

**THE ROLE OF MULTIMODAL SENSORIMOTOR INVOLVEMENT
IN THE LEARNING OF COGNITIVE SKILLS**

by

Walter Antonio Aviles

B.S. Electrical Engineering
Stanford University, 1982

Submitted to the

**DEPARTMENT OF ELECTRICAL ENGINEERING AND
COMPUTER SCIENCE**

In Partial Fulfillment of the Requirements
for the Degree of

MASTER OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June, 1996

COPYRIGHT 1996 Massachusetts Institute of Technology
All Rights Reserved

Signature of Author _____

June, April, 1996

Certified by _____

Nathaniel I. Durlach, Senior Scientist, Thesis Supervisor

Accepted by _____

Department Committee on Graduate Students

MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

JUL 16 1996

Eng.

LIBRARIES

THE ROLE OF MULTIMODAL SENSORIMOTOR INVOLVEMENT IN THE LEARNING OF COGNITIVE SKILLS

by

Walter Antonio Aviles

Submitted to the Department of Electrical Engineering and Computer Science
on April 29, 1996 in partial fulfillment of the
requirements for the Degree of Master of Science in
Electrical Engineering and Computer Science

ABSTRACT

The existence and extent of effects of small-scale haptic involvement on recall performance in an explicit memory task were investigated in two related experiments involving a total of 40 subjects. The memory task was structured in such a way that it abstracts two of the fundamental activities encountered when interacting with real-world equipment through physical control interfaces: 1) the identification of the state of the system to be acted on; and 2) the recall of the series of actions which are to be taken when the system is in that state. Furthermore, the experiments were designed as training experiments. Haptic sensorimotor involvement is evaluated as an aid to training -- it is used during practice and not used during performance.

Results from Memory Experiment 1 show (i.e., $H_0: p < 0.01$ for 5 minute, 30 minute, and 24 hour examinations) that experimental group subjects performed approximately 2.5 times better when they trained with both visual stimuli (i.e., a keypad graphic) and haptic involvement than when they trained without these visual and haptic stimuli. Memory Experiment 2 shows (i.e., $H_0: p < 0.05$ for 5 and 30 minute examinations) that experimental group subjects performed approximately 1.5 times better when they trained with the keypad graphic and haptic interaction than when they trained with the keypad graphic only. These results were obtained even though the design of the experiments was such that independent variables affected the training of only the action part of sequences (i.e., the training of the sequence label was not affected by the independent variables) and experimental group examinations required recall of sequences memorized under two different conditions. The results of the two memory experiments suggest that: 1) small-scale haptic involvement can act as a training aid for a purely cognitive task; and 2) small-scale haptic involvement can compliment visual stimuli as an aid to training in a purely cognitive task.

Thesis Supervisor : Nathaniel I. Durlach

Title: Senior Scientist, Department of Electrical Engineering and Computer

1.0 Introduction

The increased commercial availability of interface and computational devices supporting the development of multimodal, spatially-oriented, interactive human/machine interfaces has contributed to a burgeoning interest in the field of human/computer interaction (Biocca, 1992). Devices such as binocular stereoscopic helmet mounted displays, binaural stereophonic spatialization devices, haptic interaction devices, and unobtrusive position/orientation sensors are enabling the creation of human/machine interfaces for interaction with telerobots and computational models of a form and type which were previously found only in specialized research and military efforts (e.g., Freund, 1986; Shaker and Wise, 1988; Tachi, Arai, and Maeda, 1989; Aviles, Hughes, Everett, et. al., 1990). In their most extreme manifestations, systems utilizing these technologies are attempting to remove the boundaries associated with devices mediating interaction between humans and remote task environments (physical and computer generated) and, in some sense, project the human into these remote or virtual environments. Virtual environment (VE) systems in particular are experiencing a high degree of interest from the military, commercial, and private sectors. Current and foreseen application domains for VE systems include: (1) design, manufacturing, and marketing; (2) medicine and health care; (3) entertainment; and (4) training (Durlach and Mavor, 1995). This discussion will focus on research issues associated with using VE systems and technology for training.

From a training perspective, some key features of virtual environments are that they are: 1) interactive and adaptive; 2) reconfigurable in software; 3) multimodal; and 4) can generate supernormal situations (Durlach, Pew, Aviles, DiZio, and Zeltzer, 1992). These features allow VE-based training systems to be developed which can be precisely tailored to specific tasks and individuals. In many ways, however, VE-based training systems may be viewed as natural outgrowths of previous simulation-based training systems (e.g., flight trainers). In some form, all four aforementioned key VE features apply to these "classical" trainers. In what ways, therefore, are VE-based systems unique?

First of all, in contrast to classical simulation systems, VE-based systems can be highly reconfigurable for BOTH the far-field (i.e., objects and interactions out of reach) and the near-field (i.e., objects and interactions within reach). Classical flight simulators, for example, only reconfigure the scenery and environment outside of the airplane (i.e., the far-field) and do not easily allow flexibility in the airplane display and control layout (i.e., the near-field). In contrast, a VE-based training system can simulate a variety of near- and far-field configurations while utilizing the same human/machine interface devices. The second key characteristic of VE-based training systems is the potential for allowing multimodal (i.e., visual, auditory, and haptic) interactions. Unlike classical computer-based simulation systems, near-field simulations can be generated which may be seen, felt, and heard in a spatially-oriented manner. Virtual environment systems, therefore, are unique in that they can generate reconfigurable multimodal sensorimotor near-field interactions.

A large and growing percentage of the tasks encountered in military, commercial, and private enterprises, however, are predominantly cognitive in nature. For example, many of these tasks require mainly that the human operator control and monitor a system through the use of a control panel. Although the main interaction with such control panels is through the sensorimotor system, the requirement to know what controls to operate and when to operate them is largely a cognitive task.

One viewpoint is that, given the aforementioned trends, training for manipulative skills is unimportant. Strong proponents of this view argue that the majority of IMPORTANT skills, independent of frequency of occurrence, are largely cognitive in nature (Welford, 1976). By extension, it is argued that, since sensorimotor skills are not the main skills which must be trained, sensorimotor involvement in VE-based training systems is not required or is of secondary importance. In other words, it is assumed that sensorimotor involvement is not needed for the training of cognitive skills. In the research described in this thesis, these assumptions were tested and the effect of multimodal sensorimotor involvement in the training of predominantly cognitive tasks was examined in a task which requires subjects to: 1) identify a discrete system state, and 2) identify a series of discrete actions which must be taken in response to that state. Training effectiveness is measured and compared between 1) "classical" text-based training methods and 2) training methods which provide visual stimuli and which require haptic sensorimotor involvement.

2.0 Background

Intellectual or cognitive skills "link perception and action and are concerned with translating perceptual input into a skilled response by using appropriate decisions."

Colley and Beach, 1989, p. 2

From a training research viewpoint, VE-based systems are interesting only to the extent that training efficiency, skill performance, or skill retention are influenced and/or psychophysical insight may be gained. From a practical viewpoint, issues of cost (e.g., training system cost, time/cost required to train to a certain level of proficiency, and resource use costs), safety (i.e., limiting trainee exposure to threatening situations), portability (i.e., ability to effect training at a variety of sites) and reconfigurability (i.e., ability to use the same system to train a variety of skills) are of import. Lower training effectiveness measures may be tolerable in certain situations if the system's practical factors evaluate positively for a given training system. Nonetheless, a crucial factor in the evaluation of any training methodology or system is its training effectiveness.

To date, however, research relevant to assessing the quantitative impact of multimodal sensorimotor involvement in the training of predominantly cognitive tasks has been minimal. Current models that are directed towards the learning of cognitive skills often fail to consider the possible effects of sensorimotor involvement in the learning of these cognitive skills (e.g., Anderson 1990, Posner, 1989; Preece, Rogers, Sharp, Benyon, Holland, and Carey, 1994; Benyon and Murray, 1993; Frederiksen and White, 1993; Roberts, 1993). Most relevant research has occurred in the use of

subject performed tasks (SPTs) and experimenter performed tasks (EPTs) in event/action recall memory experiments (e.g., Cohen, 1981; Cohen, Peterson, and Mantini-Atkinson, 1987; Koriat, Ben-Zur, and Nussbaum, 1990; McAndrews and Milner, 1991). These experiments, however, explore the effects of multimodal sensorimotor involvement on the recall of tasks which are inherently sensorimotor in nature. To the knowledge of the author, principled examination of the impact of multimodal sensorimotor involvement on the training of predominantly cognitive skills is non-existent.

Why pursue work in this area? First of all, insight into the role of multimodal sensorimotor involvement in training will have strong implications for the practical design of VE-based training systems (e.g., assessing if the added complexity and cost of haptic and auditory components is worthwhile for a given training task). Second, experimental results may aid in extending current cognitive training theory to encompass sensorimotor involvement. The question still remains, however, as to what evidence or plausible mechanisms exist which indicate that multimodal sensorimotor involvement may have an effect on the training of cognitive tasks?

Changes in training performance, like changes in any cognitive process, can be brought about by effects on human perception, attention, or memory (Anderson 1990). Effective memory, in turn, may be promoted by: 1) organizing, relating, and searching for meaning in material; 2) forming a visual image of the material to be remembered; and 3) memorizing material in a manner which promotes cues for easier retrieval (Gazzaniga, 1988; Schacter, 1989; Bartlett, 1932; Paivio, 1971; Tulving and Donaldson, 1972; Tulving, 1983). Many of the potential mechanisms by which multimodal sensorimotor involvement may affect cognitive training involve one or more of these routes to promoting memory. The following are some, not necessarily independent, ways in which multimodal sensorimotor involvement may increase training effectiveness through effects on human perception, attention, and/or memory:

1. Dimensionality of Stimuli

As discussed in Miller's classic paper (Miller, 1956), increased dimensionality of stimuli leads to increased span of absolute judgment and immediate memory. Multimodal interactions, where each sensory modality may be viewed as at least a single dimension, therefore, may potentially influence training through effects on perception and memory.

2. Context Isomorphisms

A large body of research (e.g., see Anderson, 1990; Godden and Baddeley, 1975; Smith, Brown, and Toman, 1978) indicates that matching the training context to the recall context positively influences memory. Inasmuch as full multimodal feedback and sensorimotor involvement during training allows the training context to more closely approximate the recall or performance context, memory may be influenced.

3. Depth or Elaborateness of Processing

Memory research has shown that increasing the amount or elaborateness of mental processing during encoding increases recall (Anderson, 1990; Craik and Lockhart, 1972; Slamecka and Graf, 1978). Sensorimotor action has been shown to require significant cognitive processing (Jordan and Rosenbaum, 1989) and each modality which must be processed also increases the mental load (Anderson, 1990). It is possible, therefore, that adding multimodal sensorimotor involvement to cognitive tasks may influence recall. The previously mentioned SPT/EPT studies did precisely that and have shown increased event/action recall.

Although the mechanisms are not clear, Koriat, Ben-Zur, and Nussbaum's (1990) work is particularly interesting. Subjects were given a series of verbal instructions (e.g., move the cup) and their recall was tested on the basis of expected and actual performance mode (i.e., verbal versus action). Subjects evidenced superior memory in the case that the performance mode was through action and in the case that it was EXPECTED that the performance mode would be through action.

4. Memory Cueing

Perceptual-motor skills can be retained after very long periods of time without practice (Annett, 1989). This contrasts sharply with the retention of verbal material which may decline quite rapidly. A classic experiment showing the retention of perceptual motor skills was performed by Hill (1934, 1957). In this study, one day of typing practice was sufficient to achieve a level of skill which required 27 days of practice when originally learned 25 years in the past. If the cognitive components of a task could be related to a perceptual-motor skill in a robust manner, then multimodal sensorimotor involvement could influence cognitive skills training through memory mechanisms (i.e., by perceptual-motor memories providing strong cues to cognitive memories).

Although the list of possible ways in which multimodal sensorimotor involvement may influence training can be further elaborated and organized, and issues concerning theory addressed more fully, clearly there is reason to suspect that multimodal sensorimotor involvement may have an effect on the training of cognitive skills.

Key questions which must be answered concerning the role of multimodal sensorimotor involvement on cognitive skills training include: Can sensorimotor involvement (i.e., using a multimodal human/machine interface) aid in training cognitive skills and tasks? Which skills? Which tasks? What aspects of training are affected (i.e., efficiency, performance, and/or retention)? What is the relative impact of the various modalities? What are the key intramodal features? How can current theory be extended to account for multimodal effects?

The research described in this subsequent sections of this thesis takes the smallest of steps along the path towards answering these questions. The EXISTENCE and EXTENT of the impact of multimodal sensorimotor involvement on training effectiveness are examined within the context of an explicit memory task.

3.0 Research

"It seems, then, that we owe to memory almost all that we either have or are; that our ideas and concepts are its work, and that our every-day perception, thought and movement is derived from this source."

Hering, 1920, p. 75

Memory plays a crucial role in most any cognitive process. In fact, memorization may be viewed as the most elemental of cognitive tasks (Schacter, 1989). A simple explicit or intentional memory task, therefore, was chosen to begin to explore the effect of multimodal sensorimotor involvement on the training of cognitive tasks. Furthermore, the effect of small-scale haptic interaction as an aid to training was chosen as the focus of the work described in this thesis. There were three major reasons for this choice: 1) there is little or no work on the effect of haptic involvement on the learning of cognitive tasks; 2) requiring haptic interaction significantly affects the complexity and cost of a training system; and 3) systems providing small scale haptic interaction (i.e., force feedback component) are commercially available.

