Are Programs of Assessment and Continuous Improvement Really Worth the Effort?

R. M. Richards, Ph.D.

ABSTRACT: *Assessment and programs of continuous improvement are becoming widespread in higher education. Several of the questions about assessment are concerned with its costs as compared to its benefits. This paper provides the results of an assessment effort for the introductory computer course in a large college of business and shows that assessment can be a valuable experience because of the information the process can provide and its ability to demonstrate, objectively, that learning is occurring.*

The members of a large introductory computing class were given "before" and "after" examinations to measure directly the increase in their level of knowledge about computers and their level of computing expertise over the semester term. They were also examined as to attitudes, specifically their personal preferences for risk, a possible metaphor for their increase in computing experience and expertise. The results of the two "snapshots" showed that much learning occurred during the course and that certain preferences for risk changed as the students gained information technology skills. The assessment effort provided input into the development of the course and identified areas in which the course could be improved.

KEYWORDS: *Assessment, innovative education, risk preferences, computer literacy.*

INTRODUCTION

Assessment and programs of continuous improvement are becoming more popular in the academic world as administrators and faculty alike seek new, more effective methods of imparting knowledge and skills in the classroom [1, 2, 3]. Many resources such as time, money, and equipment are being invested in the process. Is it worth the effort? Are assessment and programs of continuous improvement actually working? Should the resources devoted to assessment be reassigned to other activities? These are only a few of the questions that are being asked.

Assessment in the introductory computer course at this university has been going on for several years. The results from Spring semester, 1993, are presented here. This study was designed to be a baseline assessment to determine if/how much learning was occurring in the course, individually and as a class, and to identify any other changes that might occur as a result of the learning experience.

At the beginning of the semester, after the class roll had stabilized, but before any substantial changes in level of knowledge could take place, the students in this large section (140 students) were given an exercise to benchmark their level of knowledge and proficiency with computing. They were also asked to respond to a number of questions concerning their preferences for careers, their records as students, and their attitudes about scholarly life and life in a career. Other questions examined their preferences for risk and their prospects in life. They were not told that they would be examined again. At the end of the semester, the same battery of instruments was administered to those who were still enrolled in the course. (Attendance was mandatory, although not perfect.) The individual records of the students for the first and second sampling were then matched and subjected to statistical analysis (Wilcoxon Signed Ranks, a=0.05) to determine if there were any differences between the first and second sampling. The results proved to be very interesting. It was found that, 1) the class did increase significantly in both knowledge and proficiency, 2) that every student who completed the course learned something, and 3) that the class achieved a satisfactory level of knowledge on the average, the three goals of the course. Even more interesting are the results of the analyses of their personal preferences for careers and risk in role.

EFFECTS OF INFORMATION TECHNOLOGY ON PERCEPTION OF RISK IN ROLE

It is intuitive that a student who is learning in the class will increase in both knowledge and proficiency in computing. While such increases can be measured using examinations and other testing procedures over time, the effects of gaining that additional knowledge and proficiency on the student's outlook are somewhat more difficult to predict. The goal of computing is to provide more and better information for the decision maker. If such information is available, uncertainty decreases and the individual's perception of a given situation may change [4]. With less uncertainty, there is less variation of actual from expected, and perceived risk should decrease. If the learning experience is successful for the computing student, perception of risk should change.

One might ask, "Why would learning computing knowledge and skills change an individual's perception of risk?" The answer is in the information the computer is able to provide. An example is appropriate. Imagine trying to create a complex spreadsheet by hand. You would need a piece of paper, a calculator, and a pencil with a very large eraser. What if something changed or you made a mistake and you had to recalculate the whole thing over again? The process is time-consuming

Table 1: Results of Introductory Computer Course Assessment Project

Category	Mean #1	Mean #2	Significant? (∂=0.05)	Probability
Knowledge/Proficiency	31.74	74.35	Yes	$p<0.0001$
Computer Experience	3.19	4.21	Yes	$p<0.0001$
GPA	3.02	3.33	Yes	$p=0.0003$
Planning Horizon	3.34	3.11	No	$p=0.6043$
Time Span	3.45	3.32	No	$p=0.1014$
Organization Size	3.11	3.25	No	$p=0.0903$
Organization Type	3.89	3.99	No	$p=0.2660$
Success (Scholarly)	3.55	3.75	Yes	$p=0.0461$
Gaming Preferences	2.74	3.22	Yes	$p=0.0002$
Utility for Risk	3.04	2.89	Yes	$p=0.0241$
Prospects	3.63	3.70	No	$p=0.0857$

and prone to error. With an electronic spreadsheet on a computer, the task becomes infinitely simpler. The spreadsheet will recalculate automatically allowing "what if" types of questions to be asked and answered quickly and easily. The electronic spreadsheet on the computer is able to provide more information, more timely information, and more accurate information to the decision maker. This additional information results in different decisions than before, decisions about which there is less uncertainty through the use of information technology (IT).

