DOCUMENT RESUME

ED 272 529 TM 860 442

AUTHOR	Klapp, Stuart T.
TITLE	Memory and Processing Limits in Decision-Making.
INSTITUTION	California State Univ., Hayward.
SPONS AGENCY	Air Force Human Resources Lab., Brooks AFB, Texas.
REPORT NO	AFHRL-TR-85-60
PUB DATE	May 86
NOTE	48p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	College Students; *Decision Making; Higher Education; *Learning Processes; *Memory; Performance Factors; *Problem Solving; Recall (Psychology); Recognition (Psychology); Retention (Psychology); Short Term Memory; Simulation

ABSTRACT

According to the classical working memory perspective, tasks such as command and control decision-making should be performed less effectively if extraneous material must be retained in short-term memory. Only marginal support for this prediction was obtained for a simulation involving scheduling trucking and transportation missions, although robust memory loading did occur in the case of more abstract tasks. This pattern of results is not consistent with the classical single system view of working memory, and suggests that working memory may have multiple subsystems. Apparently, some memory demands can be isolated from the memory system used to accomplish the trucking and transportation task. Increased understanding of the nature of working memory subsystems and of isolation of task components into separate subsystems should aid in design of command and control task environments. (Two pages of references conclude the document.) (Author)

AFHRL-TR-85-60

AIR FORCE HUMAN RESOURCES LABORATORY

ED272529

MEMORY AND PROCESSING LIMITS IN DECISION-MAKING

Stuart T. Klapp

Department of Psychology
California State University, Hayward
Hayward, California 94542

LOGISTICS AND HUMAN FACTORS DIVISION
Wright-Patterson Air Force Base, Ohio 45433-6503

May 1986

Final Report for Period July 1983 - September 1985

Approved for public release; distribution is unlimited.

TM 860 442

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235-5601

NOTICE

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 Public Affairs Office 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.

LAWRENCE E. REED
Contract Monitor

BERTRAM W. CREAM, Technical Director
Logistics and Human Factors Division

DENNIS W. JARVI, Colonel, USAF
Commander

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFHRL-TR-85-60
6a. NAME OF PERFORMING ORGANIZATION California State University, Hayward 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Logistics and Human Factors Division
6c. ADDRESS (City, State, and ZIP Code) Hayward, California 94542	7b. ADDRESS (City, State, and ZIP Code) Air Force Human Resources Laboratory Wright-Patterson Air Force Base, Ohio 45433-6503
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Air Force Human Resources Laboratory 8b. OFFICE SYMBOL (If applicable) HQ AFHRL	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F33615-83-K-0039
8c. ADDRESS (City, State, and ZIP Code) Brooks Air Force Base, Texas 78235-5601	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 62205F; PROJECT NO. 3017; TASK NO. 08; WORK UNIT ACCESSION NO. 07

11. TITLE (Include Security Classification)
Memory and Processing Limits in Decision-Making

12. PERSONAL AUTHOR(S)
Klapp, Stuart T.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM Jul 83 TO Sep 85	May 1986	46

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP
05	09	
05	10	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
command and control
short-term memory

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

According to the classical working memory perspective, tasks such as command and control decision-making should be performed less effectively if extraneous material must be retained in short-term memory. Only marginal support for this prediction was obtained for a simulation involving scheduling trucking and transportation missions (Experiments 3 and 4), although robust memory loading did occur in the case of more abstract tasks (Experiments 1 and 2). This pattern of results is not consistent with the classical single system view of working memory, and suggests that working memory may have multiple subsystems. Apparently, some memory demands can be isolated from the memory system used to accomplish the trucking and transportation task. Increased understanding of the nature of working memory subsystems and of isolation of task components into separate subsystems should aid in design of command and control task environments.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Nancy A. Perrigo, Chief, STINFO Office	22b. TELEPHONE (Include Area Code) (512) 536-3877; 22c. OFFICE SYMBOL AFHRL/TSR

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

SUMMARY

According to classical theory, people have limited immediate memory in which information can be retained while working on a problem, such as decision-making for command and control. This theory predicts that, when extraneous material must be retained in memory along with relevant information, the presence of the extraneous material will "load" memory, and thereby reduce performance on the problem-solving task. As a test of this prediction, a simulation involving scheduling trucking and transportation missions was designed, and performance was tested with and without the presence of the extraneous information. Only marginal support for the predicted interference was obtained, although corresponding memory loading did occur in the other, less realistic, situations that were also studied. Apparently some memory demands can be isolated from the memory system used to accomplish the trucking and transportation task.

PREFACE

The research reported herein was performed under Contract F33615-83-K-0039 for the Air Force Human Resources Laboratory (AFHRL), during the period July 1983 - September 1985. Dr. Lawrence E. Reed was the AFHRL contract monitor. The author thanks Dr. Reed for his insightful advice throughout the project. The author also thanks Stephanie Milgrom, systems analyst in private practice, who programmed the trucking and transportation simulation, and the following students at California State University, Hayward: Ed Dum, who designed the "games" and collected the data of Experiment 3, Warren Shea who designed the "games" and collected the data of Experiment 4, and Piyale Comert, Joseph Schwarz, and Bob Sanchez, who collected the data of Experiments 1 and 2.

TABLE OF CONTENTS

Table of Contents (Continued)

8

LIST OF FIGURES

LIST OF TABLES

MEMORY AND PROCESSING
LIMITS IN DECISION-MAKING

I. INTRODUCTION AND OVERVIEW

Modern command and control, decision and planning tasks often require that the planner keep track of large amounts of information of only temporary utility. The information would consist of facts applicable to the particular mission under consideration (mission requirements and resources currently available). The necessity of keeping track of and using a large amount of information that is subject to change creates the possibility of overload of the decision maker's temporary memory.

This report describes a set of experiments with the goal of analyzing decisions and planning from the perspective known as "working memory." Working memory is a term which implies that immediate memory functions as the "site of ongoing cognitive activities -- for instance, meaningful elaboration of words, symbol manipulation such as that involved in mental arithmetic, and reasoning" (Klatsky, 1980, p. 87). Thus, working memory is the "space in which information can be stored temporarily while it is being processed" (Klapp, Marshburn, & Lester, 1983, p. 240). It is "the part of the memory system where active information processing takes place" (Chase & Ericcson, 1982, p. 40). According to this hypothesis, overloads in working memory may significantly limit human performance in reasoning and decision-making situations. This viewpoint is proclaimed in most current textbooks in Cognitive Psychology (see Klapp et al., 1983 for a review), and in current texts in Human Factors (Kantowitz & Sorkin, 1983; Wickens, 1984).

Surprisingly, given the nearly universal acceptance of the general working-memory framework, there have been no demonstrations that limitations in working memory actually do produce problems in decision-making and planning. The research has been focused instead on interference between "loads" imposed in short-term memory and certain laboratory tasks done concurrently with retention of the load. Experiments 3 and 4, the major effort reported herein, were attempts to demonstrate that extraneous memory loads might "fill up" space in the limited working-memory system and thereby reduce performance in a decision-making task involving planning of trucking and transportation missions.

All of the reported experiments make use of a general approach in which extraneous information was introduced into short-term memory, and other tasks (decision-making or otherwise) must be done while the extraneous memory "load" was also being retained. Performance on the "embedded" task was measured in the "loaded" condition, and in appropriate unloaded control conditions.

As indicated in the review of literature in the next section of this report, previous demonstrations of the effect of memory loading on concurrent tasks have been limited to cases in which the order of occurrence of the embedded-task memory items must be retained. By contrast, the trucking and transportation simulation (Experiments 3 and 4) does not require retention of order information. Thus, it is desirable to determine if the effects of memory loading occur in simple cases that do not require order information, before proceeding to the more complex transportation simulation. This is the goal of the first two experiments.

II. REVIEW OF PREVIOUS RESEARCH

According to the working-memory hypothesis, retention of extraneous material in memory should reduce performance on concurrent tasks that also require use of working memory. This prediction has led to experiments in which each trial is comprised of the following sequence of events: (a) present an extraneous memory "load" such as a string of letters, (b) perform an embedded task which presumably requires working memory, and (c) recall the extraneous memory load. Thus, the embedded task is performed concurrently with retention of a memory load. According to the working-memory viewpoint, the load and the embedded task should exhibit mutual interference. That is, the presence of the load should reduce performance on the task, compared to a no-load control, and the presence of the task should reduce recall of the load compared to a no-task control. Results of this form were first reported by Baddeley and Hitch (1974) and Hitch and Baddeley (1976).

