DOCUMENT RESUME

ED 373 713 IR 016 723

AUTHOR Farquhar, John D.; Regian, J. Wesley
TITLE Effects of Knowledge Representation during Computer-Based Training of Console Operation Skill.
PUB DATE 94
NOTE 18p.; In: Proceedings of Selected Research and Development Presentations at the 1994 National Convention of the Association for Educational Communications and Technology Sponsored by the Research and Theory Division (16th, Nashville, TN, February 16-20, 1994); see IR 016 784.
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Autoinstructional Aids; *Computer Assisted Instruction; *Computer Simulation; *Expert Systems; *Feedback; *Intelligent Tutoring Systems; *Military Training; Teaching Machines

ABSTRACT

Console-operation skill is procedural knowledge of control panel actions. The console operator must select appropriate sequences of steps (e.g., setting dials, pressing buttons, etc.) as mandated by desired goals and subgoals. A computer-based training (CBT) environment can simulate an environment where knowledge acquisition is efficiently and effectively achieved. In the training of console-operation skill, a CBT environment can simulate control-panel responses and device actions while providing instructional support. The instructional support delivered during practice consists of a set of intervention strategies. This paper discusses an ongoing investigation of CBT approaches in the acquisition of console-operation skill. Two variables of the intervention include the content of feedback and the timing of feedback. Preliminary results from a laboratory experiment suggest that feedback which elaborates on the goal structure of the console-operation skill is more effective than feedback that simply corrects errors. (Contains 42 references.) (Author/JLB)

Title:

Effects of Knowledge Representation during Computer-Based Training of Console Operation Skill

Authors:

John D. Farquhar
Department of Instructional Technology
University of Georgia

J. Wesley Regian
Air Force Armstrong Laboratory
Brooks AFB, TX 78235

ATHENS, GA 30602

Funds for this research were provided by the
U.S. Air Force, Technical Training Research Division
Brooks Air Force Base, San Antonio, Texas

and the

Air Force Office of Scientific Research
Bolling Air Force Base, Washington, D.C.

ABSTRACT

Console-operation skill is procedural knowledge of control panel actions. The console operator must select appropriate sequences of steps (e.g. setting dials, pressing buttons, etc.) as mandated by desired goals and subgoals. In the training of console-operation skill, a Computer-Based Training (CBT) environment can simulate control-panel responses and device actions while providing instructional support. The instructional support delivered during practice consists of a set of intervention strategies. This paper discusses an ongoing investigation of CBT approaches in the acquisition of console-operation skill. Two variables of intervention described in this paper include the content of feedback and the timing of feedback. Preliminary results from a laboratory experiment suggest that feedback which elaborates on the goal structure of the console-operation skill is more effective than feedback that simply corrects errors.

INTRODUCTION

Computer-Based Training (CBT) seeks to provide an environment where knowledge acquisition is efficiently and effectively achieved. In the training of console-operation skill, the CBT environment can simulate control-panel responses and device actions while providing instructional support. This paper describes an ongoing investigation of CBT approaches in the acquisition of console-operation skill.

At a basic level, console-operation skill is the execution of control-panel actions such as the pressing of a button or the rotation of a dial. At a higher level, the skill is a complex hierarchy of goals, subgoals and discrete steps. The console operator must select appropriate sequences of steps as mandated by the desired goals and subgoals. During practice of this procedural skill, the instructional support provided by the CBT environment can be described as a set of intervention strategies. These strategies must resolve two important questions: When is intervention most appropriate, and with what message does one intervene?

Studies of human-tutoring suggest that tutors intervene at strategic points within the execution of a learning task (Merrill, Reiser, Ranney, and Trafton, 1992). Tutors often "overlook" minor errors during practice, reserving immediate feedback for errors deemed more important. In contrast, traditional CBT methods typically provide immediate feedback for all errors. Alternative intervention strategies require sophisticated decision-making processes.

The message delivered at an intervention can include an array of information types. Beyond the simple verification of incorrectness, the message can provide an explanation for the error, hints or guidance toward the correct response, and even the correct response itself. Guidance delivered at an intervention can direct the learner's attention to the goal structure of the procedure. This is often the approach of Intelligent Tutoring Systems (ITS) where the content of an intervention message goes beyond simple verification of incorrectness. The study described in this paper attempts to determine the value of these intervention approaches in terms of the efficiency of skill acquisition, quality of achievement and learner preferences.

The platform of the investigation is a computer-based, console-operations simulation called LOADER (Farquhar, 1993). Analogous to many military and industrial tasks that call for knowledge of various procedures, LOADER is a laboratory research tool developed to assess instructional strategies in the acquisition of console-operation skill. The task simulated by the LOADER environment is the operation of a remote crane control arm to load various canisters from a set of storage bins to one or more railroad cars (see, Figure 1 for the simulation interface).

In order to provide instructional guidance capable of determining when to intervene and what to say, an expert system was embedded into the delivery system of LOADER. The expert system evaluates error criticality in order to determine points of intervention. In addition, the expert system analyzes solution methods to best select the goal structure message. Essentially, LOADER behaves as an Intelligent Tutoring System (ITS) when features of the expert system are employed.

BACKGROUND

The evaluation of ITS programs and methods has only recently become a topic of discussion in the ITS literature (Mark & Greer, 1993; Murray, 1993; Shute & Regian, 1993; Winne, 1993). Reported evaluations cover a broad spectrum of analyses from investigations of the internal components of a program to its external effectiveness (Legree, Gillis, & Orey, 1993; Littman & Soloway, 1988). *Internal evaluations* focus on the effects of specific components or processes of a program, such as diagnostic accuracy and pedagogical decisions, while *external evaluations* address the overall educational impact.

In a review of ITS external evaluations conducted by Legree and Gillis (1991), ITSs deemed extensive in nature (i.e. covering a third of a college course or more than 30 hours of instruction) showed an improvement on knowledge-based tests of approximately 1.0 standard deviation. This analysis tends to support the overall effectiveness of ITSs. However, the number of extensive external evaluations conducted remains very small (four in the Legree and Gillis report). The 1.0 standard deviation improvement is also far short of the 2.0 standard deviation improvement that appears to be possible with human-tutoring interactions (Bloom, 1984).

