INDIVIDUAL DIFFERENCES IN ASSOCIATION VALUE SCALES AND PAIRED-ASSOCIATE LEARNING

by

Yurdal Topsever
Department of Psychology
University of Stirling

Ph.D. Thesis, September 1974

ACKNOWLEDGEMENTS

I should like to take this opportunity to thank Dr. P. Hamilton for his supervision over the past three years. I should also like to thank Mr. M.F. Moore for his continuous guidence and encouragement. Acknowledgement is also due to Mr. C. Dracup for reading this manuscript and pointing out unwitting grammatical errors. I should like to extend my gratitude to Dr. R.H. Ranyard, Dr. S. Ambler, Dr. R.R. MacDonald and to many other past and present staff and post-graduates who made my five years' stay in Stirling a pleasure. I would also like to mention the endurance of my wife and son who were to some extent "abandoned" during this piece of work.

Y. Topsever

CONTENTS

ABSTRACT

An attempt was made to investigate the theoretical and methodological problems involved in the study of individual differences in verbal learning. Particular emphasis was placed upon association value scales and paired-associate learning. An association value scale of the test items used was obtained for each subject collecting data on their behaviour in a number of similar learning tasks. The scales so obtained were compared with association value scales which were obtained by subjects' rating the test items on a five-point scale. The comparisons between individual and group results indicated that a group association value scale could not sufficiently control the item homogeneity for individual subjects. However, the rated association value scales obtained for individual subjects were highly correlated with their own association value scales obtained in actual learning situations, indicating that subjects' own ratings of the items can be used to yield more accurate predictions of their performance in verbal learning tasks.

In studying individual differences in paired-associate learning the emphasis was placed upon various mathematical models with different theoretical assumptions. Thus, individual differences variables were investigated in terms of parameter values of paired-associate learning models. For each subject repeated measurements were taken by running them on a number of parallel tasks which were designed according to their own association value scales. A set of parameter values was estimated for each subject seperately. Changes in the values of the parameters representing various theoretical constructs of the models used were investigated with regard to changes in subjects' overall performance. Results indicated that when fitting models to data from individual subjects, one particular model may best fit the behaviour of one group of subjects whereas the behaviour of other subjects may be better represented by other models.

GENERAL INTRODUCTION

In a recent review which covers the period from approximately 350 B.C. to 1969 A.D., with a strong emphasis on the literature published between 1967 and 1969, Tulving and Madigan (1970) depict a gloomy picture of progress made in the field of "memory and verbal learning". One of the criteria chosen by Tulving and Madigan in evaluating the literature in this field is that the function of experiments is to allow the construction, elaboration, modification, and overthrow of theories. Given such a criterion, the reasons behind the slow progress in the field of memory and verbal learning seem to be more deep-rooted than appears at first sight.

The hostility to theorizing which started at the turn of the century as a reaction against the dubious reputation of introspection and philosophical speculation still seems to dominate the conceptual framework of psychology. In other words, early American functionalism and Watsonian behaviorism, in the footsteps of Comte, urged the restriction of psychology to observed facts. As a result, "the blind gathering and intercorrelation of data has led to selective emphasis upon regions relatively accessible to observation, thus sacrificing relevence to convenience, and has opened the door to tacit presuppositions of a more theoretical kind as well" (Brunswik, 1969, p.48).

In fact, the history of psychology is full of such "sacrifices". Few other sciences in a given time period develop so many new ideas and techniques which fall into disuse, or publish so many experimental papers which are consigned to oblivion from the outset.

In many areas of psychology, an outstanding example of "sacrificing relevence to convenience" is the neglect of the study of individual differences. In verbal learning, for instance, although the importance of individual differences is recognized, there has always been a dearth of research evidence, only a few works appear in the literature (Cf.

Noble, 1961; Gagne', 1967; Jenkins, 1967; Pavio, 1970). Furthermore, as Travers (1967) puts it, most of the work on individual differences has had the outcome of discovering and rediscovering the nature of the difficulties involved in research of this kind.

The aim of the present study is to investigate the theoretical and methodological problems involved in the study of individual differences in verbal learning. The problem has been treated on two different levels. At the first level, individual difference in verbal learning is viewed as a variable which simply expresses itself as variance, hence the problem is considered as a methodological one (e.g. Underwood, mentioned in Jenkins, 1967). In all research where the inferences are based on group results, this kind of approach tacitly assumes that each individual has the same psychological mechanism for the phenomenon under study. Then the differences between subjects are assumed to be quantitative differences in their performance, not qualitative differences in the underlying psychological mechanism.

At a second level, postulating different psychological mechanisms for each individual or for groups of individuals can be viewed as a theoretical problem. It was not long ago that a deep concern about the inferences based on group curves brought up the issue of representation of individual performance in group results (e.g., Sidman, 1952; Bakan, 1954; Estes, 1956). The issue is simply a mathematical one. It is possible to obtain a group function (curve) quite unlike those describing the individual observations themselves. It is also possible to obtain different functional forms for each individual which are in no way similar to grpup results (Cf. Audley and Jonckheere, 1956). If there is no alternative explanation, then the possibility of obtaining different forms of functions for different individuals might simply mean that each individual has a different psychological mechanism for the same phenomenon. In this case, the differences between subjects are assumed to represent qualitative differences in the underlying psychological mechanism, not just quantitative differences in their performance.

Given the assumption that each individual has the same function

(Case I), problems arising from the group function being different from the individual observations themselves can easily be handled within the general rules of statistical inference: depending upon given assumptions various hypotheses can be formulated about the form of the group function and tested against the obtained data.

The second case where different individuals are represented by different functional forms does not immediately give a unique testable hypothesis, since the assumption of obtaining different functions for different individuals means that there is an infinite number of possible hypotheses to be formulated. This, of course, represents an extreme case. Alternatively, some intermediate solution such as obtaining a small number of groups each of whose members show the same functional form would be much easier to handle.

Before mentioning the implications of these type of approaches, a brief review of individual differences in verbal learning will be presented. References to the above arguments will be given as they appear in this review.

Individual Differences in Verbal Learning.

In the early days when the definition of intelligence as "ability to learn" had many supporters among psychologists, individual differences in learning were considered as variations in IQ scores. This generally accepted view was first questioned by Woodrow(1946). Woodrow's findings were that data from laboratory and classroom experiments contradict the assumption that the ability to learn, in the sense of ability to improve with practice, is identical with intelligence. Correlations between intelligence and gain were generally insignificantly positive and often close to zero.

Woodrow interpreted his results by assuming that a score at any stage of practice consists of a general factor, G, and specific factors. He further pointed out that specific factors change with practice. As a result, there can be high correlation between the general factor and scores at all stages of practice, but it is also possible for the correlation between G and gain to be negligible when gain is the result of a high degree of specificity; this specificity results from task characteristics and individual differences in performing these tasks.

The line of work generated by Woodrow is reflected today in the psychometric, correlation-oriented studies as a result of Gulliksen's active interest in the problem. This is exemplified by the work of his students, such as Stake (1961) and Duncanson (1964). In general, this type of study was designed to investigate individual differences in certain learning tasks with reference to various aptitude and achievement tests. Data for performance on each task were intercorrelated and factor analysed.

As Glaser (1967) points out, a major inadequacy of the factor analytic-psychometric approach is the lack of theoretical framework for the selection of reference tests and learning measures. Another concern about the psychometric bias is the preoccupation with ways of getting around error variance rather than investigating the conditions which influence it.

A major approach initiated by Hull (1945) was concerned with the effects of individual differences on learning functions. As is known, he adopted the point of view of the natural sciences, of physics in particular. A scientific law is expressed in terms of an equation of a particular form, and the empirical constants in the equation are determined by experimental conditions so that they vary with individual situations but do not change the general form of the law. Hull's notion was that individual differences find expression in these empirical constants. This approach formed the basis of a number of illustrative studies. The findings of these studies of psychomotor performance (Reynolds and Adams, 1954; Zeaman and Kaufman, 1955) and verbal learning (Noble, Noble, and Alcock, 1958; Carroll and Burke, 1965), and many others generally support Hull's notion that individual differences affect the constans of a functional relationship rather than its form.

All of these studies were conducted with subjects placed into subgroups on the basis of either their initial performance on the same task or their performance on some other relevant task. The classification was usually in terms of slow versus fast learners. Although the individual cases in these sub-groups have generally been neglected, the assesment of subjects in terms of their initial state measurements has yielded some interesting research findings. In general, the idea of initial baselines requires an intensive screening and classification of subjects prior to experimentation as is done in experimental genetics. And it is hoped that with increased attention to initial baselines, our experimental methodology can change.