Much of this research seems to overlook the implications of the finding that rehearsal early in the retention interval protects short-term memory from subsequent loss from distraction. For example, Dillon and Reid (1969) used easy versus difficult distraction tasks during the retention interval to study the relative importance of early and late rehearsal. It was assumed that the easy distractor would permit more rehearsal than would the difficult distractor. Retention was better with the easy distractor first followed by the difficult distractor than with the reverse order, supporting the view that early rehearsal is more important than late rehearsal. Similarly, the original Peterson and Peterson (1959) report included the finding that recall was better when an initial rehearsal period of 3 sec. preceded the 18-sec. distracted delay than when no such initial rehearsal was provided. Furthermore, subjects report more rehearsals early than late in the retention interval (Kroll, Kellicut, & Parks, 1975). Presumably subjects rehearse more earlier than later in the interval because late rehearsal is not needed if early rehearsal has occurred. Finally, reaction time (RT) to a secondary task can be used to assess the amount of rehearsal, if one assumes that RT is longer when rehearsals are in progress. Consistent with the other studies, secondary task RT was found to be longer early than late in the retention interval (Stanners, Meunier, & Headley, 1969). Thus, there is ample reason to suppose that much of the short-term memory retention loss in brief distracted intervals can be prevented by a few rehearsals early in the interval and that, if given the opportunity, people will engage in such early rehearsals.

This conclusion has important implications for interpreting findings from the memory load paradigm. It is possible that some of the effects previously attributed to interference between <u>retention</u> of the extraneous memory load and performance of the embedded task are really due to interference between <u>initial rehearsal</u> of the load and the embedded task. If the embedded task was postponed until the initial rehearsal was complete, then there might be no interference between <u>retention</u> of the load (after completion of initial rehearsal) and the embedded task. Consistent with this interpretation, for some tasks, mutual interference between memory load and embbedded tasks occurs with no initial rehearsal interval, but vanishes if a delay for initial rehearsal is provided between the input of the memory load and the presentation of the embedded task (Klapp et al., 1983, Experiments 6, 7). Thus, for these situations, the observed interference is between the embedded task and the initial rehearsal. Passive retention of the memory load did not interfere with the embedded task.

On the other hand, mutual interference can occur even with an initial rehearsal interval when both the memory load and the embedded task require retention of the order of occurrence of events (Klapp et al., 1983, Experiment 8; Klapp & Philipoff, 1983). For these situations, retention of the memory load is incompatible with performing the embedded task. This finding shows that initial rehearsal does not cause the memory load items to enter some system (such as long-term memory) where they could not interfere with any embedded task.

Therefore, memory load studies are classified on the basis of whether a delay for initial rehearsal was provided between the memory load input and the embedded task. For those studies that did provide such a delay (and hence are valid according to the present analysis), the studies are further classified as showing or not showing interference between the load and the embedded task. Studies that did not include the rehearsal delay, and hence will not be considered in detail, include: Baddeley and Hitch (1974), Crowder (1967), Hitch and Baddeley (1976), Johnston, Greenberg, Fisher, and Martin (1970), Logan (1979, 1980), Reisberg (1983), Shulman and Greenberg (1971), and Wanner and Shiner (1976).

Some other studies are difficult to classify. For example, in the experiment by Jonides (1981) subjects initiated the embedded task on their own and, hence, could have either provided or not provided the delay. In this study, there was mutual interference between memory load and reaction time for an embedded visual attention-shift task. A way of loading memory using recall uncertanity rather than the embedded task paradigm was developed by Richardson (1984). In this procedure, two sets of items were presented simultaneously, and a cue indicated which set was to be recalled. Cues randomly occurred either in advance of or after stimulus presentation, with the latter case assumed to represent concurrent memory loading. This paradigm avoids problems of delay of processing of one task while rehearsing the other, but introduces other problems in interpretation, including the relative roles of selective attention at input versus memory loading as ways to account for the performance decrement for the delayed-cue condition.

Of greater interest are studies using the embedded task approach which did include a delay for initial rehearsal. Some of these studies show mutual interference between memory load and the embedded task, and others do not. Interference was not observed for the following four embedded tasks: First was the "Fitts' Law" target tapping (Roediger, Knight, & Kantowitz, 1977), and second was the true/ false judgments about number relations (Klapp et al., 1983, Experiment 6); neither of these embedded tasks appears to require much use of memory, and this may explain the lack of any effect of loading. However, the remaining negative instances clearly do require use of memory. The third was the modified Sternberg scanning (Klapp et al., Experiment 7); this embedded task involved the simultaneous presentation of embedded memory items, and in that sense was different from the other experiments, all of which used sequential presentation of both memory load and embedded task memory items. Finally, the fourth was the "missing scan" task where the subject reports which of the nine non-zero digits had not appeared (Klapp & Philipoff, 1983).

By contrast, mutual interference between memory load does occur for other embedded tasks (even with an interval for initial rehearsal). The embedded tasks for which interference does occur include the following two cases: First was ordered recall (Klapp et al., 1983, Experiment 8). Second was probed recall in which subjects recall the digit that had occurred just after the tested probe digit (Klapp & Philipoff, 1983). Additional instances appear in Experiments 1 and 2 reported below.

A review of these and other aspects of short-term memory limits in human performance has been prepared while the author was supported by this contract (Klapp, in press). This rather extensive previous literature is disappointing on at least three grounds. First, as stated above, all of this research involved highly artificial tasks, although the principle of working memory overloads is presumed to apply also in real-world planning and decision-making. Experiments 3 and 4 were initial attempts to demonstrate interference of memory loads with a more realistic task. Second, many of the experiments appear to be of marginal significance at best because of procedural problems, especially the lack of a delay for initial rehearsal of the memory load. Finally, even among these few experiments which do provide the delay, some show mutual interference and some do not.

Those experiments that do show mutual interference require the retention of the order of occurrence of events in both the outside load ar the embedded task. Experiments 1 and 2 below represent an attempt to extend these results to cases in which the embedded task does not require ordered recall. Specifically, Experiment 1 involved an embedded task using associative rather than order memory as the embedded task, and Experiment 2 involved an embedded task of item recognition. In both cases, the presence of a memory load involving ordered information interfered with the embedded task, which did not require ordered information. That sets the stage for the attempt to extend memory loading to the more realistic task involving trucking and transportation in Experiments 3 and 4. This task, like the embedded tasks of Experiments 1 and 2, did not require retention of order information. As it turned out, little memory loading occurred in the trucking simulation. However, from Experiments 1 and 2, it is clear that this could not be attributed to the lack of need for order information in the trucking simulation.

III. EXPERIMENT 1

Experiment 1 was designed to determine if the effect of an external memory load on task performance would occur for embedded tasks which more nearly represent command and control decision-making than has been used for the experiments previously reviewed. Loading effects have been found for embedded tasks involving ordered recall (Klapp et al., 1983, Experiment 8) and recognition of the item which had followed a probe (Klapp & Philipoff, 1983). Both of these situations require that the subject retain the order in which the embedded task items were presented. By contrast, the decision-making task of Experiments 3 and 4 does not require the retention of order information.

Thus, Experiment 1 involved a test of an associative memory embedded task. In order to provide a suitable comparison, the probe task previously found to exhibit loading was also included.

Method

Overview

Performance was measured on each of two memory tasks when done concurrently and on each task when done alone (accompanied by a control version of the other task). One memory task (embedded) was performed during the retention interval of the other (outside) task. The outside task involved memory for letters, and the embedded task involved memory for numbers. The outside task involved pronunciation of the input letters, overt vocal rehearsal, and oral recall. The embedded task required that the subject remain silent during both stimulus input and testing.

The outside task consisted of the successive presentation of six different letters in the same visual location for subsequent ordered recall. For the control outside task, the same letter was presented six times, and the subject was to pronounce this letter six times during recall.

During the retention interval for the letters (the time between their presentation and recall) an embedded task appeared. Independent groups of nine subjects received different types of embedded memory tasks. For both tasks the response required was to press one of nine switch buttons to indicate the number response.

In the probe embedded task, sequential presentation of eight digits was followed by a second presentation of one of the digits. The subject was to respond by indicating the digit which had followed this probe in the set presented. This task required that the order of the digits be remembered.

For the associative task, a sequence of four associations of the form W = 3, Y = 1, etc. were presented. The elements of each pair appeared together, but the four pairs were presented successively. To test retention, one of the letters was presented, and the subject responded by pressing the switch corresponding to the correct digit. This task required that the associations be retained, but the order of the associated pairs was not needed.

Subjects in both conditions of embedded task also received a control embedded task in which the same digit was presented eight times in succession, and the subject responded by pressing the switch corresponding to that digit.

Design

The design permitted test of the outside task performance with each embedded task, and with a control embedded task. It also permitted test of embedded task performance with the outside task present in comparision to embedded task performance with a control outside task. Independent groups of nine subjects were tested with the probe and associative embedded tasks. Alternate subjects were assigned to these conditions as they qualified for the experiment. Each subject was tested with both tasks present, with the control outside task, and with the control embedded task. The order of these three conditions was balanced across subjects. Each subject participated in four sessions. The first session was for practice and subject selection. The remaining three sessions were the scored tests. Each scored session was comprised of 16 unscored practice trials and 32 scored trials all in the same condition. Rest intervals occurred after each set of 16 trials.

On the practice session, each subject received a pretest on the outside task, with the control embedded task. The criterion was six correct out of eight trials. Subjects who could not reach criterion after three attempted blocks of eight trials were rejected. Accepted subjects then received eight trials on the embedded task with control outside task. Finally, the subject received eight trials with both tasks.

