Improving Computer Training Effectiveness for Decision Technologies: Behavior Modeling and Retention Enhancement*

Mun Y. Yi  
*Moore School of Business, University of South Carolina, Columbia, SC 29208, email: myi@moore.sc.edu

Fred D. Davis  
Sam M. Walton College of Business, University of Arkansas, Fayetteville, AR 72701, email: fdavis@walton.uark.edu

ABSTRACT

Managers and analysts increasingly need to master the hands-on use of computer-based decision technologies including spreadsheet models. Effective training can prevent the lack of skill from impeding potential effectiveness gains from decision technologies. Among the wide variety of software training approaches in use today, recent research indicates that techniques based on behavior modeling, which consists of computer skill demonstration and hands-on practice, are among the most effective for achieving positive training outcomes. The present research examines whether the established behavior-modeling approach to software training can be improved by adding a retention enhancement intervention as a substitute for, or complement to, hands-on practice. One hundred and eleven trainees were randomly assigned to one of three versions of a training program for spreadsheets: retention enhancement only, practice only, and retention enhancement plus practice. Results obtained while controlling for total training time indicate that a combination of retention enhancement and practice led to significantly better cognitive learning than practice alone. The initial difference in cognitive achievement was still evident one week after training. Implications for future computer training research and practice are discussed.


*The authors gratefully acknowledge the helpful comments and suggestions provided by the two anonymous referees and an associate editor.
INTRODUCTION

Organizations will not realize desired returns on their investments in information technologies designed to improve decision making unless users are able to use them. Multimillion dollar computer-based systems often go unused or underused largely because people do not have adequate skills to use them effectively (McCarroll, 1991; Ganzel, 1998). Empirical findings have shown a significant correlation between computer-related ability and productive use of computer resources (Nelson & Cheney, 1987; Lee, Kim, & Lee, 1995). To use decision technologies effectively, managers and analysts must increasingly master the hands-on use of interfaces to software such as spreadsheets, query languages, and modeling tools. As Compeau, Olfman, Sein, and Webster (1995) pointed out in their introduction to a recent special issue of the Communications of the ACM on end-user training and learning, "Information systems practitioners and researchers widely acknowledge that providing appropriate end-user training is critical to successfully implementing systems, and key to promoting productive use of the technology" (p. 24). Given that a third of all formal training in the U.S. is devoted to teaching employees about computers (Industry Report, 1999), it is appropriate that information systems (IS) researchers have identified computer training as a critical factor responsible for ensuring the success of computer applications (Cheney, Mann, & Amoroso, 1986; McLean, Kappelman, & Thompson, 1993; Nelson & Cheney, 1987). Although a wide variety of training methods are being used to teach computer skills (Harp, Satzinger, & Taylor, 1997; Industry Report, 1999), there remains insufficient understanding of how training inputs are systematically related to outcomes (Nelson, Whitener, & Philcox, 1995).

In their broad framework of research on end-user training, Compeau et al. (1995) delineated three main phases of the overall training process: (1) the initiation phase, which includes needs assessments and the design and development of training materials; (2) the formal training and learning phase, which is concerned with the training methods used (hands-on use, behavior modeling, exploratory learning, etc.), mode of training delivery (face-to-face, video, computer-based, etc.), and choice of training facilitator (outside consultants, in-house trainers, self-training by trainees); and (3) the post-training phase, which examines the long-term effects of training in terms of its influence on workplace behaviors. Within this framework, the present research focuses specifically on (2), the formal training and learning phase, and therefore complements research addressing the initiation phase (e.g., Nelson et al., 1995; Carroll & Rosson, 1995) and the post-training phase (e.g., Kay & Thomas, 1995). Understanding the relative effectiveness of alternative computer training techniques represents a significant value to managers, who must decide which training method to implement to realize the effectiveness gains desired from decision technologies.

To date, convergent findings across several studies suggest that behavior modeling, which consists of observation of computer skill demonstration and subsequent hands-on practice, is among the most effective computer training methods (Compeau & Higgins, 1995; Gist, Rosen, & Schwoerer, 1988; Gist, Schwoerer, & Rosen, 1989; Simon, Grover, Teng, & Whitcomb, 1996; Simon & Werner, 1996). In two field experiments, Gist and her colleagues (Gist et al., 1988, 1989) consistently found that performance of trainees in the behavior-modeling condition was
higher than that of trainees with computer-aided instruction. Compeau and Higgins (1995) compared behavior modeling with a traditional lecture-based program to confirm the effectiveness of behavior modeling in spreadsheet training. Simon and his colleagues (Simon et al., 1996; Simon & Werner, 1996) conducted similar research by comparing behavior modeling with both a lecture-based program and self-study using an inductive-style manual (exploration learning) and reported that behavior modeling was more effective than the other two training methods. Given the non-significant differences frequently found in the IS training literature (e.g., Olfman & Mandviwalla, 1994; Santhanam & Sein, 1994), the consistent findings on the effectiveness of behavior modeling in computer training is notable, warranting continued research on this topic.

Studies conducted outside of the computer training domain have found that retention enhancement, which consists of transforming key elements of modeled activities into a pattern of verbal symbols (symbolic coding) and mentally practicing the modeled activities (cognitive rehearsal), significantly enhances the effectiveness of observational learning for cognitively complex motor skills (Bandura & Jeffery, 1973; Bandura, Jeffery, & Bachicha, 1974; Gerst, 1971; Jeffery, 1976), as well as for interpersonal skills (Decker, 1980, 1982; Mann & Decker, 1984). The findings suggest that behavior modeling, which provides trainees with a chance to observe effective execution of computer operations through modeling and then assimilate the demonstrated operations through hands-on practice, can be further improved when it includes an opportunity to symbolically and cognitively process the key aspects of modeled operations. Presently it is unknown how much computer training outcomes can be improved by adding a retention enhancement intervention to behavior modeling. In this study, we compare the relative effectiveness of retention enhancement, hands-on practice, and a combination of both for training on spreadsheet software, a highly representative interface to decision technologies.