When the employment of "screening tests" is viewed as a mere convenience to reduce the variance of subjects' performance in an experiment such methodological refinements contribute little to our understanding of either individual differences or learning. As Jenkins (1967) suggests, although for statistical purposes it is possible to screen subjects on a "work sample" and classify them appropriately, it would be misleading to consider this approach as the primary goal of research on individual differences in verbal learning.

Another point of concern about individual differences has been the problem of inference from curves based on group data. A review of the literature on this subject starts with Merrell's (1931) observations on the relationship of individual growth to average growth. Later, papers by Sidman (1952), Hayes (1953), Bakan (1954) have raised serious questions on this topic, and Estes (1956) gave an excellant discussion of the problems introduced by applying models of individual behaviour to group averages. According to Estes, although the obtained group curve remains one of our most useful devices both for summarizing information and for theoretical analysis, inferences about the form of individual curves require caution. In fact, a valid treatment of averaged curves depends upon the familiar procedures of statistical inference. Therefore, any "inductive" inference from average curves to individual ones becomes impossible. But given any specified assumption about the form of individual curves, the characteristics of an averaged curve can be deduced and the predictions can be tested against obtained data.

In testing quantitative theories against averaged data the main points of concern are:

a) The form of the functional relationship,

b) The parameter values for the population of organisms sampled.

Case (a) covers all the studies operating on the tacit assumption that the form of an averaged curve will reflect the form of the individual curves. Although this assumption seems to be unwarranted, the psychological literature is full of studies which try to determine "the form of the learning curve". Case (b) is usually illustrated by attempts to determine the functional relation between the experimental treatments and the parameter values of a given learning curve. With regard to these considerations, Estes classifies functions into three types each of which requires different treatment:

1) Functions unmodified by averaging. In this case the mean curve for the group is the same as individual curves and the parameters of the group curve are simply the means of the corresponding parameter values for individual curves.

2) Functions for which averaging complicates the interpretation of parameters but leaves the functional form unchanged. In this case testing hypotheses related to the form of the function raises no difficulties, but testing hypotheses involving changes in parameter values as a function of experimental treatments requires caution because of averaging effects.

3) Functions modified in form by averaging. In some cases a function belonging to this class can be moved into class (2) or even class (1) by means of an appropriate transformation. Then, tests appropriate to the form of the function can be conducted.

Estes's comments are not meant to provide an exhaustive treatment of the problem of averaging. He tries to point out that the valid interpretation of group curves depends on the principles common to all problems of statistical inference. Therefore, the form of a group curve does not determine the forms of the individual curves, but it does provide a means of testing exact hypotheses about them. In this case, the procedure suggested by Estes is to state the hypothesis under test for the individual curve and then to derive the properties that should hold for the averaged curve if the hypothesis is correct.

With the development of stochastic learning models, individual differences have been considered as the differences in parameter values of a model type. In most applications of these learning models, it is usually assumed that the same values of parameters characterize all the subjects in an experimental group. This homogeneity assumption is not usually postulated on any theoretical grounds, but it enables the pooling of data from different subjects and the estimation of parameters from group results. But, as Sternberg (1963) points out, when this tacit assumption of individual homogeneity is made in the application of a model, that what is tested by comparisons between data and model is the conjunction of the assumption and the model type and not the model type alone. "It is usually thought that if the assumption is not entirely justified then the discrepancy will cause the model to underestimate the intersubject variances of response-sequence statistics. It is hoped (but not known) that the discrepancy will have no adverse effects"(Sternberg, 1963).

However, some attempts have been made to estimate the parameters of a certain learning model for each subject seperately (e.g., Audley and Jonckheere, 1956). Audley (1957) has also developed this procedure for a model which describes the learning behaviour of an individual subject in a two-choice situation. Audley and Jonckheere argue that "the curve fitted to the average of the individual observations at each trial may imply a mathematical function quite unlike those describing the individual observations themselves even when the latter functions are in form the same for all subjects, and differ only in the individual parameters required".

Some other suggestions for estimating individual parameters, in particular for testing the homogeneity assumption of parameter values come from Bush and Wilson (1956), Anderson (1959), Bush, Galanter, and Luce (1959), Sternberg (1959).

Research related to the variations in learning-rate parameters has drawn little attention. An early example appears in Bush and Mosteller's work (1959) where they consider a Hullian model with individual differences and a learning parameter α which has a [redacted] specific distribution. In mental test theory, logistic test models proposed by Birnbaum (1969) operate on the assumption of a logistic distribution for the ability parameter θ. Gregg and Simon's (1967) concept identification model assumes a uniform distribution on the conditioning parameter in a certain range. Their conclusion is that almost all the "fine-grain" statistics reflect mainly a random component. Hence the statistics are insensitive to individual differences, or, for that matter to any other psychological aspects of the subjects' behaviour that might be expected to affect the statistics. Contrary to this view, Offir (1972) tries to show that "fine-grain" statistics are affected greatly when a correct distribution is used to describe individual differences. By assuming a beta density function for the parameters of the single-operator linear model and the long-short model, Offir studies the effect of this heterogeneity modification on the predictions from these models. The results show that under the heterogeneity assumption the descriptive and predictive power of the models increases measurably.

Implications

As the foregoing review reveals methods adopted in the study of individual differences are as follows:

1) Factor analytic methods,
2) The Hullian approach,
3) Subjects' initial baseline studies,
4) Inferences from group curves to individual curves,
5) Testing of assumptions about the distribution of individual parameter values in stochastic learning models.

For any kind of learning process, given the assumption that each subject has the same functional form, the choice of one of these methods might be sufficient to answer some restricted set of questions related to individual differences. The general finding of these types of studies, however, does not seem to go beyond rephrasing the quite well known fact that there are individual differences in learning. What should be done next, then? The answer to this question depends on what aspects of individual differences we are interested. For instance, if we were able to seperate the components of a certain learning process by postulating different stages in learning, then the role of individual differences in these stages or the transitions between stages might become an area of interest. Similary, in factor analytic studies, investigation of individual differences in postulated factors might reveal some interesting research findings. In fact, some evidence from factor analytic studies suggests that two postulated factors, span memory and rote memory factors, play different roles in different stages of learning (Kelley, 1954). Further study by Games (1962) on this subject emphasised the need for span memory and rote memory individual difference parameters.

From the foregoing arguments it becomes clear that a systematic investigation of individual differences requires theoretical constructs for the learning process with which we want to work. The difficulty of working on individual differences firstly stems from the fact that there is a dearth of psychological theory and general laws from which appropriate deductions can be made for individual subjects. Secondly, the area of individual differences in verbal learning, in particular, poses its own methodological problems. Item homogeneity, for instance, as a confounding variable constitutes a major methodological problem, since it is difficult to determine whether the observed variance is due to the differences between subjects or the differences between the test items. In fact, the problem of item homogeneity has drawn more attention than the problem of subject homogeneity (e.g., Glaze, 1928; Hull, 1933; Noble, Stockwell, and Pryer, 1957; Noble, 1961; Battig and Spera, 1962). But, the scales of association value or meaningfulness of test items have up till now been based on group results. The relationship between item homogeneity and an individual difference parameter has never been investigated.

When taken together, the problem of subject homogeneity on the one hand and that of item homogeneity on the other, restrict our choice of a verbal learning task to test materials for which rated association or meaningfulness scales are readily obtainable. The choice of such test materials will prove to be useful when we come to compare individual meaningfulness scales with those obtained from group data.

In order to study a given learning process from the individual differences point of view, our first desideratum for related theoretical constructs seems to be very difficult to meet. However, the recent developments in mathematical learning theories have generated a theoretical framework within which appropriate deductions can be made for individual subjects. The possibility of testing assumptions about subject and item homogeneity and the further possibility of deriving testable hypotheses for individual subjects, emphasis was placed upon mathematical learning models.

When one talks about mathematical learning models probably the first thing which comes to mind is the amount of research work which went into the paired-associate paradigm. In fact, from the very beginning, not only in mathematical learning theory, but also in non-mathematical psychology paired-associate learning has drawn enormous attention. The amount of research work pertaining to this field especially in the last decade is almost prohibitive of any sort of concise review.

For the purpose of studying individual differences, the choice of mathematical models for paired-associate learning brings more problems of its own. With the choice of a mathematical learning model, the study of individual differences becomes the investigation of individual parameter values within a particular model type. Because of the theoretical nature of a particular model and its theoretical parameter values, there are problems in the estimation and identification of parameters.

