THE EFFECT OF REHEARSAL ON THE RETENTION OF A TIME-SHARED TASK

TECHNICAL DOCUMENTARY REPORT NO. AMRL-TDR-63-33

April 1963

Behavioral Sciences Laboratory
6570th Aerospace Medical Research Laboratories
Aerospace Medical Division
Air Force Systems Command
Wright-Patterson Air Force Base, Ohio

Contract Monitor: T. E. Cotterman
Project No. 1710, Task No. 171003

[Prepared under Contract No. AF 33(616)-7269 by
James C. Naylor
George E. Briggs
Donald R. Brown
Walter G. Reed
Laboratory of Aviation Psychology
The Ohio State University
Columbus, Ohio]
NOTICES

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related government procurement operation, the government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Qualified requesters may obtain copies from the Defense Documentation Center for Scientific and Technical Information (DDC), Arlington Hall Station, Arlington 12, Virginia. Orders will be expedited if placed through the librarian or other person designated to request documents from DDC (formerly ASTIA).

Do not return this copy. Retain or destroy.

Stock quantities available at Office of Technical Services, Department of Commerce $0.75.

Change of Address

Organizations receiving reports via the 6570th Aerospace Medical Research Laboratories automatic mailing lists should submit the addressograph plate stamp on the report envelope or refer to the code number when corresponding about change of address.
The research described in the report was performed by the Laboratory of Aviation Psychology, The Ohio State University, Columbus, Ohio, under Contract No. AF 33(616)-7269 during the period 1 May 1960 to 1 July 1962 for Behavioral Sciences Laboratory, 6570th Aerospace Medical Research Laboratories, Aerospace Medical Division. This is one of several laboratory research activities being carried out on the subject contract on "Techniques for Promoting the Long-Term Retention of Learned Skills." Dr. George E. Briggs is the Principal Investigator.

This contractual work was performed under Project 1710, "Training, Personnel, and Psychological Stress Aspects of Bioastronautics," Task 171003, "Human Factors in the Design of Systems for Operator Training and Evaluation," with Dr. T. E. Cotterman as Task Scientist. The contract was initiated and monitored until June 1961 by Mr. Frederick H. Kresse of the Operator Training Section, Training Research Branch, Behavioral Sciences Laboratory. Subsequent monitoring, inclusive of the studies reported here, has been carried out by Dr. Cotterman. Mr. Kresse provided valuable assistance in the development of the experimental task and Dr. Cotterman provided most helpful comments on an initial draft of this report.
ABSTRACT

Two studies were performed to investigate the influence of various methods of task rehearsal upon the retention of a time-shared task. Experiment I examined retention as a function of four rehearsal conditions (part task, whole task, simplified task, and none). Subjects in each of the groups trained for 8 days, returned 6 days later for 2 days of rehearsal, and then returned again after 7 more days for a retention test. Experiment II examined retention both as a function of 3 rehearsal methods (whole, part, and none) and amount of training (5 or 10 days). Subjects returned for 2 days of rehearsal 10 days after completion of training and then returned for retest 9 days later. Tracking performance in both studies showed significant effects due to rehearsal methods. In Experiment I part rehearsal was superior, while in Experiment II whole rehearsal was found to be best.

PUBLICATION REVIEW

This technical documentary report has been reviewed and is approved.

WALTER F. GRETHER
Technical Director
Behavioral Sciences Laboratory
### TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>METHOD</td>
<td>2</td>
</tr>
<tr>
<td>Experiment I</td>
<td>2</td>
</tr>
<tr>
<td>Apparatus</td>
<td>2</td>
</tr>
<tr>
<td>Experimental Design</td>
<td>4</td>
</tr>
<tr>
<td>Subjects and Procedure</td>
<td>4</td>
</tr>
<tr>
<td>Experiment II</td>
<td>5</td>
</tr>
<tr>
<td>Apparatus</td>
<td>5</td>
</tr>
<tr>
<td>Experimental Design</td>
<td>7</td>
</tr>
<tr>
<td>Subjects and Procedure</td>
<td>7</td>
</tr>
<tr>
<td>RESULTS</td>
<td>7</td>
</tr>
<tr>
<td>Experiment I</td>
<td>7</td>
</tr>
<tr>
<td>Tracking Task</td>
<td>7</td>
</tr>
<tr>
<td>Procedural Task</td>
<td>11</td>
</tr>
<tr>
<td>Experiment II</td>
<td>11</td>
</tr>
<tr>
<td>Tracking Task</td>
<td>11</td>
</tr>
<tr>
<td>Procedural Task</td>
<td>16</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>17</td>
</tr>
<tr>
<td>SUMMARY</td>
<td>19</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>20</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1</td>
<td>Average Tracking Performance for Each Experimental Group during Training, Rehearsal, and Retention Test in Experiment I</td>
</tr>
<tr>
<td>2</td>
<td>Average Tracking Performance for Each Group during Training, Rehearsal, and Retention Test for the One-Week Training Groups in Experiment II</td>
</tr>
<tr>
<td>3</td>
<td>Average Tracking Performance for Each Group during Training, Rehearsal, and Retention Test for the Two-Week Training Groups in Experiment II</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Conditions during Training and Rehearsal Defining the Four Experimental Groups in Experiment I</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>Schedule of Stimulus Conditions for the Nine Light Pairs (Coded 1 through 9 from Top to Bottom of Panel) Used during Training</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>Experimental Conditions Defining the Various Groups in Experiment II</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>Analyses of Variance Performed on Final Performance at the End of Training (Session 8) in Experiment I</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>Analysis of Variance Results of the Absolute Retention Test Scores for All Performance Measures in Experiment I</td>
<td>9</td>
</tr>
<tr>
<td>6</td>
<td>Analysis of Variance Results of the Difference Scores (Training Session - Retention Session) for All Performance Measures in Experiment I</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>Mean Performance Measures for Subjects during the Last Training Session and the Retention Test Session for Each of the Four Rehearsal Conditions in Experiment I</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>Analysis of Variance Results of Final Performance at End of Training (Session 5 for Groups 1-3, Session 10 for Groups 4-6) in Experiment II</td>
<td>14</td>
</tr>
<tr>
<td>9</td>
<td>Analysis of Variance Results of the Difference Scores (Training Session - Retention Test Session) for All Performance Measures in Experiment II</td>
<td>14</td>
</tr>
<tr>
<td>10</td>
<td>Mean Performance Measures for Subjects during the Last Training Session and the Retention Test Session for Each of the Groups in Experiment II</td>
<td>15</td>
</tr>
<tr>
<td>11</td>
<td>Observed Order of Efficiency of the Various Rehearsal Methods as a Function of Amount of Training and Dependent Variable</td>
<td>17</td>
</tr>
</tbody>
</table>
INTRODUCTION

Most tasks which involve control of a vehicle are multidimensional in nature, i.e., the operator is required to perform more than one subactivity. The array of subtasks making up the total or whole task may all be similar in nature or they may be widely diversified in terms of the types of activities and skills called for. Because of the increasing complexity of modern vehicular control tasks, a great deal of interest has centered on the feasibility of part-task training and part-task trainers (refs. 3, 4, 10, 17) in an attempt to reduce the time and to increase the efficiency of training individuals to a desired level of skilled performance.