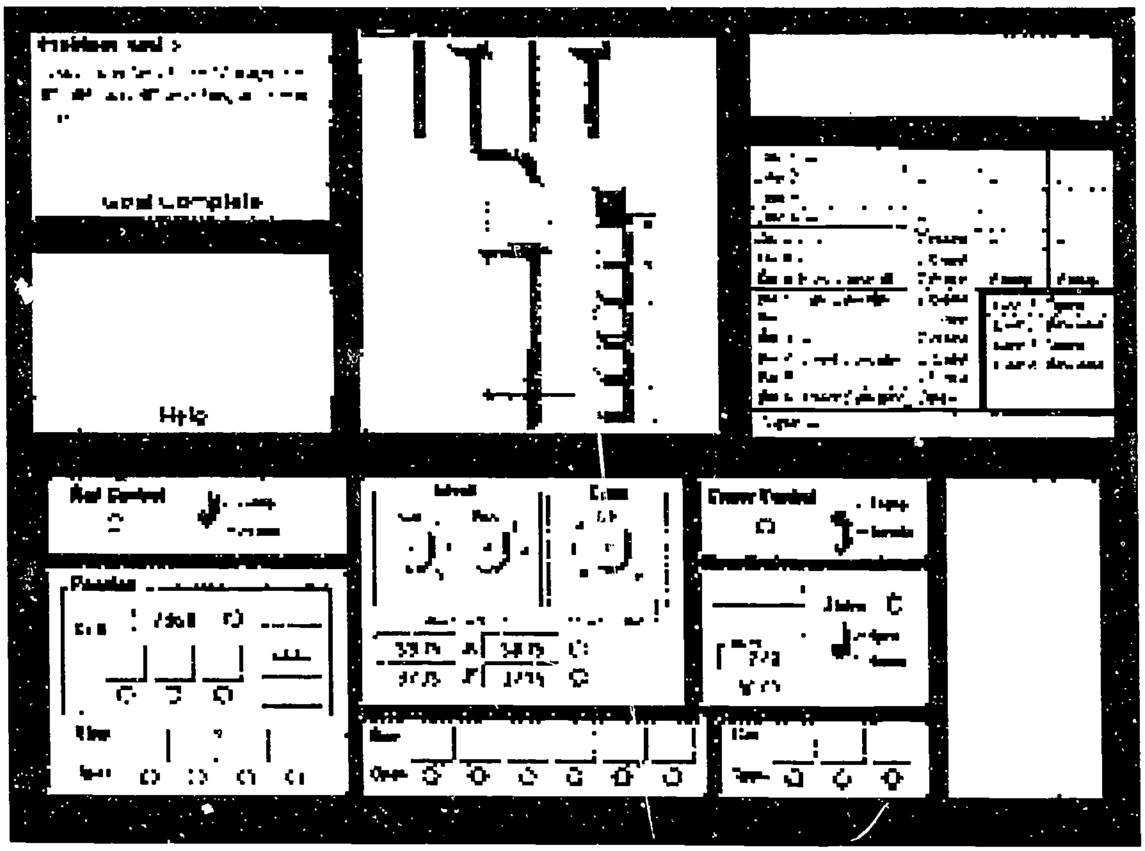

Figure 1: *LOADER interface.*

In addition to conducting more external evaluations, ITS researchers maintain that internal evaluations must also continue in order to determine the efficacy of ITS methods. One effective ITS methodology supported by the literature is *model tracing* (Anderson, 1988; Merrill, Reiser, Ranney, & Trafton, 1992). A tutor employing the model-tracing method matches each step of the student's problem-solving behavior to the possible solution paths. When the student's action deviates from a known solution path, the tutor intervenes with informative and corrective advice.

The theory driving the model-tracing methodology is Anderson's ACT* Theory (Anderson, 1983). A central assumption of the ACT* theory of cognition is that the content of error-corrective feedback is processed by learners facilitating their acquisition of knowledge. In other words, learners will gain from information-rich feedback, such as error explanations and goal descriptions. Unfortunately, some of Anderson's own studies fail to support this assumption (Corbett & Anderson, 1991; Corbett, Anderson, & Patterson, 1990).

The model-tracing methodology also relies heavily on the use of immediate feedback. While the ACT* theory does not mandate the use of immediate feedback (Corbett & Anderson, 1991; Corbett, Anderson, & Patterson, 1990), both practical and theoretical reasons exist for its implementation. Included in the theoretical reasons is the reduction of non-productive floundering. On the other hand, Schooler and Anderson (1990) have found a number of disadvantages of using immediate feedback: 1) the student grows dependent upon feedback and hesitates to act without it; 2) the student does not develop error correction and detection abilities; and, 3) the feedback competes for the student's limited short-term memory resources.

In a study designed to investigate the effects of immediate, information-rich feedback (Corbett & Anderson, 1991), forty undergraduates completed a series of Lisp programming exercises with the CMU Lisp Tutor. The treatments under investigation delivered four different types of feedback: immediate feedback, error flagging, feedback on demand, and no feedback. Of the three treatments that delivered feedback, all presented similar forms of information-rich feedback. The three feedback types differed primarily in their method of error notification.

Immediate feedback, in the Corbett and Anderson study, provided information-rich feedback upon the execution of every error. The error flagging condition notified students of an error, but did not deliver informative feedback unless requested by the student. The feedback-on-demand condition performed no error notification, delivering informative feedback only when requested. Finally, the no-feedback condition delivered neither error-notification feedback nor informative feedback.

