THEORETICAL IMPLICATIONS AND EMPIRICAL FINDINGS ON INSTRUCTIONAL CONTROL AND PART-WHOLE-TASK TRAINING

Joseph S. Mattoon
Bernell J. Edwards

HUMAN RESOURCES DIRECTORATE
AIRCrew TRAINING RESEARCH DIVISION
6001 S. Power Road, Bldg 558
Mesa, AZ 85206-0904

June 1993

Approved for public release; distribution is unlimited.
NOTICES

This technical report is published as received and has not been edited by the technical editing staff of the Armstrong Laboratory.

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Office of Public Affairs has reviewed this report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This report has been reviewed and is approved for publication.

JOSEPH S. MATTOON
Project Scientist

DEE H. ANDREWS
Technical Director

LAWN A. CARROLL, Colonel, USAF
Chief, Aircrew Training Research Division
Theoretical Implications and Empirical Findings on Instructional Control and Part-Whole-Task Training

Joseph S. Mattoon
Bernell J. Edwards

Armstrong Laboratory (AFMC)
Human Resources Directorate
Aircrew Training Research Division
6001 S. Power Road, Bldg 558
Mesa, AZ 85206-0904

Armstrong Laboratory Technical Monitor: Mr. Joseph S. Mattoon, (602) 988-6561

Approved for public release; distribution is unlimited.

A review of the literature and two experiments were conducted to examine instructional control and part-whole-task training methods applied to simulation training. The literature and results of the experiments indicate that learner-controlled systems are probably less effective than those that use computer programming to automatically adjust instructional support according to learners' progress during practice activities. Also, part-task training methods may facilitate the effectiveness of instructional control by reducing the complexity of performance measures and en route learning tasks.
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Computer-Assisted Instructional Simulation</td>
<td>1</td>
</tr>
<tr>
<td>Instructional Control: Theory and Results of Previous Studies</td>
<td>3</td>
</tr>
<tr>
<td>THE EXPERIMENTS</td>
<td>5</td>
</tr>
<tr>
<td>Criterion Task</td>
<td>6</td>
</tr>
<tr>
<td>Instructional Control and Task Difficulty: Controlling Levels of Challenge</td>
<td>7</td>
</tr>
<tr>
<td>Experiment 1: Learner Versus Program Control</td>
<td>7</td>
</tr>
<tr>
<td>Method</td>
<td>8</td>
</tr>
<tr>
<td>Subjects</td>
<td>8</td>
</tr>
<tr>
<td>Materials</td>
<td>8</td>
</tr>
<tr>
<td>Procedure</td>
<td>12</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>12</td>
</tr>
<tr>
<td>Results</td>
<td>13</td>
</tr>
<tr>
<td>Performance on Immediate and Delayed Posttests</td>
<td>13</td>
</tr>
<tr>
<td>Practice Behavior</td>
<td>14</td>
</tr>
<tr>
<td>Attitudes</td>
<td>15</td>
</tr>
<tr>
<td>Discussion</td>
<td>16</td>
</tr>
<tr>
<td>Effects of Instructional Control</td>
<td>16</td>
</tr>
<tr>
<td>Effects of Task Complexity</td>
<td>17</td>
</tr>
<tr>
<td>Attitudes Toward the Simulation Training</td>
<td>17</td>
</tr>
<tr>
<td>Experiment 2: Learner Control and Part-Whole-Task Training</td>
<td>18</td>
</tr>
<tr>
<td>Part-Whole-Task Training: Theory and Results of Previous Research</td>
<td>18</td>
</tr>
<tr>
<td>Method</td>
<td>20</td>
</tr>
<tr>
<td>Subjects</td>
<td>20</td>
</tr>
</tbody>
</table>
PREFACE

This report documents the results of a literature review and two experiments completed on specific training design issues relative to the training of Air Force aircrews using microcomputer-based training. The experiments reported herein represent a portion of the research and development (R&D) program of the Armstrong Laboratory (AL), the thrust of which is air combat training research. The general objective is to identify and demonstrate cost-effective methods and media to support the development of training materials and devices for Air Force aircrew members. The purpose of the present effort was to investigate instructional control and part-whole-task training techniques as potential enhancements for computer-based Air Force air combat training programs. The work was conducted under Work Unit 1123-25-15, by the Unit-Level Training Research Applications (ULTRA) group. The work unit monitor was Dr. Bernell J. Edwards, and the principal investigator was Mr. Joseph S. Mattoon.
THEORETICAL IMPLICATIONS AND EMPIRICAL FINDINGS ON INSTRUCTIONAL CONTROL AND PART-WHOLE-TASK TRAINING

INTRODUCTION

The Unit Level Training Research Applications group at the Aircrew Training Research Division of Armstrong Laboratory has focused on research issues concerning the design of instructional interfaces between student pilots and simulation training environments. Simulation is a practical alternative to using actual equipment which is frequently too expensive, too dangerous, or not available for training purposes. Burton and Brown (1979) predicted that simulation would have a significant impact on future education and technical training environments. It appears that this assumption is now being verified in government and private industry (Gray & Edwards, 1991; Mattoon & Thurman, 1990). In addressing the issues of simulation training, the initial goal is to provide developers with general rules and guidelines that are immediately applicable to aircrew training programs. The research strategy was to employ undergraduate student populations as subjects, and examine specific training problems that are relevant to Air Force unit-level training requirements.

The undergraduate subject pool facilitated formative evaluation of the experimental procedures and materials. The advantages of this approach are experimental control of the environment, data collection convenience, a larger number of subjects per treatment, and time/cost savings. A possible disadvantage is that undergraduate students that participated in the experiments may have possessed a different degree of motivation or may have different prior knowledge and skills than Air Force student pilots. However, instructional control and part-whole-task training have been shown to exhibit similar effects on knowledge and skill acquisition across a variety of adult learners, so it is likely that the present experimental findings can be generalized to student pilot populations. A second phase of the present research will take place in the operational training environment to verify this assumption if the laboratory experiments warrant field testing.