CONCEPTUAL BACKGROUND AND RESEARCH HYPOTHESES

Behavior-modeling training is based on social cognitive theory (Bandura, 1969, 1977, 1986), which views observational learning as a fundamental means by which humans learn new behaviors. People can acquire cognitive skills and new patterns of behavior by observing the actual performances of others and the associated consequences. People form rules of behavior by observing others, and on future occasions this coded information guides their actions. Learning is thus defined as “largely an information processing activity in which information about the structure of behavior and about environmental events is transformed into symbolic representations that serve as guides for action” (Bandura, 1986, p. 51). According to this view, observational learning is governed by four component processes: attention (observing behavioral skills), retention (transforming the observed skills into symbolic codes), production (practicing the skills physically), and motivation (getting motivated to continue using them). Attention and retention processes regulate the acquisition of observed behaviors whereas production and motivation processes govern actual enactment of behaviors. Social cognitive theory, which views all four component processes as necessary in effectively developing new
behavioral skills, places great emphasis on symbolic processing of information in the acquisition phase and physical practice of response patterns in the enactment phase (Bandura & Jeffery, 1973; Jeffery, 1976). In short, symbolic coding and cognitive rehearsal activities are effective means to organize observed component responses into appropriate temporal and spatial relationships, refine action plans, and strengthen memory traces, whereas physical practice is an effective means to smooth execution of actions, routinize response patterns, and provide performance feedback. Therefore, "the highest level of observational learning is achieved by first organizing and rehearsing the modeled behavior symbolically and then enacting it overtly" (Bandura, 1977, p. 27).

Several empirical studies conducted outside of computer training show that symbolic processing of information can make a significant contribution to learning over and above the effect of physical practice. Bandura and Jeffery (1973) found that subjects who coded the model's actions verbally or numerically and immediately rehearsed the memory codes achieved greater learning than those who physically rehearsed the modeled actions without performing symbolic coding at input. Jeffery (1976) conducted a similar experiment in which subjects observed a filmed model construct three-dimensional objects using wooden rods and joints. Subjects who were engaged in both cognitive rehearsal and physical practice were able to reproduce demonstrated skills significantly more accurately, both immediately and after a one-week delay, than were those who only physically practiced. In the context of behavior modeling, Decker (1980) found that cognitive rehearsal conducted before physical practice facilitated reproduction of modeled supervisory skills, and symbolic coding minimized reproduction decay measured one week after training. In a separate study, Decker (1982) showed that performing symbolic coding and cognitive rehearsal before skill practice produced significantly better generalization of observational learning to a novel context. These empirical results support the social cognitive theory notion that modeled behaviors are best acquired when they are first symbolically processed and then physically reproduced.

In the computer training domain, however, behavior-modeling studies have required trainees to observe the demonstration of desired computer operations and then physically perform them, but have not required trainees to symbolically encode or mentally rehearse the observed action sequences (Compeau & Higgins, 1995; Gist et al., 1988, 1989; Simon et al., 1996; Simon & Werner, 1996). Although computer skills are more cognitively complex than the skills required by the aforementioned observational learning studies, the research on observational learning and managerial skill training suggests that computer training can be further improved by adding retention enhancement processes of symbolic coding and cognitive rehearsal to the current form of behavior modeling. Most computer operations require a serial execution of component actions to perform a task. Because new sequences of actions are continuously presented by video or live demonstration, novice users can easily become cognitively overloaded as they must direct their attention to the next set of actions without sufficient time to absorb the presented material (Singer, 1980). In this situation, the symbolic coding and cognitive rehearsal processes may be a powerful intervention that can be used in conjunction with observation. By giving trainees a chance to summarize the presented material and to mentally practice the summarized actions, trainees should be able to more
deeply and meaningfully process the information required for effective use of the computer application. The findings from research on observational learning and on behavior modeling consistently indicate that the process of observational learning is more effective when the training environment promotes more meaningful symbolic transformation of the action patterns and deeper processing of information (Bandura & Jeffery, 1973; Bandura et al., 1974; Gerst, 1971; Hogan, Hakel, & Decker, 1986; Jeffery, 1976).

Observational learning of computer operations should also be more effective when symbolic coding and cognitive rehearsal of skills is followed by overt enactment of modeled behavior. Physical practice improves actual compilation of skills by speeding up the process of converting symbolic response guidance to spatial and temporal movements (Bandura, 1977). Moreover, novice computer users cannot observe and accurately remember all the details of response patterns. They must validate and correct their incomplete or inaccurate mental models on the basis of informative feedback from their performance. Physical practice allows novice users to detect initial misconceptions, identify missing component skills, organize response patterns in a more efficient way, and develop more complete and accurate mental models through an iterative process of feedback and self-correction. Thus, physical hands-on practice should also improve computer learning outcomes.

Kraiger, Ford, and Salas (1993) recommended that training effectiveness be measured by examining three dimensions of learning: cognitive, skill based, and affective. Cognitive outcomes include verbal knowledge, knowledge organization, and cognitive strategies. Skill-based outcomes include skill compilation and automaticity. Finally, affective outcomes include self-efficacy (i.e., self-perception of performance capability), goal, and attitude toward a targeted object. These dimensions are interrelated but not identical. Thus, learning may be evident from changes in any of these dimensions. As presented in Table 1, a review of the computer training literature shows that previous studies often evaluated (a) cognitive outcomes by measuring comprehension of declarative knowledge, (b) skill-based outcomes by measuring accuracy of procedural skill compilation (task performance), and (c) affective outcomes by measuring perceptions of the system's ease of use and usefulness. Similar to this three-dimensional view of learning, research on attitude has long conceptualized the attitude construct as tripartite: cognitive (knowing), conative (acting), and affective (feeling) (Allport, 1935; Katz & Stotland, 1959; Kretch, Crutchfield, & Ballachey, 1962; McGuire, 1969). Following this conceptualization, Galletta, Ahuja, Hartman, Teo, and Peace (1995) divided training outcomes into three categories of performance (consisting of both cognitive and skill-based performance outcomes), behavior, and attitude. This classification includes not only learning but also behavior as a part of training outcomes. In addition to learning and behavior, Kirkpatrick (1987) included reaction and organizational results as components of training evaluation criteria. Among these training outcomes, however, learning during training plays the central role in defining training evaluation and effectiveness (Kraiger et al., 1993) and serves as an important precursor to the behaviors on the job and desired organizational outcomes (Baldwin & Ford, 1988; Goldstein, 1991; Noe, 1986; Tannenbaum, Mathieu, Salas, & Cannon-Bowers, 1991). Thus, following the conceptual framework proposed by
### Table 1: Learning outcomes and findings of selected prior research on computer training.

