The variables of practice, incentive, instruction, and feedback were examined to determine which contributed significantly to the effectiveness of an instructional program, which was systematically developed to include the four variables. Two studies were carried out using the Aircraft Instrument Comprehension Program and Test. In the first study the effects of practice and incentive on learner posttest performance were examined, using two levels of each variable in a 2 x 2 factorial design. A second study examined the effects of instruction and feedback using the same design. Results indicated that instruction had a strong effect on learner performance in that scores on the posttest were significantly higher than on the pretest. Practice and incentive had a significant effect on the time required to perform the task. The test scores and statistical analyses are appended. (SA)
EFFECTS OF SYSTEMATIC VARIATIONS OF INSTRUCTIONAL VARIABLES IN A VERIFIED INSTRUCTIONAL PROGRAM

Robert A. Reiser


April, 1975

Session # 12.14

The research activities described in this paper were supported by the USAF Human Resources Laboratory, Williams Air Force Base, Arizona, Contract No. F41609-71-C-0027

Department of Educational Technology
Arizona State University
Tempe, Arizona
The primary focus of the present research was to determine which variables contribute significantly to the effectiveness of an effective instructional program. The variables examined in the present series of studies were practice, incentive, instruction, and feedback. While many studies have examined the effects of these variables, few studies have used an effective instructional program as the research vehicle. The effect of variables in instructional materials or other prose passages that have not been demonstrated to be effective may differ from the effects of variables in effective instructional programs.

The instructional program used in this study was the Aircraft Instrument Comprehension Program (Higgins, 1973). The Aircraft Instrument Comprehension Program is a self-instructional program designed to train learners to perform an aircraft instrument comprehension task: to identify which one of four pictures of an aircraft in flight most nearly corresponds to the position indicated on an attitude indicator and a heading indicator. Tryouts of the program indicated that it usually enabled learners to perform the task at criterion-level proficiency (Higgins & Kearns, 1973; Higgins, Kearns, & Tenpas, 1974).
The Aircraft Instrument Comprehension Program was systematically developed to include instruction (instructional cues and examples), practice, feedback, and an incentive. Two studies were designed to examine these variables. In the first study the effects of practice and incentive on learner posttest performance were examined. Two levels of practice (presence and absence) and two levels of incentive (presence and absence) were manipulated in a 2X2 factorial design. Based on the results of the first study, a second study was designed to examine the effects of instruction and feedback on learner posttest performance. Two levels of instruction (presence and absence) and two levels of feedback (presence and absence) were manipulated in a 2X2 factorial design. In both studies learner posttest performance was reflected by two measures: number of correct responses and time required to complete the posttest.

Experiment 1: Effects of Practice and Incentive

Method

Subjects. Forty-eight students enrolled in the Air Force Reserve Officer Training program at Arizona State University served as subjects in this study.

Materials. The materials used in this study were variations of the Aircraft Instrument Comprehension Program (Higgins, 1973). The program was designed to enable students to achieve the following instructional objective:

Given four illustrations of aircraft in roll, pitch, and heading, the student will identify the illustration that most nearly represents the position indicated on a compass and an artificial horizon.
Instruction in the program consisted of one instructional cue and three examples for each of the three concepts presented: roll, pitch, and heading. Eight examples were also presented in which these concepts were combined. Practice in the program consisted of one to four practice items for each concept, followed by ten additional practice items at the end of the program. All practice items required students to identify which of two or more drawings of an aircraft in flight most nearly represented the position shown on an attitude indicator and a heading indicator. Upon completion of each practice item students were able to obtain feedback, in the form of knowledge of correct response, by moving a paper slide.

Procedure. A pretest was administered to all subjects prior to initiating the experiment. The pretest consisted of nine items selected to be representative of the items on the posttest. The differences in pretest scores between treatment groups were not statistically significant.

