

**DTIC FILE COPY**



ARI Research Note 87-53

**AD-A187 002**

**OPTIMIZING THE LONG TERM RETENTION OF SKILLS:  
Structural and Analytic Approaches to Skill Maintenance**

**Alice F. Healy, K. Anders Ericsson,  
and Lyle E. Bourne, Jr.  
University of Colorado**

for

**Contracting Officer's Representative  
Judith Orasanu**

**BASIC RESEARCH LABORATORY  
Michael Kaplan, Director**

**DTIC  
ELECTE  
DEC 29 1987  
S D  
E**



**U. S. Army  
Research Institute for the Behavioral and Social Sciences**

October 1987

Approved for public release; distribution unlimited.

**87 12 14 050**

# U. S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency under the Jurisdiction of the  
Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON  
Technical Director

WM. DARRYL HENDERSON  
COL, IN  
Commanding

Research accomplished under contract  
for the Department of the Army

University of Colorado

Technical review by

Michael Kaplan



Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input checked="" type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

This report, as submitted by the contractor, has been cleared for release to Defense Technical Information Center (DTIC) to comply with regulatory requirements. It has been given no primary distribution other than to DTIC and will be available only through DTIC or other reference services such as the National Technical Information Service (NTIS). The views, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other official documentation.

UNCLASSIFIED

A187 002

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ARI Research Note 87-53	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Optimizing the Long-Range Retention of Skills: Structural and Analytic Approaches to Skill Maintenance	5. TYPE OF REPORT & PERIOD COVERED Interim Report August 86 - August 87	
	6. PERFORMING ORG. REPORT NUMBER	
7. AUTHOR(s) Alice F. Healy, K. Anders Ericsson, and Lyle E. Bourne, Jr.	8. CONTRACT OR GRANT NUMBER(s) MDA903-86-K-0155	
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Colorado Campus Box B-19 Boulder, CO 80309	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 2Q161102B74F	
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Research Institute for the Behavioral and Social Sciences, 5001 Eisenhower Avenue, Alexandria, VA 22333-5600	12. REPORT DATE October 1987	
	13. NUMBER OF PAGES 46	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) - -	15. SECURITY CLASS. (of this report) Unclassified	
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE n/a	
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) - -		
18. SUPPLEMENTARY NOTES Judith Orasanu, contracting officer's representative for back		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Cognition, Memory, Skills, Retention.		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) ➤ The program described in this research note seeks to identify characteristics of knowledge and skills which are most resistant to decay due to disuse. The program is divided into analytic and structural approaches. The first line of research for investigating skill retention using the analytic approach investigates laboratory analogues to the component skills of electronic technicians. The second approach investigates parallel natural skills learned by former college students. We have developed five laboratory methodologies, and have		

(OVER)

UNCLASSIFIED

ARI RESEARCH NOTE 87-5320. Abstract (continued)

completed the preliminary testing for each of them. We have also identified four natural skills, designed the initial questionnaires and tests, and collected preliminary data for all of them. For the structural approach, we have designed an experimental paradigm which allows us to assess the detailed encoding of new knowledge at presentation and after a delay, using verbal report techniques and chronometric measurement of retrieval components. A preliminary study of retention of vocabulary items has been completed within this paradigm. *Keywords:*

*Keywords:*

Executive Summary of Progress: August 11, 1986 to August 10, 1987  
Optimizing the Long-term Retention of Skills:  
Structural and Analytic Approaches to Skill Maintenance  
Alice F. Healy, K. Anders Ericsson, and Lyle E. Bourne, Jr.

This research program seeks to identify the characteristics of knowledge and skill which are most resistant to decay due to disuse. Our research can be divided into two complementary parts. The first part is concerned with experimental analysis of factors influencing and improving retention of skill components. The second part is concerned with analysis and assessment of the structure of acquired memory and skills and how to monitor differential retention of components. The eventual goal of both parts is to be able to make relevant recommendations about training routines for long-term skill maintenance.

### The Analytic Approach

We have developed two lines of research for investigating skill retention and maintenance using the analytic approach. The first line of research involves investigating different laboratory analogues of component skills of electronic technicians. The second complementary line of research involves investigating parallel natural skills learned by the college population during their prior education.

Laboratory skills. We have developed five laboratory methodologies, and we have completed the preliminary testing for each of them. The laboratory tasks involve (a) target detection, (b) data entry, (c) learning logical rules involved in circuit design, (d) memory for numerical calculations, and (e) temporal, spatial, and item components of memory for lists. On the basis of our preliminary studies of target detection and memory for numerical calculations, we have reached the following two conclusions. First, the perceptual skills involved in target detection, unlike verbal memory, are highly resistant to forgetting and hence easily enter permastore. Second, the crucial factor affecting memory for numerical calculations is the internal (as opposed to external) locus of cognitive operations that link the to-be-remembered material to other information in memory. In contrast, the internal (versus external) locus of stimulus production has essentially no effect on retention.

Natural skills. We have identified the following four natural skills and have developed assessment techniques for each of them: (a) mental multiplication, (b) algebra, (c) data entry, and (d) temporal, spatial, and item components of memory for class schedules. We have designed the initial questionnaires and tests and have collected preliminary data for all four of these skills.

### The Structural Approach

We have designed an experimental paradigm which allows us to assess the detailed encoding of new knowledge at presentation and a delay using verbal report techniques and chronometric measurement of retrieval components. A preliminary study of retention of vocabulary items has been completed, in which some subjects have been instructed to use the keyword method with supplied keywords and other subjects have generated their own mediating encodings. We are setting up a large-scale study of the acquisition and retention of memory skill. A review of previous studies of long-term retention for procedures, such as cardiopulmonary resuscitation (CPR), has been completed as a part of a more comprehensive review.

Annual Interim Report for the period August 11, 1986 to August 10, 1987

Optimizing the Long-term Retention of Skills:

Structural and Analytic Approaches to Skill Maintenance

We are looking for techniques to optimize the long-term retention of skilled and knowledgeable performance. This is not a new problem, but neither is it a problem that has been solved to everyone's satisfaction. In fact, we think it is a problem that often gets overlooked in the search for optimal training procedures. The goal of training is, of course, to develop a high level of performance on some category of jobs or tasks, ensuring that the requisite skills are available for application as needed subsequently in the field. Thus, ideally, any training program should be designed to induce a level of performance which is sufficient to support retention and transfer of skill after extensive periods of disuse. But, even though retention and transfer are stated goals in the design of training programs, the more immediate goal of bringing trainees to a criterion level of skill within the shortest possible training interval often takes precedence over other considerations.

It is typical that training in the military is followed by delayed and infrequent opportunities to use the skills that have been learned. But the skills in question are of the sort that when they are needed, there is typically no time for retraining. There are good reasons for adopting an acquisition efficiency criterion, given the costs associated with training and the fact that trainees are generally nonproductive during the training process. But efficient training should not come at the cost of relative neglect of the goal of permanence and maintenance over time of what has been learned. Ideally, training should provide the technician with a tool kit of stable, reliable behaviors into which he or she can dip as needed. In the same sense that hardware can be expected to be available, so should the requisite skills.

Thus, our purpose is to try to identify those conditions of training that are associated with skill permanence and availability. We use as a general

starting point the observation that a significant portion of almost any natural skill or acquired knowledge is relatively permanent. In a theoretically noncommittal way, we adopt Harry Bahrick's (1984) notion of a permastore. It is well known that a large part of acquired knowledge or skill is lost rapidly, lets say within the first several months of training. But, as Bahrick has demonstrated, a significant portion of the knowledge that a person acquires can have a life-span of 25 years or more, even if that knowledge is not rehearsed or accessed during that long interval.

This finding presents the challenging possibility of identifying conditions of learning and characteristics of the material or skills learned that contribute to the durability of acquired behaviors. The idea is to identify conditions and characteristics that distinguish between short-lived and relatively permanent knowledge and skill components. If we can identify those skill components that are most resistant to decay, we might then be able to trace back and find out what aspects of training differentiate those components and other less permanent components.

Our approach to this problem is two-fold. The first approach we call analytic. We take as a test bed for this approach the kind of training provided for and the skills developed by a military electronics technician. We analyze those skills into their components. Among those components are some that seem categorizable as front end or as perceptual, such as the detection of error signals on an oscilloscope or some other test apparatus. Some of what is learned is primarily motor, having to do, for example, with the skillful application of test equipment to possibly malfunctioning components of equipment. And part of what is learned is primarily cognitive involving, among other things, problem-solving strategies or decisions among potential tests of malfunction. We intend not only to develop laboratory analogs of these component skills but also to look at their natural or quasinatural counterparts. Studies to be undertaken involve both acquisition and retention phases (and,

down the line will also include a transfer phase). We are not disinterested in the properties of training, but our primary concern is with what gets retained, followed by post-hoc analysis of those conditions of training that distinguish between what is retained and what is not. This approach began with an extensive review of the literature followed by the development of tasks and methodologies that seemed particularly suited to answer our general questions.

The second approach we call a structural approach. It involves an analysis of the mental structures involved in natural (which we use as a contrast with laboratory) skills and the further development of methods to characterize those structures. Again, this approach began with an extensive review of the literature on the long-term retention of natural skills.

One additional comment or disclaimer should be made at this point. We are not interested in this project in building a new classification system for skills pertinent to any particular domain of technical competency. Rather, our approach has a definite process orientation. We are interested primarily in developing a conceptual framework both for understanding skill durability and for doing further basic research on skill and knowledge retention.

#### The Analytic Approach

##### Laboratory Skills

Our initial goal was to devise four laboratory analogues of component skills of electronics technicians. We have in fact developed five laboratory methodologies for this purpose, and we have either completed or initiated the preliminary testing for each of these methodologies. The laboratory tasks involve (a) target detection, (b) data entry, (c) learning logical rules involved in computing circuit design, (d) memory for numerical calculations, and (e) temporal, spatial, and item components of memory for lists.

Target detection. We completed three preliminary experiments using a new methodology that combines features of the Schneider and Shiffrin (1977; Shiffrin & Schneider, 1977) consistent-mapping automatism paradigm with a letter

detection task in which subjects press a response key every time that a target appears in a sequence of display frames shown on a computer terminal (see, e.g., Healy, Oliver, & McNamara, 1987). The first experiment involved training 32 subjects at target detection over a period of four days. The subjects were divided into two groups, those searching for the target h in a background of scrambled letters and those searching for the target 5 in a background of scrambled numbers. Three different frame sizes were employed including 2, 4, and 16 characters. In a pretest, all subjects searched for the target h in a prose passage presented on a computer terminal three words at a time. In a set of posttests, all subjects searched for the target h in a second prose passage, and then they searched for the target t in a third prose passage. We found that, although performance significantly improved over the four days of training, the subjects did not reach the criterion of automatism defined by Schneider and Shiffrin in terms of a decrease in the frame size functions; see Figure 1 for hit rates (top panel) and for reaction times (bottom panel). Nevertheless, the performance pattern for detecting letters in the prose passages was dramatically influenced by training. Both groups of subjects continued to show decreased hit rates on correctly spelled relative to misspelled test words, but the difference between the common word the and other words, which was substantial in the pretest, was greatly diminished in the posttest for both targets (see Figure 2 which provides the proportion of hits averaged across training group because training group did not prove to be a significant factor).