<table>
<thead>
<tr>
<th>Study</th>
<th>Training Intervention</th>
<th>Learning Outcomes</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gist et al. (1988)</td>
<td>BM vs. Computer-aided instruction</td>
<td>Skill: Task performance</td>
<td>BM yielded higher task performance scores for both younger and older trainees.</td>
</tr>
<tr>
<td>Davis &amp; Bostrom (1993)</td>
<td>Exploration-based vs. Instruction-based training; GUI vs. Command Interface</td>
<td>Skill: Task performance Affective: Perceived EOU</td>
<td>Trainees in GUI performed better than their command-based counterparts. No difference in EOU. No significant interaction effects between interface and training methods.</td>
</tr>
<tr>
<td>Compeau &amp; Higgins (1995)</td>
<td>BM vs. Instruction-based training</td>
<td>Skill: Task performance Affective: CSE</td>
<td>Subjects in the BM condition developed higher CSE and performed better than those in the instruction-based condition for a spreadsheet program, but not for a word-processing program.</td>
</tr>
<tr>
<td>Galletta et al. (1995)</td>
<td>Positive vs. Negative word-of-mouth</td>
<td>Cognitive: Comprehension Skill: Task performance Affective: Attitude about the software</td>
<td>Negative word-of-mouth group showed lower attitude scores and comprehension scores. No significant differences were found for task performance.</td>
</tr>
<tr>
<td>Simon et al. (1996)</td>
<td>Instruction, Exploration, and BM</td>
<td>Cognitive: Comprehension Skill: Task performance Affective: End-user satisfaction</td>
<td>BM outperformed the other two methods in all the learning outcome measures.</td>
</tr>
<tr>
<td>Lim, Ward, &amp; Benbasat (1997)</td>
<td>Self-discovery vs. Co-discovery</td>
<td>Cognitive: Inference test Skill: Task performance</td>
<td>Co-discovery had a significant effect on inference potential, which had a significant effect on task performance.</td>
</tr>
</tbody>
</table>
Table 1: (continued) Learning outcomes and findings of selected prior research on computer training.

<table>
<thead>
<tr>
<th>Study</th>
<th>Training Intervention</th>
<th>Learning Outcomes</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Venkatesh (1999)</td>
<td>Traditional vs. Game-based training</td>
<td>Affective: Perceived EOU, Perceived usefulness</td>
<td>Users who were in the game-based training had higher levels of EOU. No significant differences in usefulness across training interventions.</td>
</tr>
</tbody>
</table>

Notes:
Cognitive: Cognitive outcomes; Skill: Skill-based outcomes; Affective: Affective outcomes; BM = Behavior modeling; CSE = Computer self-efficacy; EOU = Ease of use; GUI = Graphical user interface.
Only those studies that manipulated training interventions and compared their relative effects on computer learning outcomes are included in the table.

Kraiger et al., the present study examines the relative effects of retention enhancement, hands-on practice, and a combination of both on the cognitive, skill-based, and affective learning outcomes in the context of computer skill training.

Based on the preceding discussion, we expect that the combination of retention enhancement and hands-on practice will be more beneficial than either retention enhancement alone or hands-on practice alone in acquiring computer skills. More specifically, it is hypothesized that combining retention enhancement with hands-on practice will produce significantly better cognitive outcomes than practice alone due to the cognitive advantages associated with symbolic processing of information, and significantly better skill-based outcomes than retention enhancement alone due to the skill acquisition advantages associated with physical hands-on practice. Learning outcomes are not only multidimensional, they are also interrelated (Kraiger et al., 1993). Subjects with more complete cognitive understanding and fluid compilation skills of the target system are expected to perceive the system as easier to use, and subjects with higher ease-of-use perceptions are expected to perceive the system as more useful, as suggested by the findings from the technology acceptance literature (Adams, Nelson, & Todd, 1992; Davis, Bagozzi, & Warshaw, 1989; Mathieson, 1991; Taylor & Todd, 1995; Venkatesh, 2000; Venkatesh & Davis, 2000). In sum, we hypothesize the following:

**H1:** Combining retention enhancement with hands-on practice will produce better cognitive learning outcomes than practice alone, controlling for total training time.

**H2:** Combining retention enhancement with hands-on practice will produce better skill-based learning outcomes than retention enhancement alone, controlling for total training time.
H3: Subjects with higher cognitive and skill-based outcomes will perceive the system as easier to use.

H4: Subjects with higher ease-of-use perceptions will perceive the system as more useful.

RESEARCH METHODOLOGY

Subjects

A three-hour training program on a popular type of software package commonly used for building decision models (i.e., Microsoft Excel for Windows) was set up at a large university in the eastern United States. Subjects were recruited on a voluntary basis from an introductory computer course. As a motivational incentive, students were promised and later received confidential feedback regarding their performance compared to their peers. The skills covered by the training were also required to complete a term project assigned to each individual. The lab experiment setting was preferred over a field experiment setting in order to manipulate the training components of retention enhancement and physical practice under controlled and unconfounded conditions.

A total of 111 students (43% female and 57% male) completed the experimental procedure. Participants' ages ranged from 18 to 35, with a mean of 21.8. In response to a question regarding their weekly usage of spreadsheets, most subjects indicated that they were not regular users of the spreadsheet program as follows: 54 (49%) never used, 43 (39%) used less than one hour, 10 (9%) used one to three hours, and 4 (4%) used more than four hours. Eighty-five participants (77%) had work experience and many (59 participants) had worked for companies more than a year.