How might decisions be expected to differ when IT is used in the process? Suppose a decision maker is faced with a venture in which he could lose millions of dollars and about which he knows virtually nothing. The decision maker is likely to avoid such a risk [5]. As the value at risk decreases, the decision maker would perceive less risk but would probably still avoid the venture. The problem is a lack of information about the venture. If, through computer skills, the decision maker is able to get more information about the venture, uncertainty decreases. The more information, the less uncertainty. At some point, the decision maker might have enough information to decide to undertake the venture, even at relatively high dollar amounts [6]. The amount of good, timely, pertinent information available affects the decision making process by reducing uncertainty and therefore reducing perceived risk in the decision making role.

Information systems theory tells us that the purpose of the computer is to provide information that is relevant, timely, accurate, and understandable. Risk theory tells us that as information increases, uncertainty and perceived risk decrease. Practically, this study bears that out. There were detectable shifts in risk preferences among the students in this computing class from the beginning of the course when they exhibited a low level of computer skills to the end of the course when they showed statistically significant improvement in their computer knowledge and skills.

RESULTS OF THE STUDY

The results discussed below are presented in Table 1. The column labeled MEAN #1 contains the data from the initial class assessment, and the column labeled MEAN #2 contains the data from the second class assessment. The third column shows whether the differences in the two assessments were significant at the 0.05 level and the fourth column shows the actual p-value returned by the statistical test.

Assessment of Knowledge/Proficiency

On the initial knowledge assessment exercise, the class mean was only 31.74% (out of 100%), indicating a relatively low average computer literacy level at the outset with a maximum of 72.5% and a minimum of 2.5%. As hoped, the second assessment exercise indicated that, at the end of the semester, the students had greatly improved. The mean had increased to 74.35% with a maximum of 97.50% and a minimum of 45.00%. The test for difference produced a p-value less than 0.0001, indicating a high level of statistical significance. The students did learn on the whole. Also, the analysis indicated that there were no cases showing less knowledge at the end than at the first. Every student learned something. Indeed, the high score increased twenty-five points so even the best prepared students learned a significant amount of knowledge, indicating a high, widespread learning experience was occurring among the members of the class.

It might be argued that this is what is expected to occur and that this result does not seem to be all that significant. For example, if one was tested for knowledge of Chinese geography before a Chinese geography course, a low score would be expected, with a higher score expected at the end of the course, assuming that the course was well taught and that the student attended class and studied. However, such results cannot be proved unless such tests are given. The major significance of this method might not be that we expect students to learn and they did. It might be that, at least in this case, the old accusation that "students are not learning anything in college" can be refuted with hard evidence. The ability to prove student learning and quantify it is a powerful asset in proving to our stakeholders that, indeed, we are accomplishing our goals of student learning in the classroom.

Computing Experience

The students reported that they had, as a group, done some computing but were primarily novices. At the end of the semester, they indicated that they now understood the computer and had experience with several packages. This is to be expected and again, the difference from sample 1 to sample 2 is significant ($p<0.0001$). This serves to confirm the notion that students are learning in the course as seen above in the section on knowledge/proficiency. We can prove that learning is occurring rather than resorting to intuition that learning is occurring at a quantifiable level.

Grade Point Average (GPA)

GPA was initially included as a demographic so that it could be compared with other variables. However, the results of the questions on GPA turned out to be somewhat unexpected. Theoretically, GPA would not change from the beginning of a semester to the end of the semester. Grades from the previous semester are known and no others have been earned. The student's GPA should be the same for both samplings. Such was not the case. The students reported a statistically significant increase in their GPA's from the beginning of the semester to the end of the semester ($p=0.0003$). Sixty-three of the students reported no change in GPA but 29 of them reported the change. This phenomenon needs to be studied further but could arise from a more optimistic view of expected GPA, including this semester, and a large number of students who were in their first semester at the university. It might also be indicative of a more optimistic estimate of their scholastic experience, supporting the data on success in their scholarly life, which also showed a significant difference in a positive direction. It should be noted that the analysis was conducted as paired tests so that the data presented do not include students who dropped the course. The members of the samples are identical, matched pairs for the two samples.

Planning Horizon, Time Span of Discretion (TSD), Organization Size (OS), and Organizational Type (OT)

Several classical indicators of risk preference were used in this study. Measures of the amount of money a person is comfortable in dealing with (planning horizon) and the preferred time between reviews by a supervisor (time span of discretion [7]) showed no significant difference between the samples. There were also no significant differences in preferred organizational size or organization type.

