1975

Effects of instructional variables on the parameter values of verbal discrimination learning models

Edward Otto Marks

The University of Montana

Follow this and additional works at: https://scholarworks.umt.edu/etd

Let us know how access to this document benefits you.

Recommended Citation
Marks, Edward Otto, "Effects of instructional variables on the parameter values of verbal discrimination learning models" (1975). Graduate Student Theses, Dissertations, & Professional Papers. 4922. https://scholarworks.umt.edu/etd/4922

This Thesis is brought to you for free and open access by the Graduate School at ScholarWorks at University of Montana. It has been accepted for inclusion in Graduate Student Theses, Dissertations, & Professional Papers by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.
THE EFFECTS OF INSTRUCTIONAL VARIABLES
ON THE PARAMETER VALUES OF VERBAL
DISCRIMINATION LEARNING MODELS

By

EDWARD OTTO CHARLES MARKS

A thesis submitted in partial fulfillment
of the requirements for the degree of

MASTER OF ARTS IN PSYCHOLOGY

UNIVERSITY OF MONTANA

1975

Approved by:

[Signatures and dates]
ABSTRACT

Marks, Edward Otto Charles, M.A., August 1975, Psychology

The Effects of Instructional Variables on Parameter Values of Verbal Discrimination Learning Models (55 pp.)

Director: Dr. James R. Ullrich

The experiment investigated the effects of instructional variables, rehearsal, and imagery, on the parameter values of verbal discrimination learning models. Three instructional groups, a rehearsal and two imagery groups, of 21 subjects each were tested on a standard verbal discrimination task using an anticipation procedure. The instructional variables were found to have a considerable effect on the parameter values of verbal discrimination models. Overall group performance was a reliable indicator of the rank order of parameter values between groups for a given model with 2 or fewer parameters. Also the complexity of the model required to explain the data was found to be dependent on the instructions given the subjects. Finally, it was noted that regardless of the specific instructions, all groups performed better than would be expected from uninstructed groups.
ACKNOWLEDGEMENTS

The author would like to express his appreciation to the members of his committee, Dr. Laurence Berger, Dr. Andrew Lee, and Dr. Rudy Gideon, for their assistance and consideration. Special gratitude is extended to my chairman, Dr. James Ullrich, for his continual support and his helpful direction of the thesis research and the writing of this paper. Finally, the writer is indebted to his friend, Carol Heinen, whose help and encouragement greatly facilitated the production of this paper.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Introduction to the Problem</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Strategies and Models</td>
<td>12</td>
</tr>
<tr>
<td>II.</td>
<td>METHOD</td>
<td>16</td>
</tr>
<tr>
<td>III.</td>
<td>RESULTS</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Parameter Estimation</td>
<td>20</td>
</tr>
<tr>
<td>IV.</td>
<td>DISCUSSION</td>
<td>30</td>
</tr>
<tr>
<td>I.</td>
<td>BIBLIOGRAPHY</td>
<td>33</td>
</tr>
<tr>
<td>APPENDIXES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A.</td>
<td>SUBJECT INSTRUCTIONS</td>
<td>35</td>
</tr>
<tr>
<td>B.</td>
<td>PARAMETER ESTIMATING PROGRAM FOR LS-2, LS-3, ONE-ELEMENT &amp; SINGLE-OPERATOR MODELS</td>
<td>44</td>
</tr>
<tr>
<td>C.</td>
<td>PARAMETER ESTIMATING PROGRAM FOR AS MODEL</td>
<td>52</td>
</tr>
<tr>
<td>D.</td>
<td>WORD-PAIRS</td>
<td>55</td>
</tr>
</tbody>
</table>
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Total Number of Errors Across Trials (2-15)</td>
<td>19</td>
</tr>
<tr>
<td>2. Number of Errors on Trial 5</td>
<td>22</td>
</tr>
<tr>
<td>3. Observed Frequencies for Given Four-Tuple Response Sequences (trials 2-5)</td>
<td>23</td>
</tr>
<tr>
<td>4. Parameters Estimates</td>
<td>25</td>
</tr>
<tr>
<td>5. Minimum Chi-square values</td>
<td>26</td>
</tr>
<tr>
<td>6. Lag Length Used in the Parameter Estimates For the AS Model</td>
<td>28</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

Modern mathematical learning theory begins about 1950 and considerably extends the probabilistic approach to behavior theory introduced earlier by Thurstone (1930), Gulliksen (1934) and others. Whereas earlier approaches typically looked at simple arithmetic means of errors, trials to criteria, number of correct responses, etc. to make inferences about the nature of learning, modern mathematical models have demonstrated that other aspects of the data (distribution of responses, distribution of sequences, etc.) require description and prediction. The development of these modern approaches have taken two concurrent and related paths, operator models and state models. The first class of models postulates an infinite number of stages of learning; the second class postulates only a finite number of stages. Both developments, however, formulate the probabilistic nature of learning in a similar manner. The learning process is conceived of as a sequence of discrete trials each of which involves the presentation of a stimulus situation to the subject. The subject responds to the situation by selecting from a set of responses in accordance with the associated probability. Following a response is an outcome which may produce changes in response probabilities before the next trial. All such models attempt to describe these changes in response probability from trial to trial and the resulting distribution of response sequences.
The operator approach which assumes infinite learning stages was initiated by Bush and Mosteller (1951, 1955). Part of the motivation leading to this approach resulted from the consideration of some general experimental conclusions. Typically, in a learning task the probability of a correct response increases as the number of reinforced trials increases, there is some upper limit beyond which the probability cannot increase, and the biggest changes in probability occur in the earlier trials. The two principles assumed by Bush and Mosteller which lead to a class of linear operator models are the combining-classes condition and the independence-of-path condition. The combining-classes condition asserts that the choice of how to partition the sets of response alternatives is not unique. For example, in a probability learning experiment with three lights (red, white, blue) the condition states that it does not matter if the experimenter decides that there are two alternatives, red and not red, before he calculates the probability change to the red light on the next trial. That is, the probability of the response "not red" on the next trial is the same whether he calculates the probability of the response "blue" and of "white" separately and adds these together or whether he adds them together first and calculates the effect of "not red" of the complimentary operator to "red". The independence-of-paths assumption asserts that the probability of a response on the next trial is dependent only on its probability of the preceding trial and on the event that occurred. Bush (1960) showed that these two conditions lead to a linear operator of the form:

\[ \phi p = \alpha p + a \]

\[ 0 \leq \alpha \leq 1 \]

\[ -a \leq \alpha \leq 1-a \]
where $\phi$ is the operator on $p$ (the probability of a correct response) and $a$ and $\alpha$ are constants.

The state approach is easily understood by considering Bower's (1962) one-element learning model. In this model, each item of the to-be-remembered set may be represented by a single stimulus which is sampled on every trial. The stimulus element is in either of two states: $U$ (unconditioned state) or $C$ (conditioned state). On each trial, the probability of transition from state $U$ of state $C$ is a constant, $c$, which is independent of the outcomes of previous trials and trial numbers. The probability of remaining in the conditioned state once it is entered, is $1$. If an element is in state $U$, the probability of being correct is $g$ and if an element is in state $C$, the probability of a correct response is $1$. At the start of the first trial, the subject is assumed to be in state $U$ with a probability $\alpha$, usually assumed to be $1$. The above description has the following transition matrix, response probability vector, and starting vector:

\[
\begin{array}{ccc}
1-\alpha, & \alpha \\
\hline
\text{Trial}_n & C & U \\
\end{array}
\quad
\begin{bmatrix}
1-c, & c \\
0, & 1-c \\
\end{bmatrix}
\quad
\begin{bmatrix}
1 \\
g \\
\end{bmatrix}
\]

In the transition matrix, the entries represent the probability of moving from a given row state on trial $n$ to a given column state on trial $n+1$. The response probability vector represents the probability of a correct response for each row state. The entries in the starting vector give the probability of the subject starting the experiment in a given row state. It can be seen that this model assumes an all-or-none learning process.
The purely mathematical analysis of general classes of learning models has proved valuable in the investigation of the interaction between theory and data. An example of such a study in the area of verbal learning and memory is Atkinson and Crothers' (1964) paper. They compared the predictions of linear operator models and finite state Markov models for a general paired-associate learning paradigm. One of the important contributions of this paper is the formal introduction of forgetting as a Markovian process. The motivation for the introduction of forgetting resulted from some earlier research by Suppes and Ginsberg (1963) where it was found that there is a tendency for the probability of a correct response given an error on the previous trial, to increase over trials prior to the last error and not simply remain constant as predicted by an all-or-none learning model. Suppes and Ginsberg proposed a model to account for the non-stationary effect by introducing three conditioning states with three different probabilities for a correct response. However, Atkinson and Crothers found it necessary to reject Supper and Ginsberg's (1963) model for two reasons. First, there seemed to be no convincing experimental interpretation of the added state, and secondly, certain predictions of the model were inaccurate.