In a second preliminary study, we varied the amount of training from zero to four days and included a one-month retention test of the target detection skill. The target was the letter H for all 36 subjects. Subjects who received the greatest amount of training did show a decrease in their frame size functions for both hit rates and reaction times, indicating that they had reached some degree of automatism. However, all subjects, despite the amount of

training, showed no loss in performance during the retention interval. Figure 3 presents hit rates and Figure 4 reaction times for limited training (top panel) and extensive training (bottom panel). Figure 5 presents the hit rates (top panel) and the reaction times (bottom panel) for the control group as well as the limited and extensive training groups at the one-month retention test. Further, on the letter detection tasks involving prose passages, all subjects, including those given no training, showed a change when comparing pre- and posttest performance similar to that found in Experiment 1 (see Figure 6).

We have also completed a study involving the intensive target detection training of a single subject (A.G.). This subject was given 12 sessions of acquisition training with the consistent mapping procedure. After the last session there was a six-month retention interval followed by an initial retention test, which in turn was succeeded by an additional nine-month retention interval followed by a second retention test. Throughout acquisition the subject showed dramatic improvements in both hit rates and reaction times. Further, the frame size functions indicated that performance became significantly more automatic as training progressed; see Figure 7 for hit rates (top panel) and reaction times (bottom panel). Most crucially, the initial retention test indicated that some forgetting had occurred during the six-month retention interval. However, the single refresher training provided by the initial retention test was sufficient to yield performance on the second retention test nine months later as good as performance at the end of acquisition training. Hence, on this perceptual learning task there was essentially perfect maintenance of highly skilled performance throughout a 15-month period with only one refresher training session intervening between acquisition and retention testing. These results, along with those from the second preliminary experiment, imply that like motor skills, but unlike verbal skills, perceptual skills are not subject to rapid forgetting and hence easily enter permastore.

To follow up these initial experiments, we have begun a long-term retention study in which we plan to train a total of 24 subjects, each for four sessions. Half of the subjects will be exposed to the consistent mapping training, and the remaining half will be exposed to a new varied mapping training in which the target changes across blocks. All subjects will be recontacted after approximately six months. On the basis of the hypothesis that automatism is necessary for entry into permastore, it is expected that retention will be superior for those subjects given training with consistent mapping. Alternatively, on the basis of Battig's (1972, 1979) intratask interference principle, it is expected that retention will be superior for those subjects given the more difficult training with varied mapping. More generally, the results should enable us to determine whether training in a detection skill should be consistent or varied in order to optimize its long-term maintenance.

Data entry. We developed a new methodology for measuring the acquisition and retention of the motor skill of entering number sequences on a keypad of a computer terminal. In our first major study of the data entry skill, we trained 36 subjects, each of whom was exposed to three training sessions and a subsequent or month retention test in which they learned to type three-digit sequences. The subjects were divided into three groups depending on the extent and pattern of repeated digit sequences during training. For the control group no sequences were repeated. For the massed group each of five blocks of ten digit sequences was repeated five times in a row. For the spaced group each block was presented five times but in this case with the other four blocks intervening between repetitions. During the final retention test subjects were given blocks of digit sequences shown during acquisition as well as blocks of new sequences not given earlier. According to Battig's (1972, 1979) intratask interference principle, subjects in the massed group should show superior performance during the initial acquisition phase but should show inferior performance at the retention test. In fact, we found little difference between

the three groups of subjects in terms of overall reaction times. Most interesting, however, was that for all three groups, including the control group in which each block was shown only once, reaction times at the retention test one month later were significantly faster for the old blocks relative to the new blocks that had not been seen during acquisition (see Figure 8). We plan to do further work to determine whether this facilitation is due solely to procedural memory as distinct from episodic memory.

Learning logical rules involved in computing circuit design. We have developed a new methodology to assess the acquisition and retention of logical principles. Subjects are shown pictorial displays with two binary inputs and one binary output, and their task is to learn the rule relating inputs and outputs. In our initial set of preliminary studies, we tested 57 subjects (9 of these subjects did not complete the task and will not be considered further here) and assessed their abilities to determine the output when shown two inputs for four logical rules corresponding to AND, OR, NOT AND (NAND), and NOT OR (NOR) gates in computing circuitry. Each logical rule was represented by a different geometric design (see Figure 9). Two groups of subjects were tested with pretraining followed by variable or consistent presentation of stimuli, and two other groups were tested with only variable or consistent presentation of stimuli and no pretraining. During pretraining, subjects were presented with one rule at a time until that rule was learned. During variable testing, the order of stimulus presentation was random. Subjects continued responding until all stimuli were learned. During consistent testing, the rules were presented in a particular order which was consistent across trials for each subject. Again, subjects continued responding until all stimuli were learned.

Our initial findings (see Figure 10 for the average number of errors made to each rule in each of the four groups) indicate that subjects in these tasks do not learn the stimuli as logical rules (in contrast to previous rule learning behavior by subjects in concept formation studies with simple geometric figures;

see, e.g., Bourne, 1974). Bourne found that subjects exhibited a particular ordering of difficulty in learning logical rules. The AND rule was the easiest to learn, followed by OR, followed by NOT AND and NOT OR.

Subjects in the current task did not exhibit the ordering of rules found by Bourne (1974), as can be seen in Figure 10. The number of errors made per rule interacted with what group the subject was in (i.e., how the subject was tested). Also, whether the top input was a plus or a minus sign influenced how many errors were made to a rule. If subjects were learning the stimuli as rules, it seems unlikely that the method of testing or the sign of the top input would influence the ordering of rule difficulty.

Other evidence substantiates that subjects did not learn the stimuli as logical rules. When subjects completed the task, they were asked to give a retrospective report of what they had done in order to "figure out" the task. These reports indicate that instead of treating the stimuli as logical rules, subjects adopted complex memory strategies which consisted of grouping the stimuli by input configuration and figure. That is, the representations were probably not rule based, but more stimulus-configuration specific.

Further, these representations resulted in nearly errorless performance by subjects who were retested after a two-day interval. The mean number of errors for retested subjects from the pretraining/variable group was 4.5 ( $n=6$ ;  $s=5.5$ ). Only one subject in this group, however, had more than five errors (this subject had 15 errors); excluding this subject, the mean number of errors was 2.4 ( $s=2.1$ ). The mean number of errors for retested subjects from the pretraining/consistent group was 0.88 ( $n=8$ ,  $s=0.83$ ). No subject in this group made more than two errors.

Memory for numerical calculations. We investigated the "generation effect" (see, e.g., Slamecka & Graf, 1978) in an experiment which compared recall of the answers to multiplication problems studied under four different conditions, two of which (read and verify) involved experimenter generation, or production, of

the answers and two of which (generate and calculate) involved subject production of the answers (see Table 1 for an illustration of each of the four tasks). Contrary to the findings from most previous studies of the generation effect, experimenter vs. subject stimulus production was not a crucial factor determining performance levels, but the locus of the multiplication operations (which was internal for the verify and generate conditions but external for the read and calculate conditions) seems to be the critical factor in influencing retention levels (see Table 2).

In a subsequent study we extended our work along two dimensions. First, we used a forced-choice recognition test as well as a free recall test. Second, we assessed retention after two-day and seven-day intervals as well as immediately after study. For both testing methods and all three retention intervals we found, as in our initial study, that the crucial factor affecting memory is the internal (as opposed to external) locus of cognitive operations that link the to-be-remembered material to other information in memory. In contrast, the internal (versus external) locus of stimulus production has essentially no effect on retention. Specifically, although performance significantly decreased in all conditions as a function of retention interval, we found superior memory performance for the generate and verify tasks and little difference between the generate and verify or between the calculate and read tasks (see Table 3 for the recall test and Table 4 for the recognition test).

Temporal, spatial, and item components of memory for lists. In earlier work (see, e.g., Healy, 1982) we devised a methodology to compare the short-term retention of temporal, spatial, and item information for lists of consonants. We have now extended this paradigm to the long-term retention of lists of words. The first question we will examine is whether the three types of information show similar retention functions over short and long intervals. For example, in the studies of short-term retention, temporal information shows a steeper time course of forgetting than spatial information. We will attempt to determine

whether this difference in decay holds for intervals considerably longer in duration.

In an initial laboratory study, 24 subjects learned a list of 20 common nouns, each beginning with a different consonant from the alphabet. The words were presented one at a time in a vertical array on a computer terminal, with each word occurring in a different location within the vertical array. Each word was displayed for two seconds. At the termination of the list presentation, subjects were instructed to recall the words by writing them on a sheet of paper. A trial consisted of one presentation and one recall attempt.

Eight subjects were required to recall the words according to the temporal sequence of presentation, eight subjects were required to recall them according to their spatial location within the vertical array during presentation, and the remaining eight subjects were required to choose the twenty words that had been presented from an alphabetically organized list of the critical words intermixed with distractor words. After recall another trial was started with the words being presented in the exact same order and location. This process continued until a criterion of correct recall on three successive trials was met.

Subjects returned after a one-week delay and were asked to recall the twenty words as they had during the first session. The subjects were not told during the first session that they would be required to recall the words in the second session. After an initial recall attempt, the presentation/recall trials were resumed as in the first session, and this process continued until a criterion of correct recall on three successive trials was met.

Our analyses revealed fastest acquisition of item information, followed by temporal, and then spatial information. The results also indicated that for the week-long delay, recall of item information was better than that for temporal information, which in turn was better than that for spatial information (see Table 5).

#### Natural Skills

A second goal of this project, a goal that complements our work with the five laboratory skills, was to identify four natural skills or skill components that we can expect the college population to have learned at some specifiable time during their prior education. We further sought to develop techniques to assess the characteristics of what remains in permastore. We have identified and begun to investigate the following four natural skills: (a) mental multiplication, (b) algebra, (c) data entry, and (d) temporal, spatial, and item components of memory for class schedules.

Mental multiplication. For the skills involved in mental multiplication, we have developed a two-part methodology. The first part will involve a questionnaire, similar to that used by Bahrck (1979, 1984) to assess when, how, and to what extent a student was originally trained in the skill of mental multiplication. The questionnaire will also assess the type and amount of maintenance that the skill received after the original training and the variability of the context in which the skill was maintained. The second part will involve a laboratory assessment of the students' speed and accuracy at performing simple multiplication problems. Following Bahrck, we will attempt to predict performance on the laboratory task from various indices of skill acquisition and maintenance as derived from the questionnaire. Further, following the approach we have adopted for the laboratory tasks, we will attempt to determine whether automatism is related to skill performance and long-term retention. More specifically, we will systematically vary multiplication problem difficulty and use as an index of automatism the size of the functions denoting the relationship between accuracy or speed of solution and problem difficulty.

In preliminary work we tested two groups of nine subjects who were shown three repetitions of the complete set of single digit multiplication problems, one problem at a time on a computer terminal. The subjects in one group responded by typing the answer on the keypad, as in the data entry task. The

subjects in the second group responded orally, and their reaction time was measured by means of a voice key attached to the computer. These initial data allow us to assess differences in problem difficulty and to examine improvements in speed and accuracy as a function of practice. See Figures 11 (keypad group) and 12 (voice key group) for the overall mean log reaction times and error rates (top panel) and the mean log reaction times as a function of practice block (bottom panel).