Computer-Assisted Instructional Simulation

Conventional computer-assisted instruction (CAI) is commonly used in schools and colleges to teach skills that primarily consist of declarative knowledge which is demonstrated by one's ability to provide verbal answers to test questions. Computer-assisted instructional simulation (CAIS) may be viewed as a hybrid form of CAI which
is especially suited for teaching the type of complex, dynamic skills that are essential to the operation of sophisticated electronic equipment within aviation environments.

The type of instructional materials and skills taught by CAIS differ from those taught by conventional CAI (Borsook & Higginbotham-Wheat, 1991). In conventional CAI, learners typically acquire knowledge by reading text material, and their learning progress and achievement can usually be measured by multiple-choice questions. Learning objectives are based on declarative knowledge (Munro, Fehling, & Towne, 1985). In contrast, training simulations graphically reproduce dynamic environments, and learners are taught complex skills by manipulating simulated objects within these environments (Mattoon & Thurman, 1990).

CAIS is more suited for teaching complex tasks that involve the simultaneous execution of multiple subtasks (parts of the whole task) to accomplish a single goal. The skills necessary to master such tasks consist of a combination of intellectual, perceptual, and motor abilities (Schneider, 1985). For example, student pilots can be taught how to interpret radar displays using CAIS. Students monitor the dynamic flow of information shown on simulated displays and arrange graphic objects on a computer screen to represent their interpretations of a variety of in-flight situations. The CAIS monitors students' responses, compares them to the display settings, and provides feedback on the accuracy of their interpretations and the speed of their actions. The skill objectives in this type of learning environment are made up of a complex of procedural knowledge, intellectual skill, and cognitive strategies applied to specific operations or mission goals.

Flight simulators have traditionally been designed with an emphasis on faithfully representing the flight environment rather than focusing on the effective learning and retention of skill (Allesi, 1988; Andrews, 1988). Consequently, developers of CAIS may overlook the importance of instructional events and materials under the assumption that simulations that closely resemble the real cockpit environment will automatically produce effective learning. However, the lack of instructional support within a training system will likely increase the time learners need to acquire skills and may also hinder retention of skills over time. Instructional support is "... a set of events external to the learner which are designed to support the internal processes of learning" (Gagné & Briggs, 1979, p. 155). Thus, training components such as performance feedback, practice on rule memorization, and the teaching of
strategies are all forms of instructional support. These events are core components of all effective educational and training efforts. Without instructional support, students must either painstakingly correct their own errors and misconceptions or rely on the assistance of instructors, thus, defeating the purpose of automated training.

Instructional Control: Theory and Results of Previous Studies

Microcomputers provide an opportunity to control instructional support using a variety of strategies. The goal is to control instructional materials (e.g., pace, feedback, and practice) in a way that fulfills the unique learning needs of each individual. There are two basic types of strategies, learner control and program control. Learner-controlled strategies enable the individual to alter the type and amount of instruction presented by the program. In contrast, program-controlled strategies deliver a predetermined body of content information and practice activities (fixed program control), or regulate instruction based on each individual’s performance (adaptive program control). Fixed program control presents the same sequence of instructional events for all learners (Merrill, 1984), while adaptive programs adjust the type and amount of instructional material based on each individual’s initial ability and ongoing performance improvement during practice (Borsook & Higginbotham-Wheat, 1991). Research has produced mixed findings concerning the effects of each type of control strategy on learners’ achievement and motivation (Steinberg, 1977; 1989).

Instructional control has been found to affect learners’ achievement of knowledge and skill (Carrier, Davidson, & Williams, 1985; Gay, 1986; Goetzfried & Hannafin, 1985; Munro, Fehling, & Towne, 1985; Ross & Rakow, 1981), their management of study time (Johansen & Tennyson, 1983; Tennyson, 1981, Tennyson & Buttrey, 1980), and their motivation and intrinsic interest in the topic material (Kinzie & Sullivan, 1989; Kinzie, Sullivan, & Berdel, 1988). The most important factors identified by research on learner control are associated with learners’ ability to make effective control decisions, the time and effort they expend on instructional activities, and the information made available to learners for the purpose of facilitating their control decisions (Steinberg, 1989).

Because of the high degree of motivation and level of achievement among student pilots, learner-controlled CAI may appear to be a positive component of training success in aircrew training. However, empirical evidence from a broad population of young and adult learners indicates that program-controlled systems
probably have greater advantages and are more effective overall (Steinberg, 1977, 1989). Because of the additional load on working memory during the performance of complex tasks, program-controlled strategies, which preclude the need for trainees to engage in additional decision-making tasks during learning, may be especially important to the design of effective CAIS.

Ross and Rakow (1981) demonstrated that adaptive program-controlled instruction enabled undergraduates to achieve higher math posttest scores than those who completed either a learner-controlled or fixed program-controlled version of the instruction. Goetzfried and Hannafin (1985) found that program-controlled CAI was more effective even when subjects were provided with advisement on each control decision. These studies emphasized learners' lack of ability to make good control decisions when acquiring initial knowledge and skills.

Some theorists suggest that the freedom to control instruction increases learners' intrinsic motivation for instruction (Laurillard, 1987; Lepper, 1985; Milheim & Martin, 1991). However, even studies that show that learners preferred learner- to program-controlled instruction have not been able to demonstrate higher achievement via learner control (Kinzie & Sullivan, 1989; Kinzie, Sullivan, & Berdel, 1988; Lahey, Hurlock, & McCann, 1973). Other studies have failed to detect a difference in motivation for the two instructional control methods (Klein, 1988), and some have shown that learners preferred program control (Gray, 1987).