Two experiments, focusing on the role of haptic sensorimotor involvement as a training aid in an explicit memory task, are discussed in this thesis. The common features of these experiments are described in Section 3.1 The specific characteristics and results of the experiments are discussed in Sections 3.2 and 3.3. A discussion of the results is found in Section 4.0.

3.1 General Method

This section describes the memorization task, subjects, facilities, pre-trial procedure, experimental procedure, and scoring and analysis methods used in the experiments outlined in Sections 3.2 and 3.3.

3.1.1 Task

Subjects are asked to memorize simple **sequences** of letters and numbers. A **sequence** consists of **elements**. The first element of a sequence is known as the **sequence label**. The sequence label consists of an 'S' followed by another letter. All other elements in a sequence consist of a letter and a number.

Two examples of a typical sequence are:

SC A2 B3 A1
SM B2 A2 C1 C3

This particular memorization task was selected for two reasons. First, at face-value, it is a pure memorization task. Subjects simply attempt to learn and recall a sequence of letters and numbers under conditions of experimental interest. Second, the process which subjects go through in the memorization task mirrors two of the major activities encountered when interacting with real-world equipment through physical control panels (e.g., Figure 3.1-1) or graphical user interfaces: 1) the identification of the state of the system to be acted on; and 2) the recall of the series of actions which are to be taken in that state. In other words, the memorization task was designed to be an abstraction of some of the major actions taken when interacting with real-world devices or systems.

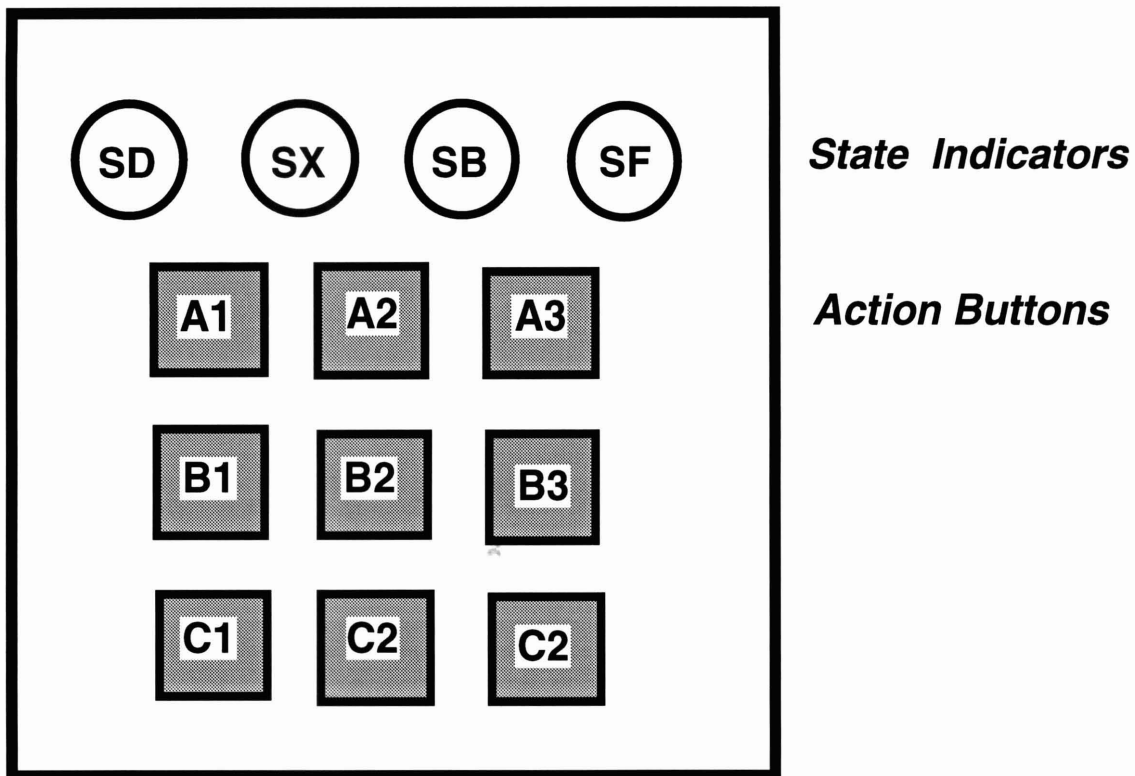


Figure 3.1-1
Example Control Panel

Specifically, the sequence label can be thought of as the system state (i.e., the **state part** of the sequence) and the other elements of the sequence as the series of discrete actions which must be performed when the system is in that state (i.e., the **action part** of the sequence). Given this view, the experiments in this thesis attempt to explore methods and conditions which may affect the learning of this state/action association. Of course, in real-systems, the process required to identify system state, and the specifics of the actions to be taken, are likely to be much more complex. The chosen memorization task is representative of situations where a single indicator unambiguously identifies system state and there is a set sequence of actions which should be taken when the system is in a given state. In a real-system, the expression of

these actions may have some sensorimotor component (e.g., pushing a series of buttons), but the component of immediate interest is the recall of these actions (i.e., which actions to take, not how these actions are to be taken).

Pilot trials helped to determine the length and number of sequences used in the experiments. The training of eight sequences of length five (i.e., a sequence label and four other elements) was determined to be sufficiently challenging for experimental purposes. These eight sequences are organized into two *sets* of four sequences. In addition, there are a total of nine possible distinct elements in the action part of the sequence. These sequence elements are: A1, A2, A3, B1, B2, B3, C1, C2, and C3.

The elements in a given sequence are randomly generated. The sequence label, as well as the first element of the action part of the sequence, are unique (i.e., generated without replacement) across all the sequences used in the experiments. All other sequence elements are randomly generated with replacement. It is desirable, for scoring purposes, that there be a unique association between the state part and the action part of a sequence. The sequence label, therefore, must be unique within a given experiment. A letter, rather than a number, was chosen for the second character in the sequence label so that: 1) the number of unique sequence labels was sufficiently large to support many experiments; and 2) to differentiate a sequence label from the rest of the elements in the sequence. The motivation for the unique first element in the action part of the sequence was to maximize the separation of sequences in haptic stimulus space. It was deemed desirable that there be a unique initial haptic stimulus (i.e., a different point in haptic stimulus space) for each sequence. The experiments vary the training conditions only for the action part of the sequence. Therefore, the first element of the action part of all the sequences is unique. Furthermore, it was felt that this minor manipulation would have little or no effect on the nature of the task itself.

3.1.2 Subjects

A total of 40 paid subjects, ranging in age from 17 to 47 years old, were drawn from the student, staff, and associated population of the Massachusetts Institute of Technology. Naive subjects were used for every experiment. No subject was used in more than one of the experiments.

3.1.3 Facilities, Tools, and Apparatus

The two experiments described in this thesis used written materials for presentation and examination purposes. Subjects control their progress through the materials by using a foot-activated electronic timer. The timer controls a light which is activated for a pre-determined time interval and resets automatically for the next experimental phase. Experiments were run in a quiet laboratory setting and subjects' adherence to protocol was monitored during the experiments.

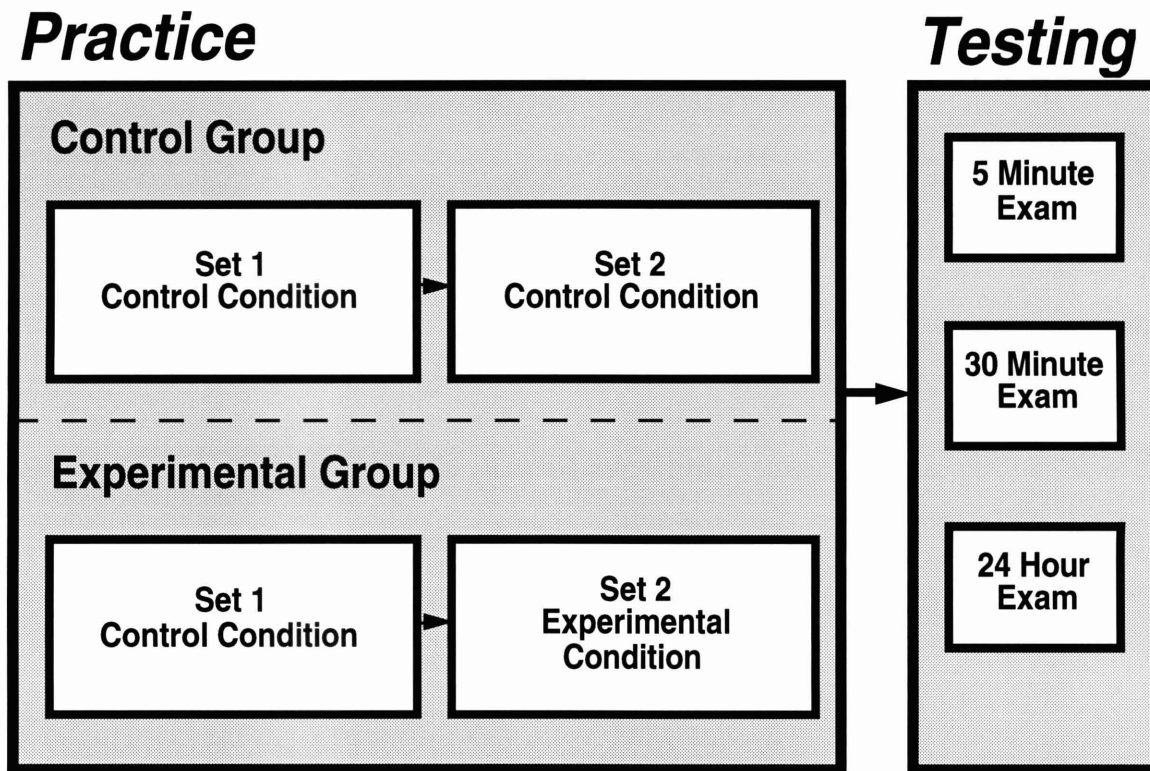
3.1.4 Pre-Trial Procedure

Subjects are given a written pre-trial briefing as well as verbal instructions emphasizing key points contained in the written materials. They are informed that the experiments are aimed at providing a better understanding of human memory and cognition. In addition, they are briefed on the form of the sequences to be memorized and the procedure for examination of their recall. In addition, administrative issues, such as pay and informed consent, are handled during the pre-trial period.

Subjects are familiarized with the experimental materials and devices using a single test sequence of length 4. Material pertaining to a particular experimental condition is only introduced immediately prior to training under that condition. This is done to avoid influencing subjects with organizational principals that may be inherent in the materials until immediately before they will train using these materials.

3.1.5 Procedure

Overall Design. A schematic representation of the overall procedure for a prototypical experiment is shown in Figure 3.1-2. An experiment consists of: 1) a **practice session**, during which a subject attempts to learn sequences under controlled experimental conditions; and 2) a **testing session**, during which the subject's ability to recall the sequences memorized during the practice session is measured. Two collections or **sets** of sequences are presented to each subject during the practice session. Subjects in the **control group** learn both sets of sequences under the same condition (i.e., the control condition). Subjects in the **experimental group** learn the first set of sequences under the control condition and the second set of sequences under the experimental condition of interest (i.e., a condition involving haptic interaction). The time that a subject practices on a set is controlled and is the same for all sets and experimental conditions. Identical written examinations of recall performance are given 5 minutes, 30 minutes, and 24 hours after practice on all the sequences has been completed. In the written examination, a subject is provided with the sequence labels from both sets of sequences in random order and asked to provide the sequence elements which correspond to a given sequence label. Thus, for the experimental group, sequences memorized under different conditions are intermixed in the examination. There is also a subjective survey given at the completion of the experiment.



**Figure 3.1-2
Prototypical Experiment**

An experiment requires a total of approximately 2 hours for two sessions over two consecutive days. The first session is approximately 1.5 hours in length. During this session: 1) a subject is introduced to the experiment; 2) administrative details are taken care of; 2) the subject practices (i.e., trains) on both sets of sequences; and 3) the 5 minute and 30 minute post practice test are administered. During the second session, which lasts approximately one-half hour, the 24 hour test is given, the subjective survey is taken, and any questions which subjects might have concerning the experiment or their performance are answered.

Details of Practice Session. Figure 3.1-3 provides further information on the structure of the practice portion of a prototypical experiment. As previously mentioned, during practice, subjects are exposed to the sequences in one set (i.e., Set 1) under training Condition A and they are exposed to the sequences in the second set (i.e., Set 2) under training Condition B. For a given condition, a subject is exposed to a sequence three times or ***cycles*** before moving on to the next sequence in the set. The subject is exposed to the entire set in this way twice in a row.

Within each cycle, the sequence is presented to the subject for T_p seconds. This ***presentation period*** is followed immediately by an ***examination and feedback period*** T_{ef} seconds in length. During the examination and feedback period, the active recall of the sequence is encouraged by requiring the subject to write down the entire sequence (including the sequence element) to which they were just exposed. The subject is then provided with the correct answer and asked to grade themselves.