Algebra. For the algebra skills, we also developed a questionnaire to assess students' previous history of algebra use and knowledge. Further, we developed a 60-item algebra proficiency test and gave this test to 30 introductory psychology students. On the basis of the results from this pilot test and advice from an algebra professor (Dr. Robert Ellingwood), we constructed two versions of a modified test and administered them to a class of 86 students who were completing a course in first-year college algebra. All of these students have agreed to be recontacted for testing at a future date in order to determine the extent of their long-term retention of these skills. We will retest these subjects at intervals of three months and one year after the completion of their first-year algebra course. We have completed an initial factor analysis of the 60-item test based on the scores of the 86 students. We will use this analysis to help guide our interpretation of the data we collect later concerning long-term retention.

Data entry. For the data entry skills, we will be employing the laboratory methodology we developed for the use of the keypad on the computer console. We have already identified one individual who has made extensive use of the keypad for data entry in her semester-long job of entering students' social security numbers into the computer. Immediately after the conclusion of her job and again six months later, we tested her proficiency at our related task, and we intend to test her at various further delays. Our initial findings indicate no loss in speed and a significant decrease in error rate over the six-month

retention interval. Thus, these preliminary findings for the data entry skill learned under natural conditions agree with our findings for the same skill learned in the laboratory; in both cases, retention was essentially perfect.

We also hope to find a group of students who have similar job experiences with keypad data entry, perhaps as cashiers. We intend to investigate the effects of changing context and changing task as well as retention interval on the speed and accuracy of data entry.

Temporal, spatial, and item components of memory for class schedules. For the temporal, spatial, and item components of memory for class schedules, we have begun to develop a methodology to assess differences in these three components of memory in a more natural environment than we are using in the laboratory to study the same memory components. More specifically, we propose to question students about their current and previous classroom schedules to determine how much they remember about the course number, room number, and instructor's name (item information), hour, duration, and day of course (temporal information), and building, floor, and room location (spatial information).

We devised a questionnaire and began preliminary testing with 7 subjects. These preliminary data have allowed us to refine our questionnaire for use in a larger study. Our goal for the larger study is to test three groups of students. In the initial phase of the experiment, each group will be tested about either their immediately preceding semester, their second-to-last semester, or their third-to-last semester. In the second phase, students will be called back for retesting. This procedure will allow us to assess retention using both cross sectional and longitudinal methods.

#### The Structural Approach

The structural approach uses methods such as verbal reports (Ericsson & Simon, 1984) along with other types of assessment to provide a detailed description of knowledge and skill at the time of learning. Against such a

detailed description, specific losses or decrements can then be specified by a similar detailed analysis at retest. The relevance of this approach to decay of skill and knowledge is clear: If certain components decay differently than other components a more detailed structural description should allow one to chart the different components and how they decay. The structural approach was developed to analyze one or a few individuals at a time, but we proposed extending it so that larger numbers of subjects could be studied.

Our attempt to extend this methodology during the first year of our project has focused on vocabulary learning and more specifically on a mnemonic method of vocabulary learning known as the keyword method. Before discussing the keyword method itself, we should mention why vocabulary learning was selected. First, Bahrick's work on long-term retention and the concept of permastore has dealt with the retention of vocabulary (Bahrick, 1984; Bahrick & Phelps, 1987). Second, in a number of real-world contexts (e.g., the military, the medical profession, and in almost any type of training situation), individuals must learn new vocabulary or terminology. Though different methods of vocabulary learning have been studied, what is really needed is a more general approach to studying vocabulary learning and retention which allows one to look at any method of vocabulary learning and determine which components or aspects of the method are either poorly or well retained.

The keyword method seemed an ideal vocabulary method to study for a number of reasons. First, it has been demonstrated to be extremely effective in learning not only new foreign vocabulary (e.g., Atkinson & Raugh, 1975) but also in learning new English vocabulary (Sweeney & Bellezza, 1982) as well as in learning technical terms and their definitions (e.g., Jones & Hall, 1982). Secondly, it seemed that the method might lend itself quite well to a componential analysis and the use of verbal reports to trace the sequence of processing the various components. In the keyword method, individuals learn vocabulary items in a two-step process. In the first step, an English word,

similar in sound to the foreign word (the 'keyword'), is learned by noting its similarity to the foreign word. In the next step, an interactive image is formed using the keyword and the English equivalent of the foreign word. For example, to learn the Spanish word 'codo,' which means 'elbow,' a subject first learns the keyword 'code' so that later when 'codo' is seen the word 'code' should come easily to mind. The subject then forms an interactive image between 'code' and 'elbow': for example, an image of a secret code written on someone's elbow.

In all of the experiments conducted so far, we have tested subjects individually, using 30 to 60 Spanish words, gathering acquisition data as well as immediate and delayed retention data. Our general methodology for the experiments has been as follows. Subjects initially learn a set of 30 vocabulary items using the keyword method with the help of a computer presentation program which presents the Spanish word, its associated keyword, and the English equivalent. During this acquisition phase subjects are asked to think aloud as they learn the items, providing a record of the encoding process that can later be compared to the verbal reports gathered at test. In the next phase, a dropout phase, subjects are tested to one correct trial on all of the vocabulary items using a cued recall task in which they see the Spanish word and must say the English word. Response times are recorded using a voice key system. Feedback is provided on missed items by reshowing the Spanish word, keyword, and English word. No verbal reports are taken during this phase.

Following the first two phases, subjects are then tested on the vocabulary items, again using a cued recall task. However, during the test phase, no feedback is provided. Subjects are asked to respond by saying the correct word into a microphone and after each trial they are asked to give a retrospective verbal report. During the test phase, three types of retrieval tasks are used to test the various components of the retrieval process that we would expect to be associated with using the keyword mnemonic. Naturally, we test whether

subjects can recall the English word in response to the Spanish word, but we also test the two components that presumably mediate the retrieval of the English word: recall of the keyword with the Spanish word as a cue and recall of the English word with the keyword as a cue. For each of the three retrieval tasks, we collect recall data, reaction times, as well as the verbal reports of what subjects thought about as they tried to remember a word. There are two blocks of test trials for all the different retrieval items during both immediate and delayed testing.

In Figures 13 and 14 are recall data for two groups of subjects, the first group tested immediately and after one week, the second group tested immediately and after one month. The results are quite striking. After a week, according to the recall data, subjects showed little forgetting of either the Spanish to English word or the two component steps. However, for the one-month retention group, recall dropped considerably for the Spanish to English word task and the keyword to English word component. Recall for the keyword given the Spanish word, though, remained quite high. The implication is that the keyword component is retained quite well, while the image component has decayed after a month. This picture is sharpened if we consider the reaction time data as well. In Figures 15 and 16, we have averaged the reaction times across the three different probe types for the one week and one month retention groups. The reaction times clearly show decay at one week and at one month, though the decay at one month is obviously much greater. The important point is that the reaction time data provide a much more sensitive measure of the decay function: With the reaction time measures, it is clear that there is some decay on the first test block after even a week. If we then consider the times for the three different retrieval tasks an even more interesting picture emerges (see Figures 17 and 18). The image retrieval task and the full retrieval task show considerable decay after even a week but the keyword retrieval task shows little if any decay. After a month, though, the keyword task shows some decay, but the

decay for the other two tasks is even greater still. The crucial implication of these results is that it is important to consider more than a single source of data in assessing retention functions.

Along with the reaction time and recall data, we have the protocol data for which we have developed an encoding scheme. The encoding consists of a set of five variables corresponding to the following five components of the retrieval process: retrieval of the Spanish word, retrieval of any mediator associated with the Spanish word and the keyword; retrieval of the keyword (itself a mediator); retrieval of the mediator between the keyword and the English word (either an image or some other type of semantic relation); and retrieval of the English word. For each protocol, we can encode whether or not each of the component steps was present as well as other information to capture the sequence or duration of steps for that retrieval task. Thus, for example, on the image retrieval task, in which the keyword is provided and the subject retrieves the English word, the subject might report looking at the keyword, remembering his or her interactive image, and then retrieving the English word. We have just completed the encodings for one of our subjects from the one-week retention experiment and will soon enter his data into our analysis of his results.

The protocol data should enable us to model even more precisely the retrieval process. With the protocol data and the reaction time data we can fill in the picture provided by the recall data. A subject forgets one of the English words. So far, the reaction time data suggest that the most likely reason he or she forgets will be due to the image (or semantic relationship) encoded between the keyword and English word. The protocols seem to support this account: Subjects report a fair amount of interference between their images, so that an aspect of one image interferes with another image. The protocols should provide the most detailed account of the retrieval process and combined with the other two types of data should provide a comprehensive account of not only which components decay but how they decay.

The practical implications of our results to date are that when using the keyword method it might be advantageous to spend more effort developing ways to improve the second component of the mnemonic. In our present experiments, our keywords were generally quite close to the Spanish words, usually formed from a stem of the Spanish word (e.g., 'door' for 'doronico'). Using less similar looking or sounding keywords might alter the picture somewhat.

In addition to our work with the keyword method and vocabulary learning, during the first year, we have completed a rather careful analysis of cardiopulmonary resuscitation. Our analysis suggests that as with the vocabulary task, there may be certain components of this task that decay more easily, and we plan eventually to test how the different components decay and how differences in knowledge may mediate this retention. Again, we plan to use verbal reports to provide a detailed account of the process.

We have also completed some retesting of one of our waiters from our study of a year ago concerning waiters' expert memory for restaurant orders, and we plan to retest him again. The results suggest little decay of his skill after a year.

References

- Atkinson, R.C., & Raugh, M.R. (1975). An application of the mnemonic keyword method to the acquisition of a russian vocabulary. Journal of Experimental Psychology: Human Learning and Memory, 104, 126-133.
- Bahrick, H.P. (1979). Maintenance of knowledge: Questions about memory we forgot to ask. Journal of Experimental Psychology: General, 108, 296-308.
- Bahrick, H.P. (1984). Semantic memory contents in permastore: Fifty years of memory for Spanish learned in school. Journal of Experimental Psychology: General, 113, 1-29.
- Bahrick, H.P., & Phelps, E. (1987). Retention of vocabulary over 8 years. Journal of Experimental Psychology: Learning, Memory, and Cognition, 13, 344-349.
- Battig, W.F. (1972). Intratask interference as a source of facilitation in transfer and retention. In R.F. Thompson & J.F. Voss (Eds.), Topics in learning and performance. New York: Academic Press.
- Battig, W.F. (1979). The flexibility of human memory. In L.S. Cermak, & F.I.M. Craik (Eds.), Levels of processing in human memory (pp. 23-44). Hillsdale, NJ: Erlbaum.
- Bourne, L.E., Jr. (1974). An inference model for conceptual rule learning. In R.L. Solso (Ed.), Theories in cognitive psychology: The Loyola Symposium. Potomac, MD: Erlbaum.
- Ericsson, K.A., & Simon, H.A. (1984). Protocol analysis: Verbal reports as data. Cambridge, Massachusetts: MIT Press.
- Healy, A.F. (1982). Short-term memory for order information. In G.H. Bower (Ed.), The psychology of learning and motivation (Vol. 16, pp. 191-238). New York: Academic Press.
- Healy, A.F., Oliver, W.L., & McNamara, T.P. (1987). Detecting letters in continuous text: Effects of display size. Journal of Experimental Psychology: Human Perception and Performance, 13, 279-290.