Design

The experimental design utilized a $3 \times 3$ Latin square to isolate the effects of training time and trainer effects. There were three training sessions (9:00 a.m., 12:30 p.m., and 4:00 p.m.), and in each session there were three training workshops, each workshop run by a trainer. Thus, trainers, computer labs, and times of day were orthogonally counterbalanced across training conditions to control for any potential confounding effects. Participants were randomly assigned to one of the nine training workshops offered on the same day, and were not told that different training conditions were being tested.

Two professional trainers were hired to run the workshops with one of the authors. The trainers performed the role of facilitator who told trainees what to do next using scripts developed and pretested in a pilot study. The trainers also provided individual help when a trainee asked for it. There were no significant differences across trainers in any of the learning outcomes. Before the main experiment, the hired professional trainers had visited the training site several times to get acclimated to the training facilities and materials. The professional trainers were blind to the hypotheses of the study.
Procedure
In a computer lab, following the prepared scripts, trainers first introduced themselves, distributed and collected pretest questionnaires, and then implemented the assigned training conditions using prepared scripts and stopwatches. After training procedures were completed, each trainee filled out a posttest questionnaire, took the first comprehension test (5 minutes) and first task performance test (20 minutes), continued hands-on practice for 20 minutes, took the second task performance test (20 minutes), and was thanked and dismissed. The second comprehension test (5 minutes) was administered, without warning, a week later in class. The first comprehension and task performance tests were designed to assess trainees' immediate learning, and the second tests were intended to assess delayed learning.

Training Material
The videotape used in the training was a commercial product, provided by a third-party vendor that specializes in computer training. The tape consisted of four segments, each of which focused on one specific topic (i.e., basic formatting (8 minutes), formulas (8 minutes), functions (12 minutes), and expanded formulas (12 minutes)). In each section, the same middle-aged male model illustrated various features of the software showing specific steps of operations. At the end of each section, the model summarized key learning points of the demonstration. Trainees had access to the computer during the workshop except when the video was in play. A spreadsheet exercise file was installed on the computer. The file contained the same rows and columns of initial numbers as presented at the beginning of the video. Trainees used the exercise file to start their hands-on practice.

Training Conditions
To examine the effects of retention enhancement, hands-on practice, and a combination of both, three different training conditions were used for this study. The first condition (practice only) consisted of modeling and hands-on practice. No retention enhancement activities were performed in this condition. This condition represents the current practice of behavior modeling in the IS training literature. The second condition (retention enhancement only) consisted of modeling, symbolic coding, and cognitive rehearsal. Trainees were not allowed to practice the demonstrated skills on the computer until they finished taking the initial set of tests. The third condition (retention enhancement plus practice) consisted of modeling, symbolic coding, cognitive rehearsal, and hands-on practice. The three conditions were identical except that the first condition included 26 minutes of hands-on practice, the second condition included 26 minutes of retention enhancement, and that the third condition included 16 minutes of practice and 10 minutes of retention enhancement. Thus, the total training time across the training conditions was held constant, eliminating any possible interaction between additional time and training condition. The design represents a conservative test of the effects of the combined condition, because it contains 10 minutes less practice than the practice-only condition and 16 minutes less retention enhancement than
the retention-enhancement-only condition. Figure 1 shows the progress of the experiment procedures and the elements of each training condition.

**Practice-only condition**

Immediately after each segment of the video, trainees in this condition practiced the demonstrated skills individually for five minutes after each of the first two segments of the video, and for eight minutes after each of the last two segments. It was decided from the responses of a pilot study that the practice time periods were more than enough for most subjects to complete the practice for the presented skills. Trainees in this condition were not asked to perform any symbolic coding or cognitive rehearsal activities.

**Retention-enhancement-only condition**

Instead of conducting hands-on practice, trainees in this condition performed symbolic coding and cognitive rehearsal for five minutes after each of the first two segments of the video, and for eight minutes after each of the last two segments. Baldwin (1992) provided a copy of learning points and then asked trainees to actively process them by reading them over, writing them down, and thinking about them. In a similar way, trainees in this condition received a sheet containing a trainer-prepared summary, read the key points on the summary sheet trying to recall how the model on the video demonstrated each point, and then, for each recalled demonstration, summarized the key points of the demonstration in their own words on the blank paper labeled with appropriate section headings. The trainer-provided summary sheet was created by the authors using the video and accompanying manual. Trainees were told that their own summary may or may not be the same as the trainer-provided summary. After creating their own summary, trainees engaged in cognitive rehearsal, mentally picturing themselves performing the computer operations for the points listed on the summary sheet they created. They were asked to repeat the mental rehearsal, as many times as possible, and record the number of times they were able to perform the rehearsal activity.

**Retention-enhancement-plus-practice condition**

Trainees in this condition performed both retention enhancement and hands-on practice at the end of each video segment. Trainees performed symbolic coding and cognitive rehearsal in the same way as trainees in the retention-enhancement-only condition did, but only for two minutes after each of the first two segments of the video, and for three minutes after each of the last two segments. Following the retention enhancement activities, trainees conducted hands-on practice using the computer in the same way as trainees in the practice-only condition did, but only for three minutes after each of the first two segments of the video, and for five minutes after each of the last two segments. Thus, the total time for this condition was the same as the other conditions.

**Measures**

Consistent with the criteria suggested by Kraiger et al. (1993), computer learning outcomes were measured along the cognitive, skill-based, and affective
Figure 1: Experimental procedure and training elements.

- **Introduction**
  - Pretest Questionnaire

- **Practice Only**
  - Video Segment 1 (8 min.)
    - Hands-on Practice (5 min.)

- **Retention Enhancement Only**
  - Video Segment 1 (8 min.)
    - Coding & Rehearsal (5 min.)