Gaming Preferences

The subjects were given several questions concerning probability and return to get an estimate of their preferences for risk. One "game" involved a fifty/fifty chance of winning or losing increasing sums of money. With their own money theoretically at risk, they preferred a bet of just under $100.00 in the first sampling and a bet just over $100.00 in the second. Those preferring a larger amount of money outnumbered those moving lower by two to one (negative and positive ranks). The difference was significant (p=0.0002) and movement was toward larger amounts of money. Due to the relatively low p-value, it is not expected that these data are erroneous, but future replicates can provide more information.

Utility for Risk

The students were also asked to choose among a set of alternatives involving equal expected values over time but with varying probabilities and values. These questions gave an estimate of risk taking/risk aversion tendencies of the subjects and resulted in a utility index for each subject. Between the two samples, the mean for the utility index decreased from 3.037 to 2.891, a movement toward the longer-odds alternatives, and was statistically significant (p=0.0241). When viewed with the results of the "even bet" question above, the respondents seemed to move toward more risk.

Perceived Prospects for the Future

One of the theories concerning risk in role is prospect theory which involves framing risk assessment measures in positive or negative terms and considering the overall attitude of the subject at the time measurement is undertaken (8, 9, 10). An index was calculated for each respondent and a slight, but not statistically significant, increase in perceived prospects was noted (p=0.0857). According to the prospect measure, there is little reason to believe that their individual prospects changed over the course of the semester, lending validation to the risk measures which do show significant changes over the course of the semester.

Note: A copy of the instruments used and complete results of this study are available from the author.

CONCLUSIONS AND INTERPRETATION

The data on changes in the level of knowledge and proficiency indicated that learning is going on in the course and at a high rate. Not only did the class achieve an acceptable level of knowledge and proficiency but every student in the class showed positive movement. Their experience with computers increased and they gained a more optimistic estimation of their grade point averages. They showed more success in their scholarly life, indicating a positive experience over the semester. It can be said that the data supported the contention that the goals of the course are being attained, at least for this section.

This study also indicated that the risk preferences for the class changed significantly toward more risk, indicating that there might be an effect between level of computer literacy and perception of risk. An examination of this relationship might be the most productive avenue for further research based on this assessment effort.

Overall, this assessment effort has been a positive and informative experience. It has helped the student by increasing the quality of the educational experience and it has helped the professor do a better job in the classroom. It has started a new strategy for analyzing the effectiveness of the classroom experience and should serve as a model for other efforts. Most importantly, it is a response to the challenges facing us for the present and the future by lending credibility to the job being done and demonstrating responsibility on the part of the institution and those persons entrusted with the task of guaranteeing quality and currency in the higher education system. Is it worth the effort? In this case, it has been well worth the effort as the next section demonstrates.

HOW THE ASSESSMENT EFFORT HAS AFFECTED THE COURSE

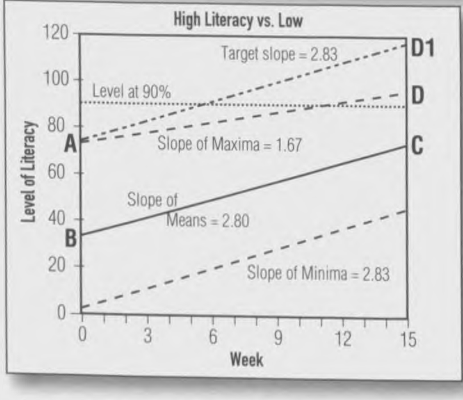
The results presented above represent the type of information available from assessment. Other sources of information, such as written comments of students and conversations with students, are also helpful. As the dynamic process of assessment goes on, information gathered and analyzed leads to other data that can add to the effort. The most important factors can be retained and additional factors can be sought. Assessment redesigns itself with every iteration of the process. It also changes the course in many ways.

Perhaps the most important change in this course is that now we can plan for development based on hard evidence rather than guesswork. An example is the class attendance policy. Early on, it was evident that there was a direct correlation between attendance and the increase in knowledge. Simply stated, if the student is not in the classroom during lecture, that student is not learning the material. One of the first results of the assessment effort was to instate a required attendance policy. A student who accumulates over six class hours of nonattendance is dropped with a failing grade, unless they withdraw voluntarily. Also, all students with perfect attendance at the end of the semester, no excuses accepted, get five points added to their final average. This positive/negative approach has lead to much better attendance with almost ninety percent of the students never missing a day. They are in class and they learn. Having them in class has made a real difference in their learning experience.

In the early stages of this study, it also became apparent that an uneven learning experience was occurring between the students who showed a relatively high level of literacy coming into the class and those who showed a relatively low level. Figure 1 shows the maximum scores went up from 72.5% to 97.5% with a slope of 1.67 (line AD) while the minimum scores (2.5%-45%) and the class mean (32%-74%, line BC) exhibited a slope of approximately 2.8 demonstrating a higher level of learning among students in the low literacy classification. This is an indication of why, under the leveling approach, the better students are frequently less challenged than others, lending to a less effective experience for the better student.