The general finding of most part-whole training comparisons has been that whole practice generally is superior (ref. 12), although Briggs, Naylor, and Fuchs (ref. 5) obtained results which indicated that for tasks of low organization, part practice may become superior if task difficulty is set at a high level. One explanation for the almost perpetual superiority of whole training is that such practice provides experience on the time-sharing aspects of the entire task complex (refs. 4, 6, 10, 11, 16).

Time-sharing, as a characteristic of complex control tasks, is imposed upon the operator by the fact that responses in two or more task dimensions must be made concurrently. Thus, not only must skill be learned for each task component, but also the operator must learn adequate ways of performing simultaneously on the several task components. While in the strict sense time-sharing may be considered to involve two or more responses performed simultaneously, in many cases it is also appropriate for describing demands created by a task having an interlacing of two or more response types over time (ref. 9). Several studies have examined the validity of the hypothesis that part training is less efficient than whole practice for tasks having time-sharing requirements. Dougherty, Houston, and Nicklas (ref. 7) failed to find evidence that such was the case. However, Adams, Hufford, and Dunlop (ref. 1), in a quite thorough examination of the question, did show the importance of learning to time-share concurrent activities.

If the time-sharing hypothesis is valid, i.e., if time-sharing does comprise a skill to be learned over and above the skills required on the several task dimensions, it follows that this skill may also be subject to forgetting as are other learned capabilities. That such is indeed the case was made apparent by Hufford and Adams (ref. 9) who found that part practice was not sufficient to re-establish completely a former skill level on a time-shared task after 10 months without practice. Only through the use of whole practice was it possible to achieve performance levels equivalent to those found at the end of initial training. Thus, it appears that the retention of a task involving time-shared responses depends upon both (a) the degree of skill loss for the separate response classes and (b) the degree of skill loss for the time-sharing aspect of the task. Therefore, any procedure designed to facilitate the retention of skill for a complex task should involve techniques designed to maintain both component and time-sharing skill.

A standard procedure for maintaining proficiency on a complex task is to provide the operator with some means of rehearsal at one or more stages during
the retention interval. This rehearsal may utilize the whole task or may involve a systematic modification of that task, e.g., part-task training, simplified task training, etc. With rehearsal, as with original training, there is the problem of determining to what degree it is necessary to duplicate the operational task in order to maintain a skill that has already been acquired. Thus, Naylor and Briggs (ref. 13), using a procedural task involving a series of discrete responses, found that whole-task rehearsal was superior to part practice on either the temporal or spatial characteristics of the task, which in turn were better than no rehearsal. Similarly, Sackett (ref. 15) showed that actual task rehearsal was better than deliberate "mental" rehearsal which in turn was more efficient in maintaining proficiency than no rehearsal at all.

Assuming the utility of task rehearsal as a means of sustaining proficiency, the question arises concerning what method of rehearsal is most appropriate for a task involving time-sharing skills. The alternative methods of rehearsal training which are available are the standard ones: whole-task rehearsal, part-task rehearsal, and "simplified" rehearsal, i.e., practice on a task which contains all the aspects of the operational task but at levels of difficulty which are less demanding. Since rehearsal is merely additional practice during the retention interval, it would seem likely that the rehearsal methods should exhibit the same hierarchy of efficiency as found in training. Thus, for a task involving a time-sharing dimension, whole rehearsal should be superior to part-task rehearsal, since such practice allows for maintaining both the time-sharing and the component skills necessary for successful task performance. Similarly, it might be expected that simplified rehearsal also would lead to better retention than part-task rehearsal, as practice on time-sharing is still present even though the difficulty of the separate task components has been reduced.

The purpose of the research reported here was to investigate the above possibilities concerning the various types of rehearsal as means of facilitating the retention of a time-shared task. In addition, the studies were designed to examine the relative loss of skill on the individual task components—one which was a continuous response dimension and one which was procedural (discrete) in nature.

METHOD

Experiment I

The first study was designed to compare three rehearsal techniques: whole-task rehearsal, part-task rehearsal, and simplified whole-task rehearsal.

Apparatus: The subjects (Ss) were required to perform a two-component task concurrently during both training and retention test. The two component tasks were independent and each involved qualitatively different responses and objectives. One task was procedural (discrete response) in nature while the other was a continuous, three-dimensional, compensatory tracking task. Both tasks have been defined in detail previously (refs. 13, 14). The procedural task consisted of a panel of nine pairs of stimulus lights (one amber and one red) with 1-inch vertical separation between pairs. To the left of each pair were three response buttons labeled "Emergency," "OK," and "Check." If an amber light occurred, S pressed the OK button; if a red light appeared, he responded by pressing the
corresponding emergency and OK buttons in that sequence; and if no light at all occurred at the specified time, S responded with the sequence of check, emergency, OK. A correct response or response sequence resulted in the light being locked into the amber or OK condition. Any failure to produce the appropriate response sequence resulted in the red light being locked in.

The procedural task panel was situated approximately 2 feet from S's left shoulder and 30° to the left of center of his frontal vision. The panel was rotated in such a manner that the plane area of the panel was maximal. By use of a 1-second stepping switch it was possible (a) to program the spatial order of the nine stimulus events, (b) to program the duration of each stimulus event, and (c) to program the time between the onsets of the stimulus events. In addition, the experimenter (E) was able to program the initial condition for each stimulus event: red, amber, or no light.

A particular level of procedural task organization was used throughout the entire experiment. The exact level of organization used was set in terms of the spatial contingency relationships of the stimulus event pairs, the spatial sequence used being 1, 5, 2, 9, 8, 3, 6, 7, 4, where 1 refers to the top pair of lights, 2 to the next pair, and so on. Since the sequence of positions involved sampling light positions without replacement, an appropriate index of the sequential relationships is the informational metric Η calculated on the basis of the relationships between changes in light positions rather than on the basis of light positions themselves. The Η for this sequence was 2.808 bits. This was identical to the low organization sequence used in the prior retention study with this apparatus (ref. 14).

The duration of each stimulus event was set at 4 seconds. The S was required to activate the appropriate response button or buttons within this time or the red light automatically locked in. There was a constant interval of 6 seconds between stimulus event onsets (4 seconds for a stimulus event plus 2 seconds delay before the next stimulus event).

The tracking task was three dimensional, providing simulation of the three attitude control dimensions of a vehicle in free flight (roll, pitch, and yaw); rate control dynamics were present in all three dimensions. The display panel was 19 x 10.5 inches and was situated directly in front of S to the right of the procedural task display. Three pairs of center-null-position meters, 4-5/8 inches wide x 4-1/2 inches high, were mounted in two rows of three each on the panel, with the upper dial of each pair providing S with attitude error and the lower dial providing rate error. The dials were labeled, from left to right: roll, pitch, and yaw. The input signal was a simple sine wave of 0.01 cycle per second as generated by a Hewlett-Packard Model 202A signal generator, and this signal was tracked in all three dimensions simultaneously. The S utilized a three-dimensional control stick for his tracking response.