Researchers have attempted to improve learner-controlled strategies by designing programs that present evaluative information and adaptive advisement to learners to assist them in making instructional control decisions during a learning activity. Evaluative information refers performance data that is summarized by statistical computations (e.g., number of errors and mean response time) and describes an individual's progress toward specific objectives. Adaptive advisement refers to a recommendation (delivered by the program) to study specific instructional content or complete particular practice activities (Santiago & Okey, 1992).

Results of CAI studies indicate that learners have difficulty in simultaneously meeting the challenge of acquiring new knowledge and skill and determining the type and amount of instruction they need even with the assistance of advisement (Goetzfried & Hannafin, 1985). Evaluative advisement does not appear to be sufficient for guiding learners' decisions (Santiago & Okey, 1992), and adaptive advisement
begs the question—instead of telling the learner what to do, why not simply control the instruction directly via program control? Additionally, in CAIS, learners are often engaged in dynamic tasks whereby the presentation of evaluative information or adaptive advisement can disrupt learning (Munro, Fehling, & Towne, 1985).

Tennyson and associates have claimed that learner-controlled CAI that provides continuously updated advisement on each individual's ability based on pretest and practice performance can be more effective than program control. This notion is based on several studies that indicated learner-controlled subjects were able to complete instructional activities in less time (prior to a posttest) than those under program control (Tennyson, 1980, 1981; Tennyson & Buttrey, 1980). Yet, none of these studies showed a higher rate of achievement for subjects under learner control with advisement compared to those under program control. In these studies, only subjects who received specific directions on which instructional options to choose (via program-controlled advisement) were able to achieve posttest scores that were statistically equivalent to those achieved by program-controlled subjects.

Tennyson and associates have suggested that subjects' personal knowledge and strategies improved on the program-generated advisement. However, if the instructional materials provided by the program-controlled treatments were equivalent to those identified and recommended by the advisement, learner-controlled subjects must have chosen to disregard some of the advice. Unfortunately, the frequency that subjects followed advice and the time they spent on advisement portions of the programs were not reported.

THE EXPERIMENTS

The underlying rationale for the two present experiments was to make preliminary assessments of training methods to examine their potential value to the Air Force prior to committing time and resources to full-scale development and evaluation. The experiments represent a first phase of research that was completed for the purpose of guiding more in-depth and applied field studies.

The two present experiments examined adaptive learner- and program-controlled strategies applied to CAIS. In the first experiment (Mattoon & Klein, in press), three types of control strategies were compared for regulating challenge (task difficulty) during practice on a complex display-interpretation task. The second experiment (Mattoon, 1992) compared part- and whole-task training methods. The two
experiments will be described and results reported in the following sequence:

1. The criterion task and its relevance to aircrew training will be discussed.
2. A brief review of the literature that is directly relevant to the independent and dependent variables of experiment 1 will be presented.
3. The method and procedures for experiment 1 will be described.
4. The results of experiment 1 will be presented and discussed.
5. A brief review of the literature that is directly relevant to the independent and dependent variables of experiment 2 will be presented.
6. The method and procedures for experiment 2 will be described.
7. The results of experiment 2 will be presented and discussed.
8. General conclusions of the literature review and results of experiment 1 and 2 will be presented.

Criterion Task

A fighter pilot’s ability to intercept and defeat enemy aircraft requires the individual to detect and track the advancement of an adversary (“target aircraft”) via radar from a distance of up to 40 or 50 miles. The pilot must correctly interpret information on the radar display in order to know the direction, speed, and course of the target, and subsequently, to decide when and how to maneuver to overtake the adversary. A substantial part of this air-intercept task is accomplished by visually monitoring cockpit instruments while the target is well beyond the pilot’s unaided vision. These instruments display words, numbers, and symbols that indicate the state of the target aircraft and surrounding airspace. Thus, a pilot’s ability to quickly and accurately interpret these displays is essential. Most of the symbology used in radar displays is highly abstract, nonintuitive, and dynamic in nature. Radar interpretation skills must become automatic and must be integrated with the “finger skills” (switchology) so that pilots are able to execute control actions in coordination with display-interpretive skills.

Interpretation of a simplified version of one of the most common radar displays was chosen as the criterion task for the present experiments. The head-up display (HUD) provides digital and analog information that is used to identify the location and direction of flight of target aircraft relative to the pilot’s aircraft. Lahey and Boyer (1990) demonstrated that subjects with no previous training could learn how to interpret the HUD by practicing with a CAIS program. However, subjects required about 400
whole-task practice exercises to attain the criterion performance.

Target estimation with the HUD is a complex and challenging task because it simultaneously presents three different target measures that consist of both digital and analog information which must be interpreted as spatial information. Additionally, the task of estimating the location of target aircraft and determining their direction of flight is a highly integrated task, because pilots' estimation of one parameter (e.g., location) can affect the accuracy of their estimation of another (e.g., direction of flight).

Instructional Control and Task Difficulty: Controlling Levels of Challenge

In referring to practice and test questions in written instruction, Klein (1988) states that "... learners may not be the best judges when given control over item difficulty" (p. 25). Task difficulty is closely associated with degree of effort expended on attaining learning objectives. Steinberg's (1989) review of instructional control literature indicates that learners tend to expend less effort than is needed to attain objectives when given control of practice activities. This indicates that learners, who are given control over task difficulty, tend to set difficulty levels that are too low for optimal skill development. In contrast, Johansen and Tennyson (1983) and Tennyson (1980) have demonstrated that learners can effectively make control decisions and be induced to expend an appropriate amount of effort and time on practice activities if they are provided with evaluative and adaptive advisement.

Because practice is a major component of complex skill training, and learners' speed and accuracy increases with practice, a common training method is to increase the level of challenge (task difficulty) as performance improves (Gropper, 1983). This provides learners with goals that are appropriate for their current level of skill. Relaxing the time pressures associated with performing a dynamic task (at the criterion level) enables learners to closely examine a complex task and develop an understanding of interactions among its components prior to attempting to execute the task at the criterion level (Mané, Adams, & Donchin, 1989). The control over levels of challenge by adjusting criteria for speed and accuracy was addressed in experiment 1.