Self-grading consists of the subject annotating their answer with a "+" next to each correctly recalled sequence element and a "-" next to each incorrectly recalled sequence element. Note that both the characters in an element need to be correct and the element must be in its proper position in the sequence for the answer to be considered correct. In addition to actively encouraging memorization during the practice period, it was also thought that information on the subject's performance during the feedback periods would give insight into their learning rate. After the presentation and examination / feedback periods, there is a waiting period of T_{wp} seconds in length before the next cycle begins.

The overall time of a cycle (i.e., presentation, examination/feedback, and waiting period for a single sequence) was set at 30 seconds ($T_p = T_{ef} = T_{wp} = 10$ seconds) based on results of pilot trials. A subject, therefore, practices on a given sequence for a total of 3 minutes and on a set for a total of 12 minutes. In addition, there is a five minute waiting period between practice on Set 1 and Set 2.

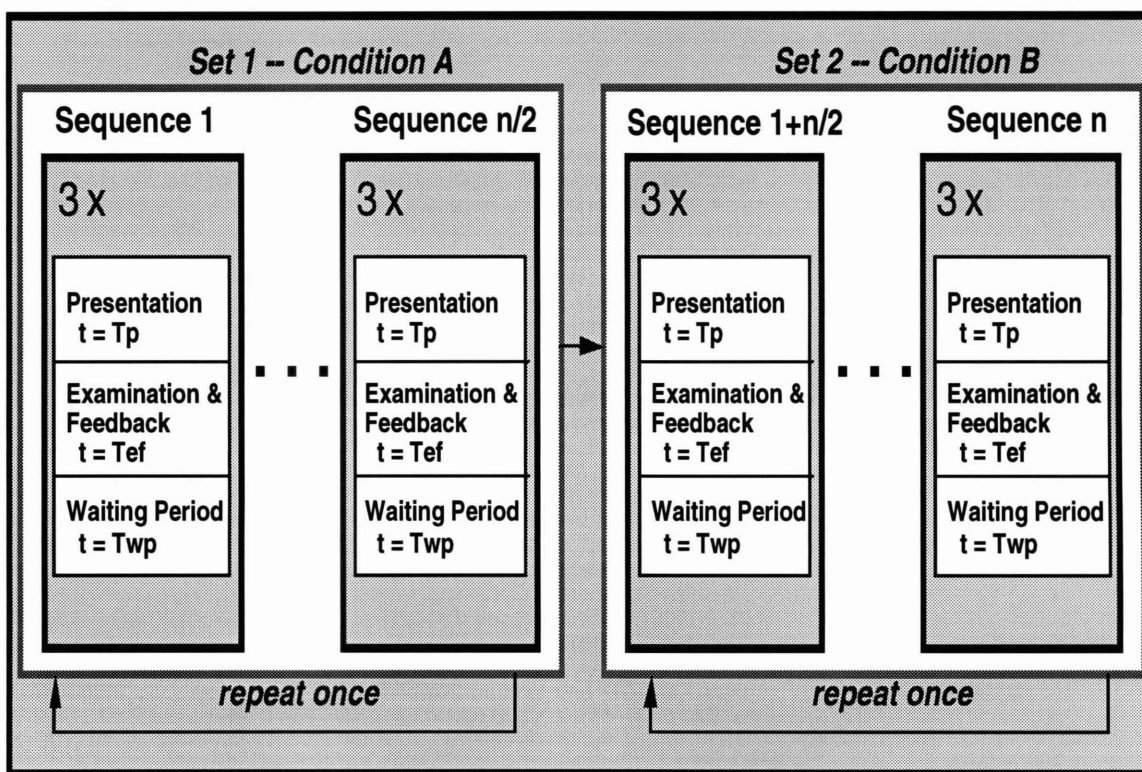


Figure 3.1-3
Prototypical Practice Period

Details of Testing Session. After a subject finishes practice sessions on both sets of sequences, his or her recall performance is measured using a written examination. In this examination, the subject is provided with sequence labels from both sets of sequences and asked to provide the sequence elements which correspond to a given sequence label. The presentation order of sequence labels in the written test is random and each sequence label from both sets of sequences appears exactly once. Subjects are asked to attempt to recall the sequence elements corresponding to a

given sequence label as best they can. The examination is not forced-choice and no time limit is used during the examination. No feedback is provided to subjects on their recall performance during these tests. The same written examination is given 5 minutes, 30 minutes, and 24 hours after the completion of practice/training.

In addition to the objective tests, subjects are also asked to fill out a questionnaire in an attempt to gain insight into individual learning strategies and the effect of the presentation conditions. Subjects, if interested, are informed of their performance after the completion of the final examination and the questionnaire.

Appendix C contains a discussion concerning the experimental impact of not using a forced-choice methodology. There is also a discussion of some shortcomings in the pre-trial procedure.

3.1.6. Scoring and Data Analysis

The results of the written examinations are scored on a per element basis. In order to be graded as correct, a sequence element must be completely recalled (i.e., both characters must be correct) and must be in the proper position within the sequence. A single point is awarded for every sequence element correctly recalled. Incorrect answers receive zero points. For scoring purposes, no differentiation is made between an incorrect answer and a failure to answer.

The experiments were designed so that the expected (and encountered) between-subject variability would not mask the relative effects of the experimental conditions. The control procedure tests the null hypothesis between subject performance on the two sets of sequences under the control condition. It verifies that neither set of sequences is inherently more difficult under the control condition. The experimental group and conditions test the null hypothesis between conditions. Results of each examination (i.e., 5 minute, 30 minute, and 24 hour) are treated and analyzed independently. The standard error bars used in many of the figures in this thesis, therefore, pertain to the data for a particular group (i.e., control or experimental) and examination. So the variability indicated by the standard error bars is completely BETWEEN subjects.

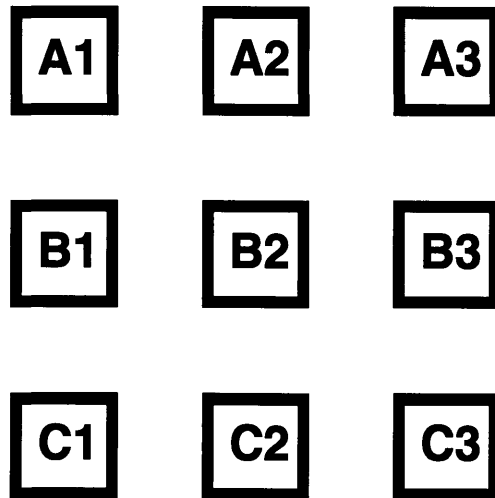
The main measures used to determine the statistical significance of performance differences are the paired and unpaired versions of the Student's T test. Since positive and negative performance differences between experimental treatments are contemplated, the two-tailed version of the test is used. For comparison of performance WITHIN a group, data samples (i.e., the examination results) are not treated as independent but as paired observations of the same individuals before and after treatment. The DIFFERENCE between performance on Set 1 and Set 2, therefore, is analyzed for statistical significance and the paired, two-tailed, version of the Student's T test is used. This method tends to remove the effects of individual differences between subjects and results in a large increase in the precision of statistical measurement. For comparison of performance ACROSS groups, however, each group's performance data is treated as an independent sample and the unpaired, two-tailed version of the Student's T test is used.

3.2 Memory Experiment 1

Overview. The first memory experiment was designed to contrast performance on the memory task between a simple, text-based, presentation condition (Condition A) and a presentation condition which requires haptic interaction (Condition B). Experimental materials were completely paper-based and the experimental apparatus was limited to a foot-switch activated timer with a lamp attached. The timer was used by subjects to pace their progress through the experimental materials. For this experiment, subjects are not made explicitly aware of the range of sequence elements or the length of sequences.

Condition A -- Text Only Presentation. The text-based presentation condition is representative of a significant portion of classical (both "manual" and computer-aided) training -- rote memorization. Subjects are simply presented with the sequences visually as a text string (e.g., SZ A3 B3 C2 A1) and asked to memorize them.

Condition B -- Visual Graphic and Haptic Interaction Presentation. The visual stimulus for this condition consists of the text string, as presented in Condition A, plus a visual graphic or keypad. The visual graphic shows the nine possible sequence elements (i.e., other than the sequence label), organized into three rows and three columns as shown below:



During presentation, subjects are required to place their index finger on the keypad square corresponding to the sequence element they are currently attempting to memorize. They are asked to slide the finger between keypad squares when transitioning from one sequence element to the next. Note that the sequence label is not included in the graphic.

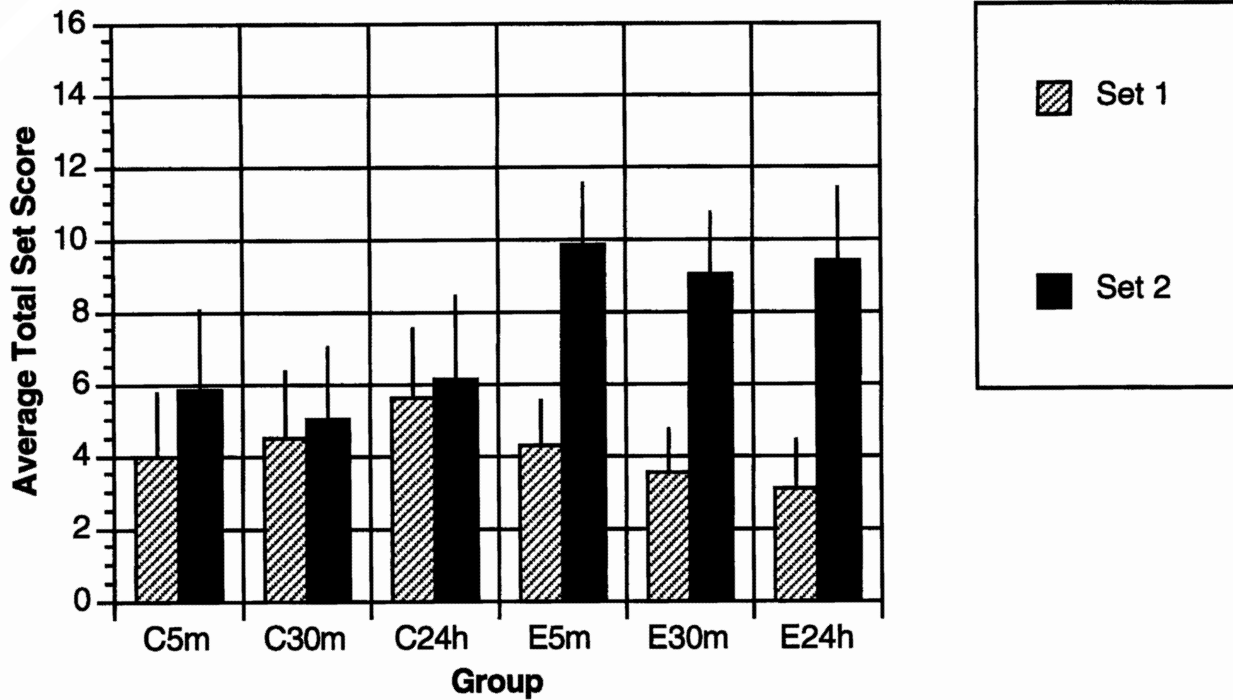
Subjects. A total of twenty subjects (excluding subjects used for pilot probes) were employed in this experiment. Eight subjects were used for the control condition (i.e., Condition A for both sets of sequences) and 12 subjects were used for the experimental condition (i.e., Condition A for Set 1 and Condition B for Set 2).

Results. Subjects' responses to the written examinations at 5 minutes, 30 minutes, and 24 hours after the completion of practice were scored. As expected, the variability of subjects' scores is quite high. Scores ranged from perfect to zero. Plots of total set scores for each subject organized by group and examination are found in Appendix A. In addition, these figures include: 1) the mean scores for the group (i.e., Avg) annotated with a standard error bar; and 2) the probability, using the paired, two-tail, version of Student's T Test, that the differences between Set 1 scores and Set 2 scores is due to chance (i.e., H_0 -- the null hypothesis).