An alternative model proposed by Atkinson and Crothers assumes four stages of learning. Learning is conceived of as an independent process of encoding the stimulus element and then associating the stimulus with a correct response. Initially, the stimulus is assumed to be in state U (uncoded); in this state, the subject is assumed to respond by randomly guessing among the response alternatives. After the element is encoded, it can be associated with the correct response. Once the association is
formed, the stimulus element is absorbed in state L (long-termed memory) and the subject makes no subsequent errors. Transitions between the intermediate states, S (short-term memory) and F (forgetting state) are assumed to occur between the encoding and association phases. While the stimulus element is in state S, the subject always responds correctly. However, while the stimulus element is in state S there is a certain probability that it will be forgotten. In which case, the stimulus element is said to be in state F. While in state F, the subject guesses randomly; however, forgetting is only partial since once a stimulus has been encoded, it cannot return to state U. The model can be represented by the following transition matrix and response probability vector:

\[
\begin{bmatrix}
L & S & F & U \\
L & 1 & 0 & 0 & 0 \\
S & a & (1-a)(1-f) & (1-a)f & 0 \\
F & a & (1-a)(1-f) & (1-a)f & 0 \\
U & a(1-a)(1-f) & a(1-a)f & 1-a \\
\end{bmatrix}
\]

where \(c\) is the probability that a stimulus element will be encoded, given that it hasn't been already; \(a\) is the probability that an already encoded element in either state S or F will go into state L; \(f\) is the probability that a stimulus item in short-term memory will be forgotten. It is assumed that \(f\) depends upon the number and type of intervening items and upon the exposure time of a given item, for this effects the repetition rate and hence the slope of the forgetting function (Atkinson and Crothers, 1964). The above model is referred to as the LS-3 model. The authors also investigated a special case of this model (LS-2), where \(c = 1\). That is,
the stimulus is in the uncoded state only on the first trial.

In the paired-associate learning paradigm used by Atkinson and Crothers, the subjects were told the responses available to them (either three or four response alternatives) at the beginning of the experiment in which each alternative occurred equally often as the to-be-remembered set. A response was then obtained from the subject on each presentation of an item and the subject was informed of the correct response following his response. The results indicate that only the models incorporating a forgetting hypothesis (LS-2, LS-3) were adequate to explain the data. One model which failed to adequately fit the data was Bower's one-element model. However, earlier experiments by Bower (1961) and Estes, Hopkins and Crothers (1960) obtained good fits with the one-element model. In an effort to reconcile this result, Atkinson and Crothers suggested that the difference lies in the fact that in Bower's and Estes' studies, only two response alternatives were available whereas in their study, three and four response alternatives were used. Subsequent research in the area has further complicated the issue. Underwood and Freund (1968) and Richardson (1969) found better subject performance with two response alternatives but Smith, Jones and Thomas (1963) and Hintzman (1967) reported worse performance. A study by Restle (1965) indicated that the majority of subjects given two response alternatives in a paired-associate learning task attended only to the stimuli associated with one of the two response alternatives. Since the adaptation of this strategy reduces the stimuli learning by one-half, the more subjects there are in a group who use the strategy, the better the overall performance. This was shown to be the case in a study by Hall and Weneroth (1970) in which
one group of subjects were instructed to use this strategum, one group not to use it and a third group was a control (instructed neither way). The first group made the fewest errors while the group instructed not to use the strategum made the most errors.

The necessity to consider different strategies of learning in the evaluation of learning models has become an important focal point in recent experiments intended to test models. Because of their natural occurrence and the ease with which subjects learn them, imagery and rehearsal are perhaps the most widely used strategies. Imagery usually refers to the fact that information is encoded in a spatial form as distinct encoding information in a verbal form (Pavio, 1971). The distinction between imagery and verbal processes has been primarily made in terms of their functions as symbolic systems. Pavio (1971) has made the distinction by looking at three functional dimensions: (1) concrete-abstract; (2) parallel-sequential; (3) static-dynamic. In the concrete-abstract dimension, images are specialized for the representation of concrete objects and events whereas the verbal system is more useful in dealing with abstract concepts, problems, and relationships. Pavio wrote about the parallel-sequential dimension as follows:

Imagery is basically a parallel system in both the spatial and operational sense. It is capable of sequential processes as well, if a response sequence is intrinsic to the imagery (e.g., imagining one's self walking down a familiar road or street passing familiar buildings and 'other signposts' in their natural sequence), or if its elements are linked to sequential operations involving the verbal system (e.g., counting corners of an imagined letter). The verbal system on the other hand functions in an operationally parallel manner as well as sequential. Imaginal and verbal systems thus overlap fully in regard to the capa-
city for operationally paired functions; they are differentiated with respect to spatial processing, which is characteristic only in imagery, and sequential processing, which is relatively more characteristic of the verbal system.

The static-dynamic dimension seems to be strongly situationally dependent with a great deal of interaction. However, in general the imagery system seems to be able to manipulate spatial transformations with greater facility than the verbal system whereas static representation is generally better in the verbal system.

The only clear experimental manipulation of memory strategies is through instructional and training variables with rehearsal and imagery instructions the most widely used. In the typical instructions for rehearsal, the subject is asked to rehearse a word or word pair aloud or subvocally. A typical experimental situation using imagery is when the subject is presented with arbitrary pairs of unrelated words (usually concrete nouns), like TABLE-DOG, and instructed to associate them by imagining a visual scene or mental picture in which the two objects are interacting in some way. After a few examples, the subject begins the experiment. Bower (1969) pointed out that it may be significant that the college students he tested with such a procedure apparently understood and carried out the instructions without the slightest difficulty. Evidence based on subjective reports indicate that encoding of the stimuli (both auditory and visual) in a spatial form is not uncommon (Pavio, 1971). Classic mnemonic systems such as the method of loci are occasionally used as instructional or training variables.

Results demonstrating the superiority of imagery have been obtained by Bower (1969); in this study, the subjects were divided into
two groups. One group was given imagery instructions where they were asked to associate a pair of concrete nouns by imagining a visual scene in which these two objects interacted in some manner. Each subject was given a five-second exposure to each of twenty pairs, followed by a cued recall test of twenty pairs immediately after the study trial. The left hand word was the cue for the literal recall of the right hand word; five seconds were given for each recall and the subject was informed of the correct response at the end of the test trial. Five lists were learned in this manner and at the end of the session, all hundred pairs were tested again. The results clearly indicate the superiority of the imagery subjects; they recalled about one and one-half times as much as the rehearsal subjects and the difference was statistically significant on all lists. Schnorr and Atkinson (1969) found similar results using a within-subjects design in which one-half of the items in a thirty-two pair list of concrete nouns were studied by rote repetition and the other half by imagery. Three such lists were presented for one study test trial and then retention was tested again one week later. Immediate recall tests showed much higher recall for pairs learned by imagery, the percent correct ranged from about 80% to 90% as compared to 30% to 40% learned by the repetition group. Imagery also resulted in significantly better long-term retention. Smith and Noble (1965) demonstrated the powerful effect of mnemonic training over a control group given no training. There is a considerable body of literature in agreement with the above studies for paired-associate learning. Hall and Wenderoth (1972) essentially replicated their earlier study (mentioned above, 1970) comparing the effects of instructional variables on the
parameter values of paired associate learning models with results in accord with the above experiments. That is to say, the more efficient recall strategy instructions (learning only the stimuli associated with one of the responses) yield data which was adequately described by a simple Bower-type all-or-none learning model whereas more complex models (LS-3, LS-2) were required to describe the results from the group instructed not to use the strategy.

There is only a small amount of research investigating the effects of imagery instructions on verbal discrimination learning. In verbal discrimination, the subjects are presented a pair of verbal units, usually words one of which has been arbitrarily designated as correct by the experimenter. The subjects' task is to learn which of these units is the correct one. Using a similar paradigm, Winograd, Karchmer, and Russell (1971) studied the effect on recognition memory of imagery vs. associative instructions. The subjects were required to recognize only the to-be-remembered words. The results indicate that imagery mnemonics are superior in recognition learning to the usual associative learning instructions. Bower (1969) found similar results while testing the hypothesis that imagery simply increases the general level of availability of the response term rather than exerting a specific associative effect suggesting that imagery would not have an effect on recognition memory. This hypothesis was disconfirmed by his experiment comparing imagery and repetition instructions where a random half of the pairs were tested for recognition and half for recall. The recognition test was a five-alternative multiple choice. Recognition was 97% for imagery subjects vs. 71% for repetition; recall was 87% for imagery subjects vs. 37% for
repetition. All differences were significant. The recognition outcome disconfirms the response availability conjecture since the imagery-repetition differences still appears when responses are made directly available by the multiple-choice test.

**Introduction to Present Problem**

The intended investigation centered around the question of how instructional variables, rehearsal and imagery, affect the parameter values of verbal discrimination learning models. The verbal discrimination task was one in which each subject was given a list of word pairs and in each pair one word was arbitrarily labeled correct and the other word labeled wrong. A standard anticipation procedure was used where the subject's task was divided into a study phase and a test phase. During the study phase, the subject saw all the word pairs on the list and which of the words have been designated as correct. After the study phase, the subject was again presented the list and asked to guess which word of each pair had been designated as correct.

Three instruction groups, a rehearsal group and two imagery groups were used. The rehearsal instructions simply asked the subject, during the study phase, to repeat the correct word three times. The imagery instructions asked each subject to form a mental image of the two words interacting in such a way that the correct one is always in some invariant spacial relation to the incorrect word (e.g., on top of, right of, etc.); or the subject was simply asked to form an image using only the correct word.
Strategies and Models

By considering how instructional variables affect learning, some expectations of how the instructions affect the parameters of some verbal discrimination models can be developed. Since the one-element and single-operator models are the simplest, they will be presented first in the discussion of how learning strategies might affect the parameters of these models. In addition to the models already covered (LS-2, LS-3, the one-element, and the single operator model), a recent model by Atkinson and Shiffrin (1969) and Atkinson, Brelsford, and Shriiffin (1967) will be discussed.