Assessment is concerned with the amount of learning that occurs, not with leveling. For those students showing a higher level of literacy, some way must be found to enable them to learn at the AD1 rate instead of the AD rate, maintaining the rapid learning rate of the less-literate student. This would greatly increase the amount of learning going on in the course and keep the best students challenged and interested. Assessment makes identification of these students possible and leads to techniques designed to increase the challenge to those students such as working with other students in the lab and tutoring students who want or need help. Trying to teach the materi-

Figure 1: Differential Learning Levels



al to other students sharpens their own knowledge of the course content and strengthens their communication skills. Assignments that require a higher level of skill, beyond the literacy parameters of the course, can also be used to keep these students interested. Recent experiences have shown that the best students do not have to feel bored and unchallenged. In fact, experience beyond the literacy parameters of the course are generally seen as being highly valuable.

Another item of information gained from this study regarded presentation of the material. We had been using overhead projectors and transparencies with live demonstrations of the software packages. Now we use presentation graphics for lectures, standardizing the presentation of material across sections of the course and reducing the "war stories." The instructor must adhere to the presentations, providing a more concise but complete learning experience. The students like them much better and they seem to be learning the material much more easily. Assessment also indicated that the level of the previous textbook was somewhat inappropriate. Gathering data about how they study led to a textbook that more closely matches their needs. The approach is straightforward and at a level the student can understand. This seems to greatly enhance the learning experience and enhances the balance of topics.

Assessment has also provided information on course content. Recently we have found that most of the students are using or acquiring computers that have Microsoft Windows. For Fall semester, 1994, Microsoft Windows was added to the course. Plans include going to integrated software such as Microsoft Office in the future.

By comparing the results of all the assessment exercises over time, the effects of various techniques to enhance learning can be compared. For example, it seems that sections containing 110 students do better than sections containing 220 students. The reason why is not clear but is being studied as a part of current assessment efforts. We have also learned that more and more students have their own computers and have changed the assumption that most of them will use the fine lab facilities in the COBA to the assumption that they will use their own machine. This has greatly affected our planning.

Another factor that can be attributed directly to the assessment effort is the success instructors are having in teaching the course. Student evaluations of the course have shown a steady increase since the assessment effort began. Ratings for both the course and for those teaching it have gone up dramatically. Professors do a better job because of assessment. That, by itself, is worth all the resources invested because it leads to a better educational experience for the student, the purpose of assessment and programs of continuous improvement.

REFERENCES

[1] The American Council on Education. Highlights from campus trends, 1992.

[2] Erwin, T. Dary. Assessing student learning and development. Josey-Bass, 1991.

[3] Gardiner, Lion F. Planning for Assessment, a guide for colleges and universities. New Jersey department of Higher Education, 1989.

[4] Greene, Mark R. Risk and Insurance. Cincinnati, OH: Southwestern Publishing Co., 1973.

[5] Friedman, Milton J., and L. J. Savage. "The utility analysis of choices involving risk." *Journal of Political Economy*, 1948, LVI(August).

[6] Swalm, Ralph O. "Utility Theory: insights into risk-taking." *Harvard Business Review*, XLVII(6).

[7] Jaques, Elliot. Time Span Handbook. London: Heinemann, 1968.

[8] Bazerman, Max J. "The relevance of Kahneman and Tversky's concept of framing to organizational behavior." *Journal of Management*, 10(3) 333-343.

[9] Kahneman, Daniel and Amos Tversky. "The psychology of preferences." *Scientific American*, 246(1) 160-173.

[10] Tversky, Amos and Daniel Kahneman. "The framing of decisions and the psychology of choice." *Science*, 211(4481) 453-458.

R. M. Richards, Ph.D.
Department of Business Computer Information Systems
University of North Texas
Denton, Texas 76203-3677
(817) 565-3167
RICHARDM@COBAF.UNT.EDU

Roy Martin Richards, Jr., Ph.D., is Assistant Dean for Assessment in the College of Business Administration and Associate Professor of Business Computer Information Systems at the University of North Texas in Denton, Texas. He earned his doctorate from the University of Georgia in 1978, a Master of Business Information Systems in 1972, and a Bachelor of Arts in 1970 from Georgia State University. He has published in various IS, education, and risk management Journals and is currently active in research into the effects of information technology on decision making and perception of risk in executive business management roles.

Information Systems & Computing
Academic Professionals

EDSIG
Serving Information Systems Educators

STATEMENT OF PEER REVIEW INTEGRITY

All papers published in the Journal of Information Systems Education have undergone rigorous peer review. This includes an initial editor screening and double-blind refereeing by three or more expert referees.

ISSN 1055-3096