Experiment 1: Learner Versus Program Control

The first experiment compared three different instructional control strategies for altering the degree of challenge. Challenge was controlled by changing the criteria by which subjects earned (or lost) points for speed and accuracy during practice on
estimating location and direction of flight of simulated target aircraft with a simplified version of the HUD. Three different adaptive CAIS treatments for instructional control were examined: (a) learner control (LC), (b) learner control with advisement (LCA), and (c) program control (PC). LC subjects were free to increase or decrease level of challenge after each target estimate. LCA subjects were also in control of challenge level but were advised to set challenge at specified levels based on their performance on practice exercises. The program automatically set levels of challenge for PC subjects. All subjects were given cumulative feedback on their speed and accuracy throughout practice. Subjects completed an immediate and a one-week delayed posttest.

The following results were expected:

1. Subjects would lose some of the skill they initially acquired for performing the criterion task after a week without practice.
2. LC subjects would set lower levels of challenge during practice than levels chosen by LCA subjects or levels of challenge assigned to PC subjects.
3. The LCA and PC groups would perform better on the posttests than the LC group.

Method

Subjects

Subjects were 78 undergraduates enrolled in an educational psychology course at a large public university in the southwestern United States. They participated in the experiment for partial course credit. Performance data were initially collected from 12 subjects (four per treatment) who participated in a pilot experiment. The remaining 64 (12 males and 52 females) participated in the main experiment.

Materials

A CAIS program was developed to run on Macintosh computers with 12-inch screens. Lesson and practice activities were composed of monochromatic graphics, text, and interactive displays that subjects controlled with a mouse. The instruction and content material were validated in two ways. First, the criterion task was analyzed and defined in a previous study via interviews with subject matter experts (Lahey & Boyer, 1990). Second, subject matter experts served as consultants during formative evaluation of the present materials. Their comments and suggestions guided revisions in the program. En route learning tasks were self-paced. For example,
subjects were able to move forward or back through the content material at their own pace and as many times as they wanted prior to the immediate posttest.

The instruction portion of the program presented the concepts, rules, and procedures for estimating target aircraft on the "overhead" (a two-dimensional plane that represented the airspace scanned by the pilot's radar system). The lesson explained how to interpret three target parameters displayed on the HUD:

1. Range was represented as the distance (nautical miles) of the target from the pilot's aircraft and was displayed as a whole number from zero to 40 on the HUD.

2. Azimuth was the angular distance of the target's location to the right or left of the pilot's aircraft. It was displayed as a whole number from zero to 60, and an arrow appeared above the azimuth value to indicate the direction (right or left) of the azimuth.

3. Aspect was the orientation of the target relative to the pilot's aircraft. Aspect was provided by the position of a triangular-shaped symbol along the "aiming reticle" (circle) on the HUD.

An interactive practice exercise was developed for practicing the target-location task. It included 75 practice exercises. Two symbols shaped like aircraft, the "locator" and the "target," played key roles in the simulation. The locator was positioned on the screen by subjects to estimate target location and rotated to estimate its direction of flight. Two "rotate buttons," shown as circular arrows on the screen, were activated by the mouse to rotate the locator. The target symbol appeared after each estimate to show the correct target location and direction of flight. Figure 1 shows the locator and target symbols, rotate buttons, and a target setting on the HUD that has a range of 25 miles, an azimuth of 30 degrees right, and a left wing aspect (target's left wing is facing the pilot's aircraft).

The program instructed subjects to position the locator on the screen, then activate a "continue" button which displayed the target symbol at the correct location and orientation, points for speed and accuracy, and two linear scales that indicated the current challenge level. The scales were labeled "speed meter" and "accuracy meter." Activating the continue button a second time began a new practice exercise by erasing the target, points, and scales from the screen, displaying new target parameters on the HUD, and resetting the locator to the home position. The feedback screen, including the target, speed and accuracy points, and speed and accuracy meters is shown in Figure 2.

9
Azimuth = 30

Aspect pointer indicates target is flying from right to left relative to the pilot's aircraft.

Range = 25

The two posttests were essentially the same as the practice activity, but challenge levels and feedback were excluded. No targets used in practice exercises were repeated on the test. Attitudes of subjects and their judgments of the effectiveness of the training were measured by bipolar statements that expressed an opinion about the instruction or criterion task. Subjects rated how much they agreed or disagreed with each statement by moving a pointer on the screen toward the left end (labeled "agree") or right end (labeled "disagree") of a scale.

After each target estimate, LC subjects could adjust the level of challenge for speed and accuracy via the speed and accuracy meters. The LCA treatment provided the same control, but subjects received advice for setting challenge levels on every fifth exercise. The advice was based on the individual's cumulative speed and accuracy points across the most recent five targets. For example, if points for speed had been earned on the last five targets, the individual was advised to increase speed challenge to a specific level that corresponded to the proportion of points earned. If the subject had lost points, advice was given for a reduction in challenge level according to the proportion of points lost. The same calculations and decision rules...
were employed in the PC treatment, but they resulted in an automatic adjustment in challenge levels every fifth target estimate.

A change in accuracy challenge corresponded to an increase or decrease in error tolerance for target estimates. Points were added to an individual's score if the error in their estimate was within the tolerance level, and points were subtracted if the error exceeded this level. Speed challenge adjusted the time allowed to estimate the target. The number of accuracy points gained or lost on each estimate was proportional to the percentage of error on each of the three target parameters. The number of speed points was computed according to the proportion of current time allowed that was used to estimate the target. Additionally, increasing challenge increased the number of points possible (gained or lost) per target estimate. Thus, an increase in challenge was associated with two factors: (a) an increased risk in point

Figure 2. Feedback with Target, Locator, Speed and Accuracy Points and Meters.
loss when accuracy or speed was below criterion and (b) an increased potential for point gain.