Subjects

The 18 subjects were students in Introductory Psychology at California State University, Hayward, who participated as one option of a course requirement, with supplemental pay for the final session in some instances. Because persons who are not native speakers of English often report translation prior to number retention, such persons were not used as subjects. It was necessary to replace 21 subjects prior to data collection for the following reasons: 2 had native languages other than English, 2 failed to return to complete the experiment, and 17 failed to meet the pretest criterion.

Apparatus

The subject and experimenter were in individual sound isolation chambers (Industrial Acoustics 400A). The experiment was controlled by a microcomputer, and all alphanumerical stimuli were displayed on a cathode ray tube (CRT) monitor at a visual angle of 0.6 deg. Subjects indicated their response to the embedded digit task by pressing one of nine switches placed just below the monitor. These switches were in a horizontal array numbered from left to right and spaced 3.25 cm. from each other.

Outside task details

For each trial, the first six letters of the alphabet were randomly ordered and then presented successively. Each letter appeared for 150 ms, followed by a 100-ms blank interdigit interval, except that after the third letter this interval was increased to 300 ms. This extra pause grouped the letters into two sets of three. Grouping is known to facilitate performance in ordered recall (Klapp et al., 1983, Exp. 3). Note that this task required only order information because the item pool was well defined and constant across trials.

During presentation of the letters, subjects were to pronounce the letters aloud as they appeared. Then they were to continue overt rehearsal during a 5-sec interval. At the signal "STOP REHEARSAL," they were to complete the current rehearsal and then remain silent.

Testing for the letters occurred after the completion of the embedded task. The display "LETTERS?" indicated that the subject was to recall the letters orally. At the end of the trial, the correct letters were displayed to the subject to maintain motivation on the outside task. Scoring was dichotomous--either the recall was correct (all letters in correct order) or else incorrect.

For the single-task control, one letter was presented six times, with the same timing as in the experimental outside task. Rehearsal and "recall" followed the same procedure and timing as in the experimental task. Thus, the experimental and control outside tasks were equivalent in timing, stimulus presentation, and overt responding, but differed markedly in memory loading. For the experimental outside task, the subject was to remember the six letters in order. For the control conditions, the subject had to remember only a single letter.

Probe embedded task details

The eight digits to be remembered were always a random ordering of the digits 1 to 8. The probe serial position excluded position 8, the final position for which there would be no following digit. Each of the other serial positions was probed equally often in each block of eight trials, except that position 5 was probed twice. The order in which the probe serial positions appeared across trials in each block of eight trials was randomized. Timing of the digit presentation corresponded to the timing of the letter input for the outside task, except that an additional long pause appeared after the sixth (as well as after the third) digit. After the digits were presented, a single digit appeared, followed by a question mark as the signal to respond. The subject responded by pressing one of the answer switches. This test of the embedded task was followed by a test of the outside task, after which the correct embedded task answer was displayed in the form "NUMBER = 5."

Associative embedded task details

The digits 1 to 4 were randomly assigned to the letters W,X,Y, and Z for each trial. Then the order in which the four pairs were to appear was determined randomly. For each block of eight trials, each of the serial positions 1 to 4 was tested twice, with the order randomized for each such block. The pairs were displayed sequentially, such that each pair appeared for 300 ms, with a 200-ms delay between pairs. Presentation of the pairs was followed by a test display of the form "W=?." The subject responded by pressing one of the switches, as in the probe condition. Then the outside task was tested, after which the embedded task answer was displayed in the form "W=2."

Control embedded task details

A single digit (randomly selected for each trial) was presented eight times with timing corresponding to that in the probe task. Presentation of the digit was followed by a test display of the form "#=?." The subject responded by pressing one of the eight response switches, as in the other conditions. Then the outside task was tested, after which the embedded task answer was displayed in the form "NUMBER = 5."

Trial sequence

Each 43-sec trial was comprised of the following event sequence:

1. Alerting. The words "GET READY" appeared for 1 sec, followed by a 250-ms blank interval.
2. Letter display (for outside task).
3. Rehearsal interval (5 sec).
4. Display of "STOP REHEARSAL" for 500 ms followed by a 1-sec further delay.
5. Presentation and testing of the embedded task.
6. Oral recall of the outside task letters upon presentation of the display "LETTERS."
7. Feedback. The word "LETTERS" and the correct letter sequence were displayed. Below that, the correct answer for the embedded task was displayed, preceded by the word "NUMBER."

Results

Outside task performance

Recall of the outside task letters was considered correct if all letters were given in the correct order. Table 1 displays the proportion of trials with correct outside task recall as a function of the type of embedded task. Note that independent groups of subjects were tested with the probe and associative embedded tasks and that all subjects were also tested with the same control embedded task. The two groups displayed equivalent outside task performance with the control embedded task, indicating that the groups were well matched. Performance on the outside letter task was degraded by the presence of the experimental rather than control embedded task for the probe embedded task, $F(1,8) = 25.1$, $p < .001$, and for the associative embedded task, $F(1,8) = 8.4$, $p < .05$. The amount of degradation did not differ significantly between the two types of embedded tasks, as indicated by the lack of a significant task-type by task-versus-control interaction, $F(1,16) = 2.4$, $p > .10$.

Embedded task performance

Embedded task performance was correct if the appropriate number switch was pressed. Table 2 displays the proportion of trials for which this response was correct as a function of which embedded task was measured ((probe or associative) for experimental (six-letter) and control (one-letter) outside task loads. Performance was degraded by the presence of the actual outside task in comparison with the control for the probe embedded task, $F(1,8) = 27.7$, $p < .001$, and for the associative embedded task, $F(1,8) = 5.6$, $p < .05$. The amount of degradation did not differ significantly between the two types of embedded tasks, as indicated by the lack of a significant task-type by outside task actual-versus-control interaction, $F(1,16) < 1$.

Table 1. Outside task performance as a function of embedded task, Experiment 1

Embedded task type	Embedded task condition Present	Control	Decrement
Probe	64%	90%	26%
Associative	75%	90%	15%

Table 2. Embedded task performance as a function of type of embedded task and of outside task condition, Experiment 1

Embedded task type	Outside task condition 6-Letter	1-Letter Control	Decrement
Probe	58%	83%	25%
Associative	63%	82%	19%

Discussion

Performance on the outside loading task was reduced when the experimental embedded task was present, in comparison to outside task performance with the control embedded task. This relationship held for both embedded tasks, probe and associative. Similarly, performance on both embedded tasks was reduced by the presence of the experimental outside task (loaded condition), compared to performance with the control outside task. Previous reports of this type of mutual interference involved embedded tasks, such as the probe task of the present experiment, which required retention of order information. The finding that mutual interference also occurred for the associative task generalized this result to a case not involving the retention of order information. This is important, because the trucking and transportation simulation (Experiments 3 and 4) appears to resemble an associative task much more than it resembles tasks requiring ordered information. In particular, the trucking and transportation task requires that the subject associate resources (trucks and drivers) and mission delivery requirements with locations on a map and remember these associations while planning the game strategy.

IV. EXPERIMENT 2

Experiment 1 demonstrated memory loading for an embedded task that required retention of associations but not order information. By contrast, there are other tasks in which order is not required and for which loading does not appear. Embedded tasks that are not loaded include the missing-digit task (Klapp & Philipoff, 1983), and modified Sternberg scanning (Klapp et al, 1983, Experiment 7). Perhaps loading occurs for tasks which require either order information or associations, but not for tasks which require only that item information be retained.

However, there are at least two reasons to be hesitant about such a conclusion. First, the missing-digit task is unusual on several grounds (Klapp et al., 1983) and may not be especially representative of item tasks in general. Second, the modified Sternberg scanning task (Klapp et al., 1983, Experiment 7) involved parallel presentation of the embedded task items, rather than sequential presentation as in the other experiments. In Experiment 2, the embedded task that was studied involves item recognition but does not use parallel presentation or the missing-digit task. A probe recognition task was included as a reference for comparison of any memory loading that may occur on the item task.

Method

Overview

As in Experiment 1, one memory task (embedded) involving digits was carried out during the retention interval of another (outside) task involving letters. Subjects were instructed to emphasize performance on the outside task. The outside task and its control were similar to those of Experiment 1.

During the retention interval for the letters (the time between their presentation and recall), an embedded task appeared. Independent groups of eight subjects received different embedded memory tasks. For both tasks, digits were presented sequentially, after which subjects responded to a yes-no question by moving a lever to the left or right. The two tasks differed with regard to the nature of the yes-no question.

For the *item recognition* condition, presentation of the digits was followed by a single test digit and a question mark (?). The subject was to respond "yes" (rightward handle movement) if that digit had appeared in the list just presented and "no" (leftward movement) if it had not appeared. The order of assignment of "yes" and "no" correct responses was randomized.

For the *order recognition* condition, presentation of the digits was followed by a digit pair and a question mark. The two digits were always among those that had been presented. On half of the trials, the digits had appeared in exactly that sequence and the correct response was "yes" (rightward movement). On the remaining trials, the digits had been presented, but not in that sequence, and the correct response was "no" (leftward movement).