The single operator linear models assume that the probability of a correct response increases according to the equation: $P_{n+1} = (1-\phi)P_n + \phi$, $0 < \phi < 1$ where $P_1$ equals $1/r$; $r$ being the number of response alternatives. From the discussion of imagery instructions, it appears that with this strategy there is a tendency for learning to take place in an all-or-none manner. Thus one might expect to encounter a steep slope for the learning curve up to near the asymptotic point. This steep slope would be reflected by a high value of $\phi$; i.e., $\phi$ close to 1. On the other hand, the rehearsal strategy was shown to lead to poorer performance and learning seemed to be typified by a more gradual learning process. Thus one might expect that $\phi$ would be lower in the rehearsal group than in the imagery groups. The one-element model assumes an all-or-none learning process, thus it is expected to adequately describe the data for the imagery instructed groups while providing a poorer description of the data from the rehearsal instructed groups. Further $c$ should be greater for the imagery groups than the rehearsal groups due to the higher rate of encoding in the imagery groups.
The LS-3 model has three parameters which may be affected by differential learning strategies. It seems reasonable that the use of imagery, when the stimulus words are concrete, would insure the value of c to be close to 1 (provided that the subject is given sufficient time to form an image). That is to say with concrete words, the stimulus element should be easily encoded on the first study trial. Furthermore, the forgetting parameter, f, should be small; i.e., once the stimulus has been encoded, it is not likely to be forgotten by some one using an imagery strategy (Schnorr and Atkinson, 1969). Thus it would seem that the LS-2 and LS-3 models would both provide good fits of the data with no significant differences between them. When the subject uses a rehearsal strategy, however, it would appear that the LS-3 model would be significantly better than the LS-2 model for describing the data; because c should be significantly less than 1; i.e., encoding is not all-or-none. Rehearsal strategies should take longer to encode the material.

Atkinson and Shiffrin (1967) have developed an alternative model assuming three memory states: a very short lived memory system called the sensory buffer; a temporary memory state called the memory (or rehearsal) buffer; and a long term storage state called LTS. It is assumed that all incoming information from the senses first enters the sensory buffer, resides there for a short time (on the order of a few milliseconds), decays, and is lost. The memory buffer is assumed to have a limited and constant capacity for homogeneous items. It is viewed as containing those items which have been selected from the sensory buffer for repeated rehearsals. If the memory buffer is full, each new item
can only enter by knocking out one of the items which are already residing in the buffer at the time. The size of the buffer $r$ depends on the nature of the items and must be estimated for each experiment. Information transfers from the memory buffer to the long term storage; that is only during the period an item resides in buffer can information about that item be transferred to LTS. Information is assumed to be transferred to LTS at a rate $\phi$ during the entire period the item resides in buffer. Thus if an item remains in buffer exactly $j$ trials, the amount of information transferred to LTS is $j\phi$. If an item is knocked out of buffer, information about that item which is stored in LTS decays at a constant proportion $Y$. Thus if an item was knocked out of buffer $j$th trial and tested on the $i$th trial, the amount of information stored in LTS would be $j\phi(Y^{i-j})$. If an item is in the memory buffer at the time of the test, the subject attempts to retrieve the item from LTS. The probability of a correct retrieval increases with the amount of information stored about the item. The probability of a correct response from LTS is given by the following equation:

$$P_{ij} = 1-(1-g)\exp[-j\phi(Y^{i-j})]$$

where $g$ is the guessing probability. $P_{ij}$ is the probability of correct retrieval from LTS given a lag $i$ (i.e., $i$ trials between presentation and test) and given it resided in buffer for $j$ trials. It is expected that $Y$ will be close to one for the imagery group because of the excellent long term retention of subjects using this strategy (Schnorr and Atkinson), while in the rehearsal group $Y$ will be less than its value in the imagery group. $\phi$ might be expected to be larger in the imagery group due
to a faster rate of information transfer in imagery subjects.
CHAPTER II

METHOD

The present experiment used a verbal discrimination task in which each of the three groups of 21 subjects were required to differentiate between the right and wrong words of a 26-pair verbal discrimination list.¹ The items to be discriminated were all printed on individual Hollerith data cards and presented to the subject via an IBM 029 keypunch (Ullrich, 1972, Ullrich and Balogh, 1972). The right and wrong words appeared in the card bed for each presentation and the subject was required to push a button corresponding to the correct position of the correct item of a particular verbal discrimination pair. After the subject responded, the trailing edge of the data card was exposed and the correct item shown to the subject. There was one card for each of the 26 pairs and a blank card separating each of the 25 trials. Each subject received a practice list before the beginning of the experiment. Stimulus items were high in rated imagery and meaningfulness (greater than 5.00 a seven point scale as scaled by Pavio, Yuille, and Madign, 1968). The list is presented in Appendix D.

Two experimental groups were given imagery instructions while the remaining group was given rehearsal instructions. The two imagery instructions varied in that one emphasized forming an image with only

¹A preliminary study indicated that all groups learned a 20-pair verbal discrimination list by the 2nd trial. Thus a 26-pair verbal discrimination was used to discriminate between groups.
the correct response (single imagery group) whereas the other group (double imagery) was told to form an image of the two words interacting in such a manner that the correct word is on top of the incorrect word. For example, the word pair "BOTTLE-CAR" (CAR correct) might be visualized by the single imagery group as a purple car riding on two wheels around a steep curve; the double imagery group might imagine the car riding over a series of broken bottles; and the rehearsal group would simply repeat the word CAR to themselves four times. Appendix A contains complete instructions for all groups.

Subjects were selected from volunteer students registered in an introductory psychology class at the University of Montana. Students were allowed to choose the times they wished to participate from a sign-up sheet that was made available to them. The subjects were assigned randomly to groups on the basis of the times at which they arrived at the experiments.
CHAPTER 3

RESULTS

Table 1 presents the total number of errors across trials 2 to 15 for each subject whose total errors were less than 50. These scores indicate that the single-imagery group's overall performance was the best while the rehearsal group's was the worst. Following the testing of each subject, a short debriefing session was given where the purpose of the experiment was revealed and each subject was asked if they encountered any difficulties with the task. Of the five subjects who scored 50 or more errors across trials 2 to 15, three reported difficulty in following the instructions. From Table 1, it can also be seen that 35 is the highest number of errors scored. It was thereby assumed that 50 or more errors occurring during the test trials indicated that the subjects did not fully understand the instructions or the task, and thus the inclusion of these subjects would adversely affect the parameter estimates. In the single-imagery (s-I) group, one subject score of 51 was discarded; scores of 50 and 80 were discarded from the double-imagery (D-I) group and scores of 70 and 135 from the rehearsal (REH) group. A one way analysis of variance was performed for both the total scores including all subjects and the scores in Table 1. Both $F(2,65)=4.86$, $p<.14$ and $F(2,60)=15.49$, $p<.01$, were significant and in the same direction. Turkey's multiple pairwise contrast test with $K=6$, $df=57$, $d=.05$, $T=(\frac{1}{4})q$ and $q=4.096$ was performed on the means in Table 1. All group means were shown to be
TABLE 1
TOTAL NUMBER OF ERRORS ACROSS TRIALS (2-15)

| SUBJECTS | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 |
|----------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|
| S-I      | 6 | 3 | 3 | 5 | 3 | 3 | 4 | 8 | 4 | 0  | 6  | 6  | 6  | 17 | 3  | 8  | 8  | 1  | 1  | 4  | 5  |
| D-I      | 3 | 23| 18| 35| 7 | 4 | 4 | 20| 21| 6  | 3  | 27 | 3  | 12 | 3  | 15 | 5  | 16 | 6  | 8  | 7  |
| REH      | 19| 15| 8 | 35| 17| 14| 35| 11| 17| 9  | 10 | 26 | 19 | 4  | 20 | 24 | 33 | 10 | 18 | 14 | 21 |

Means

S-I 4.95
D-I 11.71
REH 18.04
significantly different from each other. This suggests that some imagery memory strategies (e.g., S-I) are more efficient than others (e.g., D-I) and both imagery strategies are better than the rehearsal strategy.

Parameter Estimation

Parameter estimates were based on the Chi-square minimization procedure described by Atkinson and Crothers (1964) and by Atkinson and Shifferin (1967). The observation classes used in the parameter estimation of the linear operator one-element, LS-2 and LS-3 models were four-tuple response sequences. The 16 possible outcome sequences over trial \( n \) to trial \( n+1 \) are denoted as follows:

\[
\begin{align*}
0_1,n &= c_n e_n+1 e_n+2 e_n+3 \\
0_2,n &= c_n e_n+1 e_n+2 e_n+3 \\
0_3,n &= c_n e_n+1 e_n+2 e_n+3 \\
0_4,n &= c_n e_n+1 e_n+2 e_n+3 \\
0_5,n &= c_n e_n+1 e_n+2 e_n+3 \\
0_6,n &= c_n e_n+1 e_n+2 e_n+3 \\
0_7,n &= c_n e_n+1 e_n+2 e_n+3 \\
0_8,n &= c_n e_n+1 e_n+2 e_n+3 \\
0_9,n &= e_n e_n+1 e_n+2 e_n+3 \\
0_{10},n &= e_n e_n+1 e_n+2 e_n+3 \\
0_{11},n &= e_n e_n+1 e_n+2 e_n+3 \\
0_{12},n &= e_n e_n+1 e_n+2 e_n+3 \\
0_{13},n &= e_n e_n+1 e_n+2 e_n+3 \\
0_{14},n &= e_n e_n+1 e_n+2 e_n+3 \\
0_{15},n &= e_n e_n+1 e_n+2 e_n+3 \\
0_{16},n &= e_n e_n+1 e_n+2 e_n+3
\end{align*}
\]

Where \( c_n \) is a correct response on trial \( n \) and \( e_n \) is an error on trial \( n \). To illustrate the method of minimization let us consider the LS-3 model with parameters a, f, and c. Let \( Pr (0_i,n; a, f, c) \) denote the probability of event \( 0_{i,n} \) given the parameters a, f, and c. Further let \( N(0_{i,n}) \) denote the observed frequency of the stimulus items and let \( T = N(0_{1,n}) + N(0_{2,n}) + \ldots + N(0_{16,n}) \). Then the following function is defined:

\[
X^2 (a, f, c) = 16 \sum_{i=1}^{16} \frac{T Pr (0_{i,n}; a, f, c) - N(0_{i,n})^2}{T Pr (0_{i,n}; a, f, c)}
\]

The parameters a, f, and c were chosen to minimize this function. This was done by initially incrementing each parameter from 0 to 1 in steps
of .05 and testing each value of a given parameter with every value of the other two. The set of parameters which had a minimum $X^2$ were then chosen to be used as a reference point around which the parameters were incremented in steps of 0.01. This same process was continued until the desired degree of accuracy was achieved. For the present experiment three such internations were performed. The computer programs used for the estimations are presented in appendices B and C.