- **Behavior Modeling Plus Retention Enhancement**
  - Video Segment 1 (8 min.)
    - Coding & Rehearsal (2 min.)
    - Hands-on Practice (3 min.)

- **Video Segment 2 (8 min.)**
  - Coding & Rehearsal (5 min.)

- **Video Segment 2 (8 min.)**
  - Coding & Rehearsal (5 min.)

- **Video Segment 3 (12 min.)**
  - Hands-on Practice (8 min.)

- **Video Segment 3 (12 min.)**
  - Coding & Rehearsal (8 min.)

- **Video Segment 4 (12 min.)**
  - Hands-on Practice (8 min.)

- **Video Segment 4 (12 min.)**
  - Coding & Rehearsal (8 min.)

- **Posttest Questionnaire**
  - Immediate Comprehension Test
  - Immediate Task Performance Test
  - Hands-on Practice (20 min.)
  - Delayed Task Performance Test
  - Delayed Comprehension Test
dimensions. In addition, manipulation checks for the retention enhancement processes were measured.

**Cognitive learning outcomes**

The cognitive learning measure consisted of 10 multiple-choice test questions designed to assess trainee comprehension of the concepts and features needed to use the software program appropriately. The items were developed from the video and manual. The items included questions about copying a cell, using a formula, adjusting the size of a column, and performing a calculation. The score was the total number of correct answers. Thus, possible scores ranged from 0 to 10. To make an accurate comparison, the same measure was used to assess immediate learning and delayed learning. The test for immediate learning was administered after the last segment of the video and associated treatment in each treatment condition. The measure of delayed learning was captured by readministering the learning test one week later during regular class time. Trainees were not informed beforehand that they would be tested for retention at a later time, in order to deter intentional efforts to find answers during the intervening period.

**Skill-based learning outcomes**

A hands-on task performance measure that contained 10 computer tasks was used to capture trainee skill compilation on the target computer program. Each task required several steps of computer operations. Examples include entering a formula in multiple cells, using functions to calculate total and average amounts, calculating percent change, and changing the formats of numbers. Each trainee saved the test result in a designated directory upon the completion of the test. Each task was scored with 2 points for totally correct answers, 1 point for partially correct answers, and 0 for incorrect or missing answers. Thus, possible scores ranged from 0 to 20. For a fair comparison, the same set of tasks involving different numbers was used to assess both immediate and delayed learning of skill compilation. The test for immediate learning was administered right after the initial comprehension test, and the test for delayed learning was administered after an extended hands-on practice time of 20 minutes. The grading of the answers was handled by the spreadsheet program module developed through several stages of programming and accuracy verification.

**Affective learning outcomes**

To capture trainee’s affective outcomes, an instrument developed by Davis (1989) was used to assess the trainee’s perceptions of the system’s usefulness and ease of use. The instrument consisted of four items for the usefulness construct and four items for the ease-of-use construct, all items using an 11-point Likert-type scale where 0 = completely disagree, 5 = neither agree nor disagree, and 10 = completely agree. Some sample items from the instrument are “I find Excel easy to use” and “I find Excel would be useful in my degree program.” The internal consistency reliability (Cronbach’s alpha) was .90 for the ease-of-use construct and .95 for the usefulness construct.
Manipulation checks

The manipulation of the symbolic coding activity was checked by counting the number of trainees who actually performed any kinds of summary activities during their training workshops. More specifically, all the papers either distributed by the trainers for symbolic coding or self-supplied by trainees for note taking were collected and examined to see how many trainees actually created some sort of summary in different training conditions. The manipulation check for symbolic coding showed that all the trainees \( n = 75 \) in the conditions that included retention enhancement performed symbolic coding (created their own summary sheet of learning points), whereas no trainees in the practice-only condition \( n = 36 \) created a summary. The summaries made by the trainees were similar to the provided summaries, covering almost the same contents in the same order, but tended to be more concise than the trainer-provided summaries.

The manipulation of the cognitive rehearsal activity was checked by examining the number of times trainees performed the rehearsal activity. After each symbolic coding activity, trainees in the conditions that included retention enhancement recorded the number of times they were able to mentally rehearse the key learning points. The mean of the responses was 2.85, which means trainees in the conditions that included retention enhancement rehearsed the presented skills more than two times on average. In sum, the results indicate that the experimental manipulation of retention enhancement did result in symbolic coding and cognitive rehearsal activities as intended.

In order to confirm that we did not confound the hypothesis tests by unintentionally creating unequal training quality across training conditions, trainee reaction was measured. The reaction measure consisted of seven items with an 11-point Likert-type scale (0 = completely disagree, 5 = neither agree nor disagree, 10 = completely agree). In addition to one item that asked about overall satisfaction with the quality of training, the measure included three items designed to assess trainee’s reaction toward the training program and another three items for trainee’s reaction toward the trainer. For example, participants were asked to indicate if they were satisfied with the trainer and if they were satisfied with the training program. Because of the high correlations among the items between these two groups and with the overall satisfaction item (all higher than .43, \( p < .001 \)), all the scores were averaged to create a summary score of each participant’s reaction. The reaction measure showed an internal consistency reliability (Cronbach’s alpha) of .93 and no significant differences across training conditions \( (F = 1.95, p = .15) \), confirming equal training quality across training conditions. Also, the reaction measure scores were high \( (M = 7.96) \), indicating that the trainees were generally satisfied with the training program.

RESULTS

A set of analysis of variance (ANOVA) tests showed that participant characteristics were not significantly different across training conditions in age \( (F = .90, p = .41) \), gender \( (F = .64, p = .53) \), GPA \( (F = .61, p = .55) \), computer experience \( (F = .76, p = .47) \), spreadsheet program experience \( (F = .08, p = .92) \), work experience
Improving Computer Training Effectiveness

(F = 1.78, p = .17), English as the native language (F = .39, p = .68), pre-training motivation (F = 1.19, p = .31), or confidence in using the computer (F = .03, p = .97) or the spreadsheet software (F = .82, p = .44), thus confirming equalized experimental conditions at the outset.