Left-to-right stick movement controlled roll; front-to-back movement controlled pitch; and rotation (twisting) around the axis of the stick controlled yaw. The control stick was mounted to the floor approximately 1 foot in front of S. Height of the stick was 28 inches, and all control-display relationships were compatible, i.e., they conformed to population stereotypes. The displacement in degrees of the control for each of the control dimensions was as follows: roll, ± 60°; pitch, ± 60°; and yaw, ± 90°. The control-display gain was 0.025
inch/second of display pointer movement for each degree of arc of the control stick. The spring loading of the control was approximately 1 ounce per degree of arc.

Experimental Design: Table 1 shows the four experimental groups and the associated experimental conditions, with the conditions during rehearsal being the critical experimental distinction between groups. Whole-task rehearsal consisted of practice on both subtasks simultaneously, part-task rehearsal involved practicing each component separately, and simplified rehearsal was similar to whole-task practice except that the dynamics of the tracking task were reduced to position control instead of rate control in all three dimensions; thus, time-sharing practice was still present, but the overall task difficulty was lessened under the simplified-rehearsal condition. During rehearsal the whole and the simplified-whole groups received one-half the number of trials as the part-rehearsal group so that the number of trials on each subtask would be equated across groups.

Subjects and Procedure: A total of 60 undergraduate males served in the experiment. All were volunteers who received $1.00 per experimental session. All Ss were instructed in the operation of both the procedural and the tracking tasks. Following the instructions in the first session, S received four 70-second trials on each of the two tasks separately, i.e., part-task training schedule was employed. Part training continued through the fourth session with 12 trials per session (six per task). During part training all procedural trials were given first each day, followed by the tracking task. Whole training began with session 5 and was continued through the end of the training period (session 8). Tracking performance was scored over the last 60 seconds of each trial and the onset of scoring coincided with the first stimulus event on the procedural task, i.e., S tracked alone for the first 10 seconds of each 70-second trial. In this way scoring on the tracking coincided with scoring on the procedural task. Six days following the end of training those subjects receiving rehearsal returned for two days of additional practice. After a lapse of seven more days all Ss returned on the retest session.

<table>
<thead>
<tr>
<th>Group</th>
<th>Training</th>
<th>Retention</th>
<th>Rehearsal</th>
<th>Retention</th>
<th>Retest</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4 days PT, 4 days WT</td>
<td>6 days</td>
<td>2 days PT</td>
<td>7 days</td>
<td>1 day WT</td>
</tr>
<tr>
<td>2</td>
<td>4 days PT, 4 days WT</td>
<td>6 days</td>
<td>2 days WT</td>
<td>7 days</td>
<td>1 day WT</td>
</tr>
<tr>
<td>3</td>
<td>4 days PT, 4 days WT</td>
<td>6 days</td>
<td>2 days ST</td>
<td>7 days</td>
<td>1 day WT</td>
</tr>
<tr>
<td>4</td>
<td>4 days PT, 4 days WT</td>
<td>6 days</td>
<td>2 days NT</td>
<td>7 days</td>
<td>1 day WT</td>
</tr>
</tbody>
</table>

Note: WT = Whole-task training (both tasks together). PT = Part-task training (tracking and procedural separately). ST = Simplified-task training (simplified tracking and procedural together). NT = No task rehearsal.
Each session of whole training consisted of three blocks of four trials on the combined task. Rehearsal consisted of six whole-task trials for groups 2 and 3 at each session and twelve part-task trials (six procedural, followed by six tracking) for group 1. Retest consisted of six whole-task trials for all groups.

Integrated absolute error served as the performance metric for the tracking task and this score was taken for each of the three task dimensions separately and then summed for each S. Performance on the procedural task was evaluated in terms of three metrics: total response time for the nine stimulus events of each trial, number of commissive errors (button presses in excess of the required number and/or button presses which were incorrect for the particular stimulus condition), and number of omissive errors (number of times S failed to respond correctly at the appropriate time).

Although the spatial sequence and the time intervals between stimulus events were constant for S, the actual stimulus conditions (amber light, yellow light, or no light) were programmed by E to change from trial to trial and from day to day. The schedule was as follows:

Session 1: No off-light conditions were used; S experienced four red and five amber light conditions on each trial with the order of conditions randomized.

Session 2: No off-light conditions during the first three trials; position 2 was an off-light during the last three trials. The amber programming procedure remained the same.

Session 3: No off-lights on trials 1 and 4, two off-lights on trials 2 and 3 (positions 2 and 4) and three off-lights on trials 4 and 5 (positions 2, 4, 6). Amber same as above.

Session 4: No off-lights on trials 1 and 4, four off-lights on all other trials (positions 2, 4, 6, and 8). Amber same as above.

Session 5: Four off-lights on every trial (positions 2, 4, 6, and 8). Amber-light sequencing remained the same.

Table 2 provides a detailed summary of these conditions for the procedural task.

The Ss were matched and assigned to groups on the basis of tracking performance in session 2. The sum of their mean roll, pitch, and yaw scores was used as the basis for matching.

Experiment II

The second study examined the effects of several rehearsal procedures across several levels of amount of training. Two rehearsal techniques were used (whole and part rehearsal) and there were two levels of training (1 or 2 weeks).

Apparatus: The apparatus used was identical in all respects to that used in Experiment I, consisting of both the procedural and the continuous tracking tasks. The input and task dynamics from the first study were maintained with the tracking task.
### TABLE 2