Subjects

The 16 subjects in Experiment 2 were from the same population as in Experiment 1. It was necessary to replace four subjects during the first session, prior to data collection. One was replaced whose native language was not English, and three additional subjects were replaced because they failed the pretest.

Design

Independent groups of eight subjects each received the item-recognition and order-recognition conditions. Alternate subjects were assigned to the conditions in the order that the subjects qualified for the experiment. Within each group, all subjects received both easy (four-digit) and hard (eight-digit) versions of the embedded task. All subjects were tested on each of these versions of the embedded task with the experimental outside load of six letters and also with the control outside load of one letter repeated six times. Balancing of the order of these conditions is described below. Unlike Experiment 1, there was no control embedded task condition. Here, embedded task performance was primarily assumed to be a function of outside load.

Each subject received three sessions of 1 hr each. The first session was for practice and subject selection. Each subject first received a pretest comprised of the outside six-letter task, with no embedded task. The criterion was seven totally correct recalls (correct letters in correct order) out of 10 trials. If the subject failed this pretest on the first block of 10 trials, two more blocks of 10 trials were attempted. Subjects passing the pretest within three blocks continued with five trials of practice in each condition. The second and third sessions were comprised of two scored conditions each.

For half of the subjects in each condition, the control outside task (one letter presented six times) appeared in the first scored session, and the experimental outside task (six different letters) appeared in the second session. The remaining subjects received the reverse order. Within each session, the easy (four-digit) version of the embedded task appeared first for half of the subjects and second for the remaining subjects. Thus, each subject received four test blocks; two on each day. Each test block was comprised of 5 unscored practice trials followed by 20 scored trials. These 20 trials were comprised of two blocks of 10 trials. Each block of 10 trials was comprised of 5 for which "yes" was the correct answer and 5 for which "no" was the correct answer, with the order of yes and no trials randomized.

Outside task details

For each trial, six letters were selected randomly, without replacement, from the set BCMFRHKLG, and the selected letters were randomly ordered and then presented successively at the same physical location. Other details were as in Experiment 1.

Embedded task details

For each trial the four or eight digits to be displayed were selected randomly without replacement from the digits 1 to 9 and were displayed in a random order. For the item-recognition task, the test digit was selected at random from among those presented for the "yes" trials or at random from those not presented for the "no" trials. For the order-recognition task, one of the test digits was selected at random from among those presented, excluding the last digit in the sequence. The second digit of the test pair was the correct following digit for "yes" trials and was selected at random from the other digits that had been presented for "no" trials.

Apparatus

The apparatus corresponded to Experiment 1, except that subjects indicated their response to the embedded digit task by moving a 7.5-cm-long lever to the left or right by approximately 2 mm.

Trial sequence

Each 40-sec trial was comprised of ie following sequence of events:

1. Alerting. The words "GET READY" appeared for 1 sec, followed by a 250-ms blank interval.

2. Letter Display (for outside task). The six letters appeared sequentially at the same location on the CRT display.

3. Rehearsal interval. The CRT was blank for 5 sec.

4. End rehearsal. The display "STOP REHEARSAL" appeared for 500 ms, followed by an additional 1-sec blank interval.

5. Digit display (embedded task). The four or eight digits appeared sequentially at the same location on the CRT display.

6. Digit test (embedded task). For the item-recognition condition, a single test digit appeared followed by "?." For the order-recognition condition, a pair of digits appeared followed by "?." These remained visible until the subject had responded by moving the lever right or left, or until 5 sec had elapsed without any response.

7. Letter recall (outside task). The display "LETTERS" indicated that the subject was to recall the outside load letters orally. Either six different letters were recalled, or for the control condition, the same letter was recited six times. This display remained for 4 sec.

8. Feedback. The display "NUMBERS WERE RIGHT" (or "WRONG") appeared, and below that, the correct outside task letters appeared preceded by "LETTERS =."

9. Blank intertrial interval of 7 sec.

Results

Outside task performance

Table 3 displays performance on the outside task in terms of percentage of trials in which all six letters were correctly recalled in the correct order. First note that performance on the control outside task (one letter repeated six times) was virtually perfect in all conditions. Note next that performance for the experimental outside task did not vary significantly as a function of whether the embedded task was item recognition or order recognition, $F(1,14) = 2.0$, $p > .10$, and did not vary significantly as a function of whether the embedded task had four or eight digits, $F(1,14) = 1.9$, $p > .1$. This is consistent with the instruction to emphasize outside task recall.

Table 3. Percent correct on outside task as a function of type of outside task and type of embedded task, Experiment 2

Embedded task	Type of outside task 6-letter recall	1-letter control
Item recognition		
Four digits	83%	99%
Eight digits	79%	98%
Order recognition		
Four items	66%	100%
Eight items	61%	99%

Embedded task performance

As indicated in Table 4, embedded task performance was, of course, better when only four digits were presented compared with the eight-digit task, $F(1,14) = 63$, $p < .001$. The overall performance did not differ significantly between the item and order tasks, $F(1,14) < 1$.

Table 4. Percent correct on embedded task as a function of type of embedded task and of outside task loading, Experiment 2

Embedded task	Outside task condition 6-letter load	1-letter control	Decrement
Item recognition			
Four digits	93%	98%	5%
Eight digits	72%	86%	14%
Order recognition			
Four digits	88%	99%	11%
Eight digits	66%	88%	22%

The primary interest was in the decrement of performance on the embedded task due to loading by the outside task. Overall performance was reduced by the load, compared to the control outside task, $F(1,14) = 92.7$, $p < .001$. The interest here was in examining the size of this decrement as a function of type and difficulty of task. This was done by testing for load by condition interactions.

Overall there was more decrement measured in the harder, eight-item task, compared to the easier four-item embedded task, $F(1,14) = 11.8$, $p < .005$. Although smaller, the decrement due to outside loading was significant even for the smaller, four-item, task $F(1,14) = 21.3$, $p < .001$. Of course, the decrement was also significant for the eight-item task, $F(1,14) = 58$, $p < .001$.

There was also an interaction such that the decrement (outside task present-versus-control) was larger measured on the order-recognition task than on the item-recognition task, $F(1,14) = 6.7$, $p < .05$. Although the decrement was smaller in the item task, it was significant considering that task alone, $F(1,7) = 14.2$, $p < .01$. Of course, the decrement was also significant considering just the order-recognition task, $F(1,7) = 295$, $p < .001$.

Discussion

The presence of the outside memory load reduced performance for both types of embedded tasks, item recognition and order recognition, with more decrement for the order task. The amount of decrement was sensitive to the number of digits in the embedded task, with larger decrement for the longer strings of digits. Thus, the difference between the order- and item-recognition tasks in their sensitivity to loading may reflect differences in the difficulty of these tasks, rather than any qualitative difference between order and item tasks.

The important result here is the finding of a decrement due to memory loading in a task requiring only the retention of item information. This result appears to contradict the finding of no-loading in modified Sternberg scanning (Klapp et al., 1983, Experiment 7). However, the present experiment involved sequential presentation of the embedded task digits, in contrast to parallel presentation of the previous study.

V. EXPERIMENT 3 DECISION-MAKING

The results of Experiments 1 and 2 indicate that outside memory loads requiring ordered recall can exhibit mutual interference with embedded tasks even if the embedded task does not require retention of order information. The decision-making tasks of Experiments 3 and 4 also do not require retention of order information; thus, this result raises the possibility that decision-making also exhibits mutual interference with memory loads for ordered recall.

Subjects performed a simulated truck dispatching task implemented on an Apple II microcomputer. Performance on this simulation was tested with and without a concurrent short-term memory load. As discussed previously, time for initial rehearsal of the memory load was provided before continuation of the decision-making game. Rather than a single memory load, load input and recall events were dispersed within the decision-making game. However, for the loaded condition, a memory load was present during all phases of the embedded simulated truck dispatching task.

Method

Subjects

It was somewhat difficult to obtain subjects who were able and willing to perform well in the tr cking and transportation task. Ordinarily subjects are selected from a pool of students taking Introductory Psychology, but very few of these subjects were able to perform well on this task. Consequently, subjects were recruited from among friends and upper division students. On many occasions, attempts to gather data were abandoned while the experiment was in progress because software bugs appeared. The six subjects who were eventually used tended to be more mature and more motivated than the typical college student, but it is difficult to characterize this population further.

Software, General Description

The truck dispatching task was to require cognitive processes similar to those involved in command and control decision-making. However, the task involved a topic that should be within the understanding of subjects who are available in a university setting. The task confronting the subject was to assign resources (drivers, tractors, and trailers) to deliver goods from origins to destinations by the required time. Costs were assigned to the use of resources (driver's salary and rental fee for equipment). Pay-offs were provided for delivery of goods by the required time, and late charges were imposed for late deliveries. The subject's task was to maximize profit (pay off less expenses and late charges) on a set of missions. To do this the subject examined the mission requirements (pounds to deliver, origin, destination, and due time) and the resources available (location of resource, and type and number of units of each resource). The subject then scheduled the use of these resources on the missions. When all scheduling was completed, results (profits earned, etc.) were determined by the computer and displayed to the subject.