The results of the Corbett and Anderson study found no performance differences between groups that received feedback. The no-feedback condition performed consistently worse on a number of performance measures. Essentially, these results indicate that any feedback that serves to identify errors is superior to no feedback. The benefits of providing *information-rich* feedback (as accomplished with model-tracing) was not demonstrated. The lack of significant differences between feedback conditions in this study can be explained by the small sample size. Even if the feedback manipulation was expected to produce a "large" effect, a sample size of 68 would be required for an acceptable level of power (.80) (Keppel, 1991).

A comparison analysis of human tutoring and tutoring with model-tracing found a few notable differences (Merrill et al., 1992). Among these differences is the use of immediate feedback. Human tutors, it appears, are more flexible in determining when to provide immediate and intervening feedback (Littman, 1991; Littman, Pinto, & Soloway, 1990). Often, the human tutor will allow student's to perform low-level errors without immediate feedback. Instead the feedback for such errors is delivered at a later time, often at the completion of the problem. Therefore, human tutors make strategic decisions in the timing of feedback, reserving advice for the appropriate moment.

Although a number of ITSs employing the model-tracing methodology have been demonstrated as effective through external evaluations, the efficacy of their use of feedback is far from determined. And, since the content and timing of feedback is a central component within these tutors, it is prudent to investigate these attributes.

Content of Feedback

The use of feedback is widely accepted within educational theory and practice as a critical element of the learning process (Anderson, 1983; Bangert, Kulik, & Kulik, 1983; Grinder & Nelson, 1985; Gagné, 1984). Extensive research in feedback has been conducted on psychomotor and cognitive skill (Lee & Carnahan, 1990; Salmoni, Schmidt & Walter, 1984; Wulf & Schmidt, 1989), testing and instructional environments (Heald, 1970; Roper, 1977), tutor, text, and technology-based programs (Kulik & Stock, 1989; Schimmel, 1983), as well as laboratory and field settings. This review is limited to the use of feedback in the acquisition of cognitive skill from instructional situations. While there are a number of theoretical models that describe mechanisms of feedback, simply defined, instructional feedback is a response to students' actions.

Actions and responses to these actions creates a cycle of interactivity. A common feature of current theoretical models of instructional feedback is the cyclic changes that occur with the learner's cognitive state (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991; Kulhavy & Stock, 1989). These changes are influenced by a number of factors. Perhaps the most important is the type of feedback provided (Dempsey, Driscoll, & Swindell, 1993).

Type of feedback can be classified into the following five categories (Roper, 1977; Bangert-Drowns, et al., 1991; Dempsey et al., 1993):

1. *No Feedback* is often a control condition used in comparison to other forms of feedback.

2. *Verification Feedback* or *Knowledge of Results (KR) Feedback* simply informs the learner of a correct or incorrect response.

3. *Correct Response* or *Knowledge of Correct Response (KCR) Feedback goes* beyond Verification Feedback to inform the learner of the correct answer.

4. *Try-again Feedback* directs the learner to make one or more additional attempts at the answer.

5. *Elaborative Feedback*, may be a number of different forms of additional information that either provides an explanation for the error or reviews relevant material.

A review by Kulhavy and Stock (1989) found more than 50 studies completed since 1970 where elaborative feedback was one of the variables under investigation. Elaborative feedback, as defined by Kulhavy and Stock, consists of any information delivered to the student beyond simple verification. Thus KCR or Knowledge of Correct Response feedback would be included as a form of elaboration. Unfortunately, the manipulations of the content of elaborative feedback have been inconsistent, making simple generalizations of results difficult if not impossible. In addition, or perhaps due to this inconsistency, the results have been varied. Kulhavy and Stock cite that about half of the studies they reviewed where the manipulation of elaborative feedback consisted of increasing amounts of information produced improved performance. The remaining half showed no significant improvement.

An example, cited by Kulhavy and Stock, where "elaborative" feedback was shown to be beneficial is a study by Roper (1977). In that study, subjects who received KCR feedback within a computer-assisted instructional program demonstrated superior performance than those subjects receiving KR feedback. And KR feedback, in turn, led to improved performance over a no feedback condition. Note, however, that the manipulation in this study was simply KCR over KR and not elaborative feedback as defined as additional explanatory information.

The Kulhavy and Stock analysis points to one of the major problems with research into elaborative feedback: a consensual definition. Studies have included numerous types of elaborations from simple explanations of errors to re-presentations of the instruction (Kulhavy, 1985). Often, as with the Kulhavy and Stock report, mere presentations of the correct response is considered "elaborative".

Even though Kulhavy and Stock claim limited evidence for the benefits of elaborative feedback (Roper, 1977), they conclude: "Based on our assessment of the literature, we are unable to reach any useful conclusion regarding how the elaborative component of feedback operates" on post performance (p. 289).

In a more recent meta-analysis, Bangert-Drowns, et al. (1991) also failed to conclude that the amount of information given in feedback effects post performance despite their presumptions that "elaborate feedback would have encouraged the construction of a richer network of pathways to the desired information" (p. 234). Instead, the major conclusion of the Bangert-Drowns et al. analysis is that the primary role of instructional feedback is in the correction of errors. In their report, studies that investigated the use of *corrective* feedback (elaborative *or* KCR feedback), while controlling for major sources of feedback-related variation (such as presearch availability), demonstrated an average effect size of 0.77. While supportive of corrective feedback over verification feedback, the Bangert-Drowns et al. analysis found little support for the use of information-rich feedback.

Only one study in the Bangert-Drowns report compared the use of elaborative or explanatory feedback to correct-answer or KCR feedback (Sassenrath & Gaverick, 1965). In this study, simple KCR feedback was found to be superior to feedback that provided an explanation. A result inconsistent with an information-processing theory of learning.

The resulting picture of the use of elaborative feedback for instruction is, at best, inconsistent. There is much support for the use of corrective feedback, but beyond providing corrective information the role of information-rich feedback is far from understood.