Table 2 shows that significant intercorrelations exist among the study variables. Immediate learning and delayed learning were strongly correlated: .75 (p < .001) for comprehension and .65 (p < .001) for task performance, indicating that initial outcomes of trainee learning significantly influence delayed learning outcomes, no matter whether they are cognitive or skill-based outcomes. In addition, comprehension and task performance were significantly correlated—all higher than .32 (p < .001), suggesting that conceptual understanding and hands-on performance are closely related.

Table 3 shows means and standard deviations of learning outcomes by the experimental conditions. Given the high correlations among the dependent variables, a multivariate analysis of variance (MANOVA) was first performed to test the effects of different training conditions on training outcomes (Cohen & Cohen, 1983). The overall test was significant (Wilks’ lambda F(2, 108) = 1.96, p < .05). Then, for each training outcome, a univariate ANOVA was performed, followed by the Tukey-Kramer test to determine which differences between training conditions were statistically significant. The Tukey-Kramer test is the most preferred procedure among the multiple comparison procedures because it controls the Type I error with the highest power (Kirk, 1995).

As presented in Table 4, results of ANOVA on the immediate and delayed comprehension test scores were significant, supporting H1. The ANOVA on the comprehension score taken immediately after training showed a significant effect for alternative training conditions (F(2, 108) = 3.42, p < .05). Also, the ANOVA on the comprehension score taken one week later showed a significant effect (F(2, 108) = 3.10, p < .05). Results of the subsequent Tukey-Kramer tests revealed that the retention enhancement plus practice condition yielded significantly higher levels of comprehension than the practice-only condition (M = 6.84 vs. 5.69, p < .05) for immediate learning, and that the significance was persistent over one week period (M = 7.03 vs. 5.94, p = .05). For both immediate and delayed comprehension scores, no other comparisons were significant at the conventional significance level of .05. Figure 2 shows the mean levels of trainee comprehension as a function of training condition for both immediate and delayed learning.

The ANOVA using the task performance score as the dependent variable showed no significant difference for immediate task performance (F(2, 108) = 2.20, p = .12) or delayed task performance (F(2, 108) = .25, p = .78), contrary to H2. The retention-enhancement-plus-practice condition produced higher scores than the retention-enhancement-only condition did in both immediate and delayed task performance scores, but the scores were not statistically different at the significance level of .05. For both immediate and delayed task performance scores, no other comparisons were significant at the significance level of .05.

H3 predicted that subjects with higher cognitive and skill-based outcomes would perceive the system as easier to use. As shown in Table 2, ease-of-use perception was found significantly correlated with immediate comprehension (r = .28, p < .01) and immediate task performance (r = .19, p < .05). Delayed learning
Table 2: Intercorrelations for the study variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Immediate comprehension</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Delayed comprehension</td>
<td>.75***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Immediate task performance</td>
<td>.36***</td>
<td>.32***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Delayed task performance</td>
<td>.42***</td>
<td>.35***</td>
<td>-.65***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Ease of use</td>
<td>.28**</td>
<td>.32***</td>
<td>.19*</td>
<td>.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Usefulness</td>
<td>.04</td>
<td>.02</td>
<td>.00</td>
<td>-.04</td>
<td>.31***</td>
<td></td>
</tr>
</tbody>
</table>

N = 111

*p < .05

**p < .01

***p < .001

Table 3: Means and standard deviations of experiment variables by training condition.

<table>
<thead>
<tr>
<th>Condition</th>
<th>n</th>
<th>Compl1</th>
<th>Compl2</th>
<th>Task1</th>
<th>Task2</th>
<th>EOU</th>
<th>Useful</th>
</tr>
</thead>
<tbody>
<tr>
<td>Practice only</td>
<td>36</td>
<td>5.69</td>
<td>5.94</td>
<td>16.33</td>
<td>16.69</td>
<td>8.50</td>
<td>7.90</td>
</tr>
<tr>
<td>M</td>
<td></td>
<td>2.15</td>
<td>2.16</td>
<td>3.93</td>
<td>4.31</td>
<td>1.41</td>
<td>2.35</td>
</tr>
<tr>
<td>RE only</td>
<td>37</td>
<td>6.30</td>
<td>6.19</td>
<td>14.22</td>
<td>16.35</td>
<td>7.80</td>
<td>7.84</td>
</tr>
<tr>
<td>M</td>
<td></td>
<td>1.84</td>
<td>2.09</td>
<td>5.20</td>
<td>4.58</td>
<td>1.53</td>
<td>2.30</td>
</tr>
<tr>
<td>RE plus practice</td>
<td>38</td>
<td>6.84</td>
<td>7.03</td>
<td>15.84</td>
<td>17.03</td>
<td>8.35</td>
<td>8.41</td>
</tr>
<tr>
<td>M</td>
<td></td>
<td>1.65</td>
<td>1.62</td>
<td>4.36</td>
<td>3.86</td>
<td>1.37</td>
<td>1.88</td>
</tr>
</tbody>
</table>

RE = Retention enhancement; Compl1 = Immediate comprehension; Compl2 = Delayed comprehension; Task1 = Immediate task performance; Task2 = Delayed task performance; EOU = Ease of use; Useful = Usefulness.

Comprehension scores are the average number of correct answers out of 10. Task performance scores are the average number of correct answers out of 20. Ease of use and usefulness scores are on a scale of 0 (negative) to 10 (positive).

Table 4: Results of analysis of variance on immediate and delayed comprehension.