The task to be accomplished involved moving material from a starting point to a destination point along routes defined by a route map (Figure 1). The potential start and destination points are referred to as "nodes." The lines on the map indicate possible routes between nodes, and each route has a designated driving time. In addition to representing origin and destinations, a node could also contain resources (tractors, trailers, and/or drivers).

Other constraints on the task were as follows. First, large tractors could haul small or large trailers, but small tractors could haul only small trailers. Second, no driver would start before his "wake-up" time, a fixed parameter of each driver. Third, resources (drivers, tractors, trailers) could be moved empty to the origin of a mission if appropriate resources were not already available, but, of course, salary and rental fees were charged during the move.

The term "game" refers to a set of missions to be scheduled as a group. The details of a game are determined by the structure of the software and by two sets of parameters that can be set within the software. The so-called "permanent" parameters are changed less often than the "variable" parameters, so that subjects play several games (sets of variable parameters) with the same set of permanent parameters. The permanent parameters include the route map, the costs of resources (driver salary, rental fees), and the carrying capacity of large and small trailers. The variable parameters include mission requirements (origin, destination, due time, number of pounds, amount of pay-off, late fee) and the resources available at each node in the map (drivers with wake-up times, tractors, trailers).

Practice Games

Purpose. Three practice games were used to select subjects and to familiarize them with the use of the software. Permanent parameters were selected for simplicity, without problems of negative transfer to the test games. Variable parameters for the three games provided for increasing difficulty and the introduction of new concepts.

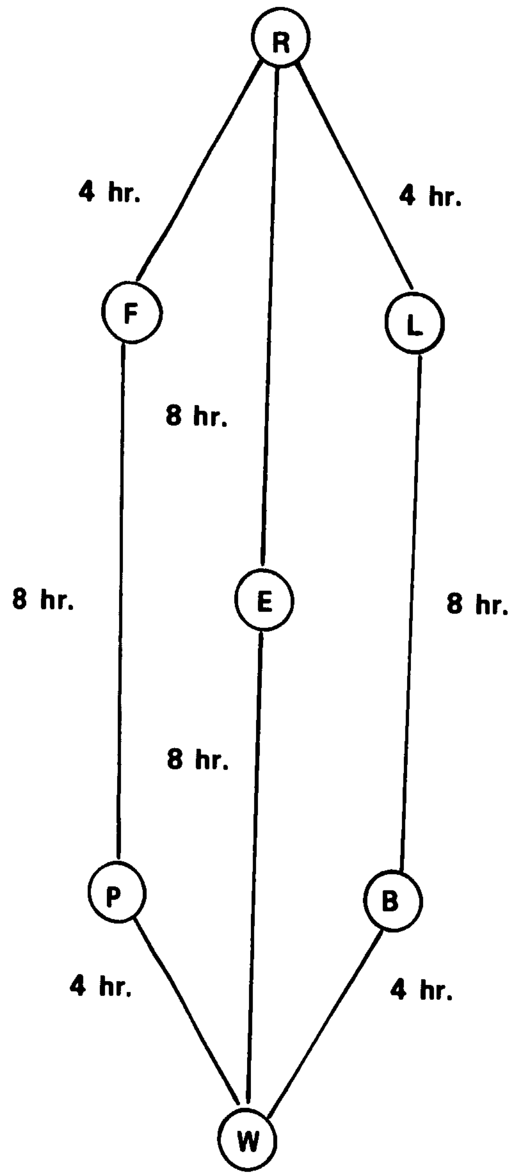

Figure 1. Route map for test games for Experiment 1.

Permanent parameters. The route map was an equilateral triangle with nodes labeled A,Y, and I, and with all three routes of 4 hours in duration. This was a simpler route structure than that used in the test games (Figure 1). The other parameters were the same as in the test games: driver's salary was $20/hr with no charge for an overnight stay, large tractor costs $10/hr, small tractor costs $5/hr, large trailer limited to 8,000 pounds, small trailer limited to 4,000 pounds, and no rental charge for trailers.

Variable parameters. Three specific games (variable parameter sets) were developed as follows:

Practice Game 1: This game had only one mission: moving from node A to node Y with a load of 8,000 pounds, and a due time of 1600 hours. The fee to be earned was $1000, and the late fee was $250. All needed resources were available at node A, along with additional unneeded resources. To avoid confusion, there were no resources at other nodes. The only difficulty was that there were two drivers at location A, the first of which has a late wake-up time and cannot complete the mission on time. Thus, the second driver (wake-up at 0800 hours) must be chosen.

This initial game was a simple introduction to the operation of the software, and it demonstrated, first, that drivers cannot be selected at random without regard to wake-up time in relation to route length and deadline and second, that a large trailer and tractor are needed for large loads.

Practice Game 2: This game included two missions, each rather simple and each originating at A. Mission 1 was to deliver 8,000 pounds to Y by 1600 hours, and corresponded to the one mission of practice game 1. Mission 2 was to deliver 4,000 pounds to I at 1600 hours. Again, all available resources were at the origin, A, so that resource selection rather than resource moves were required. There were three drivers with wake-up times of 1301, 0800, and 0356. The first driver on the list could not complete any missions on time. Subjects needed to use the large trailer and tractor for mission 1. This game was to continue to familiarize the subject with the procedure and to introduce multiple missions.

Practice Game 3: This game included three missions and required resource moves. Again, each mission was due at 1600 hours and paid $1000, with late fee of $250. Mission 1 originated at A and required delivery of 8,000 pounds to Y. Mission 2 originated at Y and delivered 4,000 pounds to A. Mission 3 began at I and delivered 4,000 pounds to Y. No resources were available at A for mission 1, so that a move was needed.

Location I had enough resouces for the mission originating there, but no extra, so that a move from I to A was not wise. The solution was to move an excess driver and other resources from Y to A. This game was more difficult in that it required a move of resources and had the possibility for making a simple strategic error.

Test Games

Purpose. The test games were used in data collection. An attempt was made to make them relatively similar to each other in difficulty, but more challenging than the practice games. Of course, assignment of experimental conditions to games was balanced, as detailed below in the design section.

Permanent parameters and other constant features. The route map for all six test games (Figure 1) had seven nodes (rather than three as in the practice games), and both 4-and 8-hour routes (rather than only 4-hour routes). This more complex route structure created more opportunities for strategic decisions than the simpler practice route structure. All test games had the following features in common: due time 1600 hours for all missions, all delivery fees were $1000, all late charges were $250, driver salary $20/hr, rental $10/hr for large tractors and $5/hr for small tractors. Trailer limits were 8,000 pounds for large and 4,000 pounds for small trailers. There was no overnight stay charge and no rental charge for trailers.

Variable parameters. Six games (variable parameter sets) were developed as follows:

Test Game 1: This had three missions, all originating from node R and all involving 8,000 pounds. The destinations were F,L,and E for missions 1,2,and 3, respectively. Missions 1 and 2 required 4 hours, but mission 3 required 8 hours. The difficulty in this game was the lack of any drivers at the common origin R, so that moves of drivers from other nodes were required. The other needed resources were available at the origin. Drivers were available as follows: at P, one driver who awakes at 0907 hours; at B, one driver at 1100 hours; at F, five drivers who awake at 0959, 0713, 0359, 0800, and 0657; and at L, four drivers with wake-up times of 1511, 0930, 1657, and 0905.

The problem was to select a strategy of moving three of the available drivers to R prior to scheduling the actual missions. Only one driver had a wake-up time early enough to be moved to R and then to complete the 8-hour route to E by the deadline. A subject who simply moved this driver (driver 3 at node F, wake-up time 0359 hours) and scheduled him on mission 1 (a 4-hour drive) would later find that this does not work out and would need to cancel this initial schedule so that this special driver would be assigned to the one long mission. Thus, this game can induce the subject to make a false move if the missions are scheduled one at a time, starting with mission 1, without regard for the overall picture. To complete this game with the least number of entries, the subject must first check all nodes for the best drivers to move to R. All of the needed drivers can be found at F. (Despite the early-appearing wake-up times at L, none of the drivers at this node can complete missions without being late, due to the long driving time to get to the origin, R).

Test Game 2: This game had three missions originating at W and all 4 hours long. Missions 1 and 2 went to P with 8,000 pounds each, and mission 3 went to B with 4,000 pounds. The common origin, W, had one small tractor, one trailer, and two drivers waking up at 0800 and 1103 hours. Thus, some resources must be moved in. Resources were available as follows: E had one of each type of trailer and tractor and one driver with a wake-up time at 1010 hours; R had none; F had one of each type of tractor and trailer and three drivers, all of whom had a wake-up time of 0800 hours; L had two drivers with wake-up times of 0800 and 1103 hours; P had two of each type of tractor and trailer and two drivers with wake-up times of 0800 and 0900 hours; and B had two of each type of tractor and trailer and one driver with wake-up time of 0800.

In order to complete this game, the subject must check more than one node for resources to move to the origin and must also use resources already at the origin. The solution includes moving in resources from B and from P. Premature and incorrect scheduling would create the need to cancel moves.