Although maximum likelihood estimators are known for the single linear operator and one-element models, they have not been determined for the other models; to maintain comparability across models, all parameter comparisons were made using the results of the $X^2$ minimization procedure.

The analysis described above was performed for trials 2 to 5. In this experiment these trials were selected because a major portion of the learning occurred during the first five trials. The inclusion of trial one added nothing to the power of the parameter estimation because all models assumed this to be a guessing trial. For trials above three virtually all responses were correct; this is indicated in table 2 where the probability of a correct response on trial 5 is presented for each group, the lowest probability was .9539.

Table 3 presents the observed frequencies of each $O_{i,n}$ event. For the S-1 group, there were 513 sequences ($O_1$) with no errors on trials 2, 3, 4, and 5; there were 4 sequences ($O_2$) with no errors on trials 2, 3, and 4, but an error on trial 5, and so on. Each group had 21 subjects with 26 word pairs; hence there were $21 \times 26 = 546$
<table>
<thead>
<tr>
<th></th>
<th>Number of Errors on Trial 5</th>
<th>Probability of a correct response on Trial 5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SINGLE</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Imagery</td>
<td>5</td>
<td>.9908</td>
</tr>
<tr>
<td><strong>DOUBLE</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Imagery</td>
<td>22</td>
<td>.9597</td>
</tr>
<tr>
<td>Rehearsal</td>
<td>35</td>
<td>.9539</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>---</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>S-I</td>
<td>513</td>
<td>4</td>
</tr>
<tr>
<td>D-I</td>
<td>438</td>
<td>16</td>
</tr>
<tr>
<td>REH</td>
<td>383</td>
<td>13</td>
</tr>
</tbody>
</table>
word pair responses sequences. There were a considerable number of response-sequence categories with 4 or fewer observations in them. As a result of this, the stability of the parameter estimates, especially for models with three parameters, is unknown and may be subject to excessive variation.

The chi-square minimization procedure was applied to the data given in table 3 for all memory models excepting Atkinson and Shifferin's (AS) model. It can be seen in table 4 that the learning parameters \((c, \theta, r)\) were generally greater for the S-I group than for the other groups. In the LS-2 model the reverse was true for the forgetting parameter \(f\).

Table 5 presents the minimum Chi-square values; i.e., the values obtained by using the parameter estimates in table 4. The interpretation of these values is limited by lack of data in some observation categories.

Also the observations in each class may not have been independently and identically distributed in that subjects may have varied in the rate at which they learned the correct response and word pairs may have differed in the rate at which they were learned. Individual subject or word pair data was not used however, because their use would further restrict the number of observations in each \(0_{1,n}\) category. In general the models with the greater number of parameters had the smallest Chi-square values but there does not seem to be any pronounced effects across groups. All models except the one-element model fit the data for the S-I group the best. The LS-2 and LS-3 models have identical Chi-square values for the S-I group due to the
### TABLE 4
PARAMETER ESTIMATES

<table>
<thead>
<tr>
<th>MODEL</th>
<th>PARAMETER</th>
<th>S-I</th>
<th>D-I</th>
<th>REH</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE-ELEMENT</td>
<td>c</td>
<td>.688</td>
<td>.460</td>
<td>.426</td>
</tr>
<tr>
<td>LINEAR OPERATOR</td>
<td>φ</td>
<td>.722</td>
<td>.462</td>
<td>.370</td>
</tr>
<tr>
<td>LS-2</td>
<td>a</td>
<td>.454</td>
<td>.370</td>
<td>.376</td>
</tr>
<tr>
<td></td>
<td>f</td>
<td>.146</td>
<td>.448</td>
<td>.714</td>
</tr>
<tr>
<td>LS-3</td>
<td>a</td>
<td>.454</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>f</td>
<td>.146</td>
<td>.056</td>
<td>.048</td>
</tr>
<tr>
<td></td>
<td>c</td>
<td>1.000</td>
<td>.698</td>
<td>.530</td>
</tr>
<tr>
<td>AS*</td>
<td>r</td>
<td>7</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>φ</td>
<td>9.947</td>
<td>1.547</td>
<td>0.499</td>
</tr>
<tr>
<td></td>
<td>y</td>
<td>.830</td>
<td>.871</td>
<td>.923</td>
</tr>
</tbody>
</table>

*Parameter estimates using lag lengths as observation classes
<table>
<thead>
<tr>
<th></th>
<th>ONE-ELEMENT</th>
<th>LINEAR OPERATOR</th>
<th>LS-2</th>
<th>LS-3</th>
<th>AS*</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-I</td>
<td>106.93</td>
<td>83.97</td>
<td>10.15</td>
<td>10.15</td>
<td>.161</td>
</tr>
<tr>
<td>D-I</td>
<td>154.67</td>
<td>209.51</td>
<td>41.75</td>
<td>40.10</td>
<td>.208</td>
</tr>
<tr>
<td>REH</td>
<td>74.42</td>
<td>242.84</td>
<td>40.22</td>
<td>38.10</td>
<td>.769</td>
</tr>
<tr>
<td>Total</td>
<td>336.02</td>
<td>536.32</td>
<td>92.12</td>
<td>88.35</td>
<td>1.78</td>
</tr>
</tbody>
</table>

* Chi-square based on only 6 observation classes
fact that \( c=1 \) in the LS-3 model (table 4).

The main dependent variable in Atkinson and Shifferins (AS) model was errors per lag. A lag is the number of word pair presentations between the first presentation of a given word pair and its second presentation. Thus, if a word pair was first presented in position 25 and next presented in position 28, the lag for that given word pair would be 2. In the minimization procedure 6 different lags of length 36, 9, 11, 15, 17, and 19 were used. The choice of these lag lengths was based on considering how items enter and leave the buffer. In the AS model the process of entering and leaving the buffer is left unspecified and is to be determined by the specific task requirements. It was assumed that due to the instructions given each subject, each new item would enter the buffer with a probability of one. The old items were assumed to be randomly removed when a new item entered the buffer. Table 6 presents the total correct out of twenty responses for a given lag. The single imagery group does best while the rehearsal group the worst. The parameter estimates presented in table 4 for the AS model were based on the data found in table 6. In table 4 the parameter most sensitive to group differences was \( \chi \), the learning rate. \( \chi \) decreases according to overall group performance. The values of \( Y \) (decay or forgetting rate) suggests that forgetting is fastest in the S-I group and slowest in the REH group. In isolation these results appear to be dissonant with overall group performance; however, when considering the interaction of all three parameters the results are consistent with overall group performance. For example, comparing
TABLE 6

LAG LENGTH USED IN THE PARAMETER ESTIMATION FOR THE AS MODEL

<table>
<thead>
<tr>
<th></th>
<th>6</th>
<th>9</th>
<th>11</th>
<th>15</th>
<th>17</th>
<th>19</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-I*</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>18</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>D-I*</td>
<td>19</td>
<td>19</td>
<td>18</td>
<td>19</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>REH*</td>
<td>16</td>
<td>18</td>
<td>18</td>
<td>19</td>
<td>15</td>
<td>15</td>
</tr>
</tbody>
</table>

* Scores represent total correct out of 20 possible
the S-I group with the REH group, it is apparent that although the forgetting is slow in the REH group the rate of information transfer to long term storage is almost twenty times faster in the S-I group. Further the size of the buffer is directly related to the probability of a correct response. The buffer size for the S-I group is 7 while only 6 for the REH group. The Chi-square values presented in table 5 for the AS model are based on a much smaller sample size (20 for each lag) and are therefore difficult to compare with the values obtained for the other models.
CHAPTER 4

DISCUSSION

Earlier predictions stated that the parameters directly related to the learning rate (e.g., c, theta, and a) would be higher in the imagery groups than in the rehearsal group and the opposite would be true for the forgetting parameters (e.g., f and y). For models with two or fewer parameters estimates were monotonic with respect to group performance. Parameters which reflect rate of learning were highest in the S-I group and lowest in the rehearsal group; the converse was true for those parameters which reflected forgetting rate. In the single parameter models (one-element and linear operator) a and φ both reflect the rate of learning and were monotonic with respect to group performance.

\[ \phi(S-I) \phi(D-I) \phi(REH) \]
\[ c(S-I) c(D-I) c(REH) \]

In the LS-2 model the forgetting parameter f, was also monotonic with respect to group performance in the direction predicted.

\[ f(S-I) f(D-I) f(REH) \]

For models with more than two parameters overall group performance was not an exact indicator of the rank order differences between parameter estimates for the different groups. This was mainly due to an interaction between a given parameter and the other two parameters. For example, in the LS-3 model f does not maintain monotonicity with respect to group performance as it did the LS-2 model. Consider the
D-I and REH groups; although $f$ was higher in the D-I group, the rate at which information is encoded ($c$) was faster in the D-I group. The overall effect is consistent with mean group performance as indicated by the small difference in the Chi-square values for the two groups.

The effects of instructional variables on the parameter values of verbal discrimination models can to some extent be predicted from the effects of the instructional variables on overall group performance. For models with 2 or fewer free parameters, overall group performance between groups. More complex models require consideration of different aspects of the effect of instructional variables most of which are undetermined at present.