**Procedure**

Results of the pilot experiment were used to set maximum and minimum levels of challenge for response time criterion (speed) and accuracy criteria. An equal number of subjects were assigned to each treatment based on the order they arrived at the computer site. They participated in groups of 12 to 21 and completed (a) a lesson and skill test for using the mouse, (b) a lesson on interpreting the HUD, (c) practice exercises, (d) an immediate test, and (e) an attitude questionnaire. Subjects were instructed to “challenge themselves” and attain the highest score they could during practice and posttest activities. They were also urged to pay close attention to the feedback during practice and use it to help them improve their performance.

Subjects returned one week later and completed the delayed test. The data for five subjects were eliminated from the analyses, four due to technical problems with the software and one due to failure to complete the delayed test.

**Data Analysis**

The experiment was a two-variable mixed design with repeated measures on test occasion. It had three levels of instructional control (LC, LCA, and PC) and two levels of test occasion (immediate vs. delayed). Multivariate analyses of variance (MANOVAs) were followed by univariate analyses of variance (ANOVA) on response time, accuracy, challenge levels, speed and accuracy points, number of changes in challenge levels, and responses to the attitude statements. Student-Newman-Keuls procedures were used to test differences in treatment group means. The significance level was set at $p < .05$ for all multivariate and univariate tests.

Measures of speed and accuracy were averaged across posttest items. Subjects’ response speed on the mouse skill test described subjects’ initial motor skill for using a mouse and was used as a covariate on all analyses of variance to partition initial motor skill from treatment effects. Challenge levels were described in percent and ranged from zero (minimum challenge) to 100% (maximum challenge). The values for responses to attitude statements ranged from -1.0 (strong disagreement) to 1.0 (strong agreement).
Results

Performance on Immediate and Delayed Posttests

Subjects practiced the target-estimation task about one hour. A positive correlation for response time on the mouse skill test with speed and accuracy on posttests was detected at the p < .05 level. An analysis of response time and accuracy revealed a main effect for test occasion (multivariate F(4, 55) = 6.02, p < .001). An ANOVA on response time detected a treatment by test occasion interaction (F(2, 58) = 4.95, p < .05). Table 1 displays the means for response times for the mouse test, practice on target estimation, and the immediate and delayed posttests.

Table 1. Mean Response Times: Covariate, Practice, and Posttests.

<table>
<thead>
<tr>
<th>Group</th>
<th>Covariate</th>
<th>Practice</th>
<th>Immediate</th>
<th>Delayed</th>
</tr>
</thead>
<tbody>
<tr>
<td>LCA</td>
<td>11.7</td>
<td>16.0</td>
<td>14.2</td>
<td>15.9</td>
</tr>
<tr>
<td>PC</td>
<td>11.9</td>
<td>15.3</td>
<td>14.5</td>
<td>14.5</td>
</tr>
<tr>
<td>LC</td>
<td>12.6</td>
<td>15.3</td>
<td>12.8</td>
<td>14.4</td>
</tr>
</tbody>
</table>

LCA = learner control with advisement
PC = computer control
LC = learner control without advisement

An analysis of simple effects was performed to examine the differences in response time. LC subjects' response time was faster on the immediate test than on the delayed test (M = 12.8 vs. M = 14.4 seconds, t(38) = 4.64, p < .05). The pattern was similar for LCA subjects but did not reach the assigned significance level (M = 14.2 vs. M = 15.9 seconds, t(38) = 2.81, p = 10). PC subjects appeared to execute the task at the same speed on both posttests (M = 14.5). No other significant differences in test performance were detected.
**Practice Behavior**

Significant treatment effects were detected for levels of accuracy challenge, frequency that challenge levels were altered, and accuracy points. An analysis of challenge levels and points revealed a significant effect for treatment (multivariate $F(4, 116) = 4.48, p < .05$). An ANOVA conducted on each factor revealed a treatment effect for accuracy challenge ($F(2, 58) = 10.60, p < .001$) and for accuracy points ($F(2, 58) = 5.90, p < .01$). Tests between means showed that LC subjects set lower levels of accuracy challenge ($M = 42\%$) than LCA and PC subjects ($M = 60\%$ and $M = 64\%$) but averaged more accuracy points ($M = 53.7$) than LCA and PC subjects ($M = 35.6$ and $M = 30.8$). An examination of trends across practice exercises showed that LC subjects raised accuracy challenge for about the first 15 exercises but stopped at about 42% of the maximum level. However, accuracy challenge for LCA and PC treatments continued to increase throughout the 75 exercises.

A MANOVA revealed a significant effect for the frequency of negative scores for speed and accuracy points (multivariate $F(4, 116) = 12.93, p < .001$). Negative scores were the number of points that were subtracted from subjects' cumulative score each time a target estimate did not meet either accuracy or response time criteria. Separate ANOVAs revealed a treatment effect for accuracy points ($F(2, 58) = 39.32, p < .001$) and for speed points ($F(2, 58) = 20.29, p < .001$). Tests between group means showed that LC subjects received negative accuracy and speed scores on about half as many practice items ($M = 23\%$) as LCA ($M = 47\%$) and PC ($M = 51\%$) subjects ($p < .05$). Similarly, LC subjects received less negative speed scores ($33\%$) than LCA ($54\%$) and PC ($55\%$) subjects ($p < .05$).

An analysis of the proportion of challenge adjustments on speed and accuracy across practice items revealed a significant effect for treatment (multivariate $F(4, 36) = 5.61, p < .001$). Univariate tests indicated that the LCA subjects altered challenge on more of the practice exercises than LC subjects for accuracy ($M = 19\%$ vs. $M = 9\%$, $t(39) = 23.4, p < .001$) and for speed ($M = 13\%$ vs. $M = 7\%$, $t(39) = 10.71, p < .01$). (PC subjects' level of challenge was adjusted at the end of every fifth exercise, so speed and accuracy challenge was altered on 19% of the 75 exercises.) The frequency that LCA subjects followed the challenge advice was 81% for speed and 73% for accuracy.
Table 2 shows the challenge levels and proportion of negative scores across practice exercises by treatment.