Test Game 3: This game had three missions, all originating from E. Missions 1 and 3 delivered 8,000 pounds to R over an 8-hour route, and mission 2 delivered 4,000 pounds to W over a different 8-hour route. The origin, E, had inadequate resources for the three missions. There were two each of tractors and trailers, and three drivers with wake-up times of 0900, 0704, and 0657. No resources were available at F, L, P, or B. Node R had one tractor and one trailer and three drivers with wake-up times of 0800, 0000 (midnight), and 0430. Node W had one of each type of resource and the driver had a wake-up time of 0110.

Problems arise because only two of the three drivers at the origin, E, wake up early enough to complete the missions on time. The player must search for a suitable driver, and the only one who wakes up early enough to get to the origin and then to the destination on time is at R. The subject must move this driver, along with a tractor and a trailer, to E. Because moving a driver and resources takes time, the subject's first choice might be to move a driver a short distance (4 hours' time), so that a player may first try nodes which are close to the origin. However, this will not work because no suitable drivers are available at these nodes.

Test Game 4: This game had three missions with different origins. Mission 1 carried 4,000 pounds over a 4-hour route from L to R. Mission 2 carried 4,000 pounds over an 8-hour route from E to R. Mission 3 carried 8,000 pounds over an 8-hour route from P to F. Resource availability was as follows: At L, there were two of each type of tractor and trailer, and three drivers waking up at 0902, 1031, and 0800 hours. At E, there were two sets of small trailers and tractors and four drivers with wake-up times of 1201,0902,0704, and 0756 hours. At P, there were one of each type of tractor and trailer but no drivers. At B, there were one of each type of tractor and trailer and one driver waking up at 0000 hours.

Scheduling of missions 1 and 2 did not require moves of resources. Mission 3, by contrast, presented difficulties. The only driver with a wake-up time early enough to get to P and then complete the mission on time was at B, 8 hours away. A large tractor and trailer must be moved along with the driver.

Test Game 5: This game had three missions, all originating from R. Mission 1 carried 4,000 pounds over a 4-hour route to L, mission 2 carried 8,000 pounds over a 4-hour route to F, and mission 3 carried 4,000 pounds over an 8-hour route to E. Resource availability was as follows: Node R had three of each type of tractor and trailer and three drivers with wake-up times of 1000, 1215, and 1111 hours. Nodes F and P had no resources. Node L had two of each type of tractor and trailer and two drivers with wake-up times of 1312 and 0759. Node B had one of each type of tractor and trailer and one driver with a wake-up time of 1700. Node W had one of each type of tractor and trailer and a driver with a wake-up time of 0800. Node E had one of each type of tractor and trailer and one driver with a wake-up time of 0957.

This game can frustrate the player with an overload of useless resources. At best, only one of the missions can be completed on time, and to do that, the driver must be moved in from location L.

Test Game 6: All three missions involved 4,000 pounds. Mission 1 drove from E to R, an 8-hour route. Mission 2 was from E to W, also 8 hours. Mission 3 was a 4-hour route from W to

B. Resource availability was as follows: Node W had one of each type of tractor and trailer and a driver with a wake-up time of 1213. Node P had one of each type of tractor and trailer and three drivers with wake-up times of 1201, 1001, and 0700. Node F had one small tractor and one small trailer and one driver with a wake-up time of 0900. Location R had only a driver, with a wake-up time of 1312. Locations L and B had no resources.

The two missions originating from E were ready to go without resource moves. Mission 3, by contrast, required a resource move from P. It was necessary to select the driver who had a wake-up time of 0700.

Load Conditions

There were three conditions with respect to the auxiliary memory load, a loaded condition and two no-load controls. These conditions differed with respect to the action needed to enter the main menu of the trucking simulation program. The main menu gave the subject a selection among the following basic functions:

1. Display resource information
2. Display mission information
3. Enter plan information
4. Run off plan

For the load condition, entry to the main menu occurred after correct recall of a "password." The password was a set of four consonants randomly selected, without replacement, from the set GMFHKLRS, and presented in random order. The password was changed after each recall of the previous password, so that a given password was used only once by a subject. Entry into the main menu required the following event sequence:

1. Response to a previous menu indicating desire to return to the main menu.

2. The message "Enter password to continue" appeared.

3. Subject entered the four-letter password on the keyboard.

4. IF CORRECT: The word "right" appeared, below which the display "New password is ....." was presented for approximately 10 sec.

IF NOT CORRECT: The message "Sorry, please try again" appeared, and control reverted to step 2 above.

IF NOT CORRECT ON SECOND ATTEMPT, the subject must pay an "exit charge." The message "Sorry, exit cost is $.."appeared. Then the new password was presented.

5. The main menu was displayed.

Several aspects of this procedure should be noted. First, the entry must have the correct four letters in the correct order to be considered correct. Second, the new password was displayed for approximately 10 sec to permit study and early rehearsal, a necessary procedure for valid loading experiments, as discussed in the above literature review. Third, the subject was loaded throughout the games, but the load changed from time to time as the game proceeded. Fourth, the load letters were presented simultaneously in a line.

The load control condition was the same as the load condition except that the password was visible at the time that entry of the password was required. No new password was displayed after entry. Hence, the required keystroke actions corresponded to those of the load condition, but no memory load was imposed.

The none condition omitted passwords, and entry to the main menu was direct.

Design

Each of six subjects was tested in six games (sets of variable parameters) each. Each subject received all three load conditions, two games per condition. Across subjects, each condition was tested twice with each game, and twice with each subject.

Balancing of the order in which the conditions and games appeared was as follows: The order of testing conditions was balanced across subjects such that each subject was tested once in each condition following the order assigned to that subject, and then was tested again with the conditions appearing in the opposite order. The orders assigned formed a Latin Square so that across each three subjects the orders were balanced. The six games appeared in numerically ascending order for half of the subjects and in the opposite order for the remaining subjects. Balancing of the assignment of games to conditions was completed for each set of three subjects.

Procedure

After an introductory description of the software, the subjects played the three practice games (no load condition) with informal interaction and supervision by the experimenter. Those subjects who were judged to be competent at this point were tested further, and the others were dismissed. Then, subjects repeated at least one practice game in the load condition and at least one in the load control condition.

Total testing time for each subject was approximately 4 hrs. Breaks were provided between games as desired and testing was spaced across 2 or more days. To provide motivation, the subjects were paid according to a formula based on earnings in the games.

Results

Three measures are reported. First, the overall number of keystrokes, removing those associated with memory password entry, gives a first approximation to performance efficiency. (The password responses were not considered because none were required in one of the control conditions.) The second measure is the "profit" earned in playing the game (earnings less the costs of resources and late fees). The third measure is the number of times subjects repeat an information-gathering response. For example, if the subject received information about a particular resource at a particular node, and then asked for this same information again, it was assumed that a human memory problem had occurred. Thus, the repeated-inquiry measure is directed toward memory problems specifically. Clock time was not recorded because of software problems.

The particular repeated-inquiry measure selected for use in this analysis assumes that once a resource has been moved, subjects may check to confirm the move was accomplished. Thus, repeated inquiries separated by a move of resources are not considered to be due to problems with a subject's working memory and are not counted. On the other hand, inquiries separated by the assignment of resources to missions are assumed to be due to problems in memory because a subject should, in principle, know to subtract resources used from the resources previously available.

Table 5 displays the three performance measures as a function of loading condition. For each measure, performance did not differ significantly as a function of condition. The statistics were computed sampling across subjects and then again sampling across games, with $F(2,10) < 1$ in every case, except $F(2,10) = 2.1$, $p = .175$ for profit analyzed across games. Note, however, that a non-significant trend appears in the profit data, with less profit in the loaded condition than in the others. This pattern appeared for only two of the six games and three of the six subjects, confirming the formal statistical conclusion that the trend is not interpretable. Similarly, there is a non-signficant trend toward loading in the repeated-inquiry measure. This trend also is not interpretable, appearing for only two of the six subjects and for three of the six games.

Table 6 displays the keystroke and inquiry measures as a function of game. (Because the games differed in possible profit, no comparison among games on the profit measure is presented.) The number of keystrokes differed reliably as a function of game, $F(5,25) = 3.9$, $p < .01$. This indicates that the keystroke measure was sensitive to the game variable and, hence, may have been sensitive enough to detect an effect of the load condition had any such effect been present. Performance did not differ significantly among games for the repeated-inquiry measure, $F(5,25) = 1.45$, $p = .24$. The keystroke and repeated-inquiry measures did not correlate significantly, $r(4) = .27$, $p > .10$.

Table 5. Performance as a function of condition, Experiment 3

	CONDITION			
Measure	Load Control	Load	None	Significance
Keystrokes	119.4	110.3	121.25	n.s.
Profit	$1847	$1757	$1815	n.s.
Repeat Inquiry	3.33	4.08	3.67	n.s.