- Jones, B.F., & Hall, J.W. (1982). School applications of the mnemonic keyword method as a study strategy by eighth graders. Journal of Educational Psychology, 74, 230-237.
- Schneider, W., & Shiffrin, R.M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. Psychological Review, 84, 1-66.
- Shiffrin, R.M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. Psychological Review, 84, 127-190.
- Slamecka, N.J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. Journal of Experimental Psychology: Human Learning and Memory, 4, 592-604.
- Sweeney, C.A., & Bellezza, F.S. (1982). Use of the keyword mnemonic in learning english vocabulary. Human Learning, 1, 155-163.

Table 1

Illustration of Sample Problem

<u>Task</u>	<u>Subject Sees</u>	<u>Calculator</u>		<u>Subject Responds</u>
		<u>Available</u>		
Read	$6 \times 8 = 48$	No		" $6 \times 8 = 48$ "
Generate	$6 \times 8 = ?$	No		" $6 \times 8 = 48$ "
Calculate	$6 \times 8 = ?$	Yes		" $6 \times 8 = 48$ "
Verify	$6 \times 8 = 48$	No		" $6 \times 8 = 48$ , correct"

Table 2

Proportion of Answers Correctly Recalled in Experiment 1 as a Function of Locus of Cognitive Operations (Internal or External) and Locus of Stimulus Production (Internal or External)

<u>Stimulus Production</u>	Cognitive Operations	
	<u>Internal</u>	<u>External</u>
	Verify	Read
External	.68	.38
	Generate	Calculate
Internal	.68	.42

Table 3

Proportion of Answers Correctly Recalled in Experiment 2 as a Function of Locus of Cognitive Operations (Internal or External), Locus of Stimulus Production (Internal or External), and Retention Interval Condition (Immediate, Two-day, Seven-day)

<u>Stimulus Production</u> <u>and Retention Interval</u>	<u>Cognitive Operations</u>	
	<u>Internal</u>	<u>External</u>
<b>External</b>	<b>Verify</b>	<b>Read</b>
Immediate	.59	.42
Two-day	.40	.24
Seven-day	.24	.10
Mean	.41	.25
<b>Internal</b>	<b>Generate</b>	<b>Calculate</b>
Immediate	.55	.34
Two-day	.49	.16
Seven-day	.40	.14
Mean	.48	.21

Table 4

Proportion of Answers Correctly Recognized in Experiment 2 as a Function of Locus of Cognitive Operations (Internal or External), Locus of Stimulus Production (Internal or External), and Retention Interval Condition (Immediate, Two-day, Seven-day)

<u>Stimulus Production</u> <u>and Retention Interval</u>	<u>Cognitive Operations</u>	
	<u>Internal</u>	<u>External</u>
<u>External</u>	<u>Verify</u>	<u>Read</u>
Immediate	.82	.81
Two-day	.76	.61
Seven-day	.72	.52
Mean	.77	.65
<u>Internal</u>	<u>Generate</u>	<u>Calculate</u>
Immediate	.81	.65
Two-day	.75	.66
Seven-day	.69	.64
Mean	.75	.65

Table 5

Mean Number of Trials to Criterion for Each Memory Condition on Day 1 and Day 2  
(After a One-Week Delay)

<u>Day</u>	<u>Memory Condition</u>		
	<u>Spatial</u>	<u>Temporal</u>	<u>Item</u>
1	12.88	8.00	5.88
2	5.12	4.50	3.50

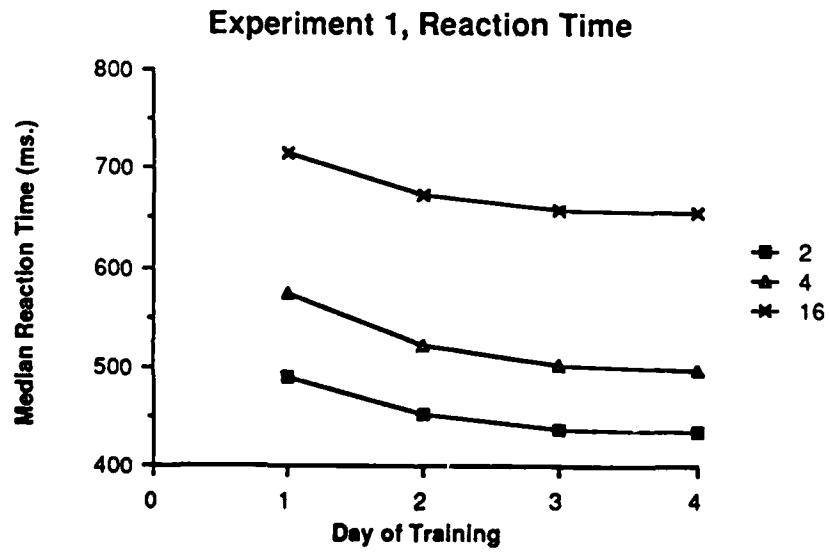
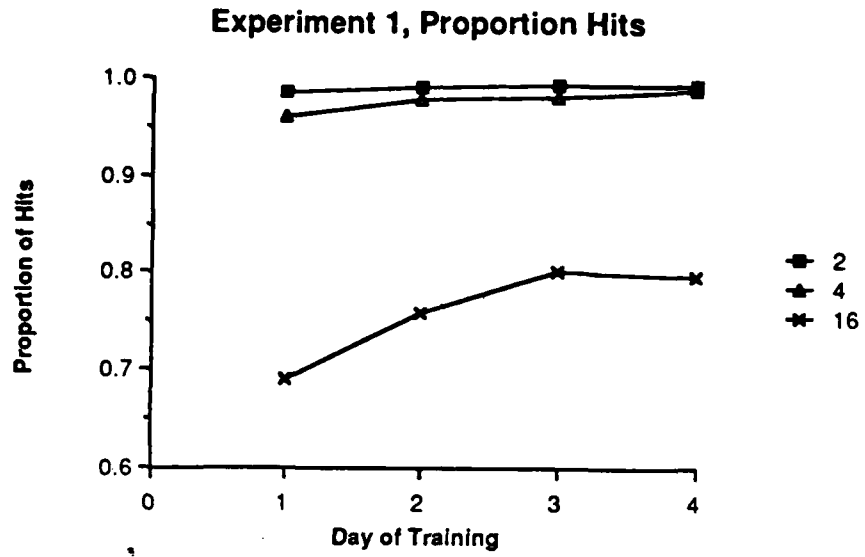
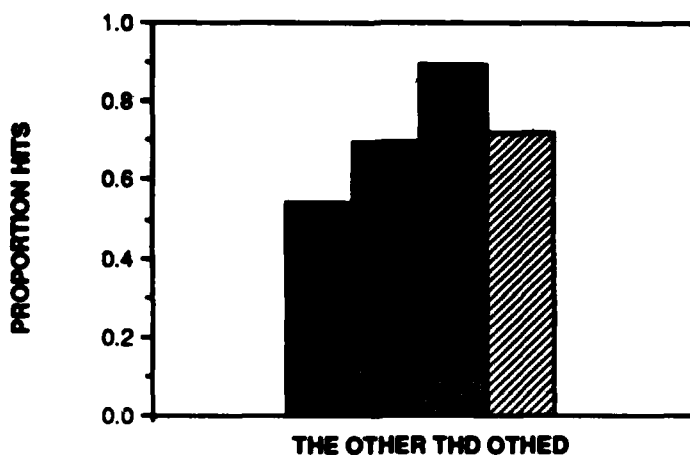
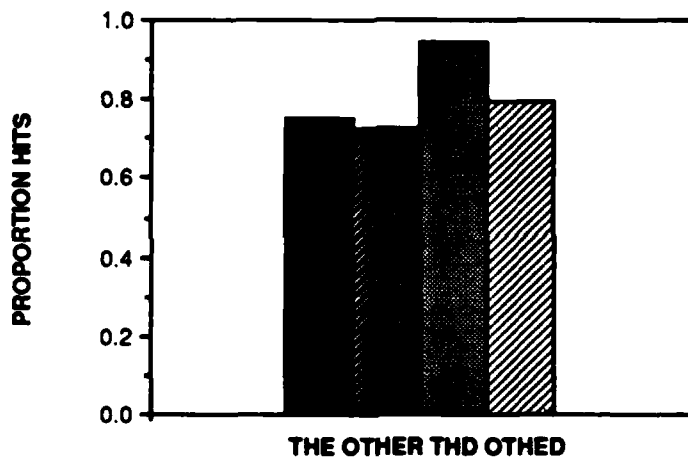