The plan of this dissertation is as follows. In the first part of the study, the association value or meaningfulness scale of the paired associate test material will be derived and presented. The obtained scale will be compared with similar scales in the literature and various comparisons will be made between individual results and group results. This first part is concluded with a presentation of individual and group serial position curves obtained under controlled presentation and testing orders. Striking differences in functional form of individual and group serial position curves and the differences between the individual results themselves would be a demonstration of various misleading generalizations which may occur when inferences are based on group results.

In the second part, by utilizing the obtained association value scale paired-associate tasks will be construsted. Data for paired-associate learning will be obtained by running the same subjects several times on the same type of tasks. The effect of such experimental variables as list length, practice, and guessing will also be investigated. Parameter values of group and individual data will be estimated seperately and the necessary comparisons made. The paired-associate models which will be used are as follows: The one-element model, the single-operator linear model, Norman's (1964) random trial increments model, Atkinson and Crother's (1964) long-short model (LS-3 and LS-2 versions). A modified version of a general four-stage Markovian model will be developed. Furthermore, problems arising from identification of parameters (in the sense of Greeno and Steiner, 1964) will be investigated.

Each seperate part has its own introduction and discussion section. A general conclusion and discussion section which covers the whole project appears at the end of the dissertation.

PART I:

INDIVIDUAL DIFFERENCES IN ASSOCIATION VALUE SCALES

INDIVIDUAL DIFFERENCES IN ASSOCIATION VALUE SCALES

Introduction

Item homogeneity is considered to be one of the prime requirements of controlled experimentation on verbal behaviour. In the literature, the problem of homogeneity of items has usually been studied with regard to the association value or meaningfulness of the test materials. Because of other desirable properties of test materials such as simplicity, numerousness, etc., consonant-vowel-consonant (CVC) trigrams have been used extensively in verbal learning studies. Therefore, much of the work on association values or meaningfulness scales have been conducted using CVC trigrams as test materials.

In the following review association value and meaningfulness will be used interchangeably at the definitional level. Following the example of Underwood and Schulz (1960) various scaling operations will be referred to by the names of the investigators who generated them.

Glaze (1928): Glaze's work represents the first systematic attempt to order CVC trigrams along a meaningfulness scale. The list which was presented to the subjects consisted of 2019 CVC trigrams. A total of 15 subjects was used. After a short practice trial with 15 trigrams, the subjects were presented with the 2019 trigrams at the rate of 252 per session. The trigrams were presented one at a time by use of a tachistoscope for about 2-3 sec. As each trigram was presented the experimenter spelt it. The subjects task was to indicate in a few words what the CVC trigram meant to him. The association value or meaningfulness of each trigram was calculated as the persentage of subjects who gave an association to the trigram.

Hull (1933): Hull used 320 CVC trigrams which were divided into 20 lists of 16 each. Twenty subjects worked with all 20 lists in a rotational scheme calculated to spread the effects of practice equally over all items. Hull's procedure corresponds to simple serial learning, since he wanted to obtain a measure of association value in an actual learning situation. Each list was presented three times at a rate of one trigram every 2 second. After the first trial subject was instructed to anticipate the trigrams and to indicate of what the trigrams made him think. In other words, the subject was not to try think of associations to a trigram, but if he did think of any, he was to report them. The association value scale was based on the number of associations reported for each trigram.

Krueger (1934): A total of 2183 CVC trigrams was used and 586 subjects, each rating only 1200 trigrams, took part in the experiment. The experimenter spelt each trigram twice taking on average about 4 sec. for both spellings. The subject's task was to write the trigram as the experimenter spelt it and indicate the idea (with a word or phrase) arouse by that trigram. The meaningfulness scale was based on the frequency of the associations for each trigram.

Witmer (1935): Witmer worked with trigrams which consisted of three consecutive consonants (CCC). A total of 4534 items was used. Each trigram was presented for 4 sec. on a memory drum. The subject's task was to spell the trigram and then state in a word or phrase what it meant to him. A special record of unverbalized associations was also obtained by instructing subjects to say merely "yes" if the trigram meant something but he could not verbalized its meaning within the 4 sec. interval. The meaningfulness scale was based on the reports of 25 subjects.

Noble (1952): In determining the meaningfulness scale of items the method employed by Noble differed from those used by previous investigators. With the so-called production method the subject is presented with an item and given 60 sec. to write all the different words elicited by the item. By this method Noble tried to determine an index of meaningfulness for dyssyllables. A total of 119 subjects was used. The meaningfulness scale was based on the mean number of responses given to each word during a 60 sec. period.

Mandler (1955): Mandler used the production method for CVC trigrams. Each printed trigram was presented in the middle of a sheet of paper ahd the subject was instructed to write his associations in 30 sec. A total of 100 trigrams was used. The scale of meaningfulness was based on the mean number of responses obtained from 34 subjects.

Noble, Stockwell, and Pryer (1957): In this study a new method for defining meaningfulness was introduced. The subjects rated CVC trigrams for the number of different things or ideas they thought each trigram suggested. A 5-point scale was used. The subject had to indicate the association value of each trigram by putting a check mark in one of the five spaces provided. A total of 100 CVC trigrams was rated by 200 subjects. The mean rating of each item was used as a basis for the meaningfulness scale.

Noble (1961): The method introduced in the Noble, et al (1957) study was applied all 2100 possible combinations of English alphabet. A total of 200 subjects was used. Different treatments of the same data yielded the following measures: association value (a), rated associations (a'), and scaled meaningfulness (m'). Since then all the research on meaningfulness and verbal learning have utilized these scales as standards.

<u>Implications</u>

In the foregoing review of scaling procedures, three clearly distinct methods of assessing meaningfulness have been used:

1) Methods which determine whether or not a CVC trigram arouses an association within a limited time interval.
2) Determination of the number of different associations actually aroused by a CVC trigram.
3) Subject's rating the number of different associations which he thinks a CVC trigram elicits.

Excepting Hull's method, all the other methods can easily be included in one of the categories given above. Hull's method differs from the others only in that it obtained an association value scale in an actual learning situation. Nevertheless, Hull's method could be included into the first category, as the final assessment of meaningfulness was based on the number of associations reported for each CVC trigram. The investigation of correlations among meaningfulness values for different scaling methods reveals that the Hull scale does not relate to other scales as highly as the other scales relate to one another (Cf. Noble, Stockwell, and Pryer, 1957). It appears that since Hull little consideration has been given to the idea of obtaining a measure of association value in an actual learning situation. On the other hand, the importance of meaningfulness scales stems from the fact that there is a high relationship between actual speed of learning and subjects' ratings of how long they think it will take them to learn the items. Furthermore, the rating given a verbal unit as to how fast it will be learned can be predicted almost perfectly from a knowledge of the number of associations that item elicits (Underwood and Schulz, 1960). Investigations have shown also that the ease of learning depends on the meaningfulness of the test items used (e.g., Underwood, Rehula, and Keppel, 1962;

Cohen and Musgrave, 1964; Carroll and Burke, 1965).

The aim of this chapter is to detail a method for determining an association value scale in an actual learning situation. As indicated earlier, Hull's association value scale was based on the number of associations an item elicits in an actual learning situation. Contrary to Hull's procedure, our association value scale will be based on the number of correct responses given to an item. In the light of the evidence given above, a high relationship between actual speed of learning and the association value of an item can be taken as empirical support for our contention that association value scales based on the number of correct responses will be measuring the same behavioral reactions as those measured by other scaling procedures.

In the present study two-digit numbers have been used instead of CVC trigrams. The reason for the choice of two-digit numbers as test materials are practical ones. The intention is to obtain association value scales for each individual as well as for the group data. Therefore, given the 2100 possible CVC combinations of the alphabet and the required number of replications for each item, the time taken to complete the experiment would be excessive. On the other hand, working with a small set of two- digit numbers, besides helping to control such experimental variables as replication, serial position, etc., takes a relatively shorter time.

The use of numbers as materials in verbal learning experiments is not uncommon. For example, employment of single or two-digit numbers in serial learning experiments, CVC trigrams-number pairs in paired associate experiments. Rated association values of all numbers from 0-100 have already been obtained by Battig and Spera (1962). Their procedure corresponds closely to those used previously to obtain similar information for CVC trigrams (e.g., Noble, 1961). A five-point scale was used. The subject was instructed to indicate his choice by rating each number with regard to "how many different things or ideas are associated with the number, and how difficult it is to think of these associations". The rated association value scale was based on the mean number of ratings for each number obtained from 95 subjects. For future reference Battig and Spera's rated association values of numbers from 0-100 are given in the Appendix I.