Table 2: Challenge Levels and Negative Scores

<table>
<thead>
<tr>
<th>Group</th>
<th>Speed</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Chal</td>
<td>Chal</td>
<td>Neg</td>
<td>Chal</td>
<td>Chal</td>
<td>Neg</td>
<td>Level</td>
<td>Chang</td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>LCA</td>
<td>28%</td>
<td>13%</td>
<td>54%</td>
<td>60%</td>
<td>19%</td>
<td>47%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC</td>
<td>30%</td>
<td>19%</td>
<td>55%</td>
<td>64%</td>
<td>19%</td>
<td>51%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LC</td>
<td>25%</td>
<td>7%</td>
<td>33%</td>
<td>42%</td>
<td>9%</td>
<td>23%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Chal Level = challenge level across exercises
Chal Chang = % exercises received change in challenge
Neg Score = % exercises received negative score

Attitudes

Treatment effects for attitudes should be viewed with caution because the initial MANOVA, which included each statement as a separate dependent measure, did not reveal multivariate effects at the assigned significance level. Individual ANOVAs conducted on each statement indicated that treatments appeared to alter subjects' attitudes toward three statements:

1. "The instruction provided me with a good understanding of how to locate targets..." ($F(2, 58 = 3.37, p < .05$). Tests between means showed that LC subjects were less in agreement with this statement than PC subjects ($M = 0.24$ vs. $M = 0.54$, $p < .05$). The means for PC and LCA groups were almost identical ($M = 0.54$ and $M = 0.55$).
2. "I would want the level of challenge for speed controlled by the computer" \((E(2, 58) = 4.00, \ p < .05)\). PC subjects tended to agree with this statement \((M = 0.20)\) while LC and LCA subjects tended to disagree \((M = -0.27 \text{ and } M = -0.021, \ p < .05)\).

3. The ratings for the corresponding statement, "I would want the level of challenge for accuracy controlled by the computer," showed the same pattern, with PC subjects tending toward agreement \((M = 0.11)\) and LCA \((M = -0.05)\) and LC \((M = -0.19)\) tending toward disagreement, but these differences failed to reach the \(p < .05\) level.

Discussion

Effects of Instructional Control

The difference in response time between the immediate and delayed posttests for LC and LCA subjects was not surprising. The longer response time on the delayed posttest exhibited by LC and LCA subjects, observations by the experimenter, and comments made by subjects during the second experimental session indicated that they had forgotten some of the concepts and rules that were essential to the criterion task. The interference of the additional task of setting challenge levels may account for this loss in skill. The consistent response time exhibited by PC subjects across the retention interval supports this assumption, because they did not have to perform challenge-control tasks. Apparently, PC subjects were better able to retain their skill due to the absence of interference created by the learner control tasks.

Challenge levels probably influenced both practice behavior and the subsequent test performance. Speed challenge level did not fluctuate much during practice. It remained at 25% to 30% across most of the practice exercises, but accuracy challenge varied considerably between treatment groups.

Observations by the experimenter indicated that LC subjects were at first confused by the task of setting challenge levels. They experimented with various levels for both speed and accuracy on the first few targets, but after the 10th target, LC subjects tended to leave accuracy challenge at about the 42% level for the remaining 65 exercises. In contrast, the accuracy challenge for PC and LCA subjects continued to increase over the 75 practice trials. PC and LCA subjects practiced estimating targets at higher levels of challenge, and this caused more frequent negative feedback compared to the LC group. This may have induced PC and LCA subjects to estimate targets more slowly on the two posttests.
Effects of Task Complexity

Task complexity is determined by the information processing demands imposed on learners by subtasks that make up the criterion task (Briggs & Naylor, 1962; Naylor, 1962). Logan (1985) indicates that subtasks which engage the same cognitive resources interfere with the development of automaticity when performed simultaneously during practice. In this view, the amount of skill acquisition is influenced by the number of subtasks and the degree to which they engage the same cognitive resources.

The decisions and actions that were necessary for the LCA and LC groups to control challenge levels can be viewed as additional subtasks that required some of the same cognitive resources used in practicing the criterion task. Even though subjects did not make control decisions and estimate targets simultaneously, controlling challenge levels probably taxed working memory which was needed for important en route tasks (e.g., using feedback to correct errors and misconceptions about the task). Schneider (1985) points out that a learner’s performance, immediately following a practice activity, can be a false indicator of learning. That is, actual learning deficits may not be observable until after a temporal delay. The slower time exhibited by LC and LCA subjects on the delayed test indicates that some of the concepts and rules of the criterion task were forgotten during the retention interval. In contrast, PC subjects did not exhibit this deficit.

Attitudes Toward the Simulation Training

LC subjects probably perceived the instruction as being less effective in "providing a good understanding" of the task, because they received no specific direction on how to adjust challenge. LCA subjects received direct advisement, and PC subjects did not have to make any control decisions. This would account for their more positive judgment of training effectiveness.

Subjects seemed to be most in favor of the type of challenge control they were assigned. Their lack of familiarity with the subject matter and the instructional medium probably made it difficult for them to conceptualize a different form of control. Additionally, the lack of interest in controlling challenge that was exhibited by LC subjects and observations by the experimenter indicated that learners were more interested in performing the task well than they were about controlling challenge.
Experiment 2: Learner Control and Part-Whole-Task Training

Results of experiment 1 indicate that a criterion task that is both complex and unfamiliar to learners may become even more difficult if they are required to make many instructional control decisions. However, the regulation of the amount of practice learners engage in prior to a test requires fewer decisions than controlling criteria for performance measures (speed and accuracy). Because this responsibility is assigned to learners in most adult training environments, the second experiment allowed subjects to decide for themselves on the amount of practice to complete prior to the posttest. The purpose of this study was to examine the effects of part- and whole-task training on skill acquisition and on learners' decisions concerning practice.