FIGURE 1

### PRETEST (h DETECTION)



### POSTTEST (h DETECTION)



### POSTTEST (t DETECTION)

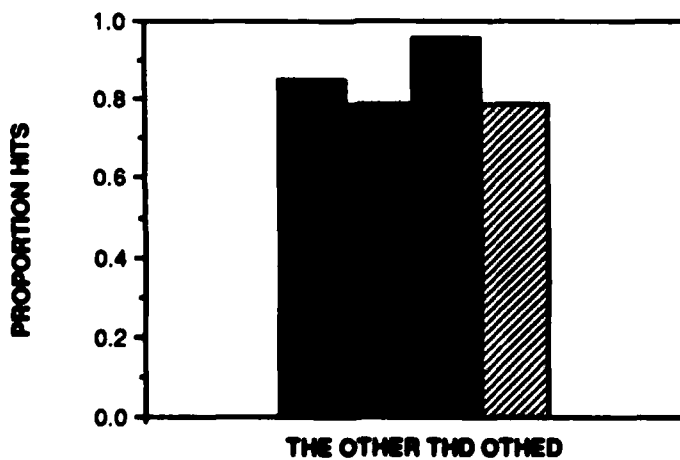


FIGURE 2

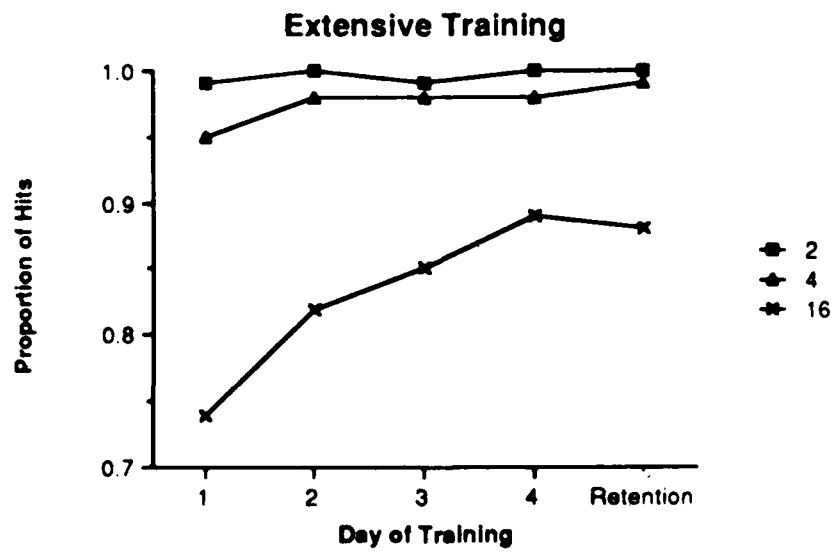
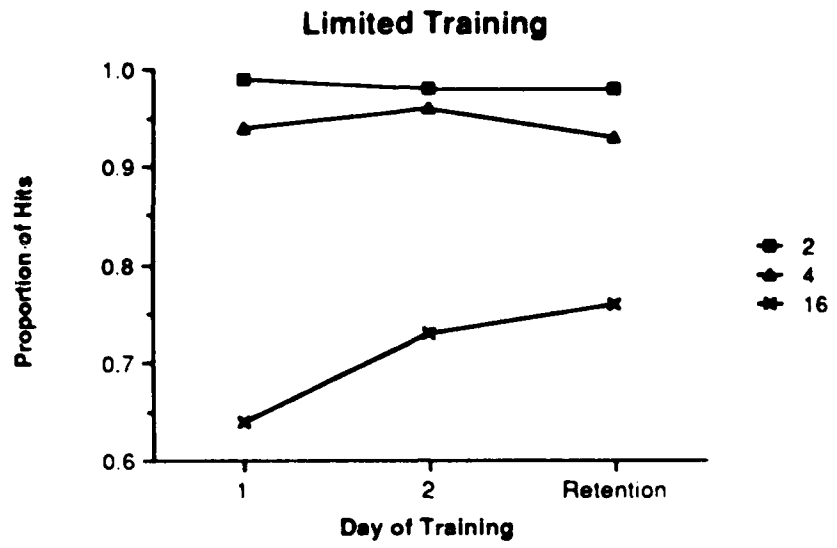


FIGURE 3

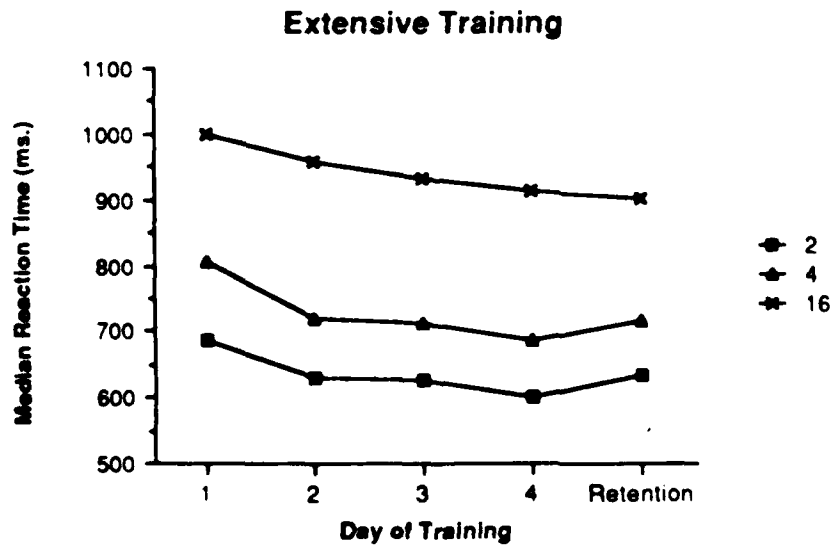
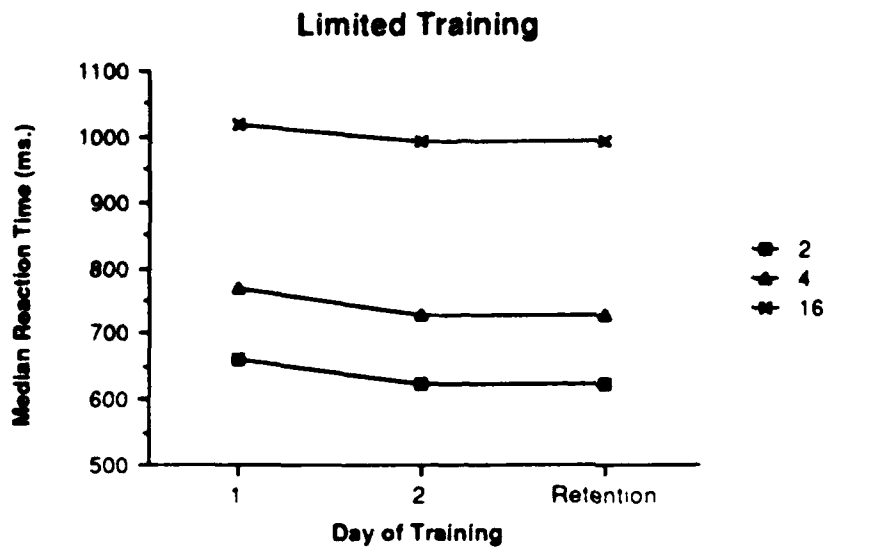


FIGURE 4

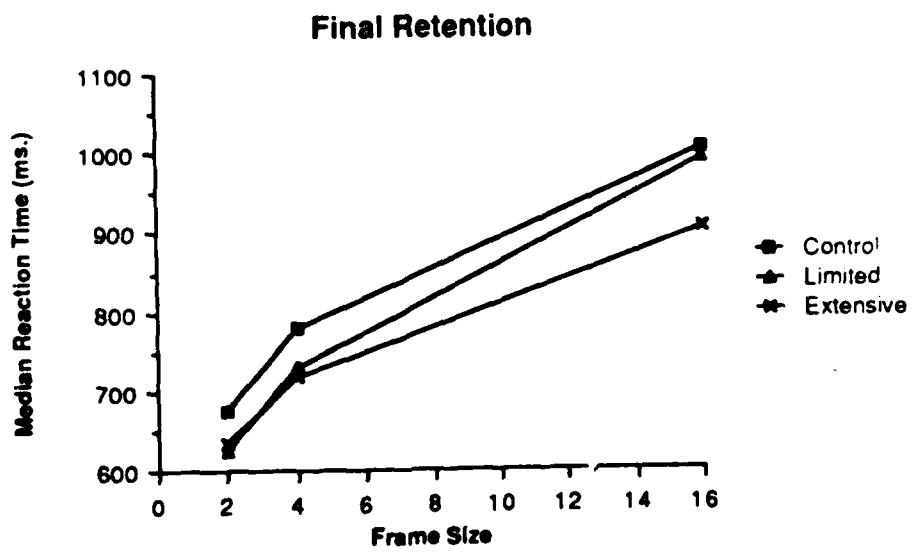
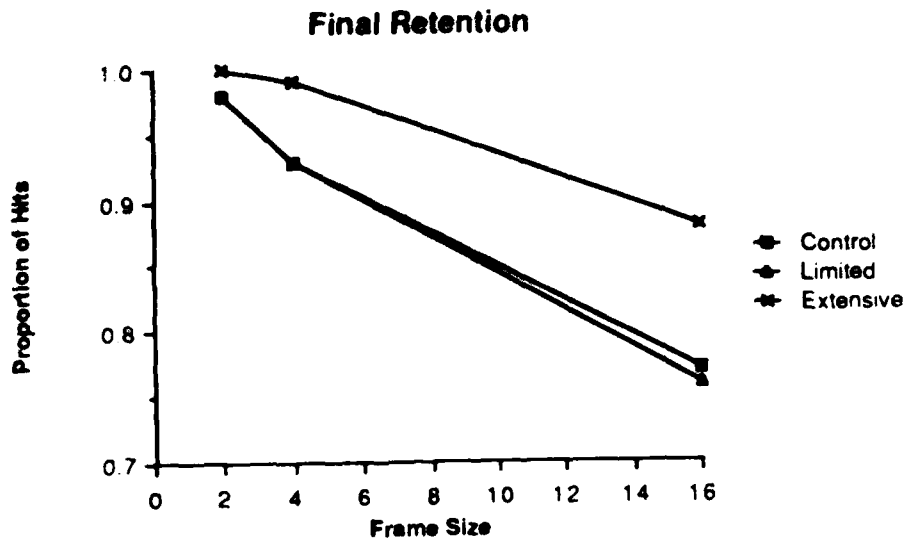


FIGURE 5

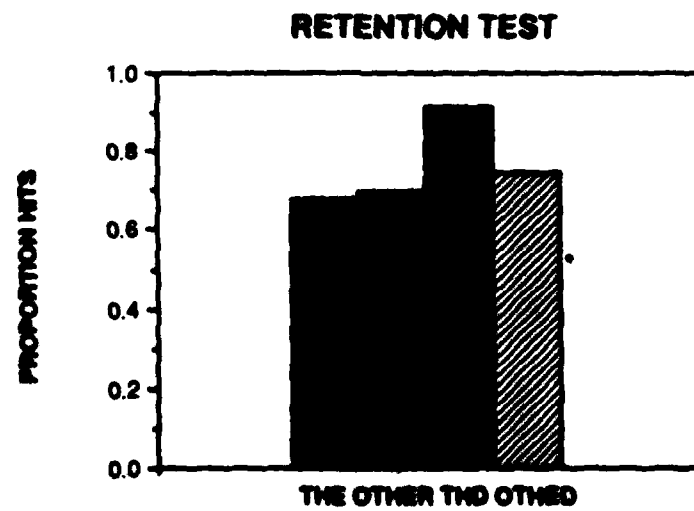
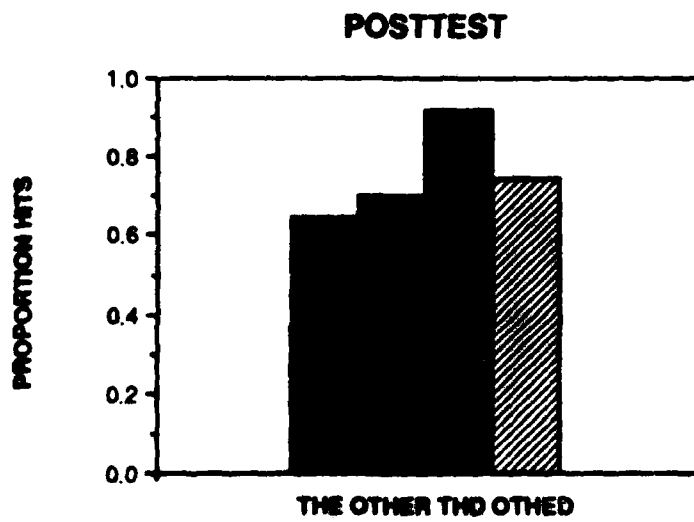
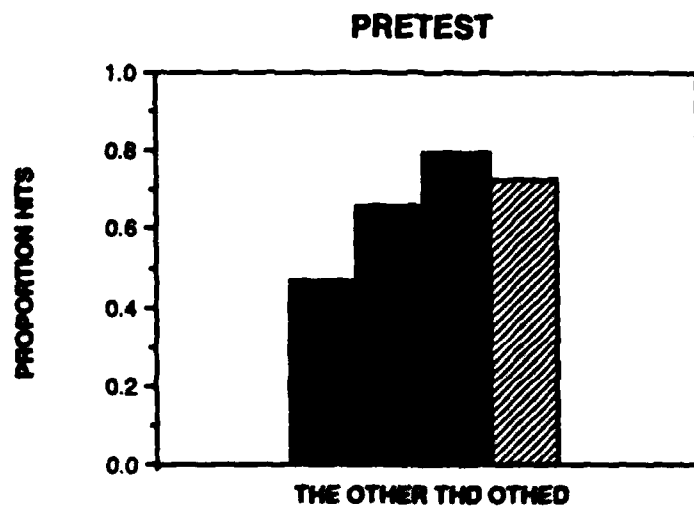