In the following experiment numbers ending with zeros (20, 30, ...) and numbers with the same two digits (22, 33, ...) are excluded. In the Battig and Spera scale these numbers usually show high associations. The present study is concerned with a task that is near to paired-associate learning: following the presentation of a list of two-digit numbers, for each randomly chosen item in the list the subject is cued with the first digit and asked to respond with the second digit which was associated with it. In a learning situation like this, association value scales for individual subjects were obtained by taking their average number of correct responses for each item.

EXPERIMENT I

Method

The stimulus material were two-digit numbers between 21 and 98 excluding the numbers ending with zeros and the numbers with the same two digits. Each list contained eight numbers which were chosen from different number categories (one from the twenties, one from the thirties, etc.). The second digits were chosen without replication. A typical sequence might be:

62
58
91
47
35
89
26
74

One-hundred and sixty tasks were constructed in this fashion.

The subjects were 24 university students. Eight of them were paid to take part in all five sessions in each of which they completed 32 tasks. The remaining 16 subjects were volunteers and only took part in the first session, each of them completed the first 32 tasks.

The apparatus used was the Wang 700B programmable electronic calculating machine with two number displays. According to a programmed schedule, 32 tasks in the first session were presented in a rotational scheme to spread the serial position effects equally over all items. This prearranged rotational scheme ensured that each item had equal oportunity of appearing in all the serial positions for different subjects. The data thus obtained were used in determining the group association value scale. For the five-session subjects the rotational presentation scheme was abondened after the first session, since over the remaining four sessions each item had equal opportunity of appearing in all possible serial positions.

The items were presented at a rate of one per second. After the presentation of all the items the single digits representing the different number categories were given according to a pre-arranged schedule. The subjects were instructed to respond with the digit which was associated with that given number category. A 30 sec. interval was given between the tasks. During each task correct and incorrect responses were recorded.

After the experiment, the subjects who took part in all the five sessions were requested to rate each number according to their association value. In other words, Battig and Spera's (1962) procedure was used to obtain rated association values of the numbers and compare them with the association value scales obtained in an actual learning situation.

Results and Discussion

In order to discover the relationship between the rated association values (obtained on a five-point scale) and the performance in an actual learning situation, Battig and Spera (1962) compared their rated association value scale with performance in paired-associate learning experiments where nonsense shapes were used as stimuli and two-digit numbers as responses (Battig, 1962; Battig and Brackett, 1961). In these studies, performance on two-digit numbers was measured by the proportion of total presentations summed over subjects on which correct responses were given. The obtained correlation coefficients were .516

and .489 for the two studies, respectively.

In the present study the association values of two-digit numbers were measured according to the following formula:

$$\text{Association value (AV)} = \frac{C_i}{\dfrac{C \quad P_i}{P}}$$

where

C_i: total number of correct responses on item i,
P_i: total number of presentation of item i,
C : total number of correct responses on all the items,
P : total number of presentations of all the items.

This formula proved to be useful in the determination of association value scales for each individual subject. In Battig and Spera's expression the ratio of total correct responses to the total presentations reflects not only the different association values of all the items, but also the different performance levels of subjects. The above formula determines the association value scales relative to the mean performance level of each subject. In effect, both formulae give the same orderings of the association values, one set of values being a linear function of the other.

Table 1 presents the association value scale obtained for group results. A direct check on the relationship of the present scale values to the Battig and Spera's rated associations gave a correlation of .487. This correlation coefficient appears to be consistent with Battig and Spera's correlation coefficients (.516 and .489) which were obtained in the same way. Battig and Spera's study also includes numbers in the teens (12, 13, ...) whereas in the present study these numbers are excluded on the

No.	AV	No.	AV	No.	AV	No.	AV
59	.59	73	.86	29	.99	87	1.19
61	.69	46	.88	84	1.00	51	1.19
65	.70	47	.89	92	1.00	42	1.19
58	.72	31	.89	76	1.00	54	1.21
85	.73	57	.90	75	1.03	95	1.24
93	.76	26	.91	74	1.05	43	1.24
83	.76	78	.92	63	1.05	41	1.24
68	.76	67	.93	56	1.05	36	1.24
62	.76	72	.94	35	1.05	27	1.24
86	.78	94	.95	97	1.07	96	1.29
79	.80	49	.95	71	1.08	23	1.34
89	.81	34	.95	69	1.10	21	1.34
38	.82	32	.95	64	1.11	25	1.35
53	.83	28	.97	81	1.11	48	1.38
39	.84	62	.98	24	1.15	98	1.43
37	.84	82	.99	45	1.17	91	1.54

Table 1: Association value scale for group results.

No.	AV	No.	AV	No.	AV	No.	AV
79	.54	46	.83	35	1.02	51	1.19
65	.63	96	.84	73	1.02	76	1.20
56	.66	68	.84	83	1.05	41	1.20
78	.67	94	.85	47	1.05	34	1.20
61	.70	23	.87	29	1.05	54	1.21
93	.70	85	.91	24	1.05	37	1.25
92	.70	62	.91	31	1.07	21	1.25
81	.70	48	.91	98	1.09	57	1.27
64	.73	84	.91	52	1.09	27	1.31
82	.74	28	.93	97	1.11	95	1.32
71	.75	69	.96	89	1.11	91	1.33
74	.76	72	.97	36	1.13	75	1.34
45	.79	43	.99	58	1.14	42	1.40
53	.80	26	1.00	25	1.15	39	1.40
63	.81	49	1.01	59	1.16	38	1.40
86	.83	87	1.02	67	1.18	32	1.40

Table 2: Association value scale for Subject 1.

No.	AV	No.	AV	No.	AV	No.	AV
58	.45	72	.74	85	1.00	35	1.21
86	.47	53	.76	75	1.01	27	1.21
65	.48	64	.77	37	1.02	46	1.26
76	.52	61	.78	41	1.04	71	1.28
84	.53	31	.81	78	1.05	96	1.29
79	.54	49	.81	68	1.09	29	1.33
32	.54	47	.81	45	1.10	59	1.34
43	.57	34	.87	94	1.13	39	1.36
89	.64	54	.88	74	1.13	81	1.39
73	.65	83	.91	95	1.14	91	1.40
87	.66	28	.92	56	1.14	67	1.40
63	.69	52	.94	26	1.15	97	1.53
36	.69	82	.95	23	1.15	21	1.79
93	.71	92	.97	98	1.21	69	1.82
62	.71	51	.97	57	1.21	24	1.82
48	.73	38	.97	42	1.21	25	1.87

Table 3: Association value scale for Subject 2.

No.	AV	No.	AV	No.	AV	No.	AV
86	.59	57	.86	39	.96	28	1.22
68	.60	35	.86	34	.98	64	1.24
32	.63	93	.86	47	1.00	73	1.25
83	.64	81	.86	62	1.01	96	1.25
84	.67	76	.86	25	1.01	97	1.26
45	.70	52	.86	74	1.03	21	1.26
75	.71	42	.86	78	1.04	94	1.28
63	.73	61	.89	26	1.06	29	1.28
53	.75	67	.90	58	1.07	36	1.30
49	.76	85	.91	54	1.09	95	1.31
56	.81	71	.91	41	1.10	37	1.35
43	.81	65	.94	92	1.11	51	1.37
91	.81	38	.94	31	1.14	48	1.37
89	.81	59	.94	87	1.17	24	1.37
72	.82	79	.95	27	1.18	23	1.39
69	.86	46	.95	82	1.19	98	1.43

Table 4: Association value scale for Subject 3.

No.	AV	No.	AV	No.	AV	No.	AV
61	.63	84	.86	81	1.04	46	1.13
69	.66	32	.88	97	1.04	94	1.13
68	.66	53	.91	93	1.05	79	1.17
82	.69	78	.92	96	1.06	48	1.19
67	.70	51	.92	36	1.07	24	1.19
37	.70	38	.92	26	1.07	28	1.20
59	.73	41	.94	83	1.07	23	1.20
58	.74	57	.96	64	1.08	87	1.20
39	.74	91	.97	54	1.08	35	1.20
31	.75	29	.99	45	1.08	72	1.21
74	.79	27	.99	25	1.08	95	1.24
65	.79	63	1.01	56	1.08	62	1.24
47	.83	85	1.01	75	1.10	98	1.25
76	.85	43	1.01	42	1.10	89	1.25
34	.85	73	1.02	21	1.11	71	1.32
92	.86	52	1.03	86	1.12	49	1.32

Table 5: Association value scale for Subject 4.