The four treatment groups which resulted from the 2x2 factorial design were practice and incentive, practice and no incentive, no practice and incentive, and no practice and no incentive. Subjects attended classes in four different sections. In order to prevent subjects in the no incentive groups from being aware of the incentive, all subjects in the first two sections were assigned to the no incentive groups. All subjects were randomly assigned to either the practice or no practice groups.

At the start of the experimental session each subject received an instructional booklet. When all subjects had received their booklets, they were instructed to begin reading the booklet that had been given to them.
Subjects in the incentive groups received written instructions located at the beginning of their instructional booklets indicating that they could improve their course grade by up to one-half of a grade point by responding correctly to the items on the posttest. In addition, the instructions indicated that the four subjects who responded correctly to the largest number of posttest items in the least amount of time would have an opportunity to fly a formation trainer at Williams Air Force Base. Reserve Officer Training candidates at Arizona State University do not get many chances to fly a formation trainer, and value any opportunity to do so. Instructions to the subjects in the no incentive groups simply stated that the subjects' grades in the course would not be affected by their participation in the experiment and that the developers of the self-instructional program would appreciate their best efforts.

The booklets for subjects in the practice groups contained all the instructional cues, examples, and practice items that originally appeared in the program. Subjects in the no practice groups received instructional booklets in which all practice items were deleted from the program. Booklets for each subject in these groups contained only the instructional cues and examples for each concept taught in the program.

The last page of each booklet instructed each subject to record the time he had finished his booklet and to raise his hand so that a proctor could collect his materials. A proctor checked the completion time the subject had listed and collected the instructional materials. The posttest was administered after all subjects had completed the instructional booklets. Subjects were again asked to record their completion times, and the test booklets were collected as each subject finished.
Criterion Measure. The Aircraft Instrument Comprehension Test: Form B (Kearns, Tenpas, & Higgins, 1973) served as the posttest. The test contained directions, a sample test item, and 36 multiple-choice items. All test items required subjects to identify which one of four drawings of an aircraft in flight most nearly represented the position shown on an attitude indicator and a heading indicator.

Data Analyses. Analyses of variance were performed to determine the individual and combined effects of practice and incentive on posttest scores and posttest items.

Results

The mean posttest scores of the treatment groups (See Table 1) were all within one point of the overall mean score of 33.31. F-ratios for incentive, practice, and the practice X incentive interaction were not statistically significant.

Insert Table 1 about here

Subjects spent differential amounts of time answering items on the posttest. Mean posttest times by treatment are reported in Table 2. Posttest time differences were analyzed using an analysis of variance, as shown in Table 3. Statistically significant differences were found for both practice, (F (1, 44) = 22.3, p < .001) and for incentive (F (1, 44) = 15.4, p < .001). The practice X incentive interaction was not significant.
In order to determine whether instruction had an appreciable effect on learner performance, subject pretest performance was compared with subject posttest performance. On the pretest, subjects answered a mean of 69.4% of the items correctly. On the posttest, subjects answered a mean of 92.5% of the items correctly. These percentages were approximately the same for all four treatment groups. Thus, it appears that the improvement in subject scores from pretest to posttest was attributable to instruction, the one variable received by all four treatment groups.

Experiment 2: Effects of Instruction and Feedback

Method

Subjects. Sixty-three students enrolled in the Air Force Reserve Officer Training program at the University of Arizona served as subjects in this study.

Materials. The materials used in this study were similar to those used in Experiment 1. The only difference in materials was that in this study subjects were able to obtain feedback, in the form of knowledge of correct response, by marking a chemically treated answer sheet.

Procedures. The four treatment groups which resulted from the 2X2 factorial design were instruction and feedback, instruction and no feedback, no instruction and feedback, and no instruction and no feedback. Performance incentives, contingent upon both the speed and accuracy of
a subject's posttest responses, were identical to those described in Experiment 1. Subjects were informed of the incentives through oral directions given at the beginning of the experimental session. The incentives were identical for all groups.

At the beginning of the experimental session a pretest was administered to the subjects. The pretest was the same pretest that was used in Experiment 1. Subjects who received a score of seven or more on the pretest were considered to be proficient on the task prior to the study and were eliminated from all the data analyses reported. The differences in pretest scores between treatment groups were not statistically significant.