FIGURE 6

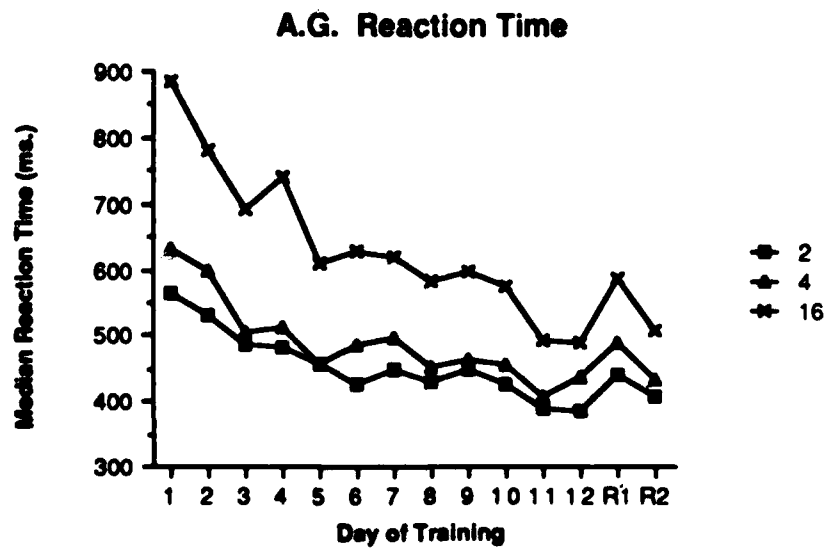
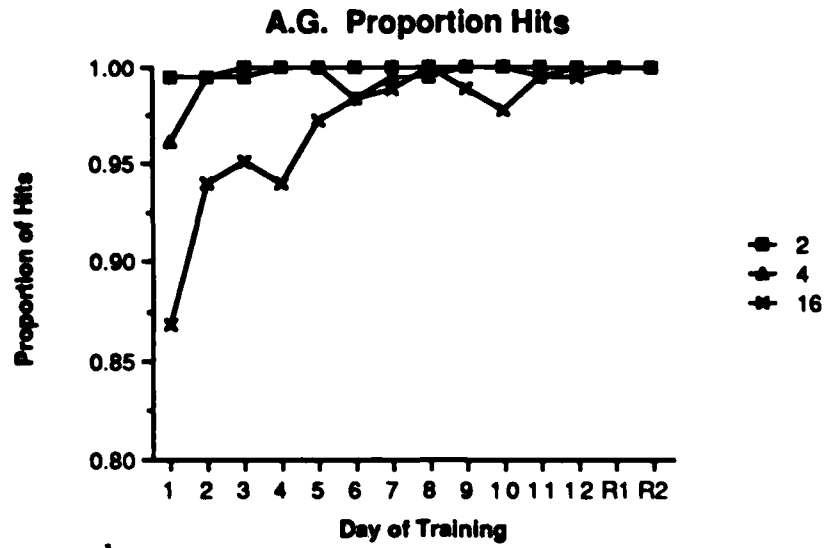


FIGURE 7

### MEDIAN RT AS A FUNCTION OF GROUP AND DAY

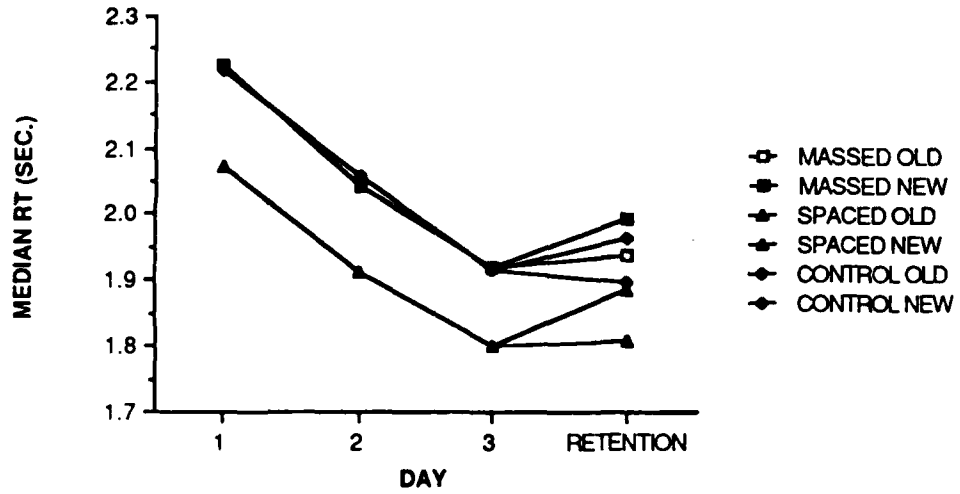


FIGURE 8

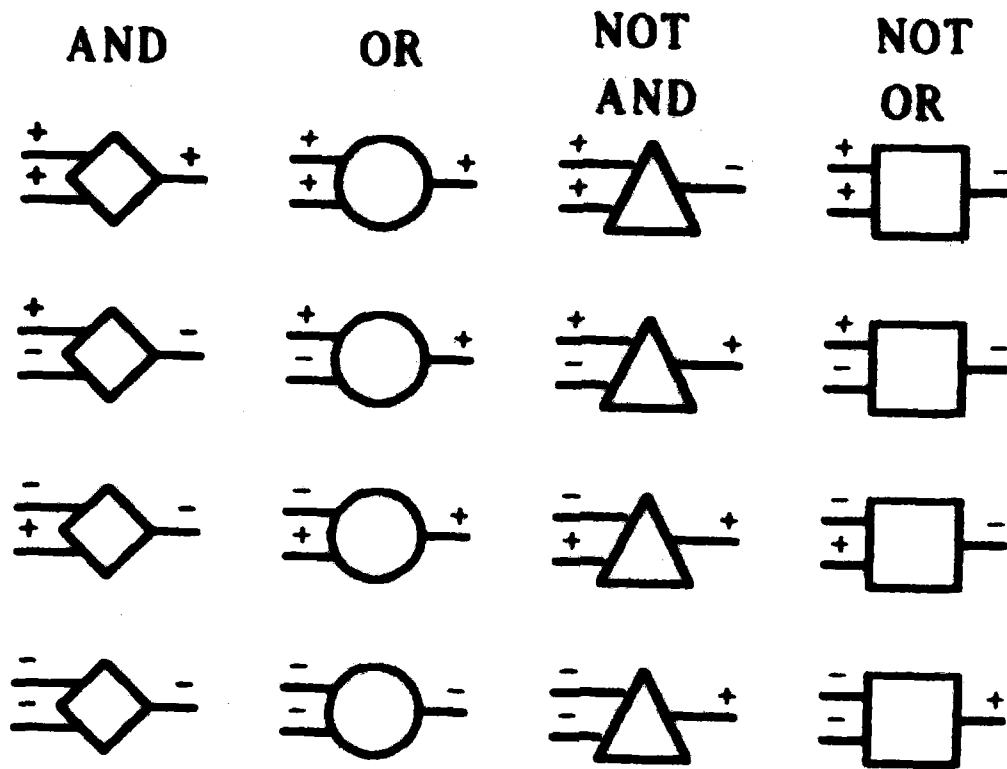


FIGURE 9

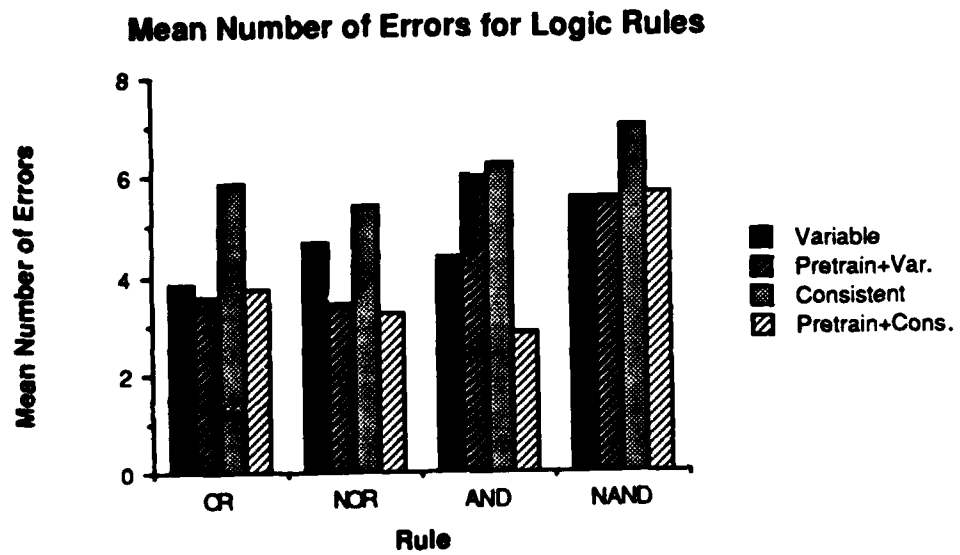
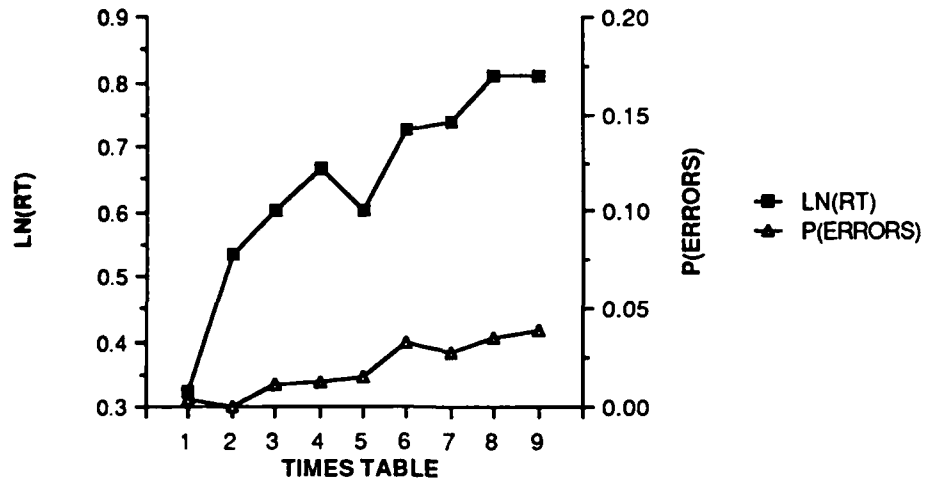