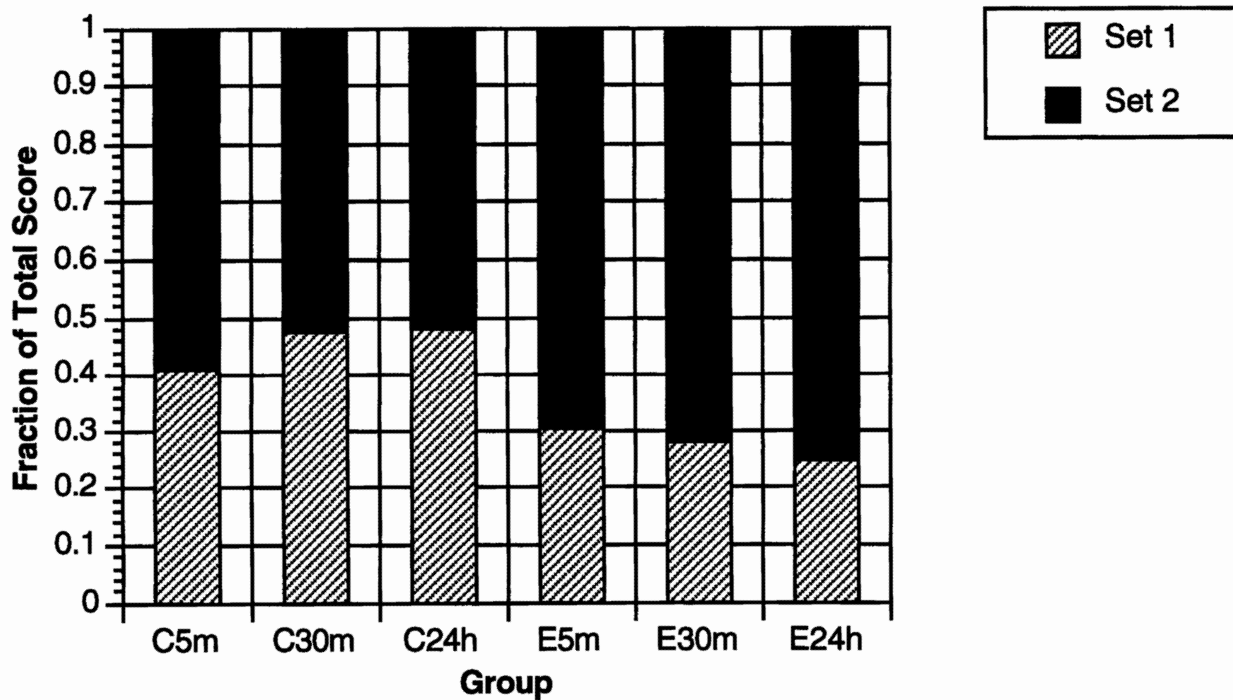
Viewing the data as a function of the difference between a given subject's performance on Set 1 and their performance on Set 2, however, is more revealing of the effect of the conditions under consideration. A paired, two-tailed, Student's T test was used to examine the statistical significance (i.e., probability of the null hypothesis H_0) of any differences between subjects' performance on Set 1 and Set 2. Applying the T test on the control group indicates that any differences between performance on Set 1 and Set 2 under Condition A are not statistically significant to the $p < 0.05$ criterion (i.e., $H_0(5m)$: $p < 0.311$, $H_0(30m)$: $p < 0.695$, $H_0(24hr)$: $p < 0.726$). Applying the T test to the experimental group, on the other hand, indicates that the null hypothesis can be rejected and that differences between Set 1 under Condition A and Set 2 under Condition B are statistically significant to the $p < 0.05$ criterion (i.e., $H_0(5m)$: $p < 0.004$, $H_0(30m)$: $p < 0.005$, $H_0(24hr)$: $p < 0.008$).

These trends may be seen in Figures 3.2-1 and 3.2-2. Figure 3.2-1 shows the average total set score (i.e., average scores summed by set) by group. Standard error bars are also included. Figure 3.2-2 plots average total set score as a fraction of total score. Data for the control group (i.e., Group C) is on the left half of each figure and the experimental group (i.e., Group E) is on the right half of each figure. There are individual columns for the results of the 5 minute (i.e., 5m), 30 minute (i.e., 30m), and 24 hour (i.e., 24h) examinations. On the average, subjects in the experimental group scored over two and one-half times better under the experimental condition than under the control condition using the results of all three examinations.

The results of Memory Experiment 1 are examined more fully in the Discussion section of this document (Section 4.0).



**Figure 3.2-1
Average Set Scores**



**Figure 3.2-2
Fraction of Total Score Using Average Set Scores**

3.3 Memory Experiment 2

Overview. The results of Memory Experiment 1 suggest that small-scale sensorimotor involvement may affect training effectiveness in the explicit memory task. The control and experimental conditions of Memory Experiment 1, however, are only able to establish that Condition B (i.e., visual graphic and haptic interaction presentation) has a statistically significant, positive, effect on recall performance. The experimental condition (i.e., Condition B), however, introduced BOTH haptic and visual stimuli. Memory Experiment 2 is designed to separate the effects of the haptic and graphic components of the presentation on recall performance. In this experiment, both the control and experimental visual stimuli are identical to those of Condition B of Memory Experiment 1. The experimental condition, however, requires haptic interaction whereas the control condition does not.

For this experiment, subjects are shown the visual graphic during the pre-trial procedure and are informed that sequence elements are chosen from the elements represented in the graphic. In all other respects, the procedures for Memory Experiment 2 are identical to those of Memory Experiment 1. Experimental materials were completely paper-based and the experimental apparatus was limited to a foot-switch activated timer with a lamp attached. The timer was used by subjects to pace their progress through the experimental materials.

Condition A -- Visual Graphic Presentation. The visual stimuli for this presentation condition are identical to that of Condition B in Memory Experiment 1. The visual stimuli consist of a text-string (i.e., the full sequence to be memorized) and the visual graphic used in Memory Experiment 1. There is no haptic interaction, however, with the visual graphic. Subjects are explicitly instructed not to use their hands on the graphic. Moreover, they must keep their hands on the arms of their chair during sequence presentation in order to help enforce this policy.

Condition B -- Visual Graphic and Haptic Interaction Presentation. This condition is identical to Condition B of Memory Experiment 1. The visual stimuli are identical to that of Condition A of this experiment. Subjects, however, do interact haptically with the visual graphic in the same manner as in Condition B of Memory Experiment 1. During presentation, subjects are required to place their index finger on the keypad square corresponding to sequence element they are currently attempting to memorize. They are asked to slide their finger between keypad squares when transitioning from one sequence element to the next.

Subjects. A total of twenty subjects (excluding subjects used for pilot probes) were employed for this experiment. Eight subjects were used for the control condition (i.e., Condition A for both sets of sequences) and 12 subjects were used for the experimental condition (i.e., Condition A for Set 1 and Condition B for Set 2).

Results. Subjects' responses to the written examinations at 5 minutes, 30 minutes, and 24 hours after the completion of practice were scored. As in Memory Experiment 1, the variability of subjects' scores is quite high. Plots of total set scores for each subject organized by group and examination are found in Appendix B. In addition these figures include: 1) the mean scores for the group (i.e., Avg) annotated with a

standard error bar; and 2) the probability, using the paired, two-tail, version of Student's T Test, that the differences between Set 1 scores and Set 2 scores is due to chance (i.e., H_0 -- the null hypothesis).

A paired, two-tailed, Student's T test was used to examine the statistical significance (i.e., probability of the null hypothesis H_0) of any differences between subjects' performance on Set 1 and Set 2. Applying the T test on the control group indicates that any differences between performance on Set 1 and Set 2 under Condition A are not statistically significant to the $p < 0.05$ criterion (i.e., $H_0(5m)$: $p < 0.715$, $H_0(30m)$: $p < 0.751$, $H_0(24hr)$: $p < 0.950$). Applying the T test to the experimental group, on the other hand, indicates that the null hypothesis can be rejected and that differences between Set 1 under Condition A and Set 2 under Condition B are statistically significant to the $p < 0.05$ criterion for the results of the 5 minute and 30 minute tests (i.e., $H_0(5m)$: $p < 0.034$, $H_0(30m)$: $p < 0.041$). Performance differences under Condition A and Condition B did not achieve statistical significance for the 24 hour test (i.e., $H_0(24hr)$: $p < 0.085$).

These trends may be seen in Figures 3.3-1 and 3.3-2. Figure 3.3-1 shows the average total set score (i.e., average scores summed by set) by group. Standard error bars are also included. Figure 3.3-2 plots average total set score as a fraction of total score. Data for the control group (i.e., Group C) is on the left half of each figure and the experimental group (i.e., Group E) is on the right half of each figure. There are individual columns for the results of the 5 minute (i.e., 5m), 30 minute (i.e., 30m), and 24 hour (i.e., 24h) examinations. On the average, subjects in the experimental group scored approximately 1.5 times better under the experimental condition than under the control condition using results from the 5 and 30 minute examinations.

The results of Memory Experiment 2 are examined more fully in the Discussion section of this document (Section 4.0).

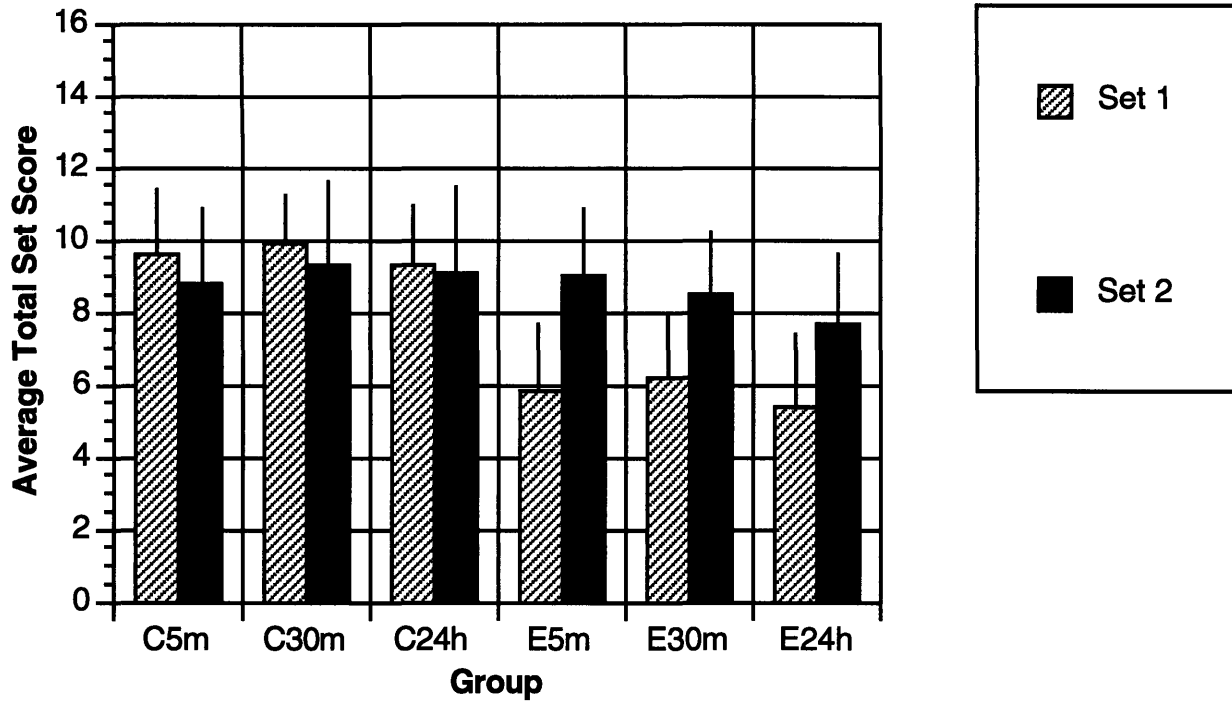


Figure 3.3-1
Average Set Scores

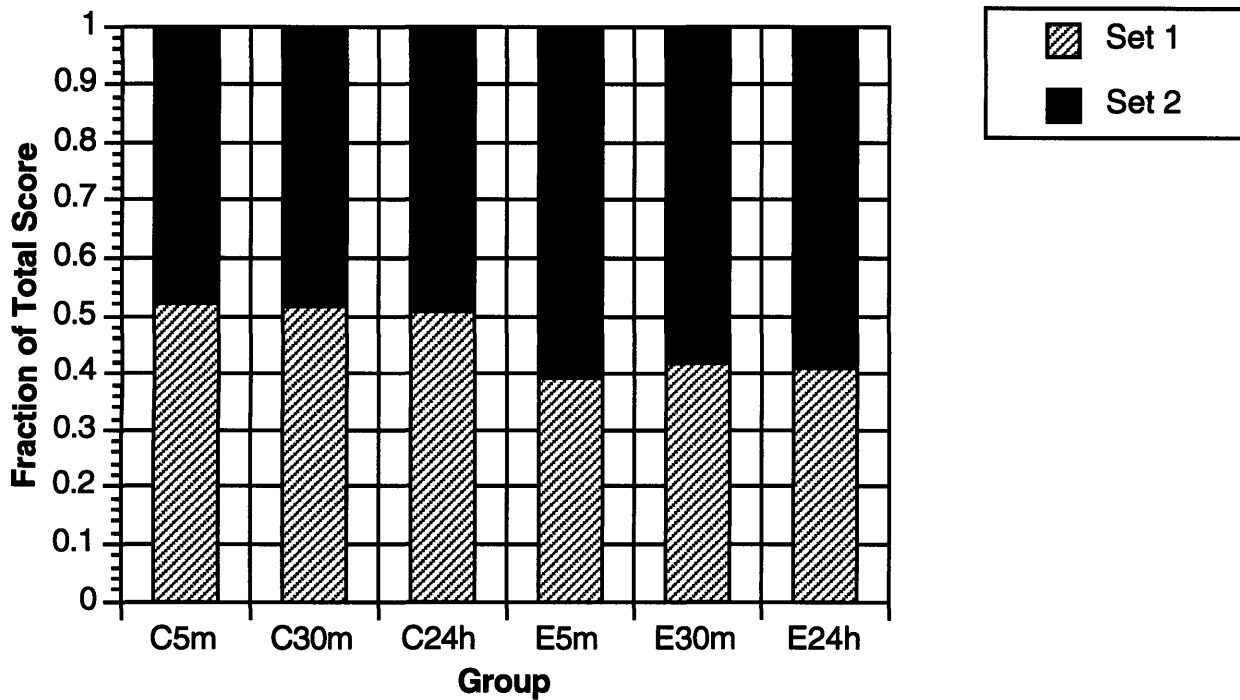


Figure 3.3-2
Fraction of Total Score Using Average Set Scores

4.0 Discussion

Results from Memory Experiment 1 show (i.e., $H_0: p < 0.01$ for 5 minute, 30 minute, and 24 hour examinations) that experimental group subjects performed approximately 2.5 times better when they trained with both visual stimuli (i.e., keypad graphic) and haptic involvement than when they trained without these visual and haptic stimuli. Memory Experiment 2 shows (i.e., $H_0: p < 0.05$ for 5 and 30 minute examinations) that experimental group subjects performed approximately 1.5 times better when they trained with the keypad graphic and haptic interaction than when they trained with the keypad graphic only. It should be noted that these results were obtained even though the design of the experiments was such that independent variables affected the training of only the action part of sequences (i.e., the training of the sequence label was not affected by the independent variables) and experimental group examinations required recall of sequences memorized under two different conditions. Furthermore, the observed results were obtained using a highly selected subject population whose everyday focus is on cognitive activities (i.e., the population of The Massachusetts Institute of Technology).