**SCHEDULE OF STIMULUS CONDITIONS FOR THE NINE LIGHT PAIRS**
(CODED 1 THROUGH 9 FROM TOP TO BOTTOM OF PANEL)

**USED DURING TRAINING IN EXPERIMENT I**

<table>
<thead>
<tr>
<th>Session</th>
<th>Trial</th>
<th>Red Lights</th>
<th>Off Lights</th>
<th>Session</th>
<th>Trial</th>
<th>Red Lights</th>
<th>Off Lights</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1-4</td>
<td>1-2-3-9</td>
<td>None</td>
<td>5</td>
<td>1-3</td>
<td>1-4-5-9</td>
<td>2-1-6-8</td>
</tr>
<tr>
<td>2</td>
<td>1-3</td>
<td>1-2-4-9</td>
<td>None</td>
<td>4-6</td>
<td>4-6</td>
<td>4-5-8-9</td>
<td>2-1-6-8</td>
</tr>
<tr>
<td></td>
<td>4-6</td>
<td>3-4-5-6</td>
<td>2</td>
<td>4-6</td>
<td></td>
<td>4-5-8-9</td>
<td>2-1-6-8</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>6-7-8-9</td>
<td>None</td>
<td>5-6</td>
<td>2-6-7-8</td>
<td>2-5-6-8</td>
<td>2-1-6-8</td>
</tr>
<tr>
<td></td>
<td>2-3</td>
<td>6-7-8-9</td>
<td>2-4</td>
<td>2-3</td>
<td>5-8</td>
<td>5-7-8-9</td>
<td>2-1-6-8</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1-2-5-8</td>
<td>None</td>
<td>4</td>
<td>3-4-5-6</td>
<td>2-5-6-8</td>
<td>2-1-6-8</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1-3-6-7</td>
<td>None</td>
<td>9-12</td>
<td>1-4</td>
<td>1-3-5-8</td>
<td>2-1-6-8</td>
</tr>
<tr>
<td></td>
<td>2-3</td>
<td>1-3-6-7</td>
<td>2-1-6-8</td>
<td>5-8</td>
<td>2-3-5-6</td>
<td>2-1-6-8</td>
<td>2-1-6-8</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>3-4-5-6</td>
<td>None</td>
<td>9-12</td>
<td>5-8</td>
<td>5-7-8-9</td>
<td>2-1-6-8</td>
</tr>
<tr>
<td></td>
<td>5-6</td>
<td>3-4-5-6</td>
<td>2-1-6-8</td>
<td>9-12</td>
<td>3-4-5-6</td>
<td>2-1-6-8</td>
<td>2-1-6-8</td>
</tr>
</tbody>
</table>

*Note:* Four lights in each series were always red and five were amber. The number of off-lights was gradually raised from zero (session 1) to the "operational" task condition of four per trial on session 4.

### TABLE 3

**EXPERIMENTAL CONDITIONS DEFINING THE VARIOUS GROUPS**
IN EXPERIMENT II

<table>
<thead>
<tr>
<th>Group</th>
<th>Training</th>
<th>Retention</th>
<th>Rehearsal</th>
<th>Retention</th>
<th>Retest</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2 days PT, 3 days WT</td>
<td>10 days</td>
<td>2 days WT</td>
<td>9 days</td>
<td>1 day WT</td>
</tr>
<tr>
<td>2</td>
<td>2 days PT, 3 days WT</td>
<td>10 days</td>
<td>2 days PT</td>
<td>9 days</td>
<td>1 day WT</td>
</tr>
<tr>
<td>3</td>
<td>2 days PT, 3 days WT</td>
<td>10 days</td>
<td>2 days NT</td>
<td>9 days</td>
<td>1 day WT</td>
</tr>
<tr>
<td>4</td>
<td>2 days PT, 8 days WT</td>
<td>10 days</td>
<td>2 days WT</td>
<td>9 days</td>
<td>1 day WT</td>
</tr>
<tr>
<td>5</td>
<td>2 days PT, 8 days WT</td>
<td>10 days</td>
<td>2 days PT</td>
<td>9 days</td>
<td>1 day WT</td>
</tr>
<tr>
<td>6</td>
<td>2 days PT, 8 days WT</td>
<td>10 days</td>
<td>2 days NT</td>
<td>9 days</td>
<td>1 day WT</td>
</tr>
</tbody>
</table>

*Note:* WT = Whole-task training (both tasks together).
PT = Part-task training (tracking and procedural separately).
NT = No task rehearsal.
The procedural task sequence used in the second study was modified somewhat from that used in Experiment I. The stimulus order (in terms of light pairs) was 1, 9, 2, 8, 3, 7, 4, 6, 5, where 1 refers to the top pair of lights, etc. The timing between stimuli and the programming of stimulus conditions on each trial (see table 2) remained the same as in the first experiment. The modification in the sequence did not affect the information in the sequence ($H = 2.808$ bits).

Experimental Design: Table 3 is a description of the six experimental groups and their associated experimental conditions. The rehearsal techniques were identical in procedure to those employed in Experiment I.

Subjects and Procedure: A total of 84 undergraduate males served in the experiment. All were volunteers who received $1.00 per experimental session. In addition, all Ss were “experienced” in that all had served in a prior study which involved learning a three-dimensional tracking task. The Ss were matched (assigned to groups) on the basis of their final tracking performance in this prior experiment. Since the earlier study had manipulated an augmented feedback variable in a transfer-of-training paradigm, it was felt that biasing effects would be slight and the advantage of good matching data outweighed any disadvantage of using experienced Ss.

The Ss trained for either 1 or 2 weeks, depending upon the group. The training procedures were similar to those in Experiment I, except that part training procedures were used on only the first 2 days, instead of for 4 days as in the earlier study. The Ss returned 10 days following the end of training for 2 days of rehearsal. As stated above, the rehearsal conditions were similar to those of Experiment I except for the procedural task sequencing. The Ss returned again 9 days after the rehearsal sessions for a single retention test session.

RESULTS

Experiment I

The data for both the tracking and the procedural tasks were examined using an analysis of variance. Separate analyses were conducted on the training data, the retention data, and on difference scores (training-retention).

Tracking Task: Figure 1 presents the tracking performance for all four groups during the 8 days of training, 2 days of rehearsal, and the retention test. Performance is given in terms of average error (in inches of display-scale displacement). Sessions 1-4 consisted of part training (see Procedure, above) and therefore are not directly comparable to the scores during sessions 5-8 or during the retention test session.

From figure 1 it is evident that the groups were not completely matched—this is especially noticeable during the first 2 days of whole training (sessions 5 and 6) where group 1 was substantially poorer in performance than were the other groups. An analysis of variance was performed on the last six trials of the final day of training (session 6). This analysis (see table 4) indicated that (a) the groups did not differ significantly one from another at the end of training, and (b) there was no increase in proficiency during the last six
Figure 1. Average Tracking Performance for Each Experimental Group during Training, Rehearsal, and Retention Test (Experiment I).

### TABLE 1

**ANALYSES OF VARIANCE PERFORMED ON FINAL PERFORMANCE AT END OF TRAINING (SESSION 8) IN EXPERIMENT I**

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Tracking MS</th>
<th>Commissive Errors MS</th>
<th>Omissive Errors MS</th>
<th>Response Time MS</th>
<th>F</th>
<th>F</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groups (G)</td>
<td>3</td>
<td>6480.63</td>
<td>7.158</td>
<td>0.367</td>
<td>6255.69</td>
<td>2.82*</td>
<td>1.58</td>
<td>2.41</td>
</tr>
<tr>
<td>Ss/Groups</td>
<td>56</td>
<td>6750.23</td>
<td>3.136</td>
<td>0.232</td>
<td>2600.52</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Trials (T)</td>
<td>5</td>
<td>982.04</td>
<td>0.336</td>
<td>0.071</td>
<td>377.64</td>
<td>1.05</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>T x G</td>
<td>13</td>
<td>529.70</td>
<td>0.634</td>
<td>0.062</td>
<td>577.52</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>T x Ss/G</td>
<td>280</td>
<td>938.29</td>
<td>0.781</td>
<td>0.105</td>
<td>656.01</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

* p < .05
training trials, indicating that the groups were beginning to asymptote in performance.