FIGURE 10

KEYPAD GROUP

LOG RT AND PROPORTION ERRORS



LOG RT AS A FUNCTION OF BLOCK

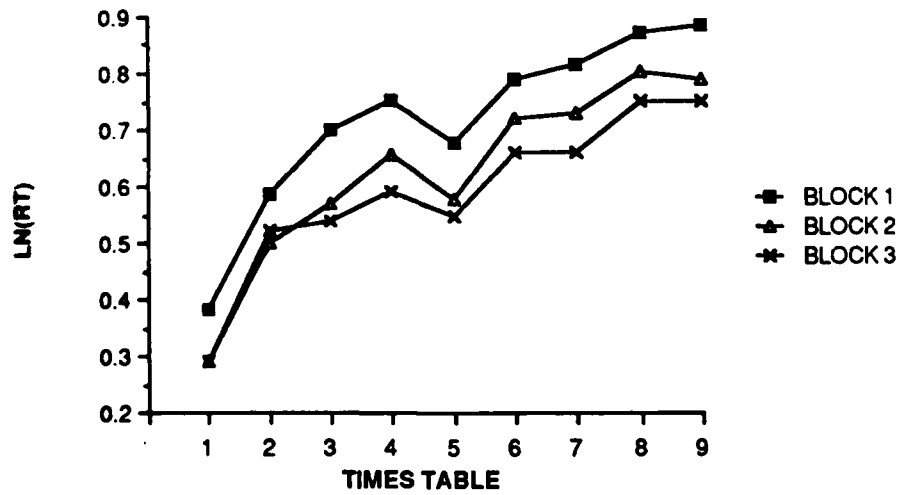
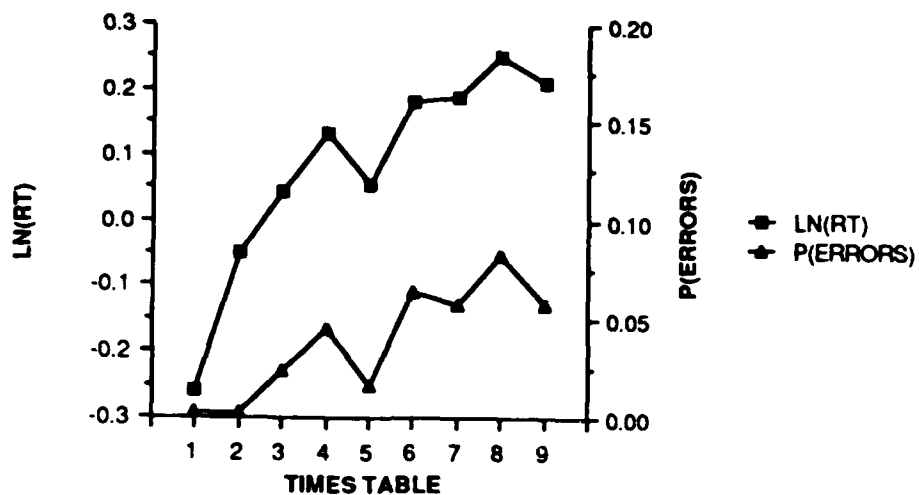