Establishing Training Effectiveness. It should be emphasized that the results of Memory Experiment 1 and 2 showed RELATIVE performance increases within the experimental groups under the experimental conditions. Experimental group subjects: 1) recalled sequences 2.5 times better when both visual stimuli (i.e., keypad graphic) and haptic interaction were used during practice; and 2) recalled sequences 1.5 times better when haptic interaction was added to the keypad graphic. Questions concerning the true impact of the experimental conditions on training effectiveness, however, still remain.

First of all, do the experimental conditions result in OVERALL increased recall performance? In other words, will subjects trained under the experimental conditions have higher recall scores than subjects trained under the control conditions? To support a claim that the experimental conditions in the memory experiments resulted in OVERALL increased recall, the Set 2 scores of experimental subjects should be higher than the Set 2 scores of control subjects. This requirement seems to be met by the results of Memory Experiment 1 but not by the results of Memory Experiment 2. The Set 2 scores of the experimental group in Memory Experiment 2 seem to be the same, if not lower, than the Set 2 scores of the control group. Across experiments, in order for the RELATIVE trends established by Memory Experiment 1 and 2 to be supported as OVERALL trends, Set 2 recall performance should obey the following order: 1) Experiment 1 control group (i.e., no keypad graphic and no haptic interaction) scores should be the lowest; 2) Experiment 2 control group (i.e., keypad graphic only) scores should be the next highest; and 3) Experiment 1 and experiment 2 experimental group (i.e., keypad graphic and haptic interaction) scores should be the highest and should be the same as each other. Although the Experiment 1 control group Set 2 scores are clearly the lowest overall scores, the other Experiment 1 and 2 Set 2 scores do not seem to obey the postulated ordering. Finally, to help clearly establish OVERALL trends, Set 1 scores, within each experiment, should be the same for the control and the experimental groups. This requirement does not seem to be met. There seems to be some depression of experimental group Set 1 scores in relationship to control group Set 1 scores. This trend is especially clear in Memory Experiment 2.

There are at least three possible, not necessarily independent, reasons for the aforementioned lack of support for increased OVERALL training effectiveness under the experimental conditions: 1) chance; 2) a depression of the experimental group's Set 1 performance by the Set 2 training condition; and 3) an overall depression of experimental group scores. First of all, it must be clearly stated that the statistical power of Memory Experiment 1 and 2 does not establish statistical significance (i.e., to the $p < 0.05$ criterion) for the observed trends. Part of the reason for this is that much more stringent statistical measures (e.g., the unpaired, two-tailed, version of Student's T test) that must be used for BETWEEN group comparisons. Nonetheless, these observed OVERALL trends do bear examination. Plausible explanations for the OVERALL trends can be put forth.

The depression of Set 1 performance by the Set 2 training condition could be caused by the Set 2 training condition being sufficiently more "vivid" than the Set 1 training condition that Set 1 performance is inhibited (i.e., not only do subjects recall sequences learned under the experimental conditions better but they also tend to forget sequences learned under the less vivid control conditions). This particular hypothesis, however, is not strongly supported by the results of Experiment 1. The results of this experiment should, given the relative inhibition hypothesis, evidence the strongest relative depression of experimental group Set 1 scores since the difference between the control and experimental conditions is the most extreme of the two experiments. So the postulated relative depression mechanism and hypothesis seems somewhat weak. On the other hand, the overall depression of experimental group scores could possibly be traced to the experimental test methodology. Testing of recall was accomplished using a written test where the sequence labels for both Set 1 and Set 2 were randomly presented and the action part of the sequence had to be supplied by the subject. In other words, sequences learned under the control and experimental conditions were intermixed. It is possible that subjects used different organizational principals to memorize the sequences under the two conditions. Experimental group subjects, therefore, must remember the sequences themselves AND identify/remember the organizational principal used in memorizing them. This added cognitive burden could constitute a form of interference and result in decreased overall recall. In Memory Experiment 1, the control and experimental conditions were sufficiently different that this form of interference might be minimal (i.e., it was less difficult to remember, or differentiate between, the organizational principals used in memorizing the sequences). Correspondingly, in Memory Experiment 1, there was not a large interference effect and the Set 2 scores under the experimental condition seem to be higher overall than under the control condition. In addition, there is a less pronounced depression in the experimental group's Set 1 scores. In Memory Experiment 2, on the other hand, there is less difference between the control and experimental conditions (i.e., they differ only in the haptic interaction component). This may make it more difficult to differentiate the organizational principal and may cause the overall depression of experimental group scores that is observed.

Continued speculation and lengthy discourse aside, however, the only way to resolve whether the observed trends are valid and, if so, give insight into their underlying cause(s), is by further experimentation. Experimental efforts along two fronts would seem to be warranted. First of all, running more subjects on both experiments may

help lend statistical support to the existence of the observed OVERALL trends. More important to establishing training effectiveness, however, would be an experiment in which subjects were trained under a condition with both the graphic keypad and haptic interaction for BOTH sets of sequences. This would help establish overall increased recall performance AND provide evidence on "depression effects."

Another major component of training effectiveness which requires examination is learning rate. A training approach which trains to a certain criterion faster than another is often of value. The results of memory experiments, however, did not offer insight into learning rate. Although the examination and feedback cycles during training practice were designed to encourage active memorization AND to provide learning rate information, in practice no useful information concerning learning rate was obtained. Subject performance in the examination and feedback cycles was essentially perfect. In other words, short-term recall (i.e., on the order of a few seconds) was not a problem with the sequences used in the experiments, so no data concerning increasing performance over time was obtained. Changing the design of the practice period so that subjects are exposed to a single cycle of each sequence in the set and examining between repetitions of the set might provide more useful information in any future versions of these experiments.

Other Issues. In addition to the work required to establish training effectiveness under the particular conditions used in Memory Experiments 1 and 2, there is a large amount of work needed to probe other dimensions of the memory task. The work described in this thesis did not probe the sensitivity and robustness of the observed training effects within the explicit memory task scenario. Factors which might affect the observed results include: 1) sequence length; 2) the organization of the haptic stimuli; 3) interference during practice and/or performance; and 4) the form and modalities of the response (e.g., written, verbal, etc.). The development of theory about the causal mechanisms and relationships for the effects of multimodal sensorimotor involvement on cognitive training will also require much further thought and experimentation. Three out of the four possible ways in which multimodal sensorimotor involvement may increase training effectiveness (i.e., dimensionality of stimuli, elaborateness of processing, and memory cueing) mentioned in Section 1.0, may have been at work in the memory experiments. Context isomorphisms were not a factor in the experiments since the training and transfer conditions were different. Questions on the underlying mechanism are particularly salient in the explicit memory task. Were effects primarily due to the haptic involvement helping subjects organize the sequences in an efficient fashion (e.g., as paths between the elements) or due to some more fundamental contribution (e.g., more elaborate processing or memory cueing)? From the subjects' surveys, it is clear that the haptic conditions helped them to organize the sequences as paths. Several subjects commented that they remembered the sequences as paths and then mapped these paths onto a mental image of the keypad graphic in order to recall the individual elements. Curiously, subjects in the Memory Experiment 2 Control Group (i.e., keypad graphic and no haptic involvement) did not report using the "path" strategy. Several Experimental Group subjects in Memory Experiment 2 commented that simply "touching" the graphic keypad during the experimental condition was useful. Again, deeper insight into underlying mechanisms will require further experimentation.

5.0 Summary

The work described in this thesis was undertaken within the context of key questions concerning the role of multimodal sensorimotor involvement on the training of cognitive tasks: 1) Can sensorimotor involvement (i.e., using a multimodal human/machine interface) aid in training cognitive skills and tasks? Which skills? Which tasks?; 2) What aspects of training are affected (i.e., efficiency, performance, and/or retention)? : 3) What is the relative impact of the various modalities?; 4) What are the key intramodal features?; and 5) How can current theory be extended to account for multimodal effects?

The existence and extent of effects of small-scale haptic involvement on recall performance in an explicit memory task were investigated in two related experiments involving a total of 40 subjects. The memory task was structured in such a way that it "abstracts" (see Lintern, 1996) activities encountered when interacting with real-world equipment through physical control interfaces -- the recall of a series of actions which are to be taken when a system is in a given state. Furthermore, the experiments were designed as training experiments. Haptic sensorimotor involvement and certain visual stimuli (i.e., a keypad graphic) were evaluated as an aid to training -- they were used during practice and not used during performance.

The results of the two memory experiments SUGGEST that: 1) small-scale haptic involvement may act as a training aid for a purely cognitive task; and 2) small-scale haptic involvement may compliment visual stimuli as an aid to training in a purely cognitive task. ESTABLISHING that the experimental conditions affect training effectiveness in the memory task, however, will require additional experimentation. Moreover, the work required to fully explore the role of multimodal sensorimotor involvement on the training of cognitive skills is staggering. The memory experiments, however, suggest that continued efforts to explore this role may be worthwhile.

6.0 Acknowledgments

This thesis, and the journey which it represents, would have been impossible without the support and influence of many people. Foremost is my partner and life-mate Barbara Fletcher. Without her untiring love and continued support, no part of this journey would have been possible, and no part of the larger journey of life would be as sweet. I would also like to deeply thank my thesis advisor Nathaniel Durlach for his intellectual guidance as well as his friendship. I have been exceedingly fortunate to find a friend and mentor such as Nat. Time has only increased the importance of my family in my life. Love and gratitude to my parents Vilma and Julio, great-aunt Mina, brother Richard, sister-in-law Lina, and niece Rachel.

I would like to thank Richard Pew and Yvette Tenney for some early discussions and ideas concerning sensorimotor involvement and memory. Finally, I gratefully acknowledge David Fowlkes, the Naval Air Warfare Center, and the Office of Naval Research for supporting this work.

7.0 References

- Annett, J. (1989). Training Skilled Performance, in Acquisition and Performance of Cognitive Skills, Colley, A. M., and Beech, J. R. (eds.), John Wiley & Sons Ltd., New York, pp. 61-84.
- Aviles, W. A., Hughes, T. W., Everett, H. R., Umeda, A. Y., Martin, S. W., Koyamatsu, A. H., Solorzano, M. R., Laird, R. T., McArthur, S. P. (1990). *Issues in Mobile Robotics: The Unmanned Ground Vehicle Program Teleoperated Vehicle (TOV)*, SPIE Vol. 1388, pp. 587-597
- Anderson, J. R. 1990. Cognitive Psychology and Its Implications, W. H. Freeman and Company, New York.
- Bartlett, F. C. (1932). Remembering, Cambridge University Press, Cambridge.
- Benyon, D., Murray, D. (1993). *Adaptive Systems: From Intelligent Tutoring to Autonomous Agents*, Knowledge Based Systems, 6(4), pp. 197-219.
- Biocca, F. (1992). *Communication Within Virtual Reality: Creating a Space for Research*, Journal of Communication, 42(4), Autumn.
- Cohen, R. L., (1981). *On the Generality of Some Memory Laws*, Scandinavian Journal of Psychology, 22, pp. 267-281.
- Cohen, R. L., Peterson, M., and Mantini-Atkinson, T. (1987). *Interevent Differences in Event Memory: Why are Some Events More Recallable than Others?*, Memory & Cognition, 15(2), pp. 109-118.
- Colley, A. M., Beach, J. R. (1989). *Acquiring and Performing Cognitive Skills*, in Acquisition and Performance of Cognitive Skills, Colley, A.M., Beach, J. R. (eds.), John Wiley & Sons, New York.
- Craik, F. I. M, and Lockhart, R. S. (1972). *Levels of Processing: A Framework for Memory Research*, Journal of Verbal Learning and Verbal Behavior, 11, pp. 671-684.
- Durlach, N. I., Pew, R. W., Aviles, W. A., DiZio, P. A., Zeltzer, D. L. 1992. Virtual Environment Technology for Training (VETT), BBN Systems and Technologies, BBN Report No. 7661, March.
- Durlach, N. I., and Mavor, A. S. eds. (1995). Virtual Reality -- Scientific and Technological Challenges, National Academy Press, Washington, D.C.
- Frederiksen, J. R., White, B. (1993). *The Avionics Job Family Tutor: An Approach to Developing Generic Cognitive Skills within a Job-Situation Context*, Proceedings AI-ED -- Artificial Intelligence in Education, Scotland, August, pp. 513-520.