Timing of Feedback

A second major variable of interest in the literature on instructional feedback is in timing. An operant theory approach to behavior modification insists upon the delivery of immediate reinforcement (Skinner, 1968). This approach to feedback was quickly challenged by what became known as the Delay Retention Effect (Brackbill, Bravos, & Starr, 1962; Brackbill & Kappy, 1962; Brackbill. Wagner, & Wilson, 1964). Using multiple-choice testing, researchers consistently found that a delay in feedback of one or more days increased performance. These findings were generally left unexplained until Kulhavy and Anderson (1972) proposed the *interference-perserveration* hypothesis.

The interference-perserveration hypothesis suggests that subjects' initial errors interfered with their ability to immediately acquire the correct response from feedback. A delay in the delivery of feedback reduced this interference because subjects essentially "forgot" their initial (erroneous) response. This hypothesis gathered further support through a meta-analysis by Kulik and Kulik (1987).

The Kulik and Kulik analysis reviewed 53 studies that compared some form of immediate to delayed feedback. The studies covered a broad range of experiments from applied studies with classroom quizzes and programmed materials to controlled experiments in list learning and learning from test materials. Such differences between studies warranted a separate analysis for type of study. Indeed, Kulik and Kulik found that immediate feedback was more effective than delayed in typical classroom settings with an average effect size of 0.28.

Experiments involving delayed feedback and list learning have demonstrated highly equivocal results. Strong support for immediate feedback as well as strong support for delayed feedback have been reported. Kulik and Kulik provide a detailed explanation for these findings. Essentially, studies reporting positive effects for the use of delayed feedback presented feedback in such a way that the feedback constituted a separate and additional practice trial. Experiments that controlled for this event (by fusing the feedback to the instruction through a reduction in the time interval or an elimination in cues) resulted in a greater effect for immediate feedback.

Consistent with this conclusion, experiments involving the delay of feedback during testing conditions showed gains over the use of immediate feedback. Again, the delay of feedback in these studies allowed subjects to essentially perform additional practice trials. Since most of the studies showing support for delayed feedback failed to control for additional instructional trials, the use of immediate feedback for most educational situations, is strongly recommended by Kulik and Kulik.

Unfortunately, the Kulik and Kulik analysis has not laid to rest the issue of immediate versus delayed instructional feedback. A number of supporters remain loyal to the use of delayed feedback for some situations (Dempsey, Driscoll, & Swindell, 1993; Cohen, 1985). Roper (1977) suggests that the delay of corrective feedback (while providing immediate verification feedback) may promote learners to actively rethink the problem, thus giving the question more attention than situations where all the feedback is given at once. Cohen (1985) lists a number conditions under which delay of feedback may be beneficial. Immediate feedback may impede the pace of learning for high-mastery students, becoming more of a distraction than an aid. Cohen also claims that delayed feedback, or end-of-session feedback, facilitates long-term retention and "has also been shown to be helpful if a student is working at the comprehension and application level of Bloom's Taxonomy" (p. 35). These claims, however, have little, if any, experimental support.

A study cited by Cohen in support of these claims (Gaynor, 1981), investigated the effect of delay of feedback in a computer-based mathematics program. Comparing immediate, 30-second, end-of-session, and no feedback treatments, the study analyzed the acquisition of matrix algebra at three levels of performance: knowledge, comprehension, and application. While noting some differences of feedback-delay for the knowledge and application levels of the taxonomy, all differences were statistically insignificant. The only significant difference reported by Gaynor were the post-test scores between high- and low-mastery students -- categories presumably determined by the subjects' overall performance on the task. The Gaynor study, therefore, indicates that delay of feedback, or even no feedback, had little effect on performance. A conclusion also reached by Mory (1992).

Surprisingly, most of the research comparing immediate and delayed feedback has not been conducted in the area of Computer-Assisted Instruction (Kulik & Kulik, 1988). Instead, applied situations tended to study the timing effects of feedback from classroom quizzes. The few CAI studies reviewed by Kulik and Kulik showed small to moderate gains for the use of immediate feedback.

While there is very little experimental support for the delay of feedback in instructional situations, there is some consensus that delay *could* be beneficial in *some* circumstances. One revealing line of research in the use of instructional feedback is in the analysis of interactions between students and human tutors (Fox, 1991; McArthur, 1990). As suggested by Bloom (1984), the effectiveness of human tutoring is very high and is often cited by ITS researchers as the ultimate process to model (Carbonell, 1970).

Merrill, Reiser, Ranney, and Trafton (1992) describe a number of common features of human tutoring that bring new insight into the use of immediate feedback for instruction. Human tutors maintain a delicate balance of responsibility, giving students great amounts of control while providing just enough guidance to keep them productive. In doing so, human tutors frequently give immediate feedback in very subtle forms. A slight hesitation in responding with confirmatory information often serves as a clue to the student to double-check for errors.

In maintaining the role of a guide, the tutor is ofttimes presented with a problem of intervention. Human tutors do not tend to intervene at every sign of error. Instead, the intervention strategy is determined by the context of the problem. Episodes of floundering or critical missteps are quickly corrected, but feedback on other types of errors may be withheld until a more appropriate moment. Therefore, human tutors modulate their responses based upon how serious the error is perceived to be. These attributes of selecting intervention strategies based upon the seriousness of errors has not been investigated by researchers of ITS.

METHOD

Independent variables of the study consist of the type of feedback and the timing of feedback for non-critical errors. Each variable has two levels, forming a 2 by 2 matrix (see, Figure 2). The types of feedback are instantiated with elaborative and corrective feedback, whereas the timing of feedback is either immediate or delayed for non-critical errors.

Elaborative feedback includes one or all of the following items: 1) a notification of an error; 2) a reason why the response was incorrect; 3) a brief description of the appropriate subgoal; 4) a detailed list of steps to complete the subgoal; 5) an indication of progress toward completion of the subgoal; and, 6) a direct identification of the very next step. Corrective feedback, on the other hand, simply gave notification of an error while identifying the correct answer.

The immediate feedback condition delivered feedback (either elaborative or corrective) upon the execution of any type of error. Conversely, the delayed feedback condition makes a decision on when to provide feedback. If the error was critical, immediate feedback was given, if the error was non-critical (identified as either inefficient or inappropriate) then the error was recorded, but not immediately reported to the student. Under this condition, all errors were reported in an end-of-problem information screen.