The prediction that $c$ (the probability that a stimulus item will be encoded in the SL-3) model would be closer to one for the imagery groups than the rehearsal group was supported by the results of the present experiment. $C$ was 1.0 for the single imagery group, .60 for the D-I group, and .53 for the rehearsal group. With $c$ equal to 1.0 in the S-I group the LS-2 and LS-3 models make the same predictions and are mathematically identical. This suggests that the complexity of the verbal discrimination model required to explain the data is dependent on the strategies employed by the subjects.

In summary, instructional variables were found to have considerable effect on the parameter values of verbal discrimination models. Overall group performance was a reliable indicator of the rank order of parameter values between groups for a given model with 2 or fewer parameters. Further the complexity of the model required to explain
the data was found to be dependent on the instructions given the subjects.
BIBLIOGRAPHY


Bower, G.H., Application of a Model to Paired-Associate Learning, 1961, 26, 255-280.


APPENDIX A

VERBAL DISCRIMINATION

This experiment is designed to investigate human memory and to determine how people remember entities such as words. Since there are a great many ways each of us can remember any particular word and since human memory is very complex we are going to ask you to remember words in only one particular way. Other subjects will be asked to remember the same words in other ways; the important thing is that you must work hard and you must devote a conscious effort to remember the words using only the method suggested below. If you use other techniques or if you do not try hard enough the experiment will not be valid; so we please ask your cooperation and conscious effort to remember the words in this experiment.

The experimenter will tell you about all phases of the experiment after you have finished. If you would like we will send you a short progress report and a listing of how well you as an individual did relative to the others in your group. Naturally, your name will remain confidential and how well you personally did will be known only to you and the experimenters. If you wish, we will send you a copy of the final written report when it is completed -- we expect that this will take approximately one year.

We would first like to describe the nature of the experimental task. We are going to present you with a list of pairs of words; in each pair one member and only one member is "correct" -- the other
member is "wrong". For example, one pair of words might be BOTTLE and CORNER, in this case CORNER might be the correct word. There is nothing special about CORNER which makes it correct nor is there anything special about BOTTLE which makes it wrong -- we have essentially flipped a coin to randomly determine which member of each pair is called correct and which member is called wrong.

This study is split into two phases -- a guess and study phase and a test phase. During the first (guess and study) phase we will show you a pair of words and ask you to guess which is correct. Immediately after you guess you will see which one of the pair is the correct one. At this time we ask you to study the pair and try to remember which of the words is the correct one.

After going through the entire list of 26 pairs of words you will be shown the same pairs in a different order. This is the test phase; here you will be asked to remember which word of each pair is the correct word. This procedure will be repeated until the list of words is exhausted. We will ask you to learn a number of independent lists; between each test period and the presentation of the next list there will always be a short break.

Since this is a fairly complex set of instructions, we will show you in the next few pages what you will see and what we would like you to do. The pairs of words in this experiment are typed on IBM cards and we are using an IBM Keypunch to record your responses. The pairs of words are typed on the left of the cards and the correct member of the pair are typed on the right of the cards. The following is a typical card:
When you first start the experiment you will see only the pair on the left -- you will not see which is the correct word as in the following illustration:

Suppose you guess CORNER is the correct member. You indicate your choice by pressing the button marked 1 to indicate you choose the 1st member of the pair. When you press the button the card moves physically to the left
and you can see which word is correct, and at the same time, you can see the next pair:

| CORNER | NURSERY SHADOW |

At this time we would like you to try to memorize which one (CORNER) is correct out of the pair (CORNER-BOTTLE). After you have tried to learn the correct word we want you to look at the next pair -- in this case NURSERY SHADOW. You don't know which one of this pair is correct, so you have to guess. Suppose you guess NURSERY, in this case you respond by pushing the button marked 1 because NURSERY is the first word of the pair. When you push the button the apparatus moves the card physically to the left and you will see the correct word and the next pair of words.

Throughout this guess and study phase of the experiment we want you to first guess at the word -- there is no way, except by chance, that you can get more than 50% of the words right. After you push the button to indicate your choice the card containing the words moves and you can see which one is correct. After you learn which one is correct -- by using the strategy for memorizing which we will suggest later -- you look at the next pair and repeat the procedure.
So much for the guess and study phase. When you have completed the list of 26 words there will be a blank card. You know when this occurs because there will be no new pair of words appearing after you have made the 26th guess. The situation will look like this:

At this point the test phase is about to begin. We want you to push either the 1 or the 2 button to start the test phase. It does not matter which button you push. After you push the button another pair of words will appear; this time the pair is not new but rather one of the preceding pairs. It might be for example:

SHADOW NURSERY
Please notice that SHADOW NURSERY was one of the pairs presented previously. Now we want to test your memory to see if you can remember if SHADOW or if NURSERY was the correct word. You indicate your choice by pushing the appropriate button. As before the card will move to the left and you can see if you remembered correctly or not. The next page has an illustration of the cards at this instant in time:

![Card Illustration]

Notice that the pairs may not appear in the same order or the words within a pair may not appear in the same order. The only way you can use to decide which is the correct one is to remember the words.
Group REH

The technique we would like you to use to memorize the correct words in this experiment is rehearsal. We would like you to rehearse -- to yourself -- the correct words in the guess and study phase. For example, the first pair of words you saw was CORNER BOTTLE with CORNER the correct member. What we would like you to do is to say to yourself --CORNER--CORNER--CORNER--three times. By repeating the word to yourself in this manner you will be able to recognize the correct word in the test phase as the one of the pair which was rehearsed. For a second example, consider the next set of words NURSERY and SHADOW with NURSERY correct. We want you to repeat to yourself NURSERY--NURSERY--NURSERY--three times. We want you to duplicate this procedure for each of the correct words in each of the pairs. Note that we do NOT want you to rehearse the wrong words -- only rehearse the correct ones.

Thus, the strategy for memory, the mnemonic device, the trick or whatever you call it that we want you to use is rehearsal. We want you to repeat each of the correct words to yourself three times during the guess and study phase. If you do not rehearse the items three times the experiment will not be valid so please do not relax and let the words pass by. It is extremely important for this experiment that you rehearse only the correct word exactly three times.

Because these instructions are long and complex we would like you to reread them to make certain you understand every detail. In addition, there will be a short practice session to acquaint you with the apparatus.
The technique that we would like you to use is that of mental imagery. In particular, we want you to take the two words from each pair and form a mental image of the two words interacting in some wild and peculiar way. For example, the first pair of words which you saw was CORNER BOTTLE. CORNER was the correct member. What we would like you to do is dream up some wild image of, say, a cube sitting on top of a bottle supported by one corner. The cube will remind you of CORNER and you can use the image of a cube sitting in a bottle to remind you that CORNER was correct. It would be even better if the mental image was such that the cube wobbled back and forth while it was supported by the CORNER in the BOTTLE. The most important part is that the image places the correct word physically on top of the incorrect word. We want this characteristic to be in every one of your images. Another example is the second set of words -- NURSERY SHADOW with NURSERY correct. Here you might form a mental picture of a baby bed (from a NURSERY) on top of a long, dark, SHADOW. In each case there was a definite physical image, and the correct word was on top. This way, when you see the pair during the test phase you will be able to recognize the correct word by nothing whether it is on top or on the bottom.

It is extremely important for this experiment that you make an active conscious effort to form a mental image of each pair of objects. If these images are not formed, we will not know how to interpret the results and the experiment will not be valid. Therefore, we please ask your cooperation and to try as hard as possible to form images for each pair of words in each of the lists.
The memory technique we would like you to use is that of mental imagery. We want you to take the correct word from each of the pairs and form some wild image in your mind of that word, and that word alone. For example, the first pair of words you saw CORNER BOTTLE with CORNER correct. We would like you to form a mental image of some corner -- be it a corner of some room or a corner of some street, or whatever. We want you to form this mental image in your mind and to think of that image as correct. That way when you come across the words in the test phase you will be able to recognize which ones are correct by noting which ones you have formed the mental image of.

Another example might be the second set of words -- NURSERY SHADOW. You might form a mental image of the correct word -- NURSERY -- by thinking of a baby bed (from a NURSERY) or you might think of flowers in some floral shop or NURSERY. In this way you can see the correct words visually and tag -- or note -- which ones are correct.

Thus, the strategy for memory, the mnemonic device, the trick, or whatever you want to call it that we would like you to use is to form a mental image of the correct word in your mind's eye to look at and to envision as correct. If you do not form the mental images or pictures of these words the experiment will not be valid so we please want you to try as hard as possible to form a brief, but concrete, image of each of the correct words in the list.

Because these instructions are long and complex we would like you to reread them to make certain you understand every detail. In addition, there will be a short practice problem to acquaint you with the apparatus.
FORTRAN IV PROGRAM USED TO ESTIMATE THE PARAMETERS IN THE LS-2, LS-3, ONE-ELEMENT AND SINGLE OPERATOR MODELS.