- Freund, J. F. (1986). *New Technologies in US Navy Undersea Vehicles*, Naval Sea Systems Command, Washington, D.C.
- Gazzaniga, M. S. ed. (1988). *Perspectives in Memory Research*, The MIT Press, Cambridge, MA.
- Godden, D. R., and Baddeley, A. D. (1975). *Context-Dependent Memory in Two Natural Environments: On Land and Under Water*, *British Journal of Psychology*, 66, pp. 325-331.
- Hering, E. (1920). *Memory as a Universal Function of Organized Matter*, in *Unconscious Memory*, Butler, S. (ed.), Jonathan Cape, London.
- Hill, L. B. (1934). *A Quarter Century of Delayed Recall*, *Pedagogical Seminars, Journal of General Psychology*, 44, pp. 231-238.
- Hill, L. B. (1957). *A Second Quarter Century of Delayed Recall, or Relearning at Eighty*, *Journal of Educational Psychology*, 48, pp. 65-69.
- Jordan, M. I. and Rosenbaum, D. A. (1989). Action, in *Foundations of Cognitive Science*, Posner, M. I. (ed.), The MIT Press, Cambridge, MA, pp. 683-726.
- Koriat, A., Ben-Zur, H., and Nussbaum, A. (1990). *Encoding Information for Future Action: Memory for To-Be-Performed Tasks Versus Memory for To-Be-Recalled Tasks*, *Memory & Cognition*, 18(6), pp. 568-578.
- Lintern, G. (1996). *Human Performance Research for Virtual Training Environments*, *Proceedings of SimTecT '96 -- Advancing Simulation Technology and Training*, Melbourne, Australia, March 25-26.
- McAndrews, M. P., and Milner, B. (1991). *The Frontal Cortex and Memory for Temporal Order*, *Neuropsychologia*, 29(9), pp. 849-859.
- Miller, G. A. (1956). *The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information*, *The Psychological Review*, 63(2), pp. 81-97.
- Paivio, A. (1971). *Imagery and Verbal Processes*, Holt, Rinehart, & Winston, New York.
- Posner, M. I. (ed.) 1989. *Foundations of Cognitive Science*, The MIT Press, Cambridge, MA.
- Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., Carey, T. 1994. *Human-Computer Interaction*, Addison-Wesley Publishing Company, New York.
- Roberts, B. (1993). *Constrained Learning Environments for Intelligent Tutoring*, *Proceedings AI-ED -- Artificial Intelligence in Education*, Scotland, August, pp. 521-528.

Shaker, S. M., and Wise, A. R. (1988). *War without Men -- Robots on the Battlefield*, Pergamon-Brassey's, New York.

Schacter, D. L. (1989). *Memory*, in *Foundations of Cognitive Science*, Posner, M. I. (ed.), The MIT Press, Cambridge, MA, pp. 683-726.

Slamecka, N. J., and Graf, P. (1978). *The Generation Effect: Delineation of a Phenomenon*, *Journal of Experimental Psychology: Human Learning and Memory*, 4, pp. 592-604.

Smith, S. M., Brown, H. O., Toman, J. E. P. (1978). *Environmental Context and Human Memory*, *Memory & Cognition*, 6, pp. 342-353.

Tachi, S., Arai, H. and Maeda T. (1989). *Development of an Anthropomorphic Tele-Existence Slave Robot*, *Proceedings of the International Conference on Advanced Mechatronics*, Tokyo, pp. 385-390.

Tulving, E., and Donaldson, W. eds. (1972). *Organization of Memory*, Academic Press, New York.

Tulving, E. (1983). *Elements of Episodic Memory*, Oxford University Press, New York.

Welford, A. T. (1976). *Skilled Performance: Perceptual and Motor Skills*, 111, Scott Foresman, Glenview.

Appendix A -- Memory Experiment 1 Set Scores

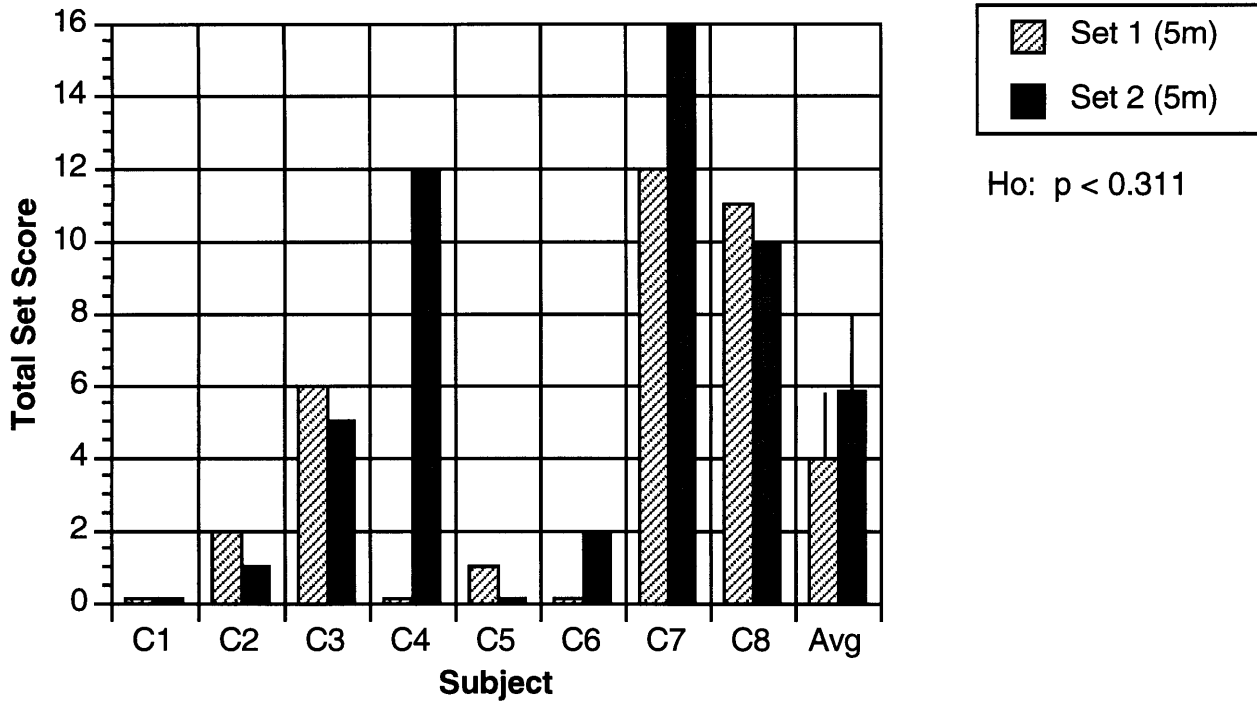


Figure A-1
Control Group Set Scores (5 Minute Examination)

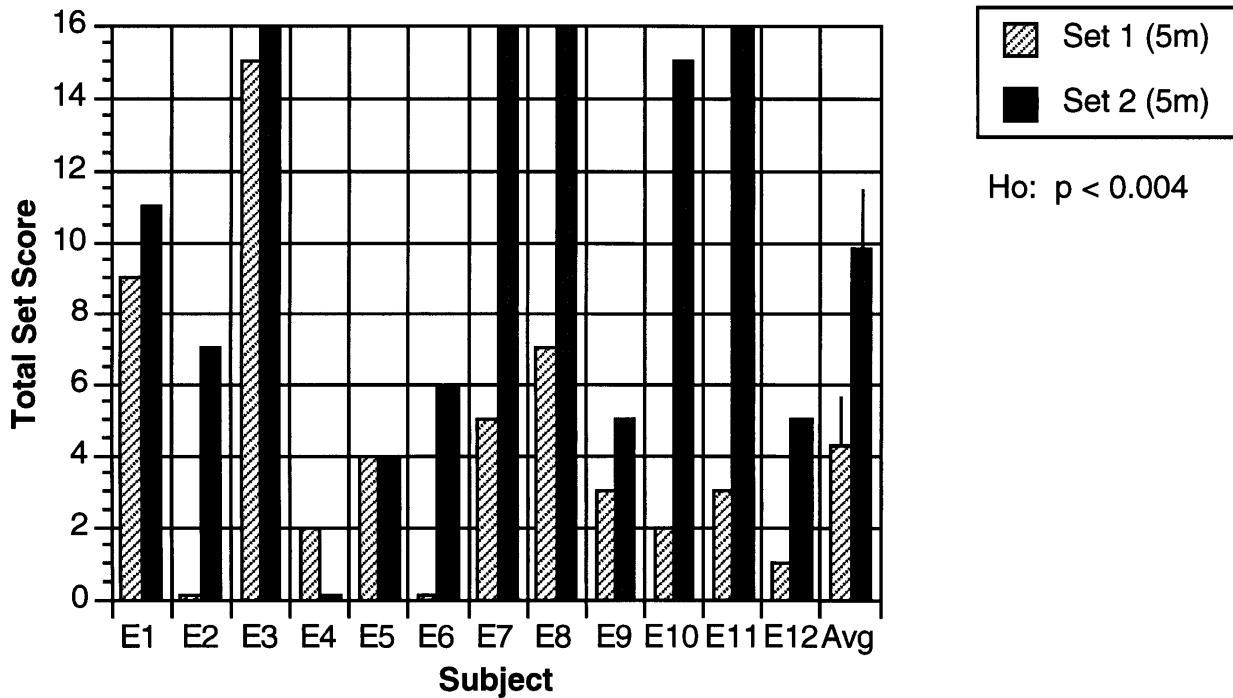


Figure A-2
Experimental Group Set Scores (5 Minute Examination)

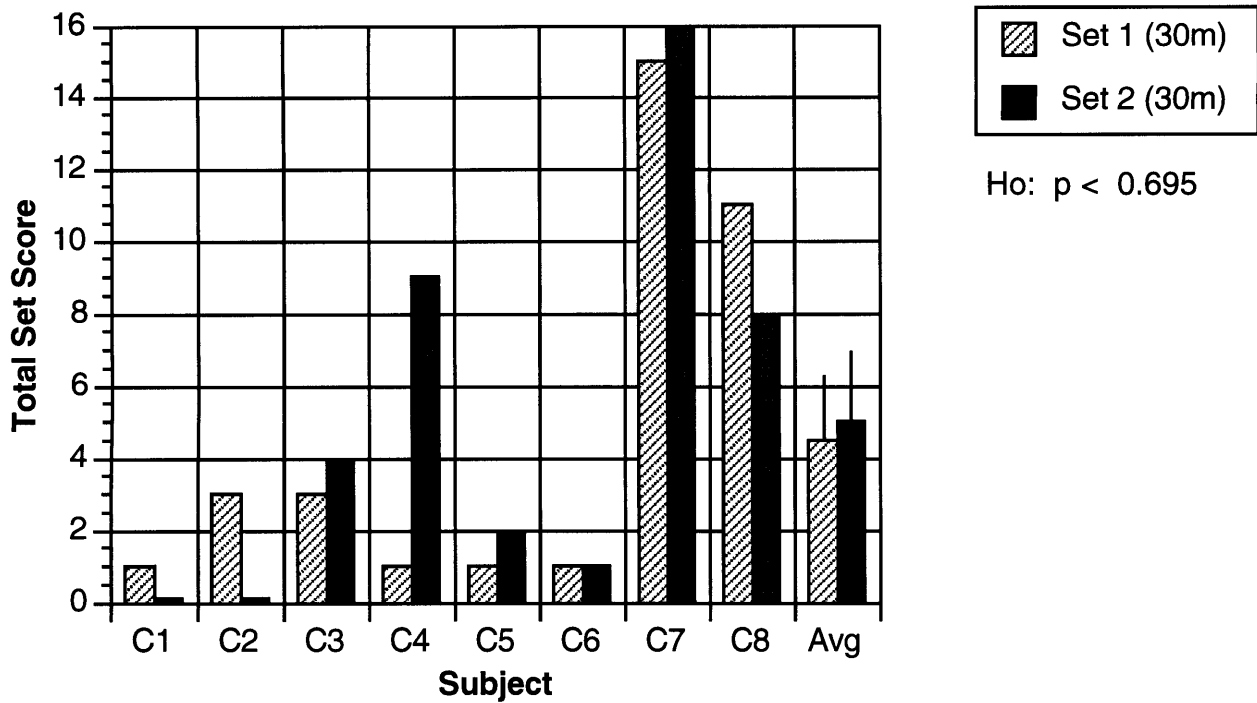


Figure A-3
Control Group Set Scores (30 Minute Examination)