Performance during rehearsal on the tracking task was as expected: the whole-task group performed least proficiently, the part-rehearsal group did noticeably better, and the simplified group, having less demanding task dynamics, attained the most proficient performance.

Retention test scores were analyzed several ways. First, an analysis of variance was performed using the actual retention test scores. This analysis is shown in table 5. No significant differences were found between groups in terms of their absolute levels of retention performance ($F = 2.47, df = 3/56$). The order of efficiency from high to low of the four groups in terms of absolute retest performance was group 3 (simplified rehearsal), group 2 (part-task rehearsal), group 4 (no rehearsal), and group 1 (whole-task rehearsal).

The lack of statistical significance, plus the order of merit itself (whole-task rehearsal being less efficient than no rehearsal) led to a question of the legitimacy of analyzing only the absolute retention values. Initially this type of analysis was felt to be justified on the basis of the earlier reported lack of significant group differences at the end of training. However, as mentioned earlier, the whole-task rehearsal group was consistently the least proficient group throughout training. Because of this, an examination of difference scores (training-retest performance) was considered as being a more appropriate measure of rehearsal efficiency. The results of this analysis are shown in table 6.

Note that in this analysis the groups do differ significantly ($p < .01$). However, this difference is in terms of relative loss in skill over the retention period. Table 7 expresses these decrements in performance in terms of the average error metric. It is interesting to examine the order of merit of the

### Table 5

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Tracking</th>
<th>Commissive Errors</th>
<th>Omissive Errors</th>
<th>Response Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MS</td>
<td>F</td>
<td>MS</td>
<td>F</td>
</tr>
<tr>
<td>Rehearsal</td>
<td>3</td>
<td>38480.3</td>
<td>2.47</td>
<td>22.65</td>
<td>3.98*</td>
</tr>
<tr>
<td>Error</td>
<td>56</td>
<td>15555.7</td>
<td>0.2877</td>
<td>29.57</td>
<td></td>
</tr>
<tr>
<td>Trials</td>
<td>5</td>
<td>40819.9</td>
<td>1.24</td>
<td>1.64</td>
<td>1.32</td>
</tr>
<tr>
<td>T x R</td>
<td>15</td>
<td>4654.1</td>
<td>1.24</td>
<td>0.2334</td>
<td></td>
</tr>
<tr>
<td>T x E</td>
<td>280</td>
<td>4818.1</td>
<td>1.24</td>
<td>0.2334</td>
<td></td>
</tr>
</tbody>
</table>

* $p < .05$

** $p < .01$
TABLE 6
ANALYSIS OF VARIANCE RESULTS OF THE DIFFERENCE SCORES
(TRAINING SESSION - RETENTION SESSION) FOR ALL
PERFORMANCE MEASURES IN EXPERIMENT I

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Tracking MS</th>
<th>F</th>
<th>Commissive Errors MS</th>
<th>F</th>
<th>Omissive Errors MS</th>
<th>F</th>
<th>Response Time MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rehearsal</td>
<td>3</td>
<td>6707.3</td>
<td>5.32**</td>
<td>1.8906</td>
<td>2.17</td>
<td>0.1302</td>
<td>1.39</td>
<td>43.9531</td>
<td>1.59</td>
</tr>
<tr>
<td>Within Cell</td>
<td>56</td>
<td>1259.5</td>
<td>0.8721</td>
<td>0.0936</td>
<td>27.6738</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

** p < .01

TABLE 7
MEAN PERFORMANCE MEASURES FOR SUBJECTS DURING THE LAST TRAINING
SESSION AND THE RETENTION TEST SESSION FOR EACH OF
THE FOUR REHEARSAL CONDITIONS IN EXPERIMENT I

<table>
<thead>
<tr>
<th>Rehearsal Group</th>
<th>Training Tracking Performance (in Inches)</th>
<th>Retention Test</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole Rehearsal</td>
<td>0.122</td>
<td>0.156</td>
<td>-0.034</td>
</tr>
<tr>
<td>Part Rehearsal</td>
<td>0.131</td>
<td>0.125</td>
<td>0.006</td>
</tr>
<tr>
<td>Simplified Rehearsal</td>
<td>0.118</td>
<td>0.118</td>
<td>0.000</td>
</tr>
<tr>
<td>No Rehearsal</td>
<td>0.112</td>
<td>0.143</td>
<td>-0.031</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rehearsal Group</th>
<th>Number of Commissive Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole Rehearsal</td>
<td>0.31</td>
</tr>
<tr>
<td>Part Rehearsal</td>
<td>0.85</td>
</tr>
<tr>
<td>Simplified Rehearsal</td>
<td>0.81</td>
</tr>
<tr>
<td>No Rehearsal</td>
<td>1.01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rehearsal Group</th>
<th>Number of Omissive Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole Rehearsal</td>
<td>0.08</td>
</tr>
<tr>
<td>Part Rehearsal</td>
<td>0.09</td>
</tr>
<tr>
<td>Simplified Rehearsal</td>
<td>0.21</td>
</tr>
<tr>
<td>No Rehearsal</td>
<td>0.03</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rehearsal Group</th>
<th>Response Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole Rehearsal</td>
<td>6.28</td>
</tr>
<tr>
<td>Part Rehearsal</td>
<td>6.56</td>
</tr>
<tr>
<td>Simplified Rehearsal</td>
<td>7.09</td>
</tr>
<tr>
<td>No Rehearsal</td>
<td>5.94</td>
</tr>
</tbody>
</table>
rehearsal procedures in terms of these difference scores. The part-practice group (group 2) experienced the least loss of skill (in fact, there is even a slight increment evident). This procedure is followed, in order of efficiency, by group 3 (simplified rehearsal), group 4 (no rehearsal), and group 1 (whole-task rehearsal). Thus, the same order of merit is observed with this performance measure. In order to determine which of the groups (rehearsal conditions) differed significantly from each other, the Duncan Multiple Range Test (ref. 6) was used to test for differences between all possible pairs of means. This analysis indicated that the means in table 7 may be grouped into two sets. Both the part-rehearsal group \( p < .01 \) and the simplified-rehearsal group \( p < .00 \) differed significantly from the whole-rehearsal and no-rehearsal groups, but did not differ from each other, and the whole-rehearsal group did not differ from the no-rehearsal group.

Analyses of the data were also performed using the measure defined as (last training session) - (first retest trial) performance. This measure is more sensitive to initial or immediate retention performance than the measure (last training session) - (retest session). No significant differences among groups were found with this measure.

**Procedural Task:** Three separate measures were obtained of procedural task performance—commissive errors, omissive errors, and response time. Table 7 shows the performance of each of the four groups on each of these measures. The performance listed is for the last six trials of training, the six retention test trials, and for the difference scores (training - retention performance).