No.	AV	No.	AV	No.	AV	No.	AV
52	.17	94	.64	74	1.00	71	1.41
49	.17	84	.65	23	1.00	36	1.42
42	.17	72	.65	59	1.03	65	1.50
62	.18	32	.66	86	1.04	51	1.50
53	.19	64	.68	29	1.05	61	1.56
25	.27	85	.70	43	1.06	63	1.57
73	.35	82	.75	76	1.07	31	1.57
83	.37	48	.75	57	1.09	89	1.57
69	.37	38	.75	97	1.10	47	1.62
82	.39	93	.75	39	1.12	26	1.71
67	.47	78	.78	27	1.12	54	1.77
87	.54	96	.80	28	1.14	21	1.89
35	.54	89	.83	56	1.23	41	1.92
58	.56	79	.83	34	1.28	45	2.04
68	.60	81	.85	75	1.37	24	2.24
37	.63	46	.89	95	1.41	91	2.68

Table 6: Association value scale for Subject 5.

No.	AV	No.	AV	No.	AV	No.	AV
69	.22	63	.75	57	1.03	65	1.31
49	.39	53	.76	96	1.04	26	1.33
72	.45	51	.78	94	1.06	23	1.33
59	.54	47	.80	28	1.08	85	1.33
58	.54	71	.82	83	1.09	56	1.33
79	.58	78	.83	39	1.09	87	1.35
73	.60	32	.84	27	1.09	35	1.35
38	.61	93	.87	64	1.11	25	1.35
67	.64	75	.87	54	1.11	84	1.36
46	.64	34	.87	29	1.13	91	1.37
52	.68	82	.91	42	1.16	36	1.41
92	.70	89	.92	24	1.22	48	1.48
74	.70	37	.92	41	1.24	21	1.56
68	.70	62	.92	61	1.25	76	1.62
86	.74	43	.92	98	1.26	45	1.66
81	.75	31	1.00	97	1.28	95	1.74

Table 7: Association value scale for Subject 6.

No.	AV	No.	AV	No.	AV	No.	AV
47	.00	27	.61	81	1.04	34	1.39
57	.11	89	.64	28	1.04	31	1.39
37	.13	86	.66	82	1.06	21	1.41
79	.14	65	.73	53	1.06	56	1.43
58	.15	93	.81	87	1.11	42	1.49
39	.15	74	.81	75	1.12	32	1.53
59	.17	36	.81	96	1.14	54	1.55
63	.23	84	.85	48	1.22	94	1.56
97	.26	61	.88	26	1.27	45	1.66
49	.27	64	.88	67	1.28	91	1.66
69	.30	46	.90	62	1.29	51	1.83
78	.32	52	.95	92	1.34	71	1.86
73	.47	68	.97	24	1.34	23	2.09
29	.49	35	1.00	98	1.35	25	2.10
38	.61	85	1.00	72	1.38	95	2.15
83	.61	43	1.00	76	1.39	41	2.26

Table 8: Association value scale for Subject 7.

No.	AV	No.	AV	No.	AV	No.	AV
75	.45	37	.82	68	1.01	47	1.17
59	.59	43	.82	24	1.01	27	1.17
52	.61	79	.87	54	1.06	85	1.19
51	.62	49	.87	73	1.08	42	1.21
56	.64	58	.88	82	1.08	76	1.22
89	.66	53	.88	78	1.08	93	1.23
61	.69	84	.88	48	1.09	36	1.26
86	.72	81	.89	32	1.10	45	1.27
62	.73	39	.97	94	1.11	98	1.30
97	.74	91	.98	28	1.11	65	1.32
67	.74	87	.99	26	1.11	63	1.33
83	.78	41	1.00	64	1.13	23	1.33
29	.78	34	1.00	57	1.13	96	1.35
25	.78	95	1.01	35	1.13	74	1.35
46	.81	71	1.01	72	1.15	38	1.40
31	.82	92	1.01	69	1.17	21	1.56

Table 9: Association value scale for Subject 8.

basis of the findings of a pilot study. The pilot study revealed that these numbers generally show high association values. If these numbers were included in the analysis, the obtained correlation coefficient would presumably be higher, since the same numbers show relatively high rated association values in Battig and Spera's scale. Furthermore, vocalization of the numbers in the teens per se do not show the same pattern as other two-digit numbers. For example, the vocalization of two-digit numbers from other categories start with a digit representing the number category whereas this order is reversed for numbers in the teens (e.g., twenty-seven, thirty-seven, etc., vs. seventeen). In a paired-associate task this characteristic means that S-R pairings are reversed, the response coming before the stimulus. For these reasons, the two-digit numbers in the teens have been excluded from the present study.

A reliability score for the group association value scale was obtained by randomly splitting the subjects into two equal groups. The obtained correlation coefficient between these two groups was .923.

Tables 2 to 9 present the association value scales for the individual subjects. Each table has been based on the performance of individual subjects who completed 160 tasks in an extended experiment. Table 10 presents the correlations between the group association value scale and the association value scales for the individual subjects. The correlations between the Battig and Spera rated association value scale and the association value scales for individual subjects are given in Table 11.

An inspection of Table 11 reveals that the correlations between the association value scales obtained in an actual learning situation and the Battig and Spera rated association

	Subject							
	1	2	3	4	5	6	7	8
r	.16	.34	.35	.37	.20	.49	.55	.26

Table 10: Correlations between the group association value scale and the association value scales for individual subjects.

	Subject							
	1	2	3	4	5	6	7	8
r	.14	.38	.17	.31	.12	.39	.23	.29

Table 11: Correlations between the Battig and Spera rated association value scale and the association value scales for individual subjects obtained in an actual learning situation.

	Subject							
	1	2	3	4	5	6	7	8
r	-.06	.55	.54	.41	.18	.54	.48	.39

Table 12: Correlations between the Battig and Spera rated association value scale and the ratings of two-digit numbers by individual subjects.

	Subject							
	1	2	3	4	5	6	7	8
r	.44	.41	.23	.54	.52	.40	.46	.58

Table 13: Correlations between the ratings of two-digit numbers by individual subjects and their association value scales obtained in an actual learning situation.

value scale are quite low and in some cases insignificant. In Table 10 it can be seen that association value scales obtained in an actual learning situation both for group data and the data from individual subjects are highly correlated when compered with the correlations given in Table 11. In Table 10 only subjects 1 and 5 gave insignificant correlations whereas in Table 11 there are four cases (Subjects 1, 3, 5, and 7) which gave insignificant correlations.

Table 12 presents the correlations between the Battig and Spera rated association value scale and the ratings of two-digit numbers by individual subjects. For each subject the ratings of two-digit numbers have been obtained by Battig and Spera's rating procedure which utilizes a five-point scale. Table 13 gives the correlations between the ratings of two-digit numbers by individual subjects and their association values obtained in an actual learning situation.

Individual differences of condirable magnitude are indicated by the varying degrees of correlation in Table 10 and 12 where under two different scaling procedures the individual results have been correlated with the group results. In Table 11, the relatively low and in some cases insignificant correlations show that the predictions from the rated association scale values to the performance on such materials will not be very accurate. Overall, the obtained results suggest that rated association value scales based on group data are far from satisfactory in controlling item homogeneity for individual subjects. On the other hand, the relatively high correlations in Table 13 indicate that subjects' own ratings of the items can be used to yield more accurate predictions of their performance in verbal learning tasks.

Serial Position Curves

Each subject serial position curve was obtained by cueing the recall of specific items according to a prearranged schedule. Since each subject had completed 160 tasks, each serial position was based on 20 independent trials. A group curve was obtained by averaging performance over subjects for each of the eight serial positions. Figure 1 presents the serial position curve for the group results. The familiar shape of a group serial position curve with differences in recall at the ends and middle of a list is demonstrated clearly. In verbal learning, studies on serial position curves are considered to be important in explaning the processes underlying acquisition of a given list. Despite the introduction of new terminology such as recency and primacy effects, as Tulving and Madigan put it, our understanding of the phenomenon has advanced little since Ebbinghaus first described it.

In the literature, no work appears to have been done on individual differences in serial position effects. Although the similarity of one serial position curve to another is no guarantee that both are consequences of one and the same set of underlying processes, this has always been the tacit assumption in the study of group serial position curves. Even if obtained group results on serial position effects contradict the traditional "bow shaped curve", existing theoretical notions usually remain untouched, instead other reasons are given why the theory should emerge unscathed from contact with the data. The present study simply illustrates the variations of serial position curves between individual subjects.