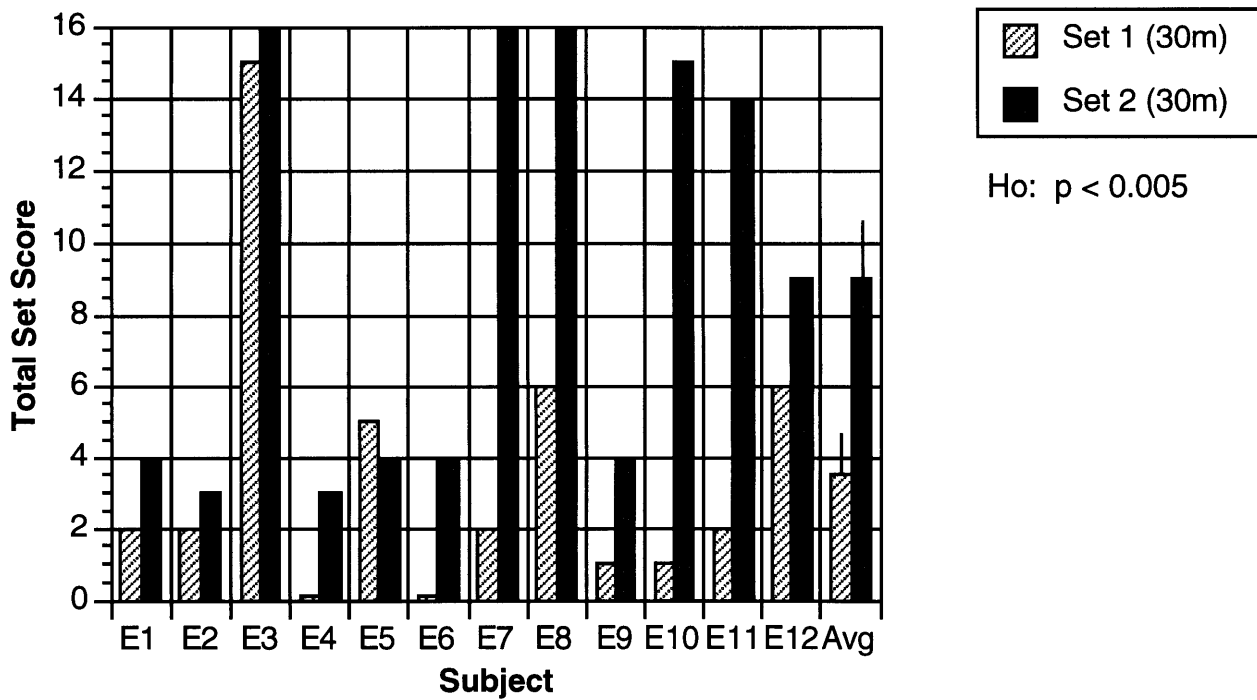


Figure A-4
Experimental Group Set Scores (30 Minute Examination)

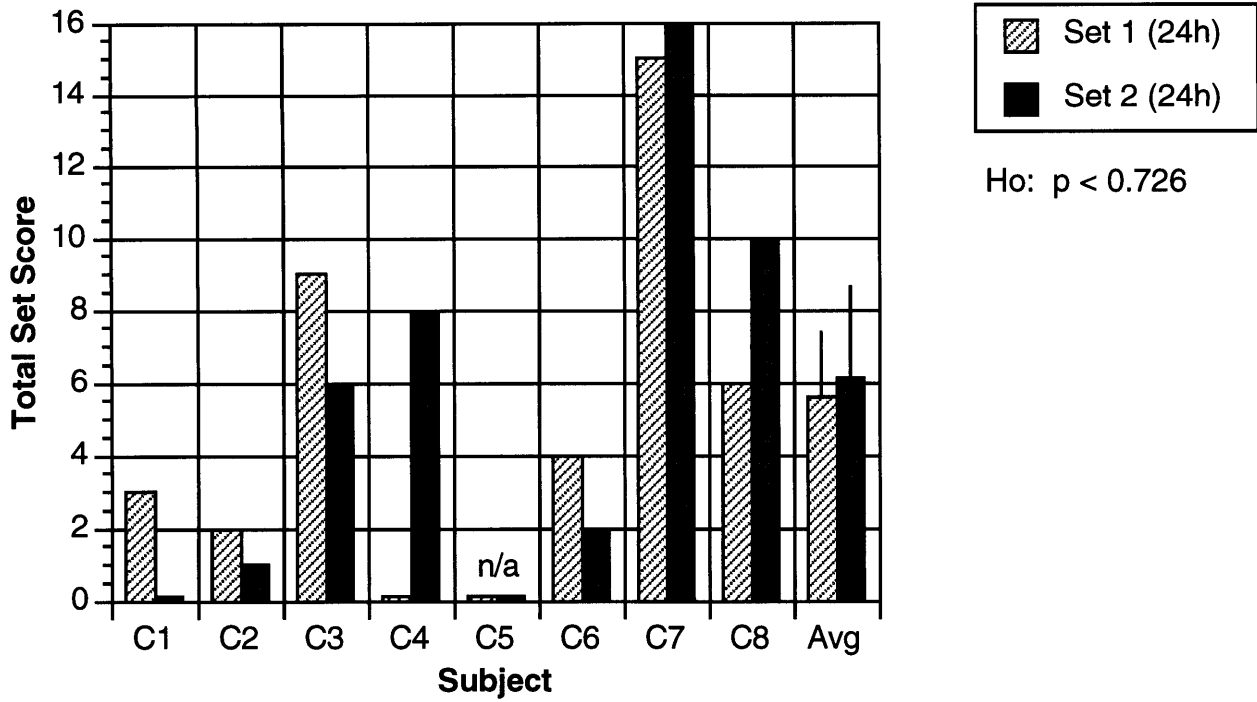


Figure A-5
Control Group Set Scores (24 Hour Examination)

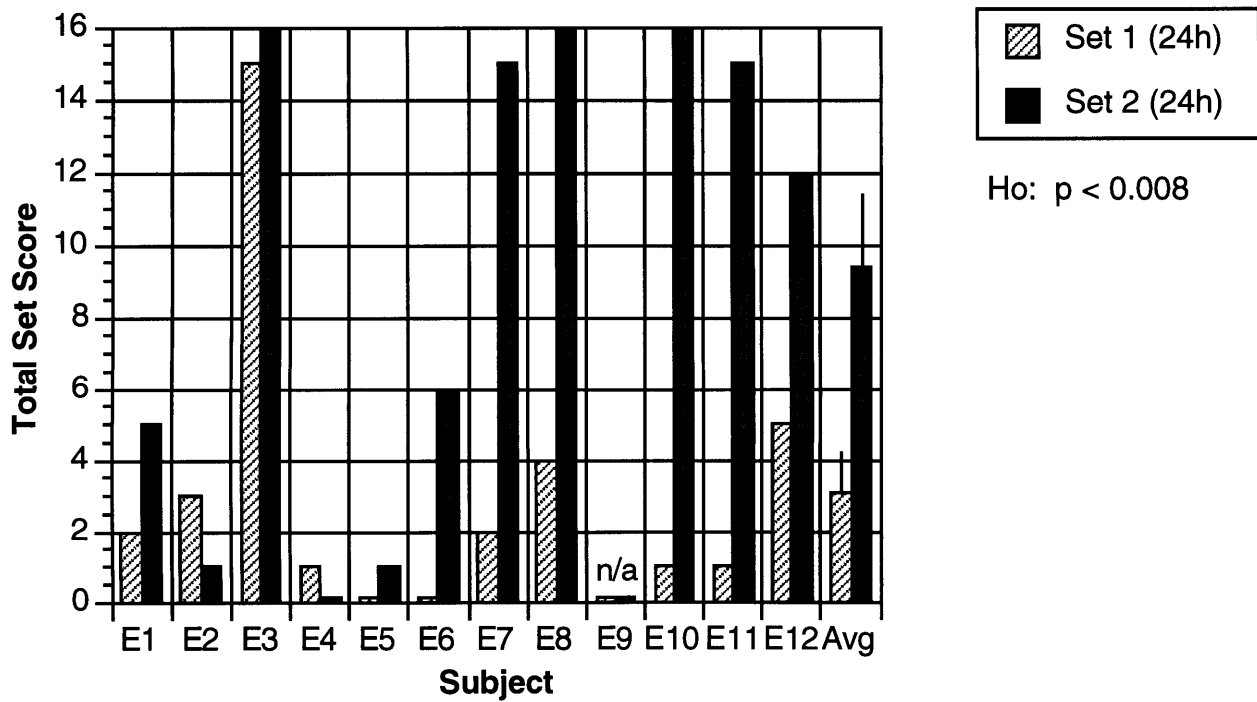


Figure A-6
Experimental Group Set Scores (24 Hour Examination)

Appendix B -- Memory Experiment 2 Detailed Scores

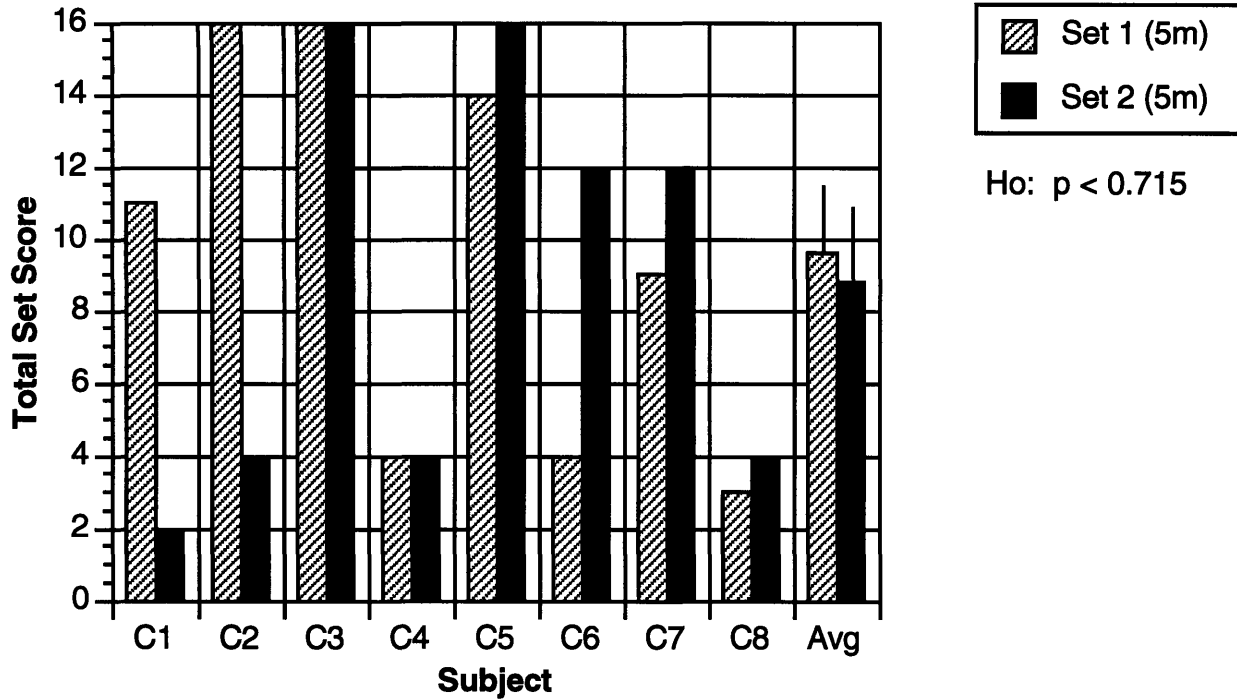


Figure B-1
Control Group Set Scores (5 Minute Examination)

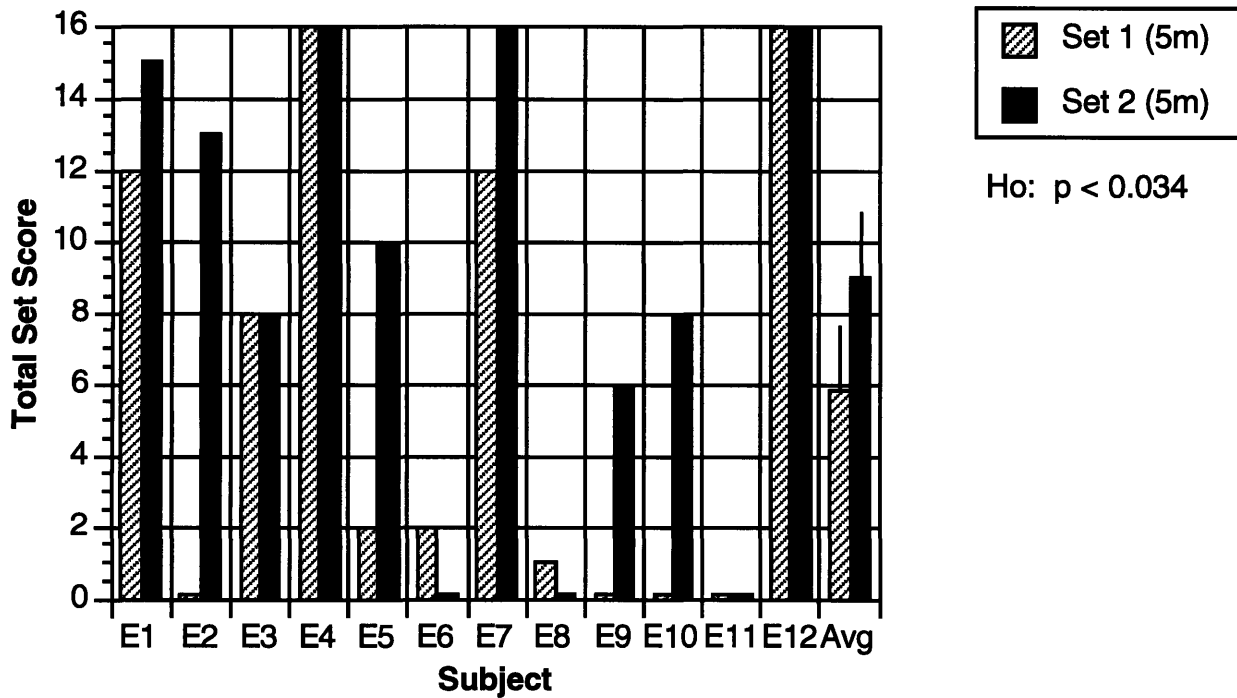


Figure B-2
Experimental Group Set Scores (5 Minute Examination)