APPENDIX B

C
C AUTHOR: JAMES P. ULLRICH, UNIVERSITY OF MONTANA
C
C SUB(1) FOR SINGLE OPERATOR LINE
C SUB(2) FOR ONE ELEMENT MODEL
C SUB(3) FOR LS?
C SUB(4) TWO ELE
C SUB(5) FOR LS3
C
DIMENSION O(16),F(16),IDEN(10),SUB(10)
COMMON O,E
REAL O
INTEGER SUB

CALL IFIL E (1, 'ACM')
99 READ (1,100) SUB
100 FORMAT (1I1)
101 READ (1,102) IDEN,N,NALT,O
102 FORMAT (10Ai,T3,I3,16F4.0)

IF (N.EQ.0) CALL EXIT
WRITE (3,103)
103 FORMAT (38H1ATKINS0-CROTHERS ANALYSIS OF VD DATA )
WRITE (3,104) IDEN,N,NALT,O
104 FORMAT (17HIDENTIFICATION = ,I3,A1,5X,14HNO TRIAL = ,I3,5X,
1 17HNO ALTERNATIVES = ,I2,8HCO(I) = ,8F6.2,8X,8F6.2)

IF (SUB(1) .NE.0) CALL SOL(NALT,N,THEATA,CHISQ)
IF (SUB(1) .NE.0) WRITE (3,135) CHISQ,THEATA
IF (SUB(1) .NE.0) WRITE (3,200) O
105 FORMAT (7H-SOL,7X,7HCHISQ = ,F12.4,5X,8WTHETA =,F12.4)
IF (SUB(2) .NE.0) CALL ONELE (NALT,N,C,CHISQ)
IF (SUB(2) .NE.0) WRITE (3,106) CHISQ,C
IF (SUB(2) .NE.0) WRITE (3,200) O
106 FORMAT (8H-ONELE,3X,7HCHISQ =,F12.4,5X,7HC =,F12.4)
IF (SUB(3) .NE.0) CALL LS2(NALT,N,A,F,CHISQ)
IF (SUB(3) .NE.0) WRITE (3,107) CHISQ,A,F
IF (SUB(3) .NE.0) WRITE (3,200) O
107 FORMAT (9H-LS2,7X,7HCHISQ =,F12.4,5X,3HA =,F12.4,5X,3HF =,F12.4)
IF (SUB(4) .NE.0) CALL TWOLE (NALT,N,GP,A,B,CHISQ)
IF (SUB(4) .NE.0) WRITE (3,108) CHISQ,GP,A,B
IF (SUB(4) .NE.0) WRITE (3,200) O
108 FORMAT (10H-TWOLE,4X,7HCHISQ =,F12.4,5X,9HP RIME =,F12.4,5X,
1 3HA =,F11.4,5X,3HB =,F12.4)
IF (SUB(5) .NE.0) CALL LS3 (NALT,N,A,F,C,CHISQ)
IF (SUB(5) .NE.0) WRITE (3,109) CHISQ,A,F,C
IF (SUB(5) .NE.0) WRITE (3,200) O
109 FORMAT (11H-LS3,7X,7HCHISQ =,F12.4,5X,3HA =,F12.4,5X,3HF =,F12.4,
15X,3HC =,F12.4)
GO TO 99
200 FORMAT (8H0E(I) = ,8F8.4/4X,8F8.4)
END

SUBROUTINE SOLCSO (O,E,CSO)
C O = OBSERVED FREQUENCIES
C E = EXPECTED FREQUENCIES
C CSO = TOTAL FREQUENCY
DIMENSION O(16),E(16)
REAL O
CSO = 0
DO 100 I=1,16
- 44 -
SUBROUTINE SOL(NALT, NT, PARI, CHISO)
CALL SUBCSO(O,T,F,CHISO)
RETURN
END

SUBROUTINE SINGLE OPERATOR LINEAR
NALT = NUMBER OF ALTERNATIVES
N = TOTAL NUMBER OF FIRST TEST TRIAL
PARI = THETA
DIMENSION E(16), O(16)
COMMON O,E
INTEGER TIME
REAL INCRE
PARI = 0.5
INCRE = 0.05
T = 3
DO 10 L = 1,16
10 T = T + O(L)
P1 = 1.0/FLOAT(NALT)
DO 300 TIME = 1,4
CHISO = 10.0E+30
DO 200 I=1,12
IF (((TIME .EQ. 4) .AND. (I .NE. 11)) .OR. (I .EQ. 11)) GO TO 200
PIT = PARI + INCRE*FLOAT(I-11)
IF (PIT .LT. 0) GO TO 200
ALPHA = 1.0 - PIT
ON = 1.0 - PIT
IF (NT .EQ. 0) ON = ALPHA*ON
PN = 1.0 - ON
ON1 = ALPHA * ON
ON2 = ALPHA * ON1
ON3 = ALPHA * ON2
PN1 = 1.0 - ON1
PN2 = 1.0 - ON2
PN3 = 1.0 - ON3
DO 100 N = 1,2
IF (N .EQ. 1) PN = PN
IF (N .EQ. 2) PN = ON
DO 100 N1 = 1,2
IF (N1 .EQ. 1) PN1 = PN1
IF (N1 .EQ. 2) PN1 = ON1
DO 100 N2 = 1,2
IF (N2 .EQ. 1) PN2 = PN2
IF (N2 .EQ. 2) PN2 = ON2
DO 100 N3 = 1,2
IF (N3 .EQ. 1) PN3 = PN3
IF (N3 .EQ. 2) PN3 = ON3
DO 100 N = 1,2
F(K) = PN*PN1*PN2*PN3
100 CONTINUE
CALL SUBCSO(O,T,F,CHISO)
IF (CHISO .GE. CHISO) GO TO 200
TOTAL = PIT
CHISO = CHISO
CONTINUE
RETURN
END

SUBROUTINE SONEE(OALT,N,PARO,CHISO)
DIMENSION E(16), O(16)
COMMON O,E
INTEGER TIME
REAL INCR
PARC = 0.0
INCRE = 0.05
T = 0.0
DO 10 L = 1,16
  10 T = T + O(L)
G = 1.0/LFLOAT(NALT)
DO 300 TIME = 1,4
CHISO = 10.0E+30
DO 200 I = 1,21
  IF (((TIME.EQ.4),.AND.,(I.NE.11)) .GO TO 200
  C = PARC + INCRE * FLOAT (I-11)
  IF ((C.LT.0.0).OR.(C.GT.1.0)) .GO TO 200
  FN = (1.0-C)**(N-1)
  SN = 1.0 - FN
  DO 160 LG = 1,2
    IF (LG.EQ.2) GO TO 110
    FNO = G
    SNO = 1.0
    GO TO 115
  110 FNO = 1.0 - G
       SNO = 0.0
       DO 150 LG = 1,2
         IF (LG.EQ.2) GO TO 120
         FN1 = G
         SN1 = 1.0
         GO TO 125
       120 FN1 = 1.0 - G
              SN1 = 0.0
              DO 150 LG = 1,2
                IF (LG.EQ.2) GO TO 130
                FN2 = G
                SN2 = 1.0
                GO TO 135
              130 FN2 = 1.0 - G
                     SN2 = 0.0
                     DO 150 LG = 1,2
                       IF (LG.EQ.2) GO TO 140
                       FN3 = G
                       SN3 = 1.0
                       GO TO 145
                    140 FN3 = 1.0 - G
                            SN3 = 0.0
                            DO 150 LG = 1,2
                              IF (LG.EQ.2) GO TO 150
                              FNO = G
                              SNO = 1.0
                              GO TO 155
                           150 FNO = 1.0 - G
                                                SNO = 0.0
                                                DO 150 LG = 1,2
                                                  IF (LG.EQ.2) GO TO 155
                                                  FN0 = G
                                                  SN0 = 1.0
                                                  GO TO 160
                                               160 CONTINUE
       M = (L0-1)*A + (L1-1)*4 + (L2-1)*2 + L3
       F(M) =
       1  FN * (1.0-C)**3
       2  +FN * ((1.0-C)**2) * C
       3  +FN * (1.0-C) + C
       4  +FN * C
       5  +S
       CALL SUBCSQ (0,T,F,CHISQ)
       IF (CHISQ.GE.CHISO) .GO TO 203
       PARCT = C
       CHISO = CHISQ
       200 CONTINUE
       PARC = PARCT
       300 INCRE = INCRE/5.0
SUBROUTINE LS2 (NALT, N, PARI, PARJ, CHISO)
C
SUBROUTINE LONG-SHORT - 2 (Atkinson-Crothers)
C
NALT = NUMBER OF ALTERNATIVES
C
N = TRIAL NUMBER OF FIRST TEST TRIAL
C
PARI = A
C
PARJ = F
DIMENSION C(16), E(16)
COMMON 0., E
REAL INCRE, 0.
INTEGER TIME
PARI = 0.0
PARJ = 0.0
INCRE = 0.05
C = 1.0
T = 0
DO 10 L = 1, 16
10 T = T + 0.05
G = 1.0 / FLOAT (NALT)
DO 300 TIME = 1, 4
CHISO = 10.0 / 0.05
DO 200 I = 1, 21
IF (TIME.EQ.4.0).AND. (I.NF.I1) GO TO 200
PIT = PARI * INCRE * FLOAT (I-1)
IF (PIT.LT.0.0) GO TO 200
IF (PIT.GT.1.0) GO TO 200
A = PIT
DO 200 J = 1, 16
IF (TIME.EQ.4.0).AND. (J.NF.I1) GO TO 200
PJ T = PARJ * INCRE * FLOAT (J-1)
IF (PJ T.LT.0.0) GO TO 200
IF (PJ T.GT.1.0) GO TO 200
F = PJ T
X = (1.0-A) * (1.0-F + F*G)
Y = (1.0-A) * (1.0-G)*F
A1 = A + X*(1.0 - Y)
A2 = Y - Y
A3 = Y * (1.0 - Y)
A4 = Y * Y
B1 = (1.0 - C) * (A*C + C*X*(1.0-Y) + G*(1.0-C)*
C*(1.0-C) + G*(1.0-C))
B2 = (1.0-C) * (C*X + G*(1.0-C) * (1.0-C) + (1.0-C)*G*(1.0-C) +
1.0-C) + G*(1.0-C))
B3 = (1.0-C) * (G*X*X + G*(1.0-C) * (1.0-C) + (1.0-G)*
G*(1.0-G) + (1.0-C) + (1.0-C) + G*(1.0-G)
B4 = (1.0-C) *((1.0-C) + (1.0-C) + (1.0-C) +
(1.0-C) + (1.0-C) + G*(1.0-G))
UN = (1.0-C)**(N-1)
TN = (C *(1.0-A) / (C - A)) * ((1.0-A)**(N-1) -
1.0-C)**(N-1)
FN = F * TN
SN = (1.0 - F) * TN
E(1) = (1.0+S+F-UN) + (SN + G*FN) * (A + Y*A1) +
G*UN*(C*(A + Y*A1) + G*B1)
E(2) = (SN + G*FN)*Y*A2 + G*UN*(C*Y*A2 + G*R2)
E(3) = (SN + G*FN)*Y*A3 + G*UN*(C*Y*A3 + G*R3)
E(4) = (SN + G*FN)*Y*A4 + G*UN*(C*Y*A4 + G*B4)
E(5) = (SN + G*FN)*Y*A5 + G*UN*(C*Y*A5 + (1.0-G)*R5)
E(6) = (SN + G*FN)*Y*A6 + G*UN*(C*Y*A6 + (1.0-G)*R6)
E(7) = (SN + G*FN)*Y*A7 + G*UN*(C*Y*A7 + (1.0-G)*R7)
E(8) = (SN + G*FN)*Y*A4 + G*UN*(C*Y*A4 + (1.0-G)*A4)
E(9) = (1.0-G)*FN*(A*X*A1) + (1.0-G)*UN*(C*A1 + G*A1)
E(10) = (1.0-G)*FN*Y*A2 + (1.0-G)*UN*(C*Y*A2 + G*A2)
E(11) = (1.0-G)*FN*Y*A3 + (1.0-G)*UN*(C*Y*A3 + G*A3)
E(12) = (1.0-G)*FN*Y*A4 + (1.0-G)*UN*(C*Y*A4 + G*A4)
E(13) = (1.0-G)*FN*Y*A1 + (1.0-G)*UN*(C*Y*A1 + (1.0-G)*A1)
E(14) = (1.0-G)*FN*Y*A2 + (1.0-G)*UN*(C*Y*A2 + (1.0-G)*A2)
E(15) = (1.0-G)*FN*Y*A3 + (1.0-G)*UN*(C*Y*A3 + (1.0-G)*A3)
E(16) = (1.0-G)*FN*Y*A4 + (1.0-G)*UN*(C*Y*A4 + (1.0-G)*A4)