Figures 2 to 9 give the serial position curves of eight different subjects. Table 14 summarizes the results of the analysis of variance which was carried out according to a mixed

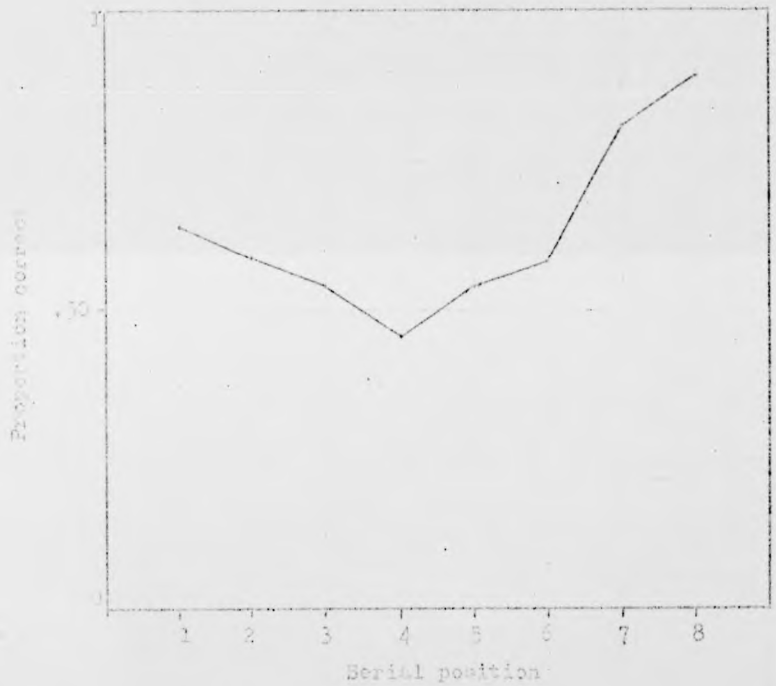

Fig.1: Serial position curve for group results

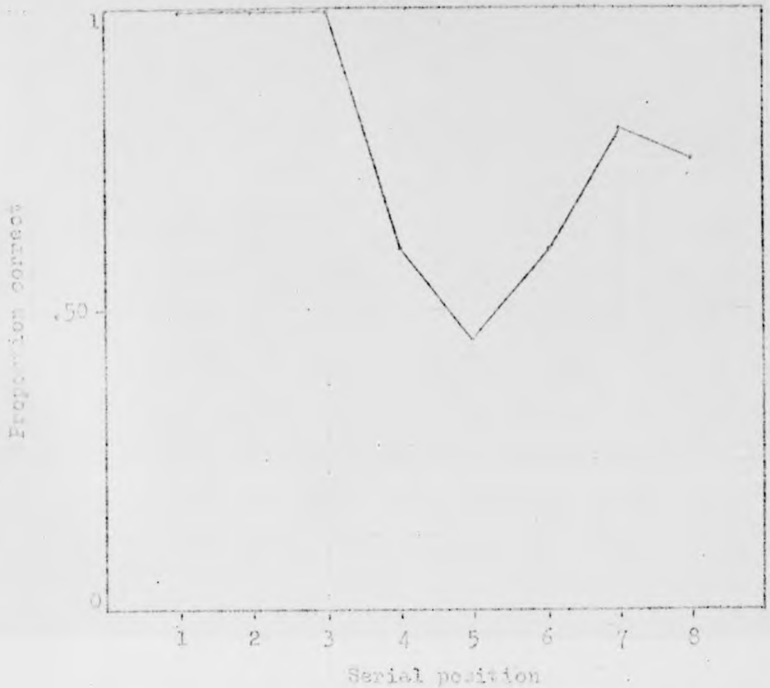

Fig.2: Serial position curve for Subject 1

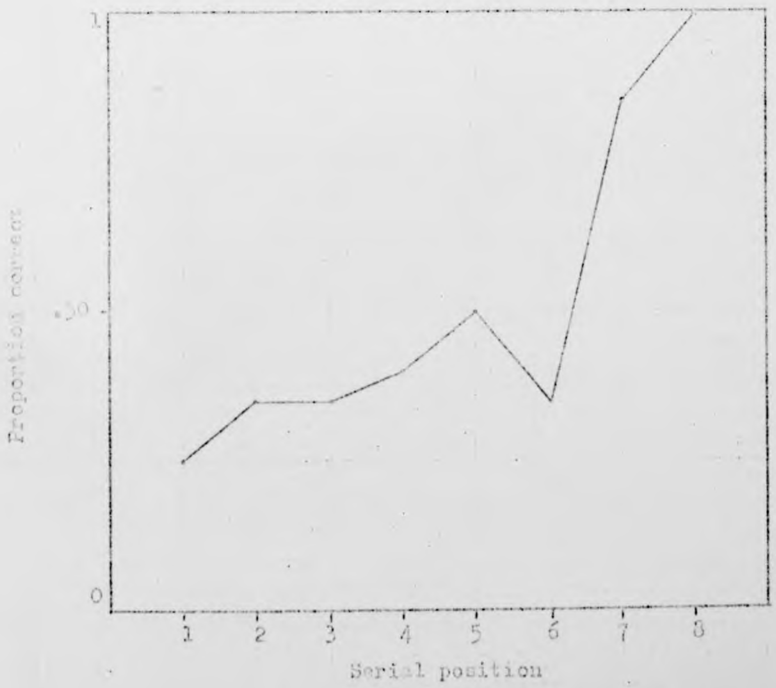

Fig.3: Serial position curve for Subject 2

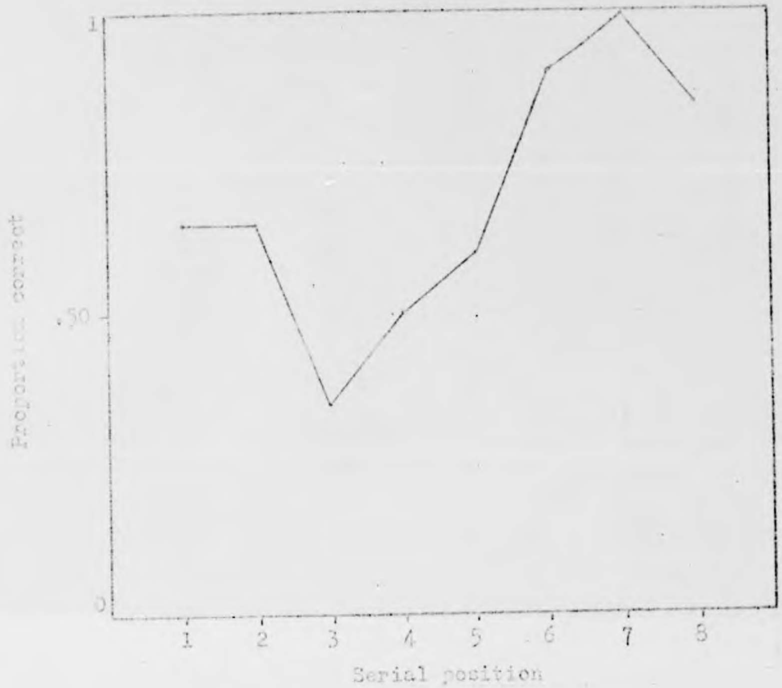

Fig.4: Serial position curve for Subject 3

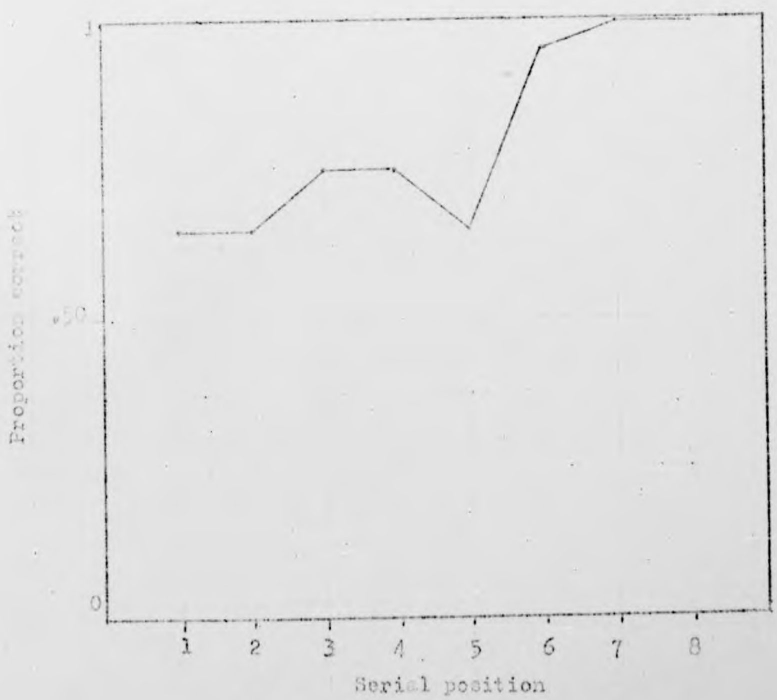

Fig.5: Serial position curve for Subject 4

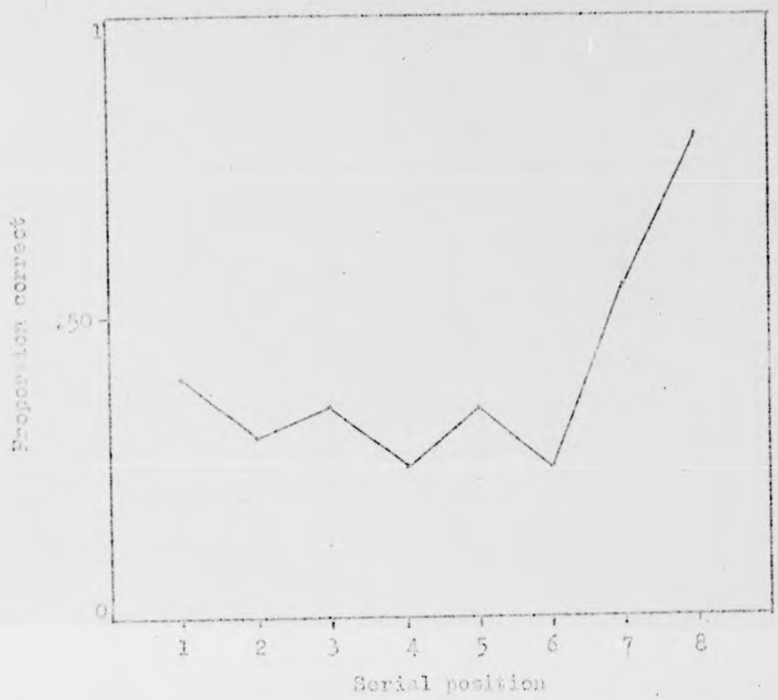

Fig.6: Serial position curve for Subject 5

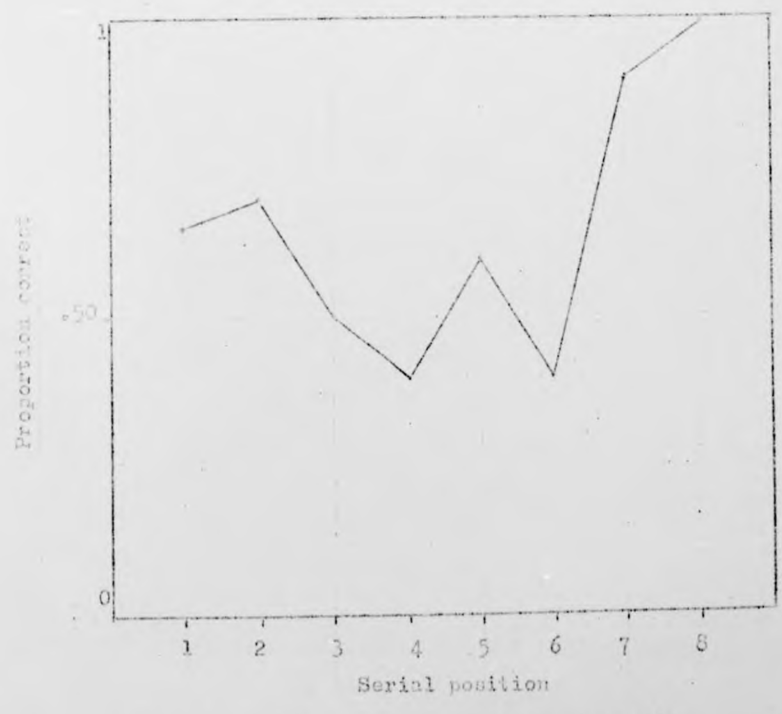

Fig.7: Serial position curve for Subject 6

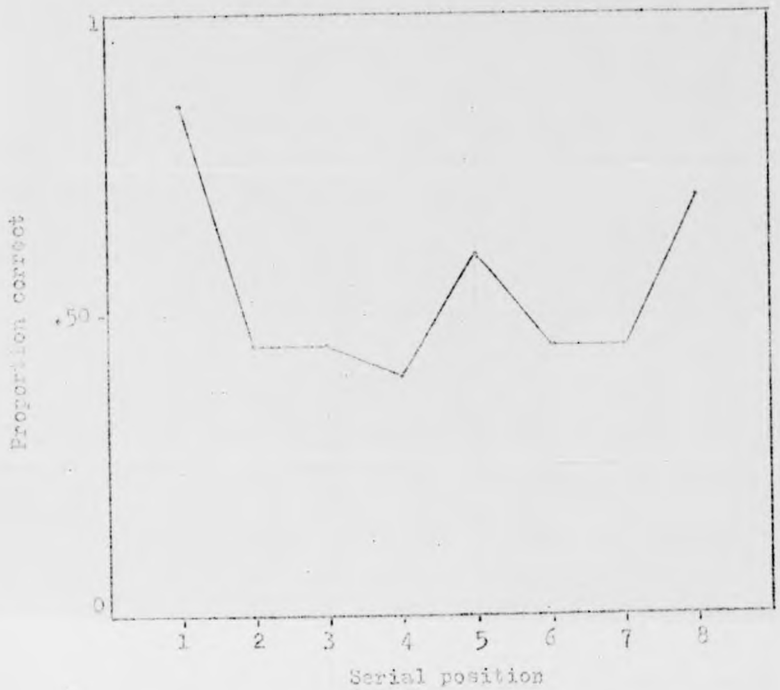

Fig.8: Serial position curve for Subject 7

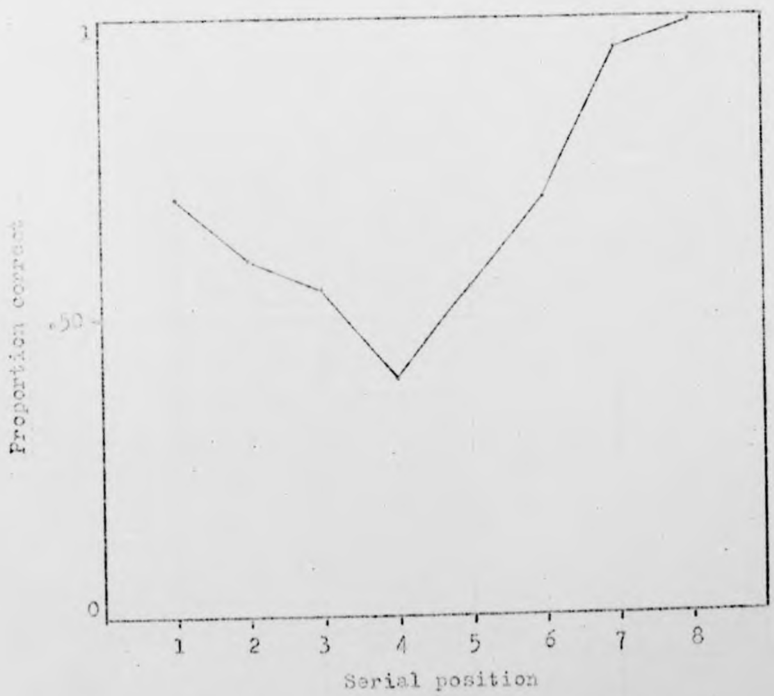

Fig.9: Serial position curve for Subject 8

Source	SS	df	MS	F	p
Rows (Subjects)	206.18	7	29.45	14.56	$p < .001$
Columns (Conditions)	239.68	7	34.24	7.06	$p < .001$
Interaction (Conditions by Subj.)	237.76	49	4.85	2.40	$p < .001$
Error	129.50	64	2.02		
Total	813.12	127			

Table 14: Summary of analysis of variance results.

model design. In the analysis, for each serial position 20 independent trials were randomly divided into two equal groups and the resulting pairs were used as replications under each different conditions. Analysis of variance results give sufficient evidence to permit us to conclude that there are both serial position and subjects effects. Furthermore, the presence of interaction effects indicate that there is something about the combination of a particular subject with a particular serial position that accounts for a significant amount of variance in the data. Thus, for each subject, differences in serial positions apparently exist, but these tend to be different for different subjects.