Table 6. Performance as a function of game, Experiment 3

	GAME					
Measure	1	2	3	4	5	6
Keystrokes	137	164	79.3	119.3	92.7	109.7
Repeat Inquiry	2.83	5.17	2.67	2.67	5.3	3.5

It would seem that use of the old password and learning a new password might be aversive in the load condition. Thus, it would be expected that memory loading might induce game-playing strategies which reduce the use of the main menu to avoid the need to use the password routine. Such strategies would be less attractive in the load control condition. Indeed, there was a trend such that more password instances seem to appear in the load control condition (average 19.1) than in the load condition (13.75). However, a trend of this type held for only four out of six games. The two games for which the trend did not occur (games 4 and 6) were of average overall difficulty (Table 6). This analysis does not appear to be promising.

Discussion

There was no clear decrement in performance in the trucking and transportation task attributable to the presence of the load in short-term memory, although trends in two of the three measures suggest the possibility of some loading. This result stands in marked contrast to the clear loading effects found for abstract embedded tasks, as reviewed earlier (and in Experiments 1 and 2). The present null result, of course, could indicate that external memory loads have little or no effect on performance in realistic tasks, such as this transportation game. Alternately, the null result could be due to a lack of sensitivity of the experiment.

VI. EXPERIMENT 4

Previous loading experiments involved sequential presentation of the memory load letters. That is, the load items appeared one at a time in sequence. These experiments often demonstrated interference between the load and the embedded task (e.g., the present Experiments 1 and 2). In contrast, loading effects were not observed in the decision task of Experiment 3. However, Experiment 3 involved simultaneous presentation of the memory load items rather than the sequential presentation used in the other experiments. That is, all letters of the password appeared simultaneously. It was thought that this procedural difference might be critical. Thus, a major purpose of Experiment 4 was to examine the role of this load presentation variable. Also, whereas Experiment 3 used passwords of only four letters, Experiment 4 used longer passwords, of six and nine letters. It was assumed that a longer password would have a greater effect on the embedded decision-making task.

Method

Twelve subjects were tested under six conditions of the memory load variable: simultaneously presented loads of six and of nine letters, successively presented loads of six and of nine letters, and two controls, one using six letters and one using nine letters. The control conditions required that a password be entered on the keyboard, but the password was visible at the time of entry so that no memory load was imposed.

For the sequential presentation, each letter appeared for approximately 300 ms, followed by a 250-ms blank off period, except that after the 3rd and the 6th letter, the blank interval was extended to 500 ms. For the simultaneous presentation, the stimulus string appeared for 5 sec. In both cases, letter presentation was followed by a blank delay of 6 sec for initial rehearsal.

Each of the 12 subjects was tested on all six conditions, with the order of appearance of the conditions balanced across subjects. Six different problem "games" were played by each subject, one game at each load condition. The assignment of games to conditions was rotated across subjects such that each condition was tested equally often with each game.

The subjects were selected as in Experiment 3. It was necessary to replace two subjects due to experimenter error involving game assignment. One of the replaced subjects made no attempt to use the password correctly, which was an additional reason for his replacement. One other subject quit before completing the experiment, and one was rejected for poor performance during the practice games.

With the exception of the changes in password length and presentation, the model task software was unchanged from that used in Experiment 3. However, new problem games were generated for which the number of keystrokes needed to solve the game in an optimum manner was nearly equated across the games. Other aspects of the logic of the games were made more uniform across games. A description of these revised games follows.

Practice Games

The three practice games were identical to those used in Experiment 3.

Test Games

Permanent Parameters. These were the same as in Experiment 3, with the same route map in use (Figure 1).

Redesign of Variable Parameters. The test games of Experiment 3 had a wider range of difficulty than would be desirable, so that the games were redesigned for Experiment 4 in an attempt to hold down the between-game variability. In order to obtain a metric for assessing the difficulty of games, it was observed that the mean keystrokes required by subjects to solve the games was highly correlated with the theoretical optimum number of keystrokes, $r(4) = .889$, $p < .05$. Thus, it appeared reasonable to use optimum keystrokes as the difficulty metric for purposes of game redesign. For the redesigned games, the standard deviation of this metric among games was lowered from 11.2 (Experiment 3) to 3.6 (Experiment 4). However, the mean optimum keystrokes was held about constant: 64.5 for Experiment 3 and 62.8 for Experiment 4.

Each game was constructed by assembly of subroutines, selected from a population of eight possible subroutines. These subroutines are described in Table 7, which indicates that a subroutine is a set of logical operations. Because the logic is a set of general operations, a given subroutine may be performed at any node, and with different types of resources. Thus, two games may be comprised of the same set of subroutines, and still be different games. Table 8 identifies which subroutines appear in each of the six test games.

<u>Table 7</u>. List of subroutines for Experiment 4

Subroutine	Keystrokes	Description
1a	15	Identifies all available resources at origin, assuming both large and small loads exist.
1b	9	Identifies available resources at origin, assuming there are only small loads or only large loads.
2a	6	Identifies available drivers, large or small tractors at another location; omits check for trailer resource.
2b	9	Same as 1b except checks available resources at location other than the origin. Assumes interest in only a large or small tractor and trailer, as well as driver information.
3a	9	Transfers one driver and large or small tractor to origin, omits transfer of trailer.
3b	10	Transfers driver, large or small tractor, and trailer to origin.
4	8	Schedules one mission, driver with tractor and trailer combination.
5	1	Executes all missions scheduled

Table 8. Subroutines appearing in each game

Game	Subroutines in order								Optimum Keystrokes[a]
1	1a	2a	3a	3a	4	4	4	5	65
2&5	1[illegible]	2b	2a	3a	4	4	4	5	59
3&4	1a	2b	3a	3b	4	4	4	5	69
6	1a	2a	3a	4	4	4	5		55

[a] One keystroke is added to each game because of the response needed to exit the resource information.

Results

Profit. Of the 12 subjects, 10 showed lower profit on the decision task with a concurrent memory load compared to the control of copying the passwords without memory retention, $p < .02$, sign test. Strangely, the 2 non-conforming subjects exhibited a very strong opposite effect, so that the loading effect vanished with respect to overall mean profits. Of course, this means that there is no significant effect of loading from the viewpoint of parametric statistics. Medians of the subjects' profits (Table 9) were used to represent the data. Note that reduced profit with memory load is apparent with a nine-letter load, but not with a six-letter load. It is quite clear that the effect of memory load on profit was marginal at best.

Profit was lower with the nine-letter load than for the six-letter load for 10 of the 12 subjects, $p < .02$, sign test. Note that this relationship held for the experimental load, but not for the control. These results also give encouragement to the possibility that the presence of memory loads interfered with the decision-making task.

Finally, note that there was no systematic effect of the type of memory load input, sequential versus simultaneous. It was expected that this variable might be critical

Table 9. Median profit on decision task, Experiment 4

Condition	Password Six-letter	Nine-letter
Control	$ 1,980	$ 2,010
Sequential load	2,080	1,835
Simultaneous Load	1,940	1,890

The data on profit in Table 9 give some encouragement to the possiblity that retention of the extraneous password interferes with decision-making efficiency. However, there is an alternative interpretation that is not based on memory retention as the mechanism by which the extraneous memory load interferes with decision-making. Suppose that retention of the memory load is not interfering, but that poor performance on the nine-letter load task is frustrating. This frustration may produce lower performance on the decision-making task. This interpretation would not have been viable if interference had been observed with the six-letter memory load for which reasonable memory task performance was achieved, but it seems to be viable given the observed finding of reduced performance only for the nine-letter load. Note especially that little of the load was retained and recalled in the nine-letter condition (Table 12) for which the presence of the load task was most interfering. If retention of the load were producing the effect, more interference would be expected for the six-letter load, for which performance on the load task indicated substantial retention of the load.

Repeated inquiry. A second dependent variable was the number of times that subjects inquired on the menu-based system about resources (drivers, tractors, or trailers) which they had inquired about earlier. If working memory is loaded by the necessity of retaining the password, then such loads would be expected to increase the number of instances of repeated inquiries.

Loading led to more repeated inquiries for only 5 of the 12 subjects (with one subject tied). Thus, the slight trend in the median scores toward loading at nine letters (Table 10) must not be considered to be a reliable finding.

The fact that memory loading did not have a clear effect on this variable leads to the suspicion that the frustration interpretation of the profit data should be taken seriously. The repeated-inquiry variable should be very sensitive to any memory-specific loading, but perhaps this measure is not sensitive to general frustration.

Table 10. Median repeated inquiries in Experiment 4

	Password	
Condition	Six-letter	Nine-letter
Control	2.5	1.5
Sequential load	1.0	2.0
Simultaneous load	1.5	3.0

Keystrokes Finally, note that overall keystrokes to solution (excluding those related to password operations) were higher for load compared to control for only 6 of the 12 subjects, with one tied and 5 showing more keystrokes on the control than on the load. Median keystrokes appear in Table 11, which supports the conclusion that this measure shows no hint of loss of efficiency with loading.

Table 11. Median keystrokes to solution, Experiment 4

Condition	Password Six-letter	Password Nine-letter
Control	88	94
Sequential load	84	97
Simultaneous load	85	86

Memory performance. The above measures of performance were on the trucking and transportation simulation, with and without a concurrent memory task. Next consider performance on the memory password task itself (Table 12). The performance measure was proportion of password attempts that were correct on the first entry. It is not surprising that performance was much better with only six letters in memory than with nine, an effect which held for all 12 subjects, $p < .01$, sign test. There appears to be little difference in performance between the sequential and simultaneous conditions.