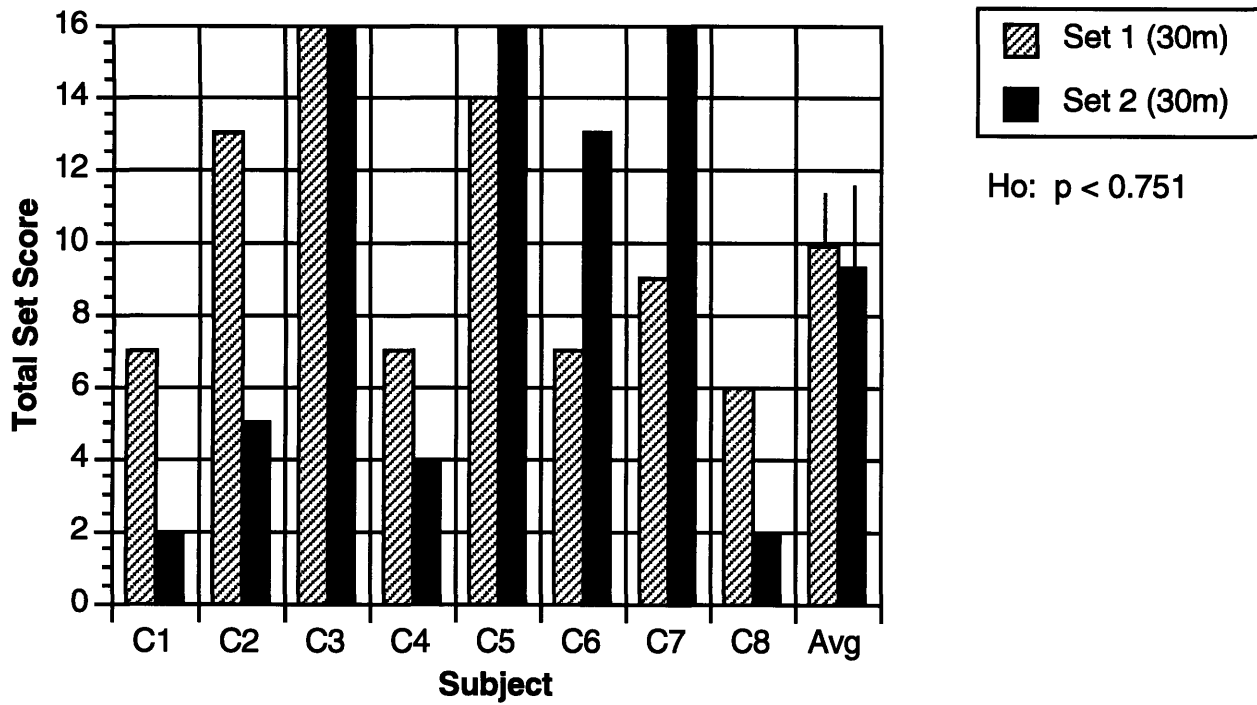


Figure B-3
Control Group Set Scores (30 Minute Examination)

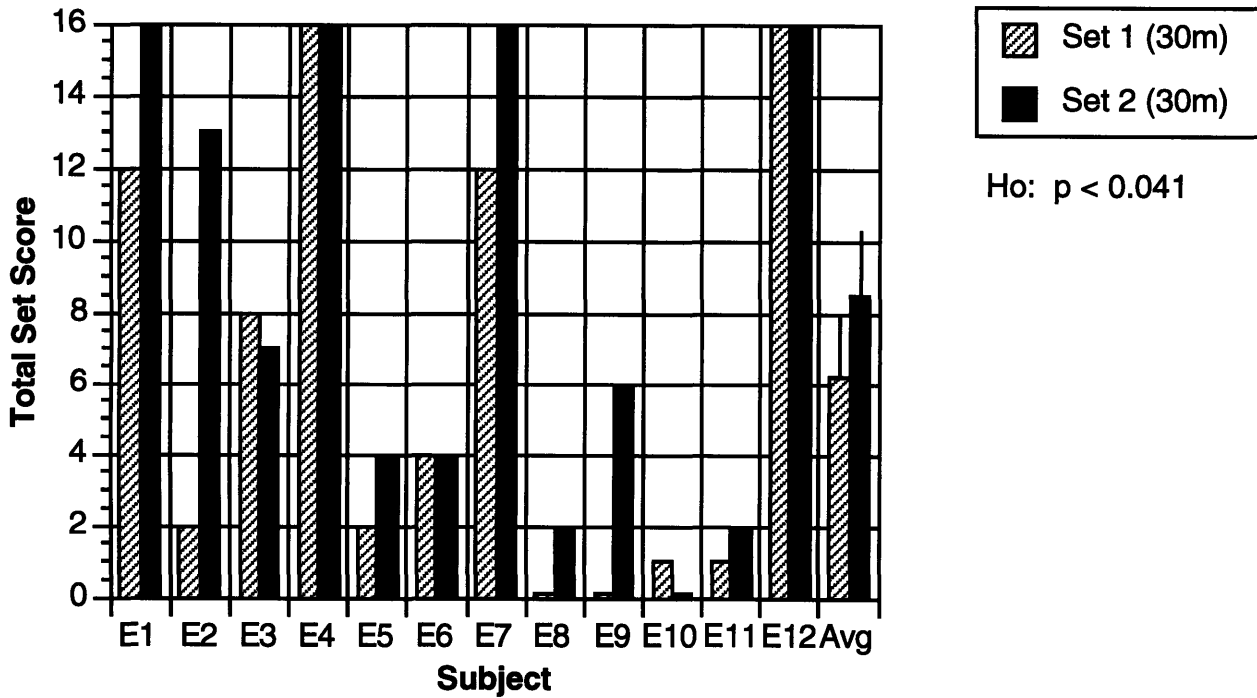


Figure B-4
Experimental Group Set Scores (30 Minute Examination)

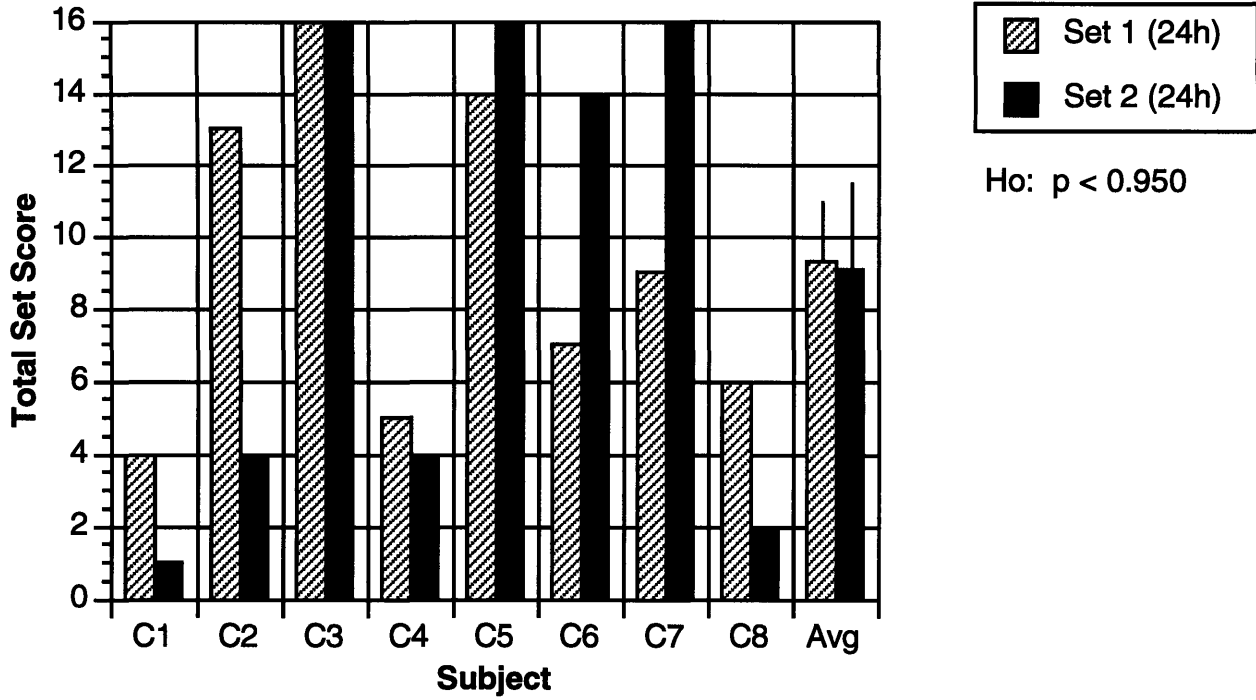


Figure B-5
Control Group Set Scores (24 Hour Examination)

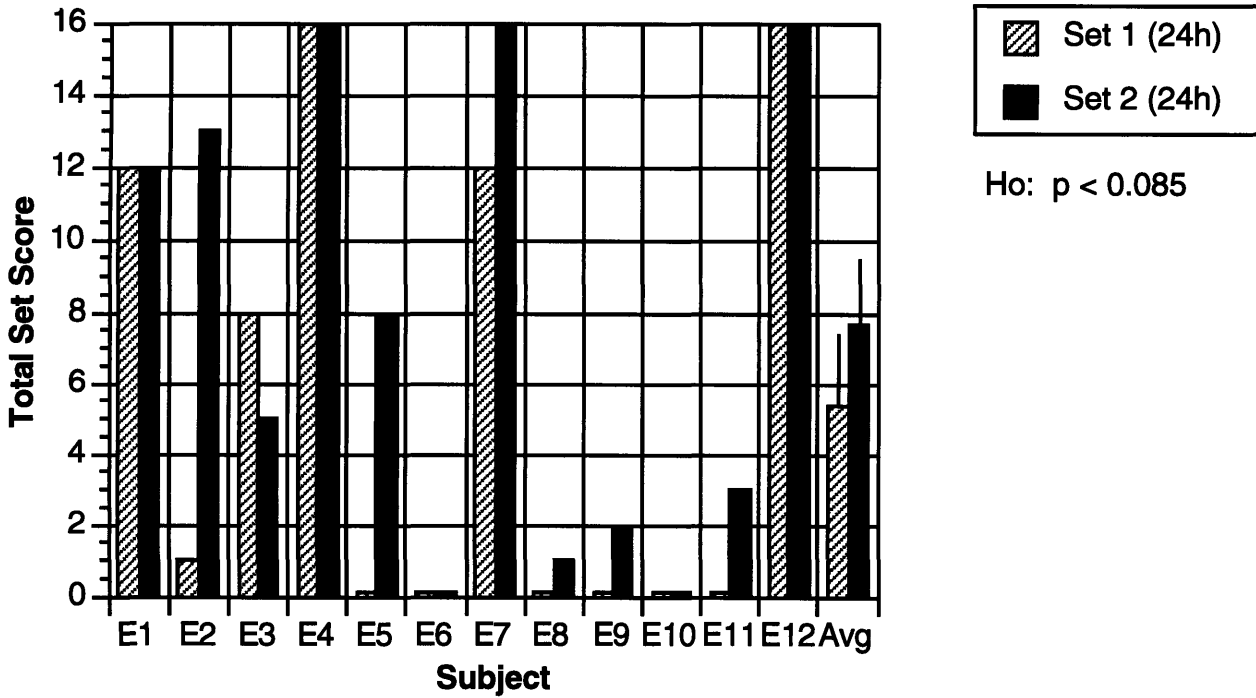


Figure B-6
Experimental Group Set Scores (24 Hour Examination)

Appendix C -- Experimental Procedure Issues

Two aspects of experimental procedure require some comment. First of all, as mentioned in Section 3.1.5, recall examinations were not forced-choice. Subjects were simply asked to do the best that they could and no specific instructions concerning test-taking strategy or scoring were given. The reason for this was to elicit the best possible performance from subjects under the least possible stress. In retrospect, this strategy was a mistake. Individual subjects simply adopted a personal strategy and thereby increased the variability of the data across subjects. Either subjects should have been instructed to use a forced-choice methodology (i.e., take their best guess if they weren't certain) or given some other explicit or implicit strategy. A manipulation and analysis of the raw data was performed, therefore, to attempt to verify that this methodological flaw did not cloud the experimental results. Specifically, any blanks left by subjects on their examination sheets were randomly assigned one of the action elements and all the examinations were re-scored. All the results from Memory Experiments 1 and 2 remained unchanged and all conclusions remained significant to the $p < 0.05$ criterion. The results of this analysis suggest that, although less than perfect, the examination strategy does not seem to have affected the results.

The second issue has to do with differences between the pre-trial procedure for Experiment 1 and Experiment 2. In both experiments, subjects were informed of the general structure of sequences (i.e., sequence label and other elements), the form of sequence elements (i.e., 'S' followed by another letter), and the form of elements in the action part (i.e., letter followed by a digit). Furthermore, subjects were not informed as to the length of sequences. In Memory Experiment 2, however, subjects WERE informed of the range of sequence labels (i.e., SA-SZ) and the range of elements in the action part (i.e., A1-3, B1-3, C1-3). Although this methodological difference does not invalidate the results of either experiment, it does make comparisons between the control groups of both experiments more tenuous.