Dependent variables of the study include time to complete each of the Session 1 trials, number of completed problems during Session 2, and, speed and accuracy of Achievement Tests 1 and 2. In addition, subjects' ratings of the treatments, from an on-line questionnaire, were collected. This paper includes an analysis of the Session 1 trials and the speed and accuracy of the first Achievement Test. Data from the remaining two sessions has been collected, but not completely analyzed at this time.

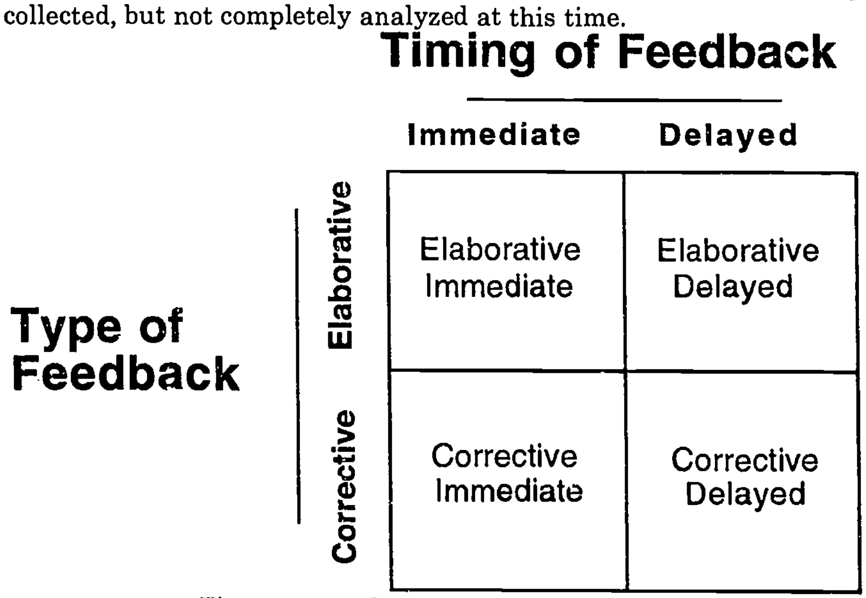

Figure 2: *Levels and variables under study.*

Subjects

This study was carried out at the USAF Armstrong Laboratory under the auspices of the Cooperative Laboratory (CoLab). A total of 119 subjects participated in the study. The Air Force selects, recruits, and hires subjects through temporary employment agencies to fulfill experimental laboratory needs. Selection criteria limited subjects to high-school graduates with no college degree.

	Activity	Data
Session 1	Tutorial & 21 Trials Achievement Test #1	Completion Time Speed & Accuracy
Session 2	Practice Trials Achievement Test #2	Completed Problems Speed & Accuracy
Session 3	Summative Evaluation	Questionnaires

Figure 3: *Schedule of Events.*

Procedure

Groups of twelve to thirty subjects entered the laboratory seating themselves at individual computer workstations. One of four treatment groups was then randomly assigned to each subject. After a brief verbal introduction to the experiment, subjects began Session 1.

Session 1 consisted of a computer-based tutorial followed by a set of 21 complex LOADER problems. The tutorial presents a series of text and graphics screens which describe the loading task. Subjects proceeded through this activity at their own pace. Immediately following the tutorial, subjects began practicing the loading task through the control-panel interface presented in Figure 1 Each treatment was presented with an identical set of practice problems. The presentation of practice problems began with each of the six basic subgoals of a larger LOADER procedure. The number of necessary control-panel actions for each subgoal ranges from five to seven steps. Three problems of each subgoal were presented. The subgoal problems were followed by a set of three problems which required performance of the complete procedure. Execution of the complete procedure is a 32-step process. During practice, feedback was presented according to the treatment types. Session 1 activities ended after subjects completed an immediate acquisition test. The acquisition test was a 32-step LOADER problem. Feedback type during the acquisition test was held constant across treatments. Representative of the "real" console panel, the feedback strategy during the test was of the corrective-delayed type with no on-line help available.

Session 2 consisted of subjects practicing additional LOADER problems with the tutor for a period of 2 hours. This practice session was again followed by an immediate acquisition test. Session 3 included a review and an evaluation of all four treatment conditions. During Session 3, subjects gave each treatment condition a rating.

LOADER: The Console-Operations Task

LOADER is a complex problem-solving task requiring the operation of a simulated control-panel console. The task simulated by the LOADER environment is the operation of a remote crane control arm to load various canisters from a set of storage bins to one or more railroad cars. Simple tasks include the loading of a single canister whereas complex tasks require the operator to analyze the availability and capacity of railroad cars in order to appropriately position and load a number of different-sized canisters. Complexity of the task depends upon variables such as the number and size of canisters to load and the position and capacity of rail cars. Simple tasks may be performed with as few as 30 steps while complex tasks may require as many as 150 steps.

Each problem can be solved through a set of six subgoals (or procedures). These procedures are: 1) Position Rail Car to Appropriate Dock; 2) Position Crane to Appropriate Storage Bin; 3) Lift Canister from Storage Bin; 4) Position Crane to Appropriate Rail Car Receptacle; 5) Load Canister into the Rail Car; and, 6) Dispatch Rail Car to Appropriate Rail Line. Each subgoal has 5 to 6 individual steps or control panel actuations. The rail yard "situation" determines the appropriate settings for various knobs and buttons as well as the correct order of sub-procedures. Complex tasks require one or more iterations through certain subgoals.

The Control-Panel Console

The control-panel console appears as set of knobs, switches, buttons, and indicators appropriately labeled and organized by function on the lower half of the LOADER interface. All interactions with the console are performed through a mouse. While most interactions occur with a simple click of the mouse, some interactions require that the mouse button be held down for several seconds until the action is complete. The panel responds with visual feedback in the form of indicator lights, knob and switch settings with each selection. For certain actions, an audible "click" indicates the completion of the action.