Following completion of the pretest, all subjects were given an instructional booklet, a marking pen, and a response sheet. Subjects in the instruction groups received a booklet with all instructional cues, examples, and practice items intact. Subjects in the no instruction groups received a booklet containing all the practice items, however, all instructional cues and examples were deleted. Subjects in the feedback groups received a special marking pen and a chemically treated response sheet which enabled the subjects to receive feedback in the form of knowledge of correct response. Subjects in the no feedback groups received a marking pen and a response sheet which did not provide any feedback.

When all subjects had received their materials, they were instructed to begin reading the booklet that had been given to them. The last page of each booklet instructed each subject to record the time he had
finished his booklet and to raise his hand so that a proctor could collect his materials. A proctor checked the completion time the subject had listed, collected the materials, and gave the subject a copy of the posttest. The time the subject began the posttest was recorded by the proctor. When a subject finished the posttest he recorded his completion time, had his material collected by a proctor, and was dismissed.

**Criterion Measure.** The criterion measure used in this study was identical to the criterion measure described in Experiment 1.

**Data Analyses.** Analyses of variance were performed to determine the individual and combined effects of instruction and feedback on posttest scores and posttest times.

**Results**

The mean score for each group on the 36-item posttest is shown in Table 4. The mean score for the instruction groups was 4.22 points higher than the mean score for the no instruction groups; a difference which is significant at the .05 level ($F (1, 37) = 4.73$). $F$-ratios for feedback and the instruction $\times$ feedback interaction were not significant (see Table 5).

Mean times required for each group to complete the posttest are reported in Table 6. The mean time for the instruction groups was 1.07 minutes less than the mean time for the no instruction groups; a
difference which is significant at the .05 level \( F(1, 37) = 4.45 \). No significant differences for either feedback or the instruction \( X \) feedback interaction were observed (see Table 7).

[Insert Table 6 and Table 7 about here]

**Discussion**

The primary focus of the present research was to determine which variables contribute significantly to the effectiveness of an effective instructional program. The results of the two experiments indicate that instruction had a strong effect on learner performance on the aircraft instrument comprehension task. In Experiment I instruction appeared to be responsible for an impressive improvement in subject scores from pretest to posttest. The effect of instruction was so powerful that there was little room for either practice or incentive to contribute significantly to subject posttest scores. In Experiment II instruction had a significant effect on posttest scores and posttest time.

The results of Experiment I indicate that practice and incentive had a significant effect on the time required to perform the aircraft instrument comprehension task. Time is an important aspect of the task in that it is an indication of the rate at which pilots read their control instruments. The ability to read instruments accurately and quickly is a highly desirable skill to acquire in pilot training. Thus, practice
and incentive were valuable because they increased the speed at which subjects were able to perform the task.

The present research indicates that instruction was the most powerful variable in an effective instructional program. Practice, incentive, and feedback, variables that have been effective in other instructional settings, did not have a significant effect on subject posttest scores in this study. These results may not hold across tasks of different degrees of complexity. In effective instructional programs that teach more complex tasks, instruction alone may not enable learners to perform the tasks at criterion-level. Other variables, such as practice, feedback, and incentive, may have to be included in the programs in order for the programs to be effective. Further research should be conducted to determine the generalizability and consistency of the present findings.
References


<table>
<thead>
<tr>
<th></th>
<th>Practice</th>
<th>No Practice</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incentive</td>
<td>33.33</td>
<td>32.42</td>
<td>32.88</td>
</tr>
<tr>
<td>No Incentive</td>
<td>33.50</td>
<td>34.00</td>
<td>33.75</td>
</tr>
<tr>
<td>Totals</td>
<td>33.42</td>
<td>33.21</td>
<td>33.31</td>
</tr>
</tbody>
</table>

n = 12 per cell

Maximum score = 36
TABLE 2
Mean Posttest Times (Minutes) by Treatment
Experiment 1