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>ω^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Immediate Comprehension</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition</td>
<td>24.35</td>
<td>2</td>
<td>12.18</td>
<td>3.42*</td>
<td>.04</td>
</tr>
<tr>
<td>Residual</td>
<td>384.42</td>
<td>108</td>
<td>3.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>408.78</td>
<td>110</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delayed Comprehension</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition</td>
<td>24.02</td>
<td>2</td>
<td>12.01</td>
<td>3.10*</td>
<td>.04</td>
</tr>
<tr>
<td>Residual</td>
<td>418.54</td>
<td>108</td>
<td>3.88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>442.56</td>
<td>110</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .05
outcomes are not appropriate in this context because ease-of-use perception was not measured after the delayed period of learning. The hypothesis was also tested by grouping trainees as high achievers or low achievers and comparing the mean ease-of-use scores of the two groups. The grouping was made based on whether or not each trainee performed better than the median score for immediate comprehension and immediate task performance. Out of 111 trainees, 34 were identified as high achievers on both measures of comprehension and task performance while 35 were identified as low achievers. The remaining trainees (42) were high only on one of the two test measures. A t-test conducted between high achievers and low achievers showed a significant difference in the system’s ease-of-use perception ($t = 2.63, p < .05$), supporting the hypothesis.

H4 predicted that subjects with higher ease-of-use perception would perceive the system as more useful. As shown in Table 2, ease-of-use perception was found to be significantly correlated with usefulness perception ($r = .31, p < .001$). As before, each trainee was grouped as a high ease-of-use perceiver or low ease-of-use perceiver with the median value of ease-of-use scores. A t-test conducted between high perceivers ($n = 52$) and low perceivers ($n = 59$) showed a significant difference in the system’s usefulness perception ($t = 2.26, p < .05$), supporting the hypothesis. Table 5 provides a summary of hypothesis testing.
Table 5: Summary of hypothesis testing.

<table>
<thead>
<tr>
<th>Learning Outcomes</th>
<th>Hypothesis</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive</td>
<td>H1: Combining retention enhancement with hands-on practice will produce better cognitive learning outcomes than practice alone.</td>
<td>Supported</td>
</tr>
<tr>
<td>Skill-based</td>
<td>H2: Combining retention enhancement with hands-on practice will produce better skill-based learning outcomes than retention enhancement alone.</td>
<td>Not Supported</td>
</tr>
<tr>
<td>Affective</td>
<td>H3: Subjects with higher cognitive and skill-based outcomes will perceive the system as easier to use.</td>
<td>Supported</td>
</tr>
<tr>
<td></td>
<td>H4: Subjects with higher ease-of-use perception will perceive the system as more useful.</td>
<td>Supported</td>
</tr>
</tbody>
</table>

DISCUSSION

Findings of the study indicate that the computer training method based on behavior modeling can be further improved by incorporating retention enhancement, which had been overlooked by previous research on computer training. Trainees who performed retention enhancement and then hands-on practice achieved significantly better cognitive learning than trainees who performed only hands-on practice. The initial difference in cognitive achievement was still evident one week after training. Specifically, the retention enhancement added to hands-on practice improved trainee comprehension by 20% for immediate learning (from 5.69 to 6.84), and by 18% for delayed learning (from 5.94 to 7.03). It should be noted that the achievements were gained without requiring any additional training time.

Contrary to expectation, the combined condition did not produce significantly better skill-based outcomes than the retention enhancement-only condition. However, the non-significant effect was also true with the comparison between the practice-only condition and retention enhancement-only condition. Further, the skill-based outcomes of the combined condition were very compatible with those of practice-only condition. To paraphrase, the combined condition was as effective as the practice-only condition in improving hands-on skills despite its substantially shorter amount of hands-on practice time (16 minutes) compared to the practice-only condition (26 minutes).

Although much prior computer training research has examined cognitive, skill-based, and affective training outcomes (e.g., Davis & Bostrom, 1993; Galletta et al., 1995; Gist et al., 1989; Olfman & Bostrom, 1991), there has been relatively little attention to the theoretical and empirical relationships among alternative training outcomes. The present research provides support for H3 and H4, which link cognitive and performance-based measures of skill acquisition to perceived ease of use and perceived usefulness, motivational determinants of technology acceptance that have been established by numerous previous studies outside the training context (for recent review, see Venkatesh & Davis, 2000). A key reason
potential users may choose not to adopt a new system is lack of sufficient skill to use it effectively. If users do not use them, decision technologies will not be able to deliver whatever performance benefits they are designed to offer. The prospects for further integrating the somewhat disparate streams of research on user training, on the one hand, with user acceptance, on the other, appear promising.

It is noteworthy that the level of immediate learning was significantly correlated with that of delayed learning for both cognitive and skill-based outcomes. The significant relationship between immediate and delayed learning was consistent across training conditions, with all correlations higher than .56 (p < .001). The initial conceptual understanding of task requirements was a significant determinant of comprehension measured after one week, and the initial level of hands-on performance was a significant determinant of skill compilation measured after the extended practice time. The results clearly show how critical it is to learn correctly from the beginning.

Limitations
Several limitations of the present study should be noted. First, the current study did not manipulate technology to examine possible interplay between training and technology. Although the training technique itself, retention enhancement, could potentially be implemented using technology, it was included as part of an instructor-led training workshop in the present study. Given that it is already common to include video-demonstrations of target skills in CD-ROM-based training materials, assessing the relative effects of retention enhancement implemented in various formats should be actively pursued. Second, many of the findings in this study are bounded by its various conditions and should be reexamined for generalizability. In this study, the training program was about three hours long, which is the same as some previous behavior-modeling studies in the computer training area (Gist et al., 1988, 1989), and longer than others (Simon et al., 1996; Simon & Werner, 1996). Also, the software application for this study was selected to be consistent with some previous behavior-modeling studies (Compeau & Higgins, 1995; Gist et al., 1988, 1989), and to be representative of software commonly used in business organizations for decision making. Most of the participants were novice users of the chosen software program, but with substantial work experience and computer experience. The assessed learning outcomes included skills that can be directly used in real work settings. Thus, the present study maintains important training conditions used by previous behavior-modeling studies in the computer training area, while capturing key trainee and computing characteristics to increase external validity. However, the findings should be validated in other settings beyond the specific conditions of this study. Finally, the study failed to support H2, that combining retention enhancement with hands-on practice will produce better skill-based learning outcomes than retention enhancement alone. One interpretation is that the hands-on practice truly exerts no better influence on skill acquisition. However, there are some rival explanations that might also be responsible for the null finding. There may have been a ceiling effect and suppression of variance in the task performance measure used. The facilitators observed that many trainees completed the performance task early, whereas others took the full 20 minutes
allocated for task completion. By not creating sufficient time pressure for task performance, this measure may have failed to discriminate adequately between subjects who achieved equivalent accuracy scores but took very different amounts of time to finish. Kraiger et al. (1993) pointed out that skill measures should assess both speed and accuracy components of performance. This is a potential shortcoming of the current measure that should be considered in future research.

