounding, one of the sources, instructor effect, was controlled for. However, this has restricted the sample to a single class, resulting in a small sample size. This has considerably constrained our analysis and the results cannot be generalized without due discretion. The performance dimensions evaluated from a single study may be an inadequate reflection of students' ability and hence continuous evaluations at different points in time may be more appropriate. The predictor variables are not totally exhaustive. Other variables, such as psychological characteristics of the student, presentation style, clarity, etc., need to be examined. The appropriateness of CAI has been implicitly assumed to be invariant across different functional disciplines in business education. The impact of course characteristics on effectiveness of CAI has to be evaluated. Future research should address these issues.

8. ACKNOWLEDGEMENTS

We are very thankful to Dr. Cathy Niden for allowing us to use her class for the experiment, Dr. Bob Nachtman for initiating the development of the system and the IBM grant to Joseph M. Katz Graduate School of Business for providing financial assistance to conduct the study.

9. REFERENCES