<table>
<thead>
<tr>
<th>Practice</th>
<th>Practice</th>
<th>No Practice</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incentive</td>
<td>5.98</td>
<td>8.13</td>
<td>7.05</td>
</tr>
<tr>
<td>No Incentive</td>
<td>7.79</td>
<td>9.60</td>
<td>8.70</td>
</tr>
<tr>
<td>Totals</td>
<td>6.88</td>
<td>8.86</td>
<td>7.88</td>
</tr>
</tbody>
</table>

n = 12 per cell

TABLE 3
Analysis of Variance: Posttest Times
Experiment 1

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F-Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incentive</td>
<td>14.40</td>
<td>1</td>
<td>14.40</td>
<td>15.42*</td>
</tr>
<tr>
<td>Practice</td>
<td>16.56</td>
<td>1</td>
<td>16.56</td>
<td>22.31*</td>
</tr>
<tr>
<td>Incentive x Practice</td>
<td>1.54</td>
<td>1</td>
<td>1.54</td>
<td>0.04 NS</td>
</tr>
<tr>
<td>Within</td>
<td>28.70</td>
<td>44</td>
<td>.65</td>
<td></td>
</tr>
</tbody>
</table>

* p < .001
NS = Not Significant
TABLE 4
Mean Posttest Scores by Treatment
Experiment 2

<table>
<thead>
<tr>
<th>Instruction</th>
<th>Feedback</th>
<th>No Feedback</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instruction</td>
<td>32.32</td>
<td>31.73</td>
<td>32.06</td>
</tr>
<tr>
<td>No Instruction</td>
<td>31.67</td>
<td>25.24</td>
<td>27.84</td>
</tr>
<tr>
<td>Totals</td>
<td>31.99</td>
<td>28.48</td>
<td>29.95</td>
</tr>
</tbody>
</table>

n = 11 per cell

Maximum score = 36

TABLE 5
Analysis of Variance: Posttest Scores
Experiment 2

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F-Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instruction</td>
<td>18.05</td>
<td>1</td>
<td>18.05</td>
<td>4.73*</td>
</tr>
<tr>
<td>Feedback</td>
<td>8.69</td>
<td>1</td>
<td>8.69</td>
<td>2.28</td>
</tr>
<tr>
<td>Instruction X Feedback</td>
<td>5.07</td>
<td>1</td>
<td>5.07</td>
<td>1.33</td>
</tr>
<tr>
<td>Within</td>
<td>140.99</td>
<td>37</td>
<td>3.81</td>
<td></td>
</tr>
</tbody>
</table>

* p < .05

NS = Not Significant
TABLE 6
Mean Posttest Times (Minutes) by Treatment
Experiment 2

<table>
<thead>
<tr>
<th>Instruction</th>
<th>Feedback</th>
<th>No Feedback</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instruction</td>
<td>7.18</td>
<td>7.11</td>
<td>7.14</td>
</tr>
<tr>
<td>No Instruction</td>
<td>8.68</td>
<td>7.50</td>
<td>8.21</td>
</tr>
<tr>
<td>Totals</td>
<td>7.93</td>
<td>7.30</td>
<td>7.61</td>
</tr>
</tbody>
</table>

n = 11 per cell

TABLE 7
Analysis of Variance: Posttest Times
Experiment 2

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F-Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instruction</td>
<td>.89</td>
<td>1</td>
<td>.89</td>
<td>4.45*</td>
</tr>
<tr>
<td>Feedback</td>
<td>.39</td>
<td>1</td>
<td>.39</td>
<td>1.95</td>
</tr>
<tr>
<td>instruction X Feedback</td>
<td>.30</td>
<td>1</td>
<td>.30</td>
<td>1.50 NS</td>
</tr>
<tr>
<td>Within</td>
<td>7.75</td>
<td>37</td>
<td>.20</td>
<td></td>
</tr>
</tbody>
</table>

* p < .05
NS = Not Significant