Implications for Future Research and Practice

The premise of the present study required a leap of faith that the retention enhancement intervention, found to be successful when added to behavior-modeling training in cognitively complex motor skills and interpersonal skills, would carry over to the context of training users of decision technologies. However, the fact that the behavior-modeling approach itself was found by previous research to carry over to the realm of computer training provided a reason for optimism. Our study was partially successful in demonstrating the effectiveness of retention enhancement for spreadsheet training among student users. The question remains how broadly the retention enhancement intervention will generalize across computer training contexts (e.g., from university to corporation, from student to business professional, from spreadsheets to other decision technologies). Two possibilities are (1) training techniques shown to be effective in one context will generalize to other contexts, or (2) it will be beneficial or even necessary to tailor training techniques to each varying context. This is a promising avenue for future research. Prior studies showed that the effects of behavior modeling were consistent between student subjects (Johnson & Marakas, 2000) and workers (Compeau & Higgins, 1995) as well as between university settings (Gist et al., 1988, 1989) and job settings (Simon et al., 1996; Simon & Werner, 1996), which may also hold true for the effects of retention enhancement.

The retention enhancement intervention studied here has some conceptual overlap with the elicitation of self-explanations, which has been shown effective in the acquisition of problem-solving skills (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Chi, De Leeuw, Chiu, & LaVancher, 1994). The self-explanation approach can be implemented by asking subjects, after they have been presented conceptual information on a subject (e.g., the circulatory system), four sets of questions designed to probe for verbatim information acquisition, comprehension, knowledge inference, and knowledge application. The process of responding to these prompt questions led to significant gains in learning among treatment subjects compared to controls, which are theorized to be a result of integration of new material with prior knowledge and the formation of correct mental models of the phenomenon (Chi et al., 1994). Thus, the mechanism underlying the self-explanation effect is similar to that of retention enhancement—creation and elaboration of a cognitive representation of knowledge. There are differences, however. Self-explanation emphasizes verbal rehearsal and elaboration of material, whereas retention enhancement emphasizes the formation, through repetition, of stored mental images of a behavior, and the creation of associated symbolic codes that assist in their retrieval from memory. Both of them facilitate learning, however, probably because they are both constructive activities of building and refining
one’s mental model through active processing of information. Overall, self-explanation is sufficiently promising to warrant continued investigation as a computer skill training method, possibly in conjunction with retention enhancement.

The current findings have practical implications. Carroll and Rosson (1987) documented that individuals face significant impediments to becoming effective users of computers, and that the skills of active users tend to asymptote at relative mediocrity. Underlying this is a “production paradox” in which users’ eagerness to accomplish actual work causes them to be unwilling to spend adequate time practicing the skills sufficiently to assure effective performance. Instead, users go into production mode as soon as they have acquired the minimal procedural knowledge to perform a task, and the repetition they experience in using the minimal skills does not expose them to the more advanced and efficient methods available. By reducing the dependence on hands-on practice, and efficiently providing the cognitive structures needed to master a new method, retention enhancement offers a desirable path toward computer mastery. This does not suggest that physical practice should be eliminated. In our study, we have found that the combined condition of retention enhancement and hands-on practice maximizes computer learning outcomes. Driskell, Copper, and Moran’s (1994) meta-analysis of the literature on mental practice concluded that, although mental practice is indeed effective for cognitive tasks, the effect of mental practice alone diminishes across retention intervals quicker than the effect of physical practice. In our study, the combined implementation of retention enhancement and physical practice produced the highest cognitive learning outcomes one week after training as well as immediately after training.

In conclusion, this research shows that the existing behavior-modeling technique can be further improved to deliver more effective computer learning outcomes, answering the calls made by several researchers to further improve behavior modeling and better understand its effectiveness in various conditions (Baldwin, 1992; Tannenbaum & Yukl, 1992; Werner, O’Leary-Kelly, Baldwin, & Wexley, 1994). In addition, no studies have examined the effects of retention enhancement, when added to behavior modeling, on computer learning outcomes. Thus, the present research offers a contribution to the literature on computer training, opening the door for other research. The results of the study also show that behavior modeling can include retention enhancement processes without requiring additional time. Computer training practitioners should benefit from employing a more effective training strategy that yields improved learning outcomes without requiring additional time. [Received: August 7, 2000. Accepted: July 31, 2001.]

REFERENCES


Improving Computer Training Effectiveness


**Mun Y. Yi** is an assistant professor of information systems at the Moore School of Business, University of South Carolina. He received his PhD in information systems from University of Maryland, College Park, and an MS in computer information systems from Georgia State University. He has taught various information technology courses at the undergraduate and graduate levels. His current research focuses on computer skill acquisition and training, information technology adoption, decision support systems, and electronic commerce. His research has previously appeared in *Information Systems Research*.

**Fred D. Davis** is a professor, David D. Glass Chair in Information Systems, and chair of the Information Systems Department at the Sam M. Walton College of Business at the University of Arkansas. Dr. Davis earned his PhD at MIT's Sloan School of Management, and has served on the business school faculties at University of Michigan, University of Minnesota, and University of Maryland. He has taught a wide range of information technology (IT) courses at the undergraduate, MBA, PhD, and executive levels. Dr. Davis has served as associate editor for the scholarly journals *Management Science* and *MIS Quarterly*, and has published extensive research on user acceptance of IT in the workplace, computer-assisted planning and decision making, management of emerging IT, and IT training and skill acquisition. Current research interests also include the motivation and retention of IT professionals, virtual teams, and technology-supported distance learning.