FIGURE 11

VOICE KEY GROUP

LOG RT AND PROPORTION ERRORS



LOG RT AS A FUNCTION OF BLOCK

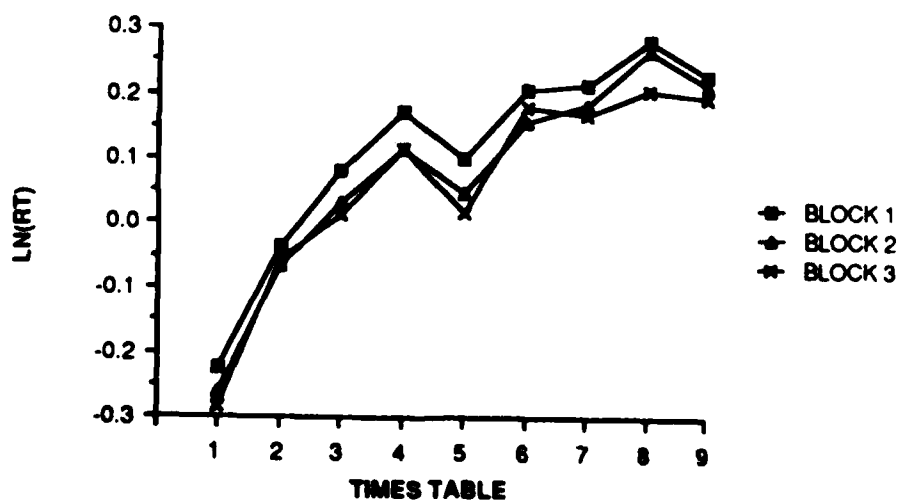


FIGURE 12

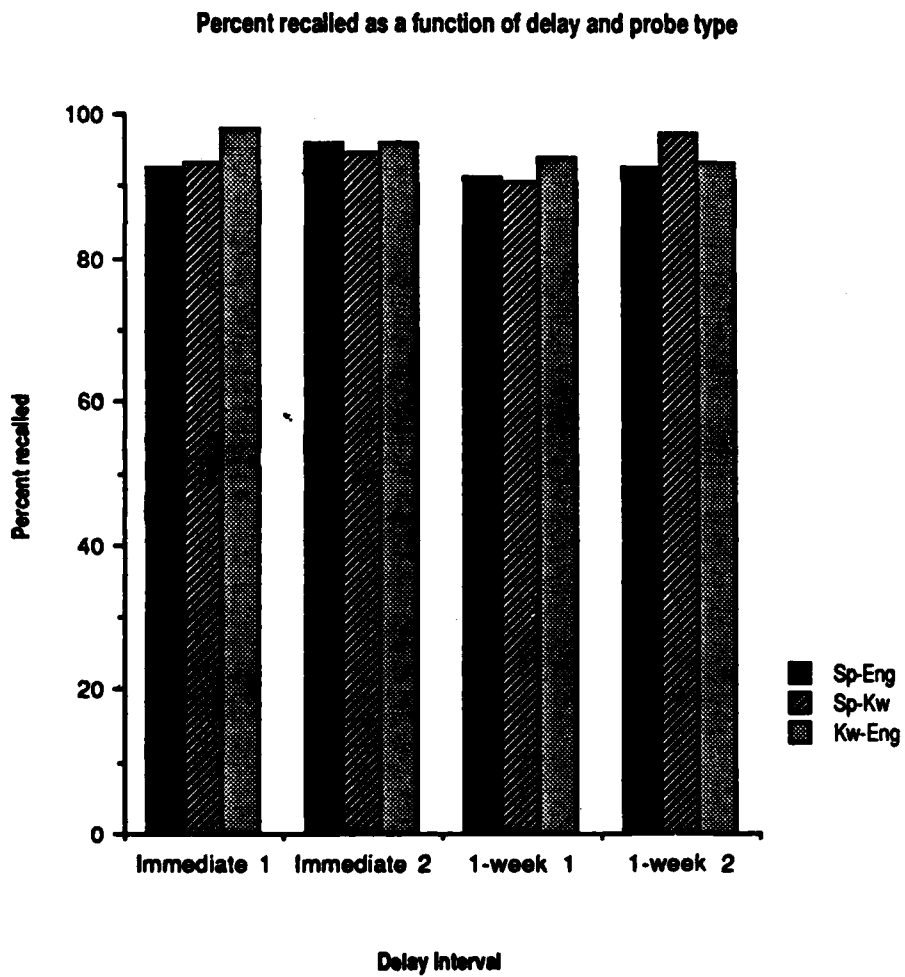


FIGURE 13

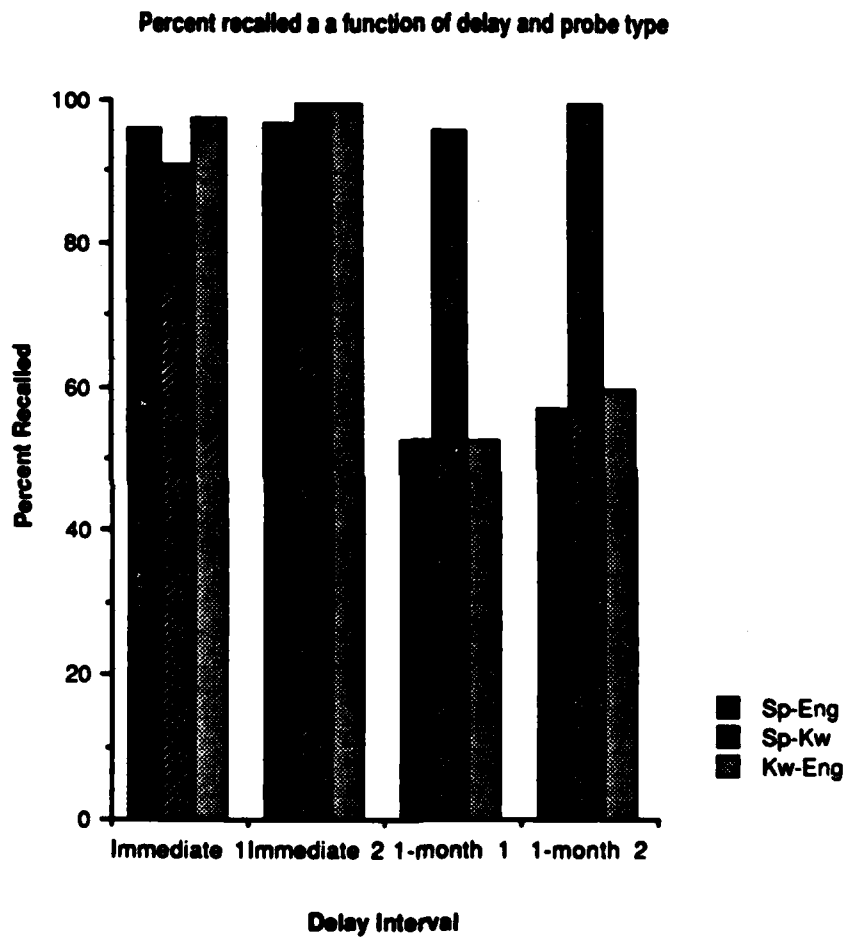


FIGURE 14

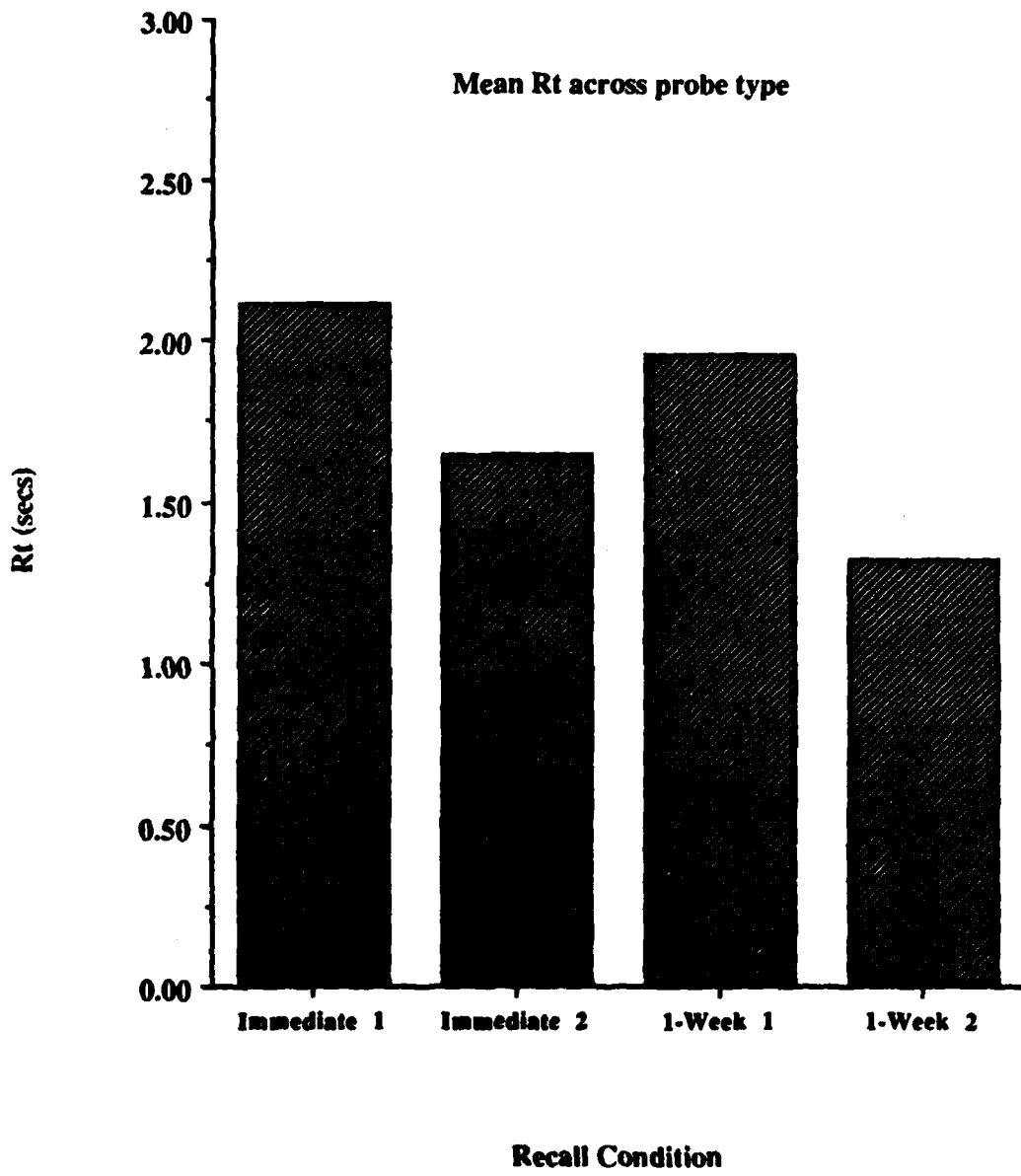
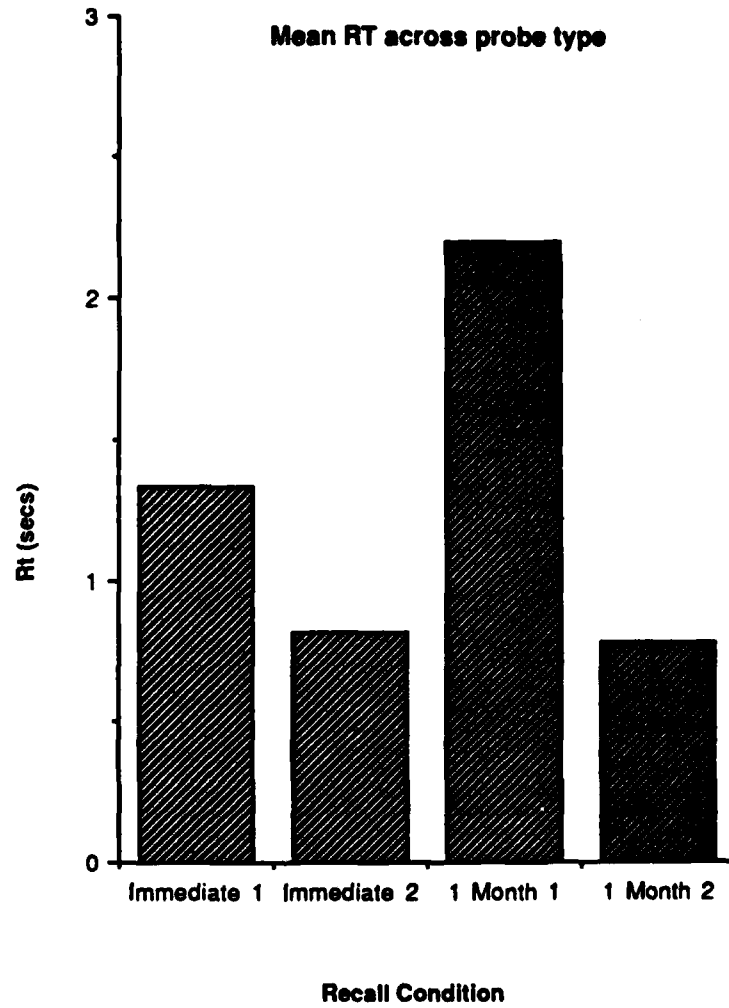


FIGURE 15



**FIGURE 16**

Mean reaction time for different recall tasks for different test occasions

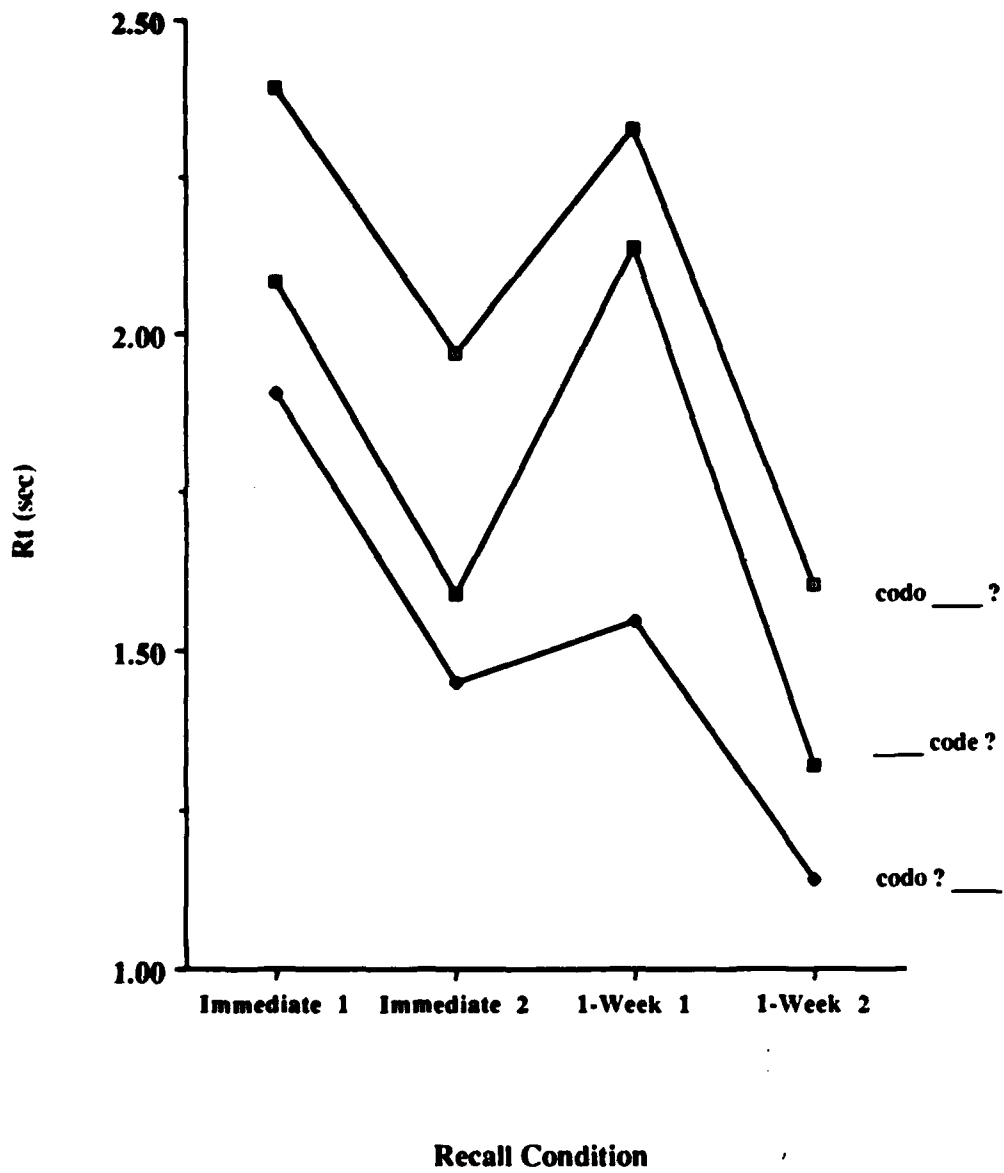
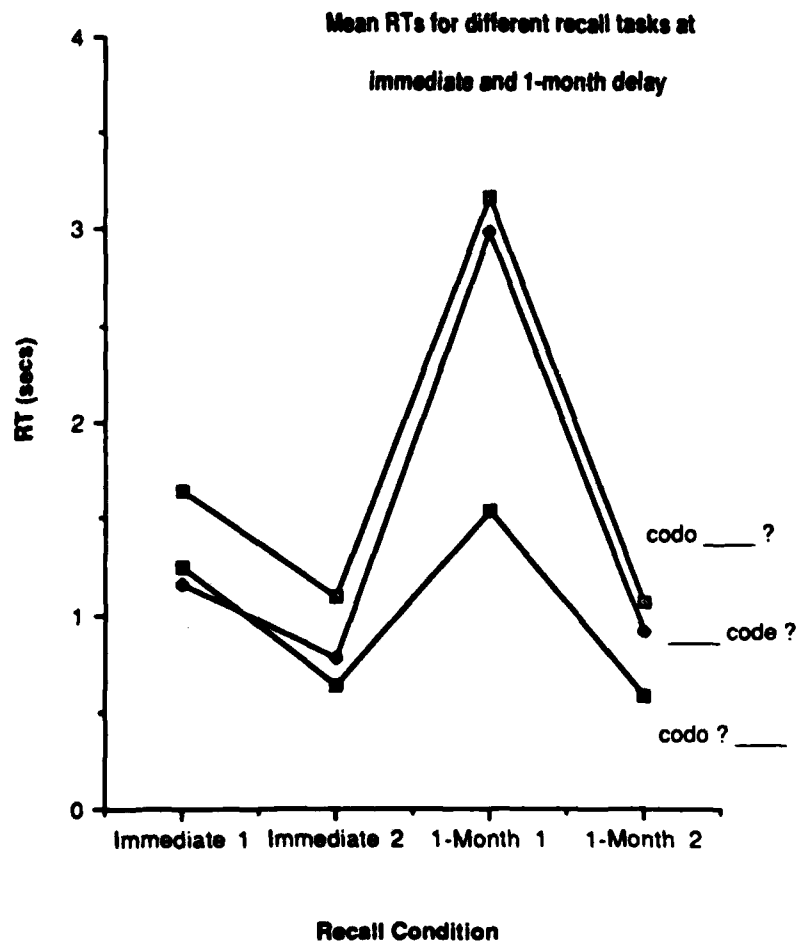


FIGURE 17



**FIGURE 18**