Table 12. Median proportion correct on password entry from memory in Experiment 4

Condition	Password Six-letter	Password Nine-letter
Sequential load	.49	0.0
Simultaneous load	.47	.18

Note that the performance on the password was generally rather low (overall mean of .31), suggesting that this aspect of the task was not greatly emphasized. This may explain, in part, the marginal effect of loading on the above measures of decision task performance.

Discussion

The results of Experiments 3 and 4 certainly were not a clear demonstration of any effect of an extraneous memory load on decision-making performance. Al hough profit appeared to be reduced somewhat by the presence f a memory load, this effect was small and quite possibly due to general frustration rather than to memory-specific interference. Had actual memory loading occurred, the effect should have appeared in the repeated-inquiry measure as well as in profit.

One possible interpretation of the lack of loading is that the experiment was not sensitive enough to detect the effects. The experiment can be improved in two general areas:

1. Possibly not enough emphasis was placed on password memory, so that subjects may have tended to ignore the memory load. This possibility seems plausible, given the overall low performance on the password. If this experiment were to be repeated, emphasis on the password could be enhanced by providing immediate display of monetary reward for correct passwords, rather than the relatively minor penalty for incorrect responses as in the present experiment. However, some device would be needed to discourage entering the main menu unnecessarily in order to ring up more profit for correct passwords. In addition, it may be desirable to use shorter passwords (e.g., four or six letters, but not nine letters) so that the task is feasible, and so that the subject can achieve a reasonable level of performance. Some subjects appeared to give up on the long passwords in the present experiment, regarding the task as impossible.

2. The sensitivity of the decision task might need improvement. In particular, the decision "games" might be improved in terms of the extent to which guessing about resources is ineffective. This might be achieved by making memory for details more essential for performance. The repeated-inquiry measure could be made more meaningful by penalizing the subject for making repeated inquiries. Possibly subjects make such inquiries as a "check" even if they know th answer. It would also be desirable to measure performance time, because this may be sensitive to delays for rehearsal of the memory load.

From a broader perspective, the illusiveness of the influences of memory loading on the decision-making task is rather remarkable. According to most textbooks (see Klapp et al., 1983 for a review), there is a single system of working memory limited to about seven "chunks." Surely a memory load of six or nine letters would fill that system. Because the decision-making task requires memory for mission goals and resources, performance on that task should be markedly reduced by the memory load, even if the measures and techniques of the experiment are not fully perfected. Such effects are not illusive in more abstract experiments, such as Experiments 1 and 2; but on the decision task of Experiments 3 and 4, the effects of loading, if present, are more illusive than might be supposed.

Perhaps the approach should be to take the suggestion that there may be multiple systems of working memory (Klapp et al., 1983; Klapp, in press) even more seriously. Perhaps memory loads can interfere with tasks that are similar to the loading task (Experiments 1 and 2), but not with tasks that are dissimilar, such as the decision-making task. Perhaps subjects are successful in isolating the memory for resources and missions in the task from the memory for the password. By putting the memory demands of these tasks into different systems of working memory, subjects may be able to overcome much of the expected interference.

This possibility needs serious study. The trucking and transportation model task program could be modified to permit comparisons of the effect of dissimilar loads (e.g., the password) with the effect of intrinsic loading due to the presence of irrelevant mission and resource information. It might be possible to demonstrate that intrinsic loading of memory is very disruptive, but extraneous information, such as the password, can be isolated from game-relevant information.

VII. GENERAL DISCUSSION

Experiments 1 and 2 showed that retention of memory loads (requiring ordered recall) can interfere with other memory tasks occurring during the retention interval even if the embedded task does not itself require ordered recall. This result was obtained even with a delay for initial rehearsal, so that the mutual interference is due to retention of the memory load, rather than to interference between initial rehearsal of the load and the embedded task (see Klapp et al., 1983). This opens up the possibility that such memory loads might also interfere with realistic tasks, such as decision-making.

Experiments 3 and 4 attempted to demonstrate an effect of memory load on decision-making in a simulated truck dispatching problem. Although there was a tendency for such interference to occur, especially in Experiment 4, the results are perhaps better interpreted as demonstrating the illusiveness of loading in this context. Certainly the sensitivity of the experimental protocol can be improved, and indeed, specific suggestions for such improvement have resulted from the experience with these initial attempts. However, the broader picture seems to be that loading of decision-making by retention of abstract passwords is illusive at best.

Taking seriously the notion of multiple systems (or multiple resources) in working memory, Klapp (in press) suggests that subjects may isolate the extraneous password from the relevant decision-making information by placing each type of information into a separate system of working memory, thereby avoiding mutual interference. This may explain why the predicted loading effects seem to be small and illusive. By contrast, memory loads within the context of the decision task (e.g., extraneous resource information) might generate a great deal of interference. This hypothesis may be tested with an improved version of the trucking and transportation model task by comparing the interference from extraneous password loads to the interference generated by loads that are intrinsically related to the decision task itself.

If the memory isolation and multiple system perspective is validated by such an experimental demonstration using a realistic decision-making task, this conclusion would indicate the desirability of basic research on the details of memory system isolation and how this can be facilitated by training and system design in operational contexts. Although representing a dramatic departure from the standard text-book view of working memory (see Klapp et al., 1983 for a review), this multiple resource perspective appears to be congruent with the emerging findings, both from abstract laboratory tasks and from the present preliminary work with decision-making.

REFERENCES

Baddeley, A.D., & Hitch, G. (1974). Working memory. In G. Bower (Ed.), The psychology of learning and motivation: Advances in research and theory (Vol. 8). New York: Academic Press.

Chase, W.G., & Ericcson, K.A. (1982). Skill and working memory. In G. Bower (Ed.), The psychology of learning and motivation (Vol. 16). New York: Academic Press.

Crowder, R.G. (1967). Short-term memory for words with a perceptual-motor interpolated activity. Journal of Verbal Learning and Verbal Behavior, 6, 753-761.

Dillon, R.F., & Reid, L.S. (1969). Short-term memory as a function of information processing during the retention interval. Journal of Experimental Psychology, 81, 261-269.

Hitch, G.J., & Baddeley, A.D. (1976). Verbal reasoning and working memory. Quarterly Journal of Experimental Psychology, 28, 603-621.

Johnston, W., Greenberg, S., Fisher, R., & Martin, D. (1970). Divided attention: A vehicle for monitoring memory processes. Journal of Experimental Psychology, 83, 164-171.

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. In J. Long & A. Baddeley (Eds.), Attention and performance IX. Hillsdale, NJ: Lawrence Erlbaum Associates.

Kantowitz, B.H., & Sorkin, R.D. (1983). Human Factors.New York: Wiley.

Klapp, S.T. (in press). Short-term memory limits in human performance. In P. Hancock (Ed.),Human factors psychology.

Klapp, S.T., Marshburn, E.A., & Lester, P.T. (1983). Short-term memory does not involve the "working memory" of information processing:The demise of a common assumption. Journal of Experimental Psychology:General, 112, 240-264.

Klapp, S.T., & Philipoff, A. (1983). Short-term memory limits in performance. Proceedings of the Human Factors Society -27th Annual Meeting, Norfolk, VA.

Klatsky, R.L. (1980). Human memory: Structures and processes. San Francisco: Freeman.

Kroll, N.E.A., Kellicut, M.H., & Parks, T.E. (1975). Rehearsal of visual and auditory stimuli while shadowing. Journal of Experimental Psychology: Human Learning and Memory, 1, 215-222.

Logan, G.D. (1979). On the use of a concurrent memory load to measure attention and automaticity. Journal of Experimental Psychology: Human Perception and Performance, 5, 189-207.

Logan, G.D. (1980). Short-term memory demands of reaction-time tasks which differ in complexity. Journal of Experimental Psychology: Human Perception and Performance, 6, 375-389.

Peterson L.R., & Peterson, M.J. (1959). Short-term retention of individual verbal items. Journal of Experimental Psychology, 58, 193-198.

Reisberg, D. (1983). General mental resources and perceptual judgments. Journal of Experimental Psychology: Human Perception and Performance, 9, 966-979.

Richardson, J.T.E. (1984). Developing the theory of working memory. Memory & Cognition, 12, 71-83.

Roediger, H.L., III, Knight, J.L., & Kantowitz, B.H. (1977). Inferring decay in short-term memory: The issue of capacity. Memory & Cognition, 5, 167-176.

Shulman, H.G., & Greenberg. D. (1971). Perceptual deficit due to division of attention between memory and perception. Journal of Experimental Psychology, 88, 171-176.

Stanners, R.F., Meunier, G.F., & Headley, D.B. (1969). Reaction time as an index of rehearsal in short-term memory. Journal of Experimental Psychology, 82, 566-570.

Wanner, E., & Shiner, S. (1976). Measuring transient memory load. Journal of Verbal Learning and Verbal Behavior, 15, 159-167.

Wickens, C.D. (1984). Engineering psychology and human performance. Columbus, Ohio: Charles E. Merrill Co.