As was the case with tracking performance, all of the procedural task measures were analyzed initially using only the retention test data. From table 5 it is apparent that two of the three measures (commissive errors and response time) indicated significant retention differences as a function of rehearsal conditions. In addition, both commissive errors and omissive errors showed a significant trial effect. This was the result of rapid reacquisition of skill on these two measures during the brief, six-trial retention session.

The analysis of variance performed on the difference scores (training-retest) did not result in the same findings, i.e., none of the procedural metrics demonstrated significant group differences as a function of rehearsal conditions (see table 6). An analysis of training-first retest trial gave identical results. An interesting pattern of consistency, however, may be observed by the order of merit evidenced (in terms of difference scores) by all three procedural task measures. These may be seen in table 7. For commissive and omissive errors the order of efficiency was identical: part rehearsal was best, followed in sequence by simplified rehearsal, whole-task rehearsal, and no rehearsal. Note that this order is almost identical to that found with the tracking task scores. A slight reversal in order was evident with the response time measure. Part rehearsal was still better, but whole rehearsal was superior to simplified, which in turn was more efficient than no practice. As mentioned above, however, none of these differences were statistically significant.

**Experiment II**

Analyses similar to those used in Experiment I were also used to examine the data of the second study.

**Tracking Task:** Figures 2 and 3 summarize the tracking performance of the six groups during training, rehearsal, and retention test. Figure 2 shows the
performance of the 1-week training groups, while figure 3 shows the performance for those groups having 2 weeks of training. Scores during the first 2 days represent part-practice performance and therefore are not directly comparable to subsequent measures obtained during training (sessions 3-5 or 3-10) or during the retention test.

The training data for the 1- and for the 2-week groups are quite similar. At the end of session 5 all groups were performing with an average error of approximately 0.30 inch. The 2-week training groups continued to improve during the additional 5 days of practice, reducing their average error scores to about 0.16 inch on session 10. The obvious implications to be drawn from figures 2 and 3 are that the 1-week groups were not trained to any approximation of asymptotic performance, while the 2-week training groups do show some indication of a leveling off in tracking skill by the tenth session. An analysis of variance performed on the training scores partially supports this conclusion (see table 8). This analysis utilized the three blocks of four trials experienced by each S on his last day of training. The significant trial-blocks effect ($F = 31.96$, $df = 2/156$) indicates that the groups were still improving in performance at the end of training, and the significant trials by groups interaction ($F = 8.84$, $df = 2/156$) indicates that the relative improvement was greater for the 1-week training groups than for the 2-week training groups. Thus, groups 1, 2, and 3 were
Figure 2. Average Tracking Performance for Each Group during Training, Rehearsal, and Retention Test for the One-Week Training Groups in Experiment II.

The training data for the 1- and for the 2-week groups are quite similar. At the end of session 5 all groups were performing with an average error of approximately 0.30 inch. The 2-week training groups continued to improve during the additional 5 days of practice, reducing their average error scores to about 0.16 inch on session 10. The obvious implications to be drawn from figures 2 and 3 are that the 1-week groups were not trained to any approximation of asymptotic performance, while the 2-week training groups do show some indication of a leveling off in tracking skill by the tenth session. An analysis of variance performed on the training scores partially supports this conclusion (see table 8). This analysis utilized the three blocks of four trials experienced by each S on his last day of training. The significant trial-blocks effect ($F = 31.96$, $df = 2/156$) indicates that the groups were still improving in performance at the end of training, and the significant trials by groups interaction ($F = 8.84$, $df = 2/156$) indicates that the relative improvement was greater for the 1-week training groups than for the 2-week training groups. Thus, groups 1, 2, and 3 were
apparently still improving, while groups 4, 5, and 6 were reaching a plateau in tracking performance.

There was a significant difference in absolute skill level at the end of training as a function of the amount of practice (F = 39.37, df = 1/78), while the nonsignificant rehearsal effect (an artificial variable at this stage since rehearsal had not yet occurred) is evidence of adequate matching of Ss within training levels at the beginning of training.

Because of the significant differences observed between groups due to amount of training, it was decided to use an analysis of difference scores (training-retention test) as a measure of retention. The analyses performed on the difference scores are shown in table 9.*

* It should be noted that the basic purpose of the difference-score analysis was to adjust the absolute retention values for final training performance. One weakness which the reader should keep in mind, however, is that because learning rate is a nonlinear function of practice (usually), unit loss of retention from a point near the asymptote may not be equal to a unit loss from some other base point.
### TABLE 8

**ANALYSIS OF VARIANCE RESULTS OF FINAL PERFORMANCE AT END OF TRAINING**

(SESSION 5 FOR GROUPS 1-3, SESSION 10 FOR GROUPS 4-6)

**IN EXPERIMENT II**

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Tracking</th>
<th>Commissive Errors</th>
<th>Omissive Errors</th>
<th>Response Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MS</td>
<td>F</td>
<td>MS</td>
<td>F</td>
</tr>
<tr>
<td>Amount of Training (A)</td>
<td>1</td>
<td>772852</td>
<td>39.37**</td>
<td>5.5299</td>
<td>18.61**</td>
</tr>
<tr>
<td>Rehearsal (R)</td>
<td>2</td>
<td>11040</td>
<td>0.2566</td>
<td>-</td>
<td>0.6270</td>
</tr>
<tr>
<td>A x R</td>
<td>2</td>
<td>13121</td>
<td>0.6045</td>
<td>2.03</td>
<td>0.3247</td>
</tr>
<tr>
<td>Subjects within Groups (Ss/G)</td>
<td>78</td>
<td>19630</td>
<td>0.2972</td>
<td>0.5004</td>
<td>22.9896</td>
</tr>
<tr>
<td>Trial Blocks (T)</td>
<td>2</td>
<td>5751.5</td>
<td>31.96**</td>
<td>0.8800</td>
<td>6.22**</td>
</tr>
<tr>
<td>T x A</td>
<td>2</td>
<td>15914</td>
<td>8.84**</td>
<td>0.9640</td>
<td>6.81**</td>
</tr>
<tr>
<td>T x R</td>
<td>4</td>
<td>453</td>
<td>-</td>
<td>0.0558</td>
<td>-</td>
</tr>
<tr>
<td>T x R x A</td>
<td>4</td>
<td>175</td>
<td>-</td>
<td>0.2166</td>
<td>1.52</td>
</tr>
<tr>
<td>T x Ss/G</td>
<td>156</td>
<td>1800</td>
<td>0.1415</td>
<td>0.1133</td>
<td>-</td>
</tr>
</tbody>
</table>