Figures 2 to 9 simply illustrate the differences in serial position curves for different subjects. Apart from random fluctuations, five out of the eight subjects seemed to give serial position curves of the same shape as the traditional group serial position curve. The serial position curves for Subjects 2, 4, and 5 clearly differ from the obtained group results. Although for all subjects there are strong recency effects, the primacy parts of the curves seem to be distorted by the individual differences.

Some introspective reports obtained from subjects indicate that the different strategies used in the process of acquisition of the list may cause the observed differences in their serial position curves. It is apparent that future research on serial position curves which includes some provision for the possibility of different strategies used by the subjects would be more benefecial than gross generalizations based on group results.

PART II:

INDIVIDUAL DIFFERENCES IN PAIRED-ASSOCIATE LEARNING

INDIVIDUAL DIFFERENCES IN PAIRED-ASSOCIATE LEARNING

Introduction

At the close of a conference on learning and individual differences (University of Pittsburgh, 1965), Melton's (1967) conclusion was that research on individual differences must be guided by theories of human learning and performance; in other words, according to Melton "what is necessary is that we frame our hypothesis about individual differences variables in terms of process constructs of contemporary theories of learning and performance".

In the field of verbal learning the S-R association theories have been concerned primarly with changes in the frequency of certain response classes as a function of experimental conditions. However, the development of a new psychology of cognition has influenced the study of verbal learning very substantially. In particular, cognitive processes such as attention, storage, and retrieval have become major areas in psychological research. Besides new experimental methods and techniques, several formal developments have provided superior tools for theoretical work. The developments in mathematical learning theory, for instance, aided the generation of new concepts for psychological processes and structures.

Since research on individual differences should emphasize the process constructs of contemporary theories of learning, the present study is an attempt to investigate the individual differences in terms of mathematical paired-associate learning models where an abundance of theoretical concepts can be found.

As Melton (1967) puts it, although the approach which emphasizes process variables is no panacea, it is a way to increase the likelihood of significant advances in the understanding of individual differences. Melton further argues that our interest in finding and manupulating individual differences in the hypothesized process will refine our analysis of the process and contribute to a "taxonomy of processes". Another point is that if there are observable individual differences in performance that can be traced directly to individual differences in a process, then the theory which identifies this process gains greatly in predictive power and acceptability; if, on the other hand, the process does not vary between individuals, or if it varies without significant correlated performance effects, there is probably something wrong with the process construct.

In the present study, individual differences in theoretical constructs such as long-term and short-term storage, acquisition, conditioning, and forgetting processes will be investigated in terms of particular model types. These paired-associate learning models and their major theoretical assumptions are given below:

The one-element model:

The one-element model, sometimes referred to as the all-or-none model, represents a special case of the more general models of Stimulus Sampling Theory. A review of the literature shows that a wide array of data on paired-associate learning has frequently been analysed in terms of an all-or-none process (e.g., Bower, 1961, 1962; Crothers, 1962; Estes, 1960, 1961; Suppes and Ginsberg, 1963).

If paired-associate learning is viewed as the learning of an association between the stimulus and response member of a given pair, then the main question is how do the subjects learn to anticipate the response member of a given pair when the stimulus member is presented alone? As its name implies, the all-or-none model assumes that a single reinforcement of a previously unlearned pair produces either complete learning of the association, or no learning whatsoever. The two principle assumptions of the model are as follows:

1) Until the stimulus element is conditioned, there is a constant probability g that the subject will respond correctly by guessing.
2) On each trial there is a probability c that the stimulus element will become conditioned to the correct response.

Thus, on trial n of an experiment the stimulus element can be regarded as being in one of two conditioning states: in state C or in state $\bar{C}$. If the stimulus element is in state C, then the element is conditioned to the correct response. If the stimulus element is in state $\bar{C}$, then the element is unconditioned and a correct response by the subject can only occur by guessing. At the beginning of the experiment, the element is in the unconditioned state $\bar{C}$ and subsequently moves to state C as specified by the transition matrix:

$$\begin{array}{ccc} & \begin{array}{cc} C(n+1) & \bar{C}(n+1) \end{array} & \text{Pr(Correct)} \\ \begin{array}{c} C(n) \\ \bar{C}(n) \end{array} & \begin{bmatrix} 1 & 0 \\ c & 1-c \end{bmatrix} & \begin{bmatrix} 1 \\ g \end{bmatrix} \end{array} \qquad (1)$$

In general, the results reported in the literature indicate a remarkably close correspondence between observed values and those predicted by the one-element model. However, there are various aspects of the data which contradict the assumptions of the model. Although the issue of all-or none learning of associations raises a number of important questions, the general consensus is, as Restle (1965) puts it, that a list entailing highly available responses and extremely strong interpair associations is learned in the all-or-none fashion.

The single-operator linear model:

As an alternative to the one-element model, the linear model views learning as a direct change in response probability from one trial to the next. Thus, the response probability on trial n+1 is obtained by a transformation of the response probability observed on trial n. The simplest version of such a model was developed by Bush and Mosteller (1955), and Bush and Sternberg (1959). This model assumes that the probability of correct response increases according to the following equation:

$$p(n+1) = (1 - \theta)\, p(n) + \theta \qquad (2)$$

where $p(1)=1/r$, and r is the number of available response alternatives.

In general, the single operator-linear model has compared unfavourably with the one-element model. The result of such comparisons revealed that the obtained goodness-of-fit to the data of the linear model is not as good that of the one-element model (Cf. Atkinson, Bower, and Crothers, 1965; Atkinson and Crothers, 1964).

The random-trial increments (RTI) model:

The random-trial increment model developed by Norman (1964) represents a combination of the one-element and linear models. The model contains two parameters and with a suitable choice of values for these the one-element model, the single-operator linear model, or a quasi-mixture of the two can be obtained. In the random-trial increments model, the probability of a correct response on trial n+1 is given by the following equation:

$$p(n+1) = \begin{cases} (1 - \theta)\, p(n) + \theta\,, & \text{with probability } c \\ p(n)\,, & \text{with probability } 1-c \end{cases} \tag{3}$$

In Eq. 3, c represents the probability of an unobservable event, called an "effective reinforcement". Thus, on each trial only one of two events can occur: with probability (1-c) no effective reinforcement occurs and no learning takes place; with probability c an effective reinforcement occurs and the response probability receives an increment described by the linear transformation given in Eq. 2. In Eq. 3, if $\theta = 1$ the process reduces to the one-element model, whereas if $c = 1$ the simple linear model is obtained.

Since much of the data on paired-associate learning seems to fall in between the one-element model and the linear model, the random-trial increments model often gives a good fit to the obtained data. Norman (1964) and Atkinson and Crothers (1964) found that the random-trial increments model did a good job of predicting data which contraindicated the one-element and linear models.

<u>The long-short model:</u>

Although the one-element model gives a good fit to the data obtained from paired-associate tasks entailing highly available responses and extremely strong interpair associations, it soon became clear that this model was insufficient for a fairly large number of learning situations. Furthermore, there is at least one contradictory aspect of the data which goes directly to the core assumption of the one-element model; i.e., appropriate statistical analyses usually reveal a nonstationary effect before the last error. In other words, there is a tendency for the probability of a correct response to increase over trials before the last error and not simply remain a constant g, as predicted by the theory. These considerations led theorists to construct more general models which account for the observed discrepancies but which can be reduced to the one-element model by a suitable choice of parameter values. The long-short model which was developed by Atkinson and Crothers (1964) is an example of this type of more general model.

The long-short model assumes four stages of learning:

1) State U: Learning is postulated to consist of the encoding of the stimulus followed by the association of the encoded stimulus with the correct response. Before encoding has occured, the stimulus is said to be in state U (uncoded); in this state the subject is assumed to respond by guessing randomly among the r response alternatives.

2) State F: This state represents forgetting of a temporary connection between the encoded stimulus and the response. If the stimulus element passes into state F, the subject guesses randomly, but the encoding of the stimulus is retained.

3) State S: This state represents a short-term memory state, expressing the notion that a temporary connection between the encoded stimulus and response may form prior to establishing the permanant association; while the association is temporarily stored the correct response occurs with probability 1.

4) State L: Once the permanant association forms the stimulus element is absorbed in state L (long-term memory) and the

subject makes no error on subsequent presentation of the item.

This model is described by its transition matrix, as follows:

$$
\begin{array}{cc}
 & \begin{array}{ccccc} L & S & F & U \end{array} \\
\begin{array}