CALL SUBROUTINE LS3 (NALT, N, PARI, PARJ, PARK, CHISO)
IF (CHISO.GE.CHISO) GO TO 203
PARIT = PIT
PARJT = PJT
CHISO = CHISO
200 CONTINUE
PARI = PARIT
PARJ = PARJT
300 INCRE = INCRE/5.0
RETURN
END

SUBROUTINE LS3 (NALT, N, PARI, PARJ, PARK, CHISO)
SUBROUTINE LONG-SHORT-3 (ATKINSON-CROTHERS)
NALT = NUMBER OF ALTERNATIVES
N = TRIAL NUMBER OF FIRST TEST TRIAL
PARI = A
PARJ = F
PARK = C
DIMENSION O(16), E(16)
COMMON O,E
REAL INCRE, O
INTEGER TIME
PARI = 0.5
PARJ = 0.5
PARK = 0.5
INCRE = 0.05
T = 0
DO 10 L = 1,15
10 T = T + C(L)
G = 1.0/FLOAT (NALT)
DO 300 TIME = 1,4
CHISO = 10.0F+30
DO 200 I = 1,21
IF (TIME.EQ.4, AND (I.NE.11)) GO TO 200
PIT = PARI + INCRE + FLOAT (I-11)
IF (PIT.LT.0.0) GO TO 200
IF (PIT.GT.1.00) GO TO 200
A = PIT
DO 203 J = 1,21
IF (TIME.EQ.4, AND (J.NE.11)) GO TO 200
PJT = PARJ + INCRE + FLOAT (J-11)
IF (PJ.TLT.0.0) GO TO 200
IF (PJ.GT.1.00) GO TO 200
F = PJT
X = (1.0-A)*(1.0-F+F*G)
Y = (1.0-A)*(1.0-G)*F
A1 = A + X*(1.0-Y)
A2 = X*Y
A3 = Y *(1.0-Y)
A4 = Y * Y
DO 200 K = 1,21
IF (TIME.EQ.0.0) .AND. (1.0.EQ.1.0) GO TO 200
PKT = PARK + INCRE * FLOAT(K-11)
IF (PKT.LT.0.0) GO TO 200
IF (PKT.GT.1.0) GO TO 200
C = PKT
A = (1.0-G)*(C*(1.0-G) + G*(1.0-C)) +
B = (1.0-C)*(G*X*Y + G*(1.0-C)*(1.0-G) - G*(1.0-C)) +
C = (1.0-C)*(1.0-G) + 1.0
D = (1.0-G)*(G*X*Y + (1.0-G) - (1.0-C)*(1.0-G) -
E = (1.0-C)*(G*X*Y + (1.0-G) + (1.0-C)*(1.0-G) -
F = (1.0-C)*(G*X*Y + (1.0-G) + (1.0-C)*(1.0-G) -
G = (1.0-C)*(G*X*Y + (1.0-G) + (1.0-C) -
H = (1.0-C)*(G*X*Y + (1.0-G) + (1.0-C) -
I = (1.0-C)*(G*X*Y + (1.0-G) + (1.0-C) -
J = (1.0-C)*(G*X*Y + (1.0-G) + (1.0-C) -
K = (1.0-C)*(G*X*Y + (1.0-G) + (1.0-C) -
L = (1.0-C)*(G*X*Y + (1.0-G) + (1.0-C) -
M = (1.0-C)*(G*X*Y + (1.0-G) + (1.0-C) -
N = (1.0-C)*(G*X*Y + (1.0-G) + (1.0-C) -
O = (1.0-C)*(G*X*Y + (1.0-G) + (1.0-C) -
P = (1.0-C)*(G*X*Y + (1.0-G) + (1.0-C) -
Q = (1.0-C)*(G*X*Y + (1.0-G) + (1.0-C) -
R = (1.0-C)*(G*X*Y + (1.0-G) + (1.0-C) -
S = (1.0-C)*(G*X*Y + (1.0-G) + (1.0-C) -
T = (1.0-C)*(G*X*Y + (1.0-G) + (1.0-C) -
U = (1.0-C)*(G*X*Y + (1.0-G) + (1.0-C) -
V = (1.0-C)*(G*X*Y + (1.0-G) + (1.0-C) -
W = (1.0-C)*(G*X*Y + (1.0-G) + (1.0-C) -
X = (1.0-C)*(G*X*Y + (1.0-G) + (1.0-C) -
Y = (1.0-C)*(G*X*Y + (1.0-G) + (1.0-C) -
Z = (1.0-C)*(G*X*Y + (1.0-G) + (1.0-C) -
CALL SUBCSQ(0,T,F,CHISO)
IF (CHISO.GE.CHISO) GO TO 200
PARIT = PIT
PARJ = PJT
PARK = PKT
CHISO = CHISO
200 CONTINUE
PARI = PARIT
PARJ = PARJ
PARK = PARK
300 INCRE =INCRE/5.0
RETURN
END
DO 300 TIME = 1.0
CHISO = 10.0E+30
DO 200 I = 1, 11
IF (((TIME.EQ.0.0).AND.(I.NE.11)) .GO TO 200
GP = PARGP + INCRE * FLOAT (I-1)
IF ((GP.LT.0.0).OR.(GP.GT.1.0)) .GO TO 200
DO 200 J = 1, 11
IF (((TIME.EQ.0.0).AND.(J.NE.11)) .GO TO 200
A = PAPA + INCRE * FLOAT (J-1)
IF ((A.LT.0.0).OR.(A.GT.1.0)) .GO TO 200
DO 200 K = 1, 11
IF (((TIME.EQ.0.0).AND.(K.NE.11)) .GO TO 200
0 = PARB + INCRE * FLOAT (K-1)
IF ((B.LT.0.0).OR.(B.GT.1.0)) .GO TO 200
IF ((A.NE.0.0).AND.(B.NE.0.0)) SN = A * (1.0 - A)**(N-1)
DO 300 K = 1, 11
IF ((N.EQ.0.0).AND.(N.NE.1)) SN = 0.0
TN = (1.0 - FN - SN)
A1 = 1.0 - A
B1 = 1.0 - B
FFFF = FN*A1*A1*A1
FFFS = FN*A1*A1*A
FFSS = FN*A1*A1*B1
FFST = FN*A1*A*B1
FSTT = FN*A*B1*B1
SSSS = SN*B1*B1*B1
SSST = SN*B1*B1*B
SSTT = SN*B1*B
TTTT = TN
do 150 L0 = 1, 2
IF (L0.EQ.0.0) .GO TO 110
FN0 = 0.0
SN0 = 0.0
TN0 = 1.0
GO TO 115
110 FN0 = 1.0 - FN
SN0 = 1.0 - GP
TN0 = 0.0
115 DO 150 L0 = 1, 2
IF (L0.EQ.0.0) .GO TO 120
FN1 = 0.0
SN1 = 1.0 - GP
TN1 = 1.0
GO TO 125
120 FN1 = 1.0 - FN
SN1 = 1.0 - GP
TN1 = 0.0
125 DO 150 L2 = 1, 2
IF (L2.EQ.0.0) .GO TO 130
FN2 = 0.0
SN2 = GP
TN2 = 1.0
GO TO 135
130 FN2 = 1.0 - FN
SN2 = 1.0 - GP
TN2 = 0.0

135 DO 150 L3 = 1, 2
    IF (L3 .EQ. 2) GO TO 140
    FN3 = G
    SN3 = GP
    TN3 = 1.0
    GO TO 145
140  FN3 = 1.0 - G
    SN3 = 1.0 - GP
    TN3 = 0.0
145  CONTINUE

M = (L0 - 1) * 8 + (L1-1)*4 + (L2-1)*2 + L3

E(M) =
1 + FFFF*FN0*FN1*FN2*FN3
2 + FFFS*FN0*FN1*FN2*SN3
3 + FSSS*FN0*FN1*SN2*SN3
4 + FSSS*FN0*SN1*SN2*SN3
5 + FSSS*SN0*SN1*SN2*SN3

IF (TN3 .EQ. 0.0) GO TO 150
E(M) = E(M)
1 + FFFF*FN0*FN1*SN2
2 + FSSS*FN0*SN1*SN2
3 + FSSS*FN0*SN1*SN2
4 + FSSS*FN0*SN1
5 + FSSS*FN0*SN1

IF (TN2 .EQ. 0.0) GO TO 150
E(M) = E(M)
1 + FFFF*FN0*SN1
2 + FSSS*FN0*SN1
3 + FSSS*SN0*SN1
4 + FSSS*SN0*SN1
5 + FSSS*SN0*SN1

150  CONTINUE
    CALL SUBCSQ (0, T, F, CHISOT)
    IF (CHISOT .GE. CHISO) GO TO 200
    PARGPT = GP
    PARA = A
    PARB = B
    CHISO = CHISOT
200  CONTINUE
    PARG = PARGPT
    PARA = PARA
    PARB = PARB

300  INCRE = INCRE .+ E,M
RETURN
END
APPENDIX C

FORTRAN IV PROGRAM USED TO ESTIMATE THE PARAMETERS IN THE AS MODEL

AUTHOR: EDWARD O. MARKS, UNIVERSITY OF MONTANA.