* *p < .05
** *p < .01

### TABLE 9

**ANALYSIS OF VARIANCE RESULTS OF THE DIFFERENCE SCORES**

(TRAINING SESSION - RETENTION TEST SESSION) FOR ALL PERFORMANCE MEASURES IN EXPERIMENT II

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Tracking</th>
<th>Commissive Errors</th>
<th>Omissive Errors</th>
<th>Response Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MS</td>
<td>F</td>
<td>MS</td>
<td>F</td>
</tr>
<tr>
<td>Amount of Training (A)</td>
<td>1</td>
<td>596,338</td>
<td>20.69**</td>
<td>5.80</td>
<td>6.13*</td>
</tr>
<tr>
<td>Rehearsal (R)</td>
<td>2</td>
<td>306,743</td>
<td>10.61**</td>
<td>0.24</td>
<td>-</td>
</tr>
<tr>
<td>A x R</td>
<td>2</td>
<td>223,284</td>
<td>7.75**</td>
<td>1.26</td>
<td>1.33</td>
</tr>
<tr>
<td>Subjects within Groups (Ss/G)</td>
<td>78</td>
<td>28,818</td>
<td>0.95</td>
<td>1.01</td>
<td>-</td>
</tr>
</tbody>
</table>

* *p < .05
** *p < .01

14
TABLE 10
MEAN PERFORMANCE MEASURES FOR SUBJECTS DURING THE LAST
TRAINING SESSION AND THE RETENTION TEST SESSION
FOR EACH OF THE GROUPS IN EXPERIMENT II

<table>
<thead>
<tr>
<th>Rehearsal Group</th>
<th>Amount of Training</th>
<th>Training</th>
<th>Retention Test</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Tracking Performance (in Inches)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whole 1 week</td>
<td>0.288</td>
<td>0.159</td>
<td>0.129</td>
<td></td>
</tr>
<tr>
<td>Part 1 week</td>
<td>0.242</td>
<td>0.187</td>
<td>0.054</td>
<td></td>
</tr>
<tr>
<td>None 1 week</td>
<td>0.275</td>
<td>0.275</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Whole 2 weeks</td>
<td>0.155</td>
<td>0.146</td>
<td>0.009</td>
<td></td>
</tr>
<tr>
<td>Part 2 weeks</td>
<td>0.158</td>
<td>0.152</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td>None 2 weeks</td>
<td>0.159</td>
<td>0.160</td>
<td>-0.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Number of Commissive Errors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whole 1 week</td>
<td>0.56</td>
<td>0.32</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>Part 1 week</td>
<td>0.69</td>
<td>0.46</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>None 1 week</td>
<td>0.41</td>
<td>0.31</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>Whole 2 weeks</td>
<td>0.26</td>
<td>0.22</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Part 2 weeks</td>
<td>0.23</td>
<td>0.29</td>
<td>-0.06</td>
<td></td>
</tr>
<tr>
<td>None 2 weeks</td>
<td>0.28</td>
<td>0.21</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Number of Omissive Errors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whole 1 week</td>
<td>0.52</td>
<td>0.22</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>Part 1 week</td>
<td>0.48</td>
<td>0.28</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td>None 1 week</td>
<td>0.27</td>
<td>0.34</td>
<td>-0.07</td>
<td></td>
</tr>
<tr>
<td>Whole 2 weeks</td>
<td>0.31</td>
<td>0.19</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>Part 2 weeks</td>
<td>0.18</td>
<td>0.16</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>None 2 weeks</td>
<td>0.21</td>
<td>0.18</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Response Time</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whole 1 week</td>
<td>11.26</td>
<td>10.16</td>
<td>1.10</td>
<td></td>
</tr>
<tr>
<td>Part 1 week</td>
<td>10.26</td>
<td>10.11</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>None 1 week</td>
<td>10.38</td>
<td>10.90</td>
<td>-0.52</td>
<td></td>
</tr>
<tr>
<td>Whole 2 weeks</td>
<td>10.51</td>
<td>9.38</td>
<td>1.13</td>
<td></td>
</tr>
<tr>
<td>Part 2 weeks</td>
<td>9.06</td>
<td>9.21</td>
<td>-0.15</td>
<td></td>
</tr>
<tr>
<td>None 2 weeks</td>
<td>9.62</td>
<td>9.36</td>
<td>0.26</td>
<td></td>
</tr>
</tbody>
</table>

Both main effect variables, amount of training and type of rehearsal, were found to have a significant effect upon retention. The 2-week training groups (groups 4, 5, and 6) showed little change in skill over the retention interval, while the groups receiving only 1 week of practice showed substantial skill increase over the retention interval. The effect of rehearsal procedure was quite evident, with whole-task rehearsal being superior to part-task rehearsal, which in turn was superior to the no-rehearsal condition. This was especially true for the groups receiving only 1 week of training, as indicated by the significant interaction between amount of training and rehearsal method (see table 10). Thus, it would appear that given a moderate amount of training the type
of rehearsal, or even its presence, is not too critical, while rehearsal is quite effective for groups receiving lesser amounts of original training. These conclusions are supported by further analysis: the Duncan Multiple Range Test (ref. 8) indicated that for the groups receiving 1 week of training significant differences (p < .01) obtained between all rehearsal conditions. In contrast, for the 2-week training groups, no significant differences (p > .05) occurred as a function of rehearsal conditions.

One obvious problem with the above findings lies in the fact that the relatively poorly trained groups retained more than the well-trained groups, indicating that they were merely continuing original acquisition during rehearsal. A more pertinent and practical question is to ask how the groups compare on final test performance per se. An analysis of variance was carried out on the first retention trial scores to examine this question. Method of training was significant (F = 10.83, df = 2/78; p < .01) as was amount of training (F = 4.13, df = 1/78; p < .05). No significant interaction was obtained. Whole training was superior to part training, which in turn was superior to no rehearsal—identical ordering to that found with the first analysis.

Procedural Task: The analyses of final training performance for all three procedural task measures are shown in table 8. Groups differed in skill as a function of amount of training in terms of both commissive errors (p < .01) and omissive errors (p < .05), although response time did not so discriminate among groups. All three measures indicated that performance was still improving during the last training session, as all three scores showed a significant trials effect. However, the significant interaction of trials by amount of training observed with commissive errors indicated that, at least for this measure, the increase in skill proficiency due to trials was primarily a function of the 1-week training groups. Plots of the learning data substantiated these conclusions. It was again decided to utilize the difference-score method of analysis, since the groups did differ at the end of training. These analyses are listed in table 9. Table 10 shows the mean scores per individual per trial at the end of training and during retention test for all groups and all measures.

The most obvious fact presented in table 9 is that there was only one significant effect observed across all three procedural task difference score measures and that difference was found with commissive errors where amount of training did influence the degree of skill loss (p < .05). This was not too surprising, since similar results had been obtained in Experiment I. Examination of these nonsignificant differences (a procedure that is never really legitimate but often quite insightful), which are given in table 10, did show that their direction had a consistency which is quite remarkable. This is true especially when the order of efficiency is found with the tracking data (which were significant) is included. Table 11 shows these orders of merit.

Thus, in all cases but one (comissive errors, 2 weeks of training) whole rehearsal was most advantageous. Exactly the same order is to be observed for all measures for the groups having only 1 week of training. Those groups having 2 weeks of training, on the other hand, are somewhat more inconsistent in terms of method efficiency. It would appear that (a) the procedural task metrics lacked sensitivity, and that (b) this insensitivity was furthered by additional training. Both (a) and (b) above tended to mask any visible influence of rehearsal.