The console is organized into eight panels associating their related functions. The rail control, line and dock panels allow the operator to position rail cars to appropriate loading docks and dispatch cars to available rail lines. Operations requiring positioning the crane, use the centrally located panels including the crane control, crane coordinate system, and the jaws panels. Along the bottom of the screen are panels that allow storage bin doors and rail car doors to be opened.

The Dynamic Model

Located on the upper half of the LOADER interface is the dynamic graphical model. This model provides a representation of the loading task from a birds-eye-view perspective. Positions of the rail cars, storage bins, canisters, doors, the overhead crane, and the line and dock selections are indicated with a number of graphical elements. Colors aid the identification of canister sizes and rail car capacities.

The dynamic model responds to control panel interactions with animated sequences representing their outcomes. For example, engaging a selected rail car south with the button labeled "South," results in the animated motion of the rail car moving down toward the rail docks. In addition to the motion of rail cars, car doors and bin doors open and close, the crane moves above the rail cars and storage bins, and the crane lowers, grasps and lifts canisters. All of the information displayed by the dynamic graphical model (minus the motion and graphical relationship qualities) is also represented on the upper half of the screen in table form.

The Expert Help Feature

Embedded within the interactions of the problem-solving task is an expert help system. The expert help provides intelligent advice to assist in the solving of the loading task. Under the *corrective* intelligence mode, expert advice consists solely of error notification while giving the correct answer. The *elaborative* intelligence mode provides levels of advice often beginning with a reason why the attempted action was incorrect (see, **Error**

Notification in Figure 4). Secondly, with **Level 1 Help**, this mode will give a description of the subgoal that the user should be attempting. By requesting additional help, the expert advice will provide **Level 2 Help** which is a list of the steps required to perform the subgoal with an indication of the steps already accomplished. Finally, if the user requests still more help, the system will give the correct answer (**Level 3 Help**). The levels of expert help (with the exception of describing the reason for an incorrect answer) are also available to the user by selecting the on-screen **Help** option.

Other Display Features

In addition to the dynamic graphical model and its accompanied table, the *LOADER* interface includes a number of other helpful display features. Starting in the upper-left corner of the screen, the name of the problem file is displayed. Below this, a verbal description of the goal for the selected problem appears. Within the same area, a **Goal Complete** button is available for the user to select to indicate that they believe they have completed the problem.

Error Notification: Critical Error -- Crane cannot be lowered through the door of a closed storage bin.

Level 1 Help: You should be attempting to: Load Canister to Receptacle 2 of Car at Dock 3.

Level 2 Help: To Load Canister to Receptacle 2 of Car at Dock 3:

* 1. Set West Car to 3
* 2. Set West Receptable to 2
* 3. Enter West Side Crane Coordinates
4. Set Crane Control to Enable
5. Execute Coordinates
6. Set Crane Control to Disable

* Denotes Action Complete

Level 3 Help: To Set Crane Control to Enable: Click at the button indicated with the red arrow.

Figure 4: *Levels of Expert Help*

Software and Hardware Specifics

LOADER is programmed entirely in OpenScript® using the ToolBook® software development environment from Asymetrix Corporation. Software requirements include MicroSoft Windows as well as ToolBook® run-time files. The program is specifically designed for '486 compatible computers with monitors of 800 X 600 screen resolution.

Results

Three dependent measures are under investigation: Session 1 completion time, Achievement Test 1 completion time, and Achievement Test error count. Analysis of variance with repeated measures on trial completion times during Session 1 revealed no significant differences between treatments [$F(2,113) = .690$, $p = .56$]. Figure 5 shows mean completion times across the 21 practice trials.