THIS PROGRAM ESTIMATES PARAMETER VALUES FOR ATKINSON, BRESFORD AND SHIFFRIN (J. MATH. PSY., 1967) MODEL OF MEMORY.

BY MINIMIZING CHI-SQUARE, THE ALGORITHM IS THAT USED BY J.R. ULLRICH, UNIVERSITY OF MONTANA, MISSOULA, MT.

CONDITIONS SET ON PARAMETERS: ALPHA = 1, I.E., THE PROBABILITY THAT A ITEM WILL ENTER THE BUFFER IS ONE; R IS THE BUFFER SIZE AND IS AN INTEGER BETWEEN 1 AND 10; THETA IS THE LEARNING RATE AND IS A REAL NUMBER BETWEEN 0 AND 1; TAU IS THE FORGETTING RATE AND IS A REAL NUMBER BETWEEN 0 AND ONE.

NCAT = NUMBER OF LAG CATEGORIES
NALT = NUMBER OF ALTERNATIVES
O = OBSERVED FREQUENCY, LAG = LAG FOR THAT OBSERVATION
E = EXPECTED FREQUENCY
TOTAL = TOTAL FREQUENCY IN OBSERVATIONS

DIMENSION IDEN(10), NAME(10), O(16), E(16)
REAL O, INCREJ, INCREK
INTEGER R, PIT, LAG(16), RT
CALL IFILE (1, 'APS')
READ (1, 1001) IPRN, NSUB, NCAT, NALT
IF (NSUB .EQ. 1) GO TO 13 021 NAME
00 10 L = 1, NCAT
READ (1, 1003) O(L), LAG(L)
10 CONTINUE
TOTAL = 24.0
R = 1
THETA = 5
TAU = 5
INCREJ = .45
INCREK = .05
C NPLACE = NO. OF PLACES
DO 500 NPLACE = 1, 4
CHISO = 10.0E+30
DO 250 J = 1, 21
IF (.NOT. (NPLACE .EQ. 0) .AND. (J .NE. 11)) GO TO 250
PJ T = THETA * INCREJ * FLOAT(J - 11)
IF (PJ T .LT. 0) GO TO 250
00 250 K = 1 - 2
IF (.NOT. (NPLACE .EQ. 0) .AND. (K .NE. 11)) GO TO 250
PKT = TAU + INCREK * FLOAT(K - 11)
IF (PK T .LT. 0) GO TO 250
00 250 L = 1, NCAT
E(L) = E(L) + 1.0
IF (NPLACE .EQ. 3) PIT = 1.0
IF (NPLACE .EQ. 4) PIT = RT
DO 200 L = 1, NCAT
E(L) = E(L) + SUM (PIT, LAG(L)) + SUM (PIT, PJ T, PK T, LAG(L), NALT)
200 CONTINUE
CALL SUBCSQ (O, TOTAL, E, CHISO, NCAT)
IF (CHISO .LE. CHISQ) GO TO 250
PT = PIT
THETAT = PIT
TAUT = PK T
CHISQ = CHISO
250 CONTINUE
R = RT
- 52 -
THETA=THETA
TAU=TAU
INCREJ=INCREJ/5.0
INCREK=INCREK/5.0

500 CONTINUE
WRITE(3,1004) IDENT, NSUB, NCAT, NALT
WRITE(3,1005) (LAG(LL), LL=1, NCAT)
WRITE(3,1006) (O(LL), LL=1, NCAT)
WRITE(3,1009) (E(LL), LL=1, NCAT)
WRITE(3,1006) NAME
WRITE(3,1007) CHISO, THETA, TAU, R

1001 FORMAT(10A1,I2,I2,I2,I1)
1002 FORMAT(/,1X, 'IDENTIFICATION: ', //)
1003 FORMAT(/,1X, 'NUMBER OF SUBJECTS: ', //)
1004 FORMAT(1X, 'NAME OF SUBJECT: ', //)
1005 FORMAT(/,1X, 'NAME OF SUBJECT: ', //)
1006 FORMAT(1X, 'OBSERVATIONS: ', //)
1007 FORMAT(1X, 'NAME OF SUBJECT: ', //)
1008 FORMAT(/,1X, 'EXPECTED PROBABILITIES: ', //)
END

* **********************************************************************
* SUBROUTINE SUBCSO (O,TOTAL,E,CSO,NCAT)
* DIMENSION O(16),E(16)
* REAL O,INCPFJ,TNCRFK
CSO=0
DO 10 I=1,NCAT
10 CSO=CSO+(TOTAL*E(I)**2)/(TOTAL*E(I))
RETURN
END

* **********************************************************************
* FUNCTION SUM1(NBUFF,NLAG)
C FUNCTION CALCULATES THE PROBABILITY THE ITEM WAS
C IN THE BUFFER AT TIME OF TEST. NOTE THAT IF THE ITEM
C IS IN THE BUFFER AT THE TIME OF THE TEST THE PROBABILITY OF A
C CORRECT RESPONSE IS ONE.
C A NEW ITEM ALWAYS ENTERS THE BUFFER AND AN OLD ITEM
C IS RANDOMLY REMOVED.
SUM1=0
RNBUFF=NBUFF
DO 10 II=1,NLAG
10 SUM1=SUM1+((RNBUFF-1)/RNBUFF)**II*(1.0/RNBUFF)
RETURN
END

* **********************************************************************
* FUNCTION SUM2(NBUFF,THETAN,TAUN,NLAG,NALT)
C FUNCTION CALCULATES THE PROBABILITY OF A CORRECT
C RESPONSE FOR A GIVEN LAG AND MULTIPLIES IT BY THE PROBABILITY
C THAT THE GIVEN LAG HAPPENED.
RNBUFF=NBUFF
G = 1.0 / FLOAT(NALT).
SUM2 = 0
DO 10 II = 1, NLAG
    SUM2 = SUM2 + (((RNUFF - 1) / RNUFF)**II) * (1.0 / RNUFF) * (1.0 - (1.0 - G) * EXP(-FLOAT(IT) * THETAN) * (TAUN** (NLAG - IT)))
10 CONTINUE
RETURN
END
<table>
<thead>
<tr>
<th>Word Pair</th>
<th>Correct Word</th>
</tr>
</thead>
<tbody>
<tr>
<td>Judge Metal</td>
<td>Judge</td>
</tr>
<tr>
<td>Nymph Yacht</td>
<td>Nymph</td>
</tr>
<tr>
<td>Ankle Woods</td>
<td>Woods</td>
</tr>
<tr>
<td>Plain Truck</td>
<td>Plain</td>
</tr>
<tr>
<td>Queen Slave</td>
<td>Queen</td>
</tr>
<tr>
<td>Shoes Candy</td>
<td>Candy</td>
</tr>
<tr>
<td>Sauce River</td>
<td>Sauce</td>
</tr>
<tr>
<td>Salad Elbow</td>
<td>Salad</td>
</tr>
<tr>
<td>Hotel Bosom</td>
<td>Bosom</td>
</tr>
<tr>
<td>Horse Paper</td>
<td>Paper</td>
</tr>
<tr>
<td>Dress Ocean</td>
<td>Dress</td>
</tr>
<tr>
<td>Shore Geese</td>
<td>Geese</td>
</tr>
<tr>
<td>Peach Snake</td>
<td>Peach</td>
</tr>
<tr>
<td>Tower Thief</td>
<td>Tower</td>
</tr>
<tr>
<td>House Coast</td>
<td>House</td>
</tr>
<tr>
<td>Storm Cabin</td>
<td>Cabin</td>
</tr>
<tr>
<td>Lemon Mucus</td>
<td>Mucus</td>
</tr>
<tr>
<td>Plant Stain</td>
<td>Plant</td>
</tr>
<tr>
<td>Grass Flask</td>
<td>Grass</td>
</tr>
<tr>
<td>Whale Cigar</td>
<td>Cigar</td>
</tr>
<tr>
<td>Swamp Death</td>
<td>Death</td>
</tr>
<tr>
<td>Brain Wench</td>
<td>Wench</td>
</tr>
<tr>
<td>Beast Array</td>
<td>Array</td>
</tr>
<tr>
<td>Arrow Pupil</td>
<td>Arrow</td>
</tr>
<tr>
<td>Woman Angle</td>
<td>Angle</td>
</tr>
<tr>
<td>Hairs Skull</td>
<td>Skull</td>
</tr>
</tbody>
</table>