Analyses were also performed on the first retention test trial scores. The results of this analysis procedure for all three procedural measures provided two
TABLE 11
OBSERVED ORDER OF EFFICIENCY OF THE VARIOUS REHEARSAL METHODS
AS A FUNCTION OF AMOUNT OF TRAINING AND DEPENDENT VARIABLE

<table>
<thead>
<tr>
<th>Amount of Training</th>
<th>Tracking Error</th>
<th>Comissive Errors</th>
<th>Omissive Errors</th>
<th>Response Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Week</td>
<td>Whole</td>
<td>Whole (Tie)</td>
<td>None</td>
<td>Whole</td>
</tr>
<tr>
<td></td>
<td>Part</td>
<td>Part</td>
<td>Part</td>
<td>Part</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>2 Weeks</td>
<td>Whole</td>
<td>None</td>
<td>Whole</td>
<td>Whole</td>
</tr>
<tr>
<td></td>
<td>Part</td>
<td>Whole</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>Part</td>
<td>Part</td>
<td>Part</td>
</tr>
</tbody>
</table>

* Differences were statistically significant.

significant \((p < .05)\) differences—a method × amount of training interaction with comissive errors and a method effect for omissive errors. The interaction effect, when examined, indicated whole rehearsal to be most efficient for the 1-week training groups, but no rehearsal differences for groups having 2 weeks of training. The method effect found with omissive errors showed whole rehearsal to be superior to part rehearsal which in turn was superior to no rehearsal. Thus, the first retention trial performance provided a more sensitive measure of rehearsal effects than did measures using the entire rehearsal session. Apparently the effects of rehearsal differences tend to disappear rapidly with additional retention trials.

DISCUSSION

There were several aspects of the findings in both studies which were anticipated from prior research. The tracking performance measures were the most sensitive indicators of performance in both studies which was not unexpected, as earlier studies had found similar results (refs. 2, 14). The lack of sensitivity of the procedural task may be a result of the task being sufficiently easy so that accurate performance becomes well established early in original training, thus being relatively unaffected by the independent variables which were applied later in the rehearsal period of the experiment. The absolute magnitude of the procedural task error scores made by Ss in both studies supports such a hypothesis as such errors were quite infrequent. During retention test in Experiment II, for example, Ss were averaging only about one comissive error in every three trials.

The powerful effect of additional training observed in the second study was also anticipated. A study by Naylor et al. (ref. 14) had shown previously that sufficient practice reduced the effects of both retention interval length and
task organization. Apparently it is also able to minimize rehearsal effects, as rehearsal had little or no influence upon retention for groups having 2 weeks of training in Experiment II. This would suggest that any long-term retention experimentation should include amount of training as one of the independent variables so that other, but potentially less influential, variables be given an opportunity to demonstrate their effects.

The major inconsistency between Experiments I and II was the reversal in relative efficiency of the rehearsal methods. The findings of Experiment I demonstrated in general that part rehearsal was superior, followed in order by simplified, whole, and no rehearsal. However, the Experiment II findings generally showed whole rehearsal best, followed by part and none, for the 1-week training groups. As mentioned above, 2 weeks of training appeared to completely negate rehearsal effects. Thus, viewing the two experiments together one finds whole rehearsal superior after 5 days of training (Experiment II), part rehearsal superior after 8 days of training (Experiment I), and no rehearsal effect at all given 10 days of training (Experiment II).

The notion that the relative efficiency of rehearsal techniques should be related to amount of practice is not surprising in itself, of course. Since rehearsal is logically just an extension of practice on the task, it is to be expected that increasing amounts of practice will tend to reduce differences between practice methods. What is surprising, however, is to find the relative order of merit among practice methods to vary as a function of amount of practice prior to rehearsal. One possible explanation for this may exist in the fact that the training procedures for Experiments I and II were not directly comparable. In Experiment I all groups received four full days of part practice prior to combined task practice, while in Experiment II only two days of component practice were given before the Ss combined the subtasks. Thus, not only did the 5- and 6-day training situations differ in quantity, they also differed in qualitative manner. The 8-day groups spent 50% of their learning on part training, as compared to 40% for the 5-day groups, who spent the rest (60%) on whole-task practice. This may indicate that rehearsal method efficiency is related not only to amount of practice, but to the kind of practice preceding it.

Such a conclusion would certainly be compatible with the assumption that a complex task involves not only the acquisition of skill on each task component, but also the acquisition of a skill in time sharing. In this task, then, the skill of time sharing (supposedly acquired during whole-task practice) was as necessary as component skill (acquired through either part or whole practice). The qualitatively different training schedules would therefore provide different degrees of training on the three aspects of total task skill. Thus, the 5-day groups, which spent 60% of their time on whole practice, responded best to whole rehearsal. Similarly, the 8-day groups, which spent a greater portion of their time on part practice, both in relative and absolute terms, responded best to part rehearsal.

In summary, it is suggested that (a) the relative effect of various rehearsal methods is inversely related to the amount of initial training received and (b) the most efficient method may depend upon the method (in addition to the amount) of original training.
Two separate experiments were conducted to study the effects of various task rehearsal methods as a means of facilitating the retention of a time-shared task consisting of a continuous component (compensatory tracking) and a discrete S-R component (procedural monitoring). Experiment I investigated four rehearsal conditions (whole-task, part-task, simplified-task, and none), while Experiment II examined three (whole-task, part-task, and none). Amount of training (5 or 10 days) was an additional variable in Experiment II. Part-task rehearsal was found to be most efficient in Experiment I, while whole-task rehearsal was more efficient in the second study.

The training procedures for Experiments I and II were not directly comparable. In Experiment I all groups received four full days of part practice prior to combined task practice, while in Experiment II only two days of component practice were given before the Ss combined the subtasks. Thus, comparisons across experiments involved qualitative as well as quantitative training differences between groups. The 8-day groups (Experiment I) spent 50% of their learning on part practice, as compared to 40% for the 5-day groups and 20% for the 10-day groups (Experiment I). This may indicate that rehearsal method efficiency is related not only to the amount of practice, but to the composition of practice methods initially used.
REFERENCES


Two studies were performed to investigate the influence of various methods of task rehearsal upon the retention of a time-shared task. Experiment I examined retention as a function of four rehearsal conditions (part task, whole task, simplified task, and none). Subjects in each of the groups trained for 8 days, returned 6 days later for 2 days of rehearsal, and then returned again after 7 more days for a retention test. Experiment II examined retention both as a function of 3 rehearsal methods (whole, part, and none) and amount of training (5 or 10 days). Subjects returned for 2 days of rehearsal 10 days after completion of training and then returned for rehearsal methods. In Experiment I part rehearsal was superior, while in Experiment II whole rehearsal was found to be best.