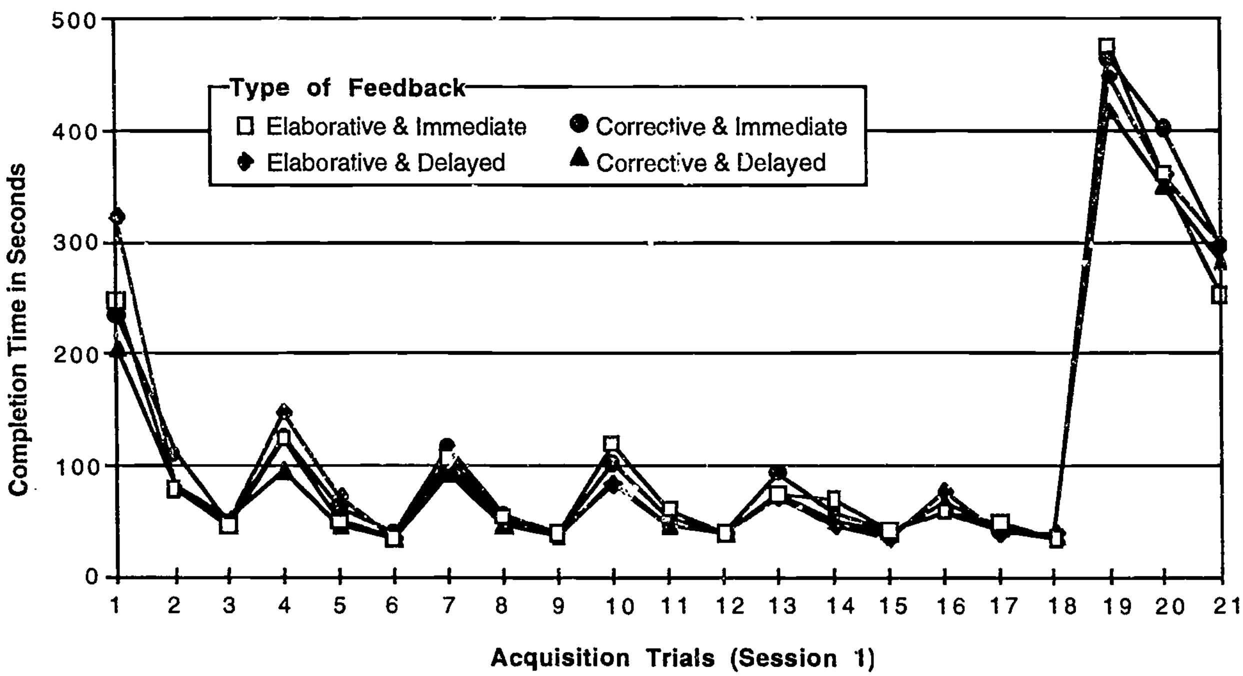

Figure 5: *Mean completion times of acquisition trials.*

Table 1 presents mean scores for the achievement test completion times. Mean error counts for the achievement test are presented in Table 2. These scores reflect an adjustment whereby outliers where removed from the analysis. Due to the nature of the laboratory setting, where subjects are paid for their time, not their performance, extremely poor outlying scores were removed. The analysis did not include six scores for completion time, and seven scores for errors . ANOVAs conducted on these means for overall effects reveal a significant difference for errors [$F(3,108) = 2.978$, $p = .0347$], and completion time [$F(3,109) = 2.709$, $p = .0487$)]. Fisher's PLSD test was used to determine which differences were significant. Of interest was the significant difference between error count means of the elaborative-immediate feedback and corrective-immediate feedback groups ($p = .0229$). All other differences between groups were either non-significant or of little interest (such as differences between elaborative-immediate and corrective-delayed groups).

Feedback Type	Count	Mean	Std. Dev.	Std. Err.
Elab/Immed	30	294.600	158.164	28.877
Elab/Delayed	27	350.741	206.678	39.775
Correct/Immed	28	402.429	196.869	37.205
Correct/Delayed	28	448.893	289.225	54.658

Table 1: *Mean completion time for Acquisition Test 1.*

Feedback Type	Count	Mean	Std. Dev.	Std. Err.
Elab/Immed	30	23.933	25.748	4.701
Elab/Delayed	26	26.269	26.784	5.253
Correct/Immed	29	42.828	33.739	6.265
Correct/Delayed	27	42.556	38.071	7.327

Table 2: *Mean errors for Acquisition Test 1.*

DISCUSSION

This study investigated the value of CBT intervention approaches in terms of the efficiency of skill acquisition, quality of achievement and learner preferences. Three general hypotheses predicted that the use of elaborative feedback and the delay of feedback for non-critical errors would result in gains over corrective feedback and immediate feedback for all errors. A discussion for each of these hypotheses is now presented.

Hypothesis 1: *Elaborative feedback and the delay of feedback for non-critical errors will result in greater efficiency of skill acquisition.* It was suggested that feedback of a more elaborative nature as well as a reduction of intervention by delaying feedback for non-critical errors would result in greater efficiency of acquisition. No differences, however, were found during the first session of training for any of the treatment groups. Differences in acquisition time may become evident with an analysis of Session 2 trial times. Nevertheless, with the data presented here, we can conclude that while elaborative or delayed feedback may not be more efficient, it is at least equivalent to other strategies.

Hypothesis 2: *Elaborative feedback and the delay of feedback for non-critical errors will result in greater performance of skill achievement.* Another measure of the benefits of an intervention approach is the post-training performance. We conjectured that the intervention strategies of elaborative feedback and the delay of feedback for non-critical errors would result in greater post-training performance. In support of that hypothesis, significant differences were found between the means of the elaborative-immediate and the corrective-immediate treatments. All other differences were not statistically significant. Therefor, support is found for the use of elaborative feedback when immediate feedback is implemented. The delay of feedback seemed to affect performance in a negative manner, although this difference was insignificant. Other differences may not become evident until the second session achievement test results are analyzed.

Hypothesis 3: *Elaborative feedback and the delay of feedback for non-critical errors will be preferred by learners.* This data was gathered following the second session of data collection. The results have not been analyzed.

In summary, we hope to reveal additional effects of CBT intervention strategies through a continued analysis of the data collected. The use of feedback which elaborates on the goal structure of the procedure is effective for console-operation training as it reduces achievement errors rates after only a brief training session. The effects of using immediate or delayed feedback for errors which are critical is less clear. These findings tend to support methods used in many Intelligent Tutoring Systems which offer elaborative feedback during practice in a variety of training programs.

REFERENCES

Anderson, J.R. (1988). The expert module. In Polson, M.C. & Richardson J.J. (Eds.). *Foundations of intelligent tutoring systems.* (pp. 21-53) Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Anderson, J.R. (1983). *The achitecture of cognition.* Cambridge, MA: Harvard University Press.

Bangert, R.L., Kulik, J.A., & Kulik, C.C. (1983). Individualized systems of instruction in secondary schools. *Review of Educational Research, 53*(2), 143-158.

Bangert-Drowns, R.L., Kulik, C.C., Kulik, & J.A., Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research. 61*(2), 213-238.

Bloom, B. (1984). The 2 sigma problem: The search for methods of Group instruction as effective as one-to-one tutoring. *Educational Researcher, 20*(4) 231-241.

Brackbill, Y. & Kappy, M.S. (1962). Delay of reinforcement and retention. *Journal of Comparative and Physiological Psychology, 55,* 14-18.

Brackbill, Y. Wagner, J.E., & Wilson, D. (1964). Feedback delay and the teaching machine. *Psychology in the School, 1,* 148-156.

Carbonell, J.R. (1970). AI in CAI: An artificial intelligence approach to computer-assisted instruction. *IEEE Transactions on Man-Machine Studies,* 11(4), 190-202.

Cohen, V.B. (1985). A reexamination of feedback in computer-based instruction: Implications for instructional design. *Educational Technology, 25*(1), 33-37.

Corbett, A.T., & Anderson, J.R. (1991, April). *Feedback control and learning to program with the CMU lisp tutor.* Paper presented at the annual meeting of the American Educational Research Association Meeting, Chicago, IL.