Part-Whole-Task Training: Theory and Results of Previous Research

There are two basic training formats for teaching complex skills, part- and whole-task training (Naylor, 1962; Stammers, 1982; Wightman & Lintern, 1985). In whole-task training, all content material must be presented at one time to enable learners to execute the entire task during practice. Although whole-task training has proven successful for developing many different skills, learners often experience difficulty during initial practice of a complex task due to overload on working memory. Part-task training reduces the amount of concepts, facts, and rules learners must recall during initial practice activities. Also, part-task instruction and measures of performance are less complicated, because each subtask is addressed separately. In part-task training, the criterion task is broken down into a number of parts, and practice is provided for each part as a prelude to attempting the whole task (Gropper, 1983; Naylor, 1962; Stammers, 1982; Wightman & Lintern, 1985).

Almost all in-flight related curricula are organized and sequenced based upon the subdivision and subsequent re-combination of complex tasks. Thus, part-whole-task methods are important to aircrew training, but research findings have failed to lead to basic principles that enable training developers to systematically choose between part- and whole-task methods.

Part-task training appears to be a logical choice for training complex tasks, but research has not always supported this notion. In an extensive review of early research on part-whole-task training, Naylor (1962) indicated that highly integrated tasks such as memorization of prose material or complex motor tasks are most suited for whole-task training. Integration refers to the degree that subtasks interact so that...
learners' performance on one subtask affects their performance on another (e.g., estimation of location and direction of flight of target aircraft with the HUD). Although most studies reviewed by Naylor favored whole-task training, he concluded that as complexity of tasks increase, the effectiveness of part-task training also increases. Also, if time on practice activities is limited as in many aircrew training programs, part-task training may be the most effective. These assumptions are supported and extended by more recent reviews on part-whole-task training (Wightman & Lintern, 1985).

Some researchers have claimed that part-task training is superior to whole-task training, even for integrated tasks (Fabiani, Buckley, Gratton, Coles, Donchin, & Logie, 1989; Frederiksen & White, 1989; Manè & Donchin, 1989). Part-task training reduces the time learners need to develop an initial understanding of the task (Naylor, 1962) and can also reduce training costs associated with simulating entire complex systems (Gray & Edwards, 1991; Reigeluth & Schwartz, 1989). Gropper (1983) has stressed the importance of part-task training for providing transitions between abstract descriptions and concrete representations (simulations) of complex tasks. He explains that when the whole criterion task is described prior to a learner's first attempt to perform it, the individual may not be able to retain enough of the information to be able to understand and effectively practice the task.

"Progressive" part-task training appears to be the most effective alternative to whole-task training (Wightman & Lintern, 1985). The progressive method specifies that each subtask is practiced until a criterion level is achieved before it is combined and practiced with other previously mastered subtasks. In progressive part-task training, one subtask is introduced, practiced, and mastered. Then, a second subtask is mastered and added to the first so they can be practiced together, and this process continues until the learner is able to execute and practice the entire task.

A group of recent studies were conducted using different CAIS treatments that taught adult male subjects how to play a game called "space fortress" (Manè & Donchin, 1989). The game was designed to involve some of the skills associated with complex tasks performed in aviation and other dynamic environments. It included two major subtasks: (a) controlling a simulated space ship and (b) protecting it from damage by the space fortress' defenses. The game required precise timing of motor and perceptual tasks as well as the application of complex rules and strategies.
Progressive part-task training produced higher levels of skill, enabled learners to understand the task and acquire skill more quickly, and produced less variance in performance outcomes compared to whole-task training (Frederiksen & White, 1989; Mané, Adams, & Donchin, 1989; Fabiani et al., 1989). The most successful whole-task method emphasized parts of the task while subjects practiced the whole task (Gopher, Weil, & Siegel, 1989). Subjects who received this type of whole-task training outperformed part-task subjects when a distracter task was combined with the criterion task (Fabiani et al., 1989).

Overall, results of these studies indicated that part-task training was superior to whole-task training across most dependent measures. Experiment 2 was designed to identify some of the problems learners may encounter when attempting to control instructional events in part- and whole-task simulation training.

Method

Subjects

Forty-one undergraduate volunteers (16 male and 25 female) from two public universities in the southwestern United States participated in the experiment for partial course credit.

Materials

The CAIS program used in experiment 1 was modified to produce one part-task and one whole-task treatment. To increase the emphasis on perceptual components of the criterion task, the additional guidance provided by the overhead (diagram that marked intervals for range and azimuth) was removed. The modified criterion task required subjects to estimate target location and direction of flight by using the pilot's aircraft as the sole reference point. This change made the learning task more like the actual task performed by pilots in an aircraft. The CAIS provided the same instructional content to all subjects prior to the practice exercises.

During practice, the program assessed the speed and accuracy of each target estimation and computed scores for each of these measures. The practice score was displayed after every fifth practice exercise and represented subjects' average degree of speed and accuracy for the most recent five target estimations.

Experimental Treatments. The part-task practice treatment included six sets of exercises that are described below in the order they occurred in the program.

1. The first set of exercises presented a HUD that displayed only target range
which was estimated by positioning the locator on a guideline that extended from the pilot's aircraft to the top of the screen. The guideline provided the correct target azimuth. The locator's orientation was fixed at 90 degrees throughout the exercise. This enabled subjects to estimate target range independent of azimuth and aspect. Figure 3 shows a reproduction of a practice exercise for a range estimate of 25 miles.

Figure 3. Practice Exercise for First Set of Exercises: Range Estimate.

2. The second set of exercises displayed only target azimuth on the HUD. Azimuth was estimated by positioning the locator on a curved guideline which provided the correct target range. Figure 4 shows a reproduction of a practice exercise for an azimuth estimate of 30 degrees right.

3. The third set of exercises presented both range and azimuth. Target estimations were made by positioning the locator on the screen at the estimated range and azimuth without the help of the guidelines.