Corbett, A.T., Anderson, J.R., & Patterson, E.G. (1990). Student Modeling and Tutoring Flexibility in the Lisp Intelligent Tuturing System. In Frasson, C., & Gauthier, G. (Eds.) *Intelligent Tutoring Systems: At the Crossroads of Artificial Intelligence and Education.* Norwood, NJ: Ablex Publishing

Dempsey, J.V., Driscoll, M.P., & Swindell, L.K. (1993). Text-based feedback. In Dempsey, J.V. & Sales, G.C. (Eds.) *Interactive Instruction and Feedback.* Englewood Cliffs, NJ: Educational Technology.

Farquhar, J.D. (1993). *LOADER: A simulation tool for the study of console operation skill.* [software]. San Antonio: TX. Air Force Armstrong Laboratory, Brooks AFB.

Fox, B.A. (1991). Cognitive and interactional aspects of correction in tutoring. In P. Goodyear (Ed.), *Teaching knowledge and intelligent tutoring*. Norwood, NJ: Ablex

Gagné, R.M. (1984). *The Conditions of Learning and Theory of Instruction.* NY, NY: Holt, Rinehart and Winston

Gaynor, P. (1981). The effect of feedback delay on retention of computer-based mathematical material. *Journal of Computer-Based Instruction, 8*(2), 28-34.

Grinder, R.E. & Nelson, E.A. (1985). Individualized instruction in american pedagogy: The saga of an educational ideology and a practice in the making. In Wang, M.C. & Walberg H.J. (Eds.). *Adapting Instruction to Individual Differences*. Berkeley, CA: McCutchen.

Heald, H.M. (1970). The effects of immediate knowledge of results and correlation of errors and text anxiety upon test performance. *Dissertation Abstracts International*, 31, 1621A.

Keppel, G. (1991). *Design and Analysis: A Researcher's Handbook.* Englewood Cliffs, NJ: Prentice-Hall.

Kulhavy, R.W. & Anderson, R.C. (1972). Delay-retention effect with multiple-choice tests. *Journal of Educational Psychology, 63*(5), 505-512.

Kulik, J.A., & Kulik, C.C. (1987). Review of recent research literature on computer-based instruction. *Contemporary Educational Psychology, 12*, 222-230.

Kulik, R.W. & Stock, W.A. (1989). Feedback in written instructon: The place of response certitude. *Educational Psychology Review, 1*(4), 279-308.

Lee, T.D. & Carnahan, H. (1990). When to provide knowledge of results during motor learning: Scheduling effects. *Human Performance, 3*(2), 87-105.

Legree, P.J., & Gillis, P.D. (1991). Product effectiveness evaluation criteria for intelligent tutoring systems. *Journal of Computer-Based Instruction, 18*(2), 57-62.

Legree, P.J., Gillis, P.D., & Orey, M.A. (1993). The quantitative evaluation of intelligent tutoring system applications: Product and process criteria. *Journal of Artificial Intelligence and Education, 4*(2/3), 209-226.

Littman, D. (1991). Tutorial planning schemas. In P. Goodyear (Ed.), *Teaching knowledge and intelligent tutoring*. Norwood, NJ: Ablex.

Littman, D., Pinto, J., & Soloway, E. (1990). The knowlege required for tutorial planning: An empirical analysis. *Interactive Learning Environments, 1,* 124-151.

Littman, D. & Soloway, E. (1988). Evaluating ITSs: The cognitive science perspective. In Polson, M.C. & Richardson J.J. (Eds.). *Foundations of intelligent tutoring systems*. (pp. 209-242) Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Mark, M.A. & Greer, J.E. (1993). Evaluation methodologies for intelligent tutoring systems. *Journal of Artificial Intelligence in Education, 4*(2/3), 209-226.

McArthur, D., Stasz, C., & Zmuidzinas, M. (1990). Tutoring techniques in algebra. *Cognition and Instruction, 7*, 197-244.

Merrill, D.C., Reiser, B.J., Ranney, M., & Trafton, J.G. (1992). Effective tutoring techniques: A comparison of human tutors and intelligent tutoring systems. *The Journal of the Learning Sciences, 2*(3), 277-305.

Mory, E.H. (1992). The use of informational feedback in instruction: Implications for future research. *Educational Technology Research and Development, 40*(3), 5-20.

Murray, T. (1993). Formative qualitative evaluation for "exploratory" ITS research. *Journal of Artificial Intelligence and Education, 4*(2/3), 179-207.

Roper, W.J. (1977). Feedback in computer-assisted instruction. *Programmed Learning and Educational Technology, 14*, 43-49.

Salmoni, A.W., Schmidt, R.A., & Walter, C.B. (1984). Knowledge of results and motor learning: A review and critical reappraisal. *Psychological Bulletin, 95*(3), 355-386.

Sassenrath, J.M. & Gaverick, C.M. (1965). Effects of differential feedback from examinations on retention and transfer. *Journal of Educational Psychology, 56*(5), 259-263.

Schimmel, B.J. (1983). *A meta-analysis of feedback to learners in computerized and programmed-instruction*. Paper presented at the annual meeting of the American Educational Researdh Association, Montreal. (ERIC Document Reproduction Service No. 233 708).

Schooler, L.J. & Anderson, J.R. (1990). The disruptive potential of immediate feedback. *The Proceedings of the Twelth Annual Conference of the Cognitive Science Society*, Cambridge, MA.

Skinner, B. F. (1968). *The Technology of Teaching*. NY, NY: Appleton-Century-Crofts

Shute, V.J. & Regian, J.W. (1993). Principles for evaluating intelligent tutoring systems. *Journal of Artificial Intelligence and Education, 4*(2/3), 245-271.

Winne, P.H. (1993). A landscape of issues in evaluating adaptive learning systems. *Journal of Artificial Intelligence and Education, 4*(2/3), 273-290.

Wulf, G. & Schmidt, R.A. (1989). The learning of generalized motor programs: Reducing the relative frequency of knowledge of results enhances memory. *Journal of Experimental Pscyhology: Learning, Memory, and Cognition, 15*(4), 748-757.