4. The fourth set of exercises showed the aspect pointer at different positions along the circumference of the aiming reticle on the HUD. The locator appeared at the correct location (range and azimuth) and required subjects to rotate it to estimate direction of flight. Figure 5 shows a practice exercise for a left wing aspect.
Figure 4. Practice Exercise for Second Set of Exercises: Azimuth Estimate.

Figure 5. Practice Exercise for Fourth Set of Exercises: Aspect Estimate.

5. The fifth set of exercises displayed the target azimuth and aspect. The practice task was to position the locator on the curved guideline (designating target range) to estimate azimuth, and rotate it to estimate direction of flight.
6. The last set of exercises displayed range, azimuth, and aspect on the HUD. The task was to estimate location and direction of flight using all three target measures (whole-task practice).

Whole-task practice consisted of practice items that were the same as the last part-task items except for the feedback. Feedback messages on the screen identified the measure (e.g., range) that was estimated with the greatest accuracy and the measure that was estimated with the least accuracy after each target estimate. Additionally, one of the guidelines (that designated range or azimuth) appeared on the screen to focus subjects' attention on the part of the estimate that exhibited the greatest error.

Posttest. The posttest was the same as that used in experiment 1 with the exception of the overhead diagram. Subjects were urged to estimate each target as "accurately and as quickly as possible." They completed three sample targets prior to the test and were encouraged to respond more quickly if their response time exceeded 30 seconds.

Procedure

Subjects were assigned to one of the two treatments in the order they arrived at the computer site. Ten to 15 subjects participated in the experiment at one time. Subjects completed (a) the mouse lesson and skill test, (b) content material on the HUD, (c) practice exercises, and (d) an immediate posttest, respectively.

Results and Discussion

The initial MANOVA on posttest scores revealed a significant effect for gender (multivariate $F(2, 37) = 4.46, p < .05$). Univariate tests indicated that males exceeded females on test score ($M = 70.7$ vs. $M = 59.0, t(38) = 7.21, p < .05$) and that males estimated targets more quickly than females ($M = 18.6$ vs. $M = 23.1$ seconds, $t(38) = 4.77, p < .05$). Because there were substantially more females than males, all additional tests were conducted on female subjects' data.

A second MANOVA revealed an effect for treatment (multivariate $F(2, 25) = 4.20, p < .05$). Univariate tests showed that the part-task group performed the task more quickly than the whole-task group ($M = 16.9$ vs. $M = 20.6$ seconds, $t(27) = 5.11, p < .05$).

An estimate of the amount of practice completed by subjects prior to the posttest indicated that part-task subjects completed about 33% more practice than whole-task
subjects. Evidently, the part-task training format, which included six different practice activities, induced subjects to engage in more practice.

It would be expected that subjects would tend to practice until they were relatively satisfied with their scores. For example, lower practice scores by part-task subjects would explain why they practiced more than whole-task subjects. However, there was no significant difference in practice scores between the two treatment groups. Also, part-task subjects tended to perform the whole task in the sixth exercise more quickly than the whole-task group ($M = 24.6$ vs. $M = 38.2$ seconds, $t(27) = 9.45, p < .01$). This suggests that the first five part-task exercises prepared subjects for the criterion task by increasing their ability and/or confidence beyond that of whole-task subjects and was likely the reason for their superior speed on the posttest.

GENERAL CONCLUSIONS

In simulation training, the degree of competition among cognitive resources and learners' perceptions of their performance are important to consider when designing instructional control strategies. Frequent learner-controlled decisions may distract learners during practice and reduce their retention of skill. Simulation training should employ program-controlled strategies when (a) there are many control decisions, (b) control decisions require the same cognitive resources needed to perform practice tasks, or (c) when performance criteria cannot be easily interpreted by learners.

Progressive part-task training should be employed for teaching most complex tasks. Such tasks can usually be analyzed to identify the essential component skills, and these skills can be acquired by learners more quickly by learners' completion of a series of part-task practice activities. This method enables learners to gradually build their skill until they can successfully perform and practice the whole criterion task. Part-task training also makes it easier to design program-controlled systems that monitor learners' performance, because it is easier to evaluate performance and provide instructional support for one subtask at a time.

Potential differences between the subject population in these experiments and student pilots have not yet been identified or defined. Thus, generalizing these findings to pilot training programs should be done with caution. However, reviews of the literature have not revealed individual differences among adult learners that limit findings on instructional control (Steinberg, 1977; 1989) or part-whole-task training (Naylor, 1962; Wightman & Lintern, 1985) to specific learner populations. This is not
surprising, because the challenge learners experience when attempting to master complex tasks is primarily dependent on general processes (e.g., motor and perceptual learning) and limitations (e.g., short-term memory load) that vary in the same way and to about the same degree across learner populations. Additionally, the greater motivation that would be expected in an aircrew training population would probably verify some of the conclusions about instructional control. For example, student pilots would probably place more value in meeting performance criteria than controlling instruction.

Prescriptions for developing effective combinations of instructional control and part-task training need to be developed, because opportunities to employ both strategies occur simultaneously in CAIS. Also, because these two training variables were examined separately in the present experiments, further research should investigate potential interactions among instructional control and part-whole-task training. Finally, the effects of gender need to be either (a) partitioned from potential treatment effects by blocking on gender in data analyses, or (b) excluded from experimental designs by selecting the most appropriate gender for subject participation.

Part-task training has been widely used in the military and other technical training environments, but there are few principles for guiding the design of part-task methods. To sharpen the focus and potential impact of this line of research, categories of aircrew tasks should be analyzed to identify those which are best suited for CAIS. Subsequently, task domains should be closely examined to identify systematic methods for decomposing tasks for part-task training and developing criterion measures and practice activities that can be effectively program-controlled. In this manner, training effectiveness and reduction of training costs through the implementation of CAIS may be realized.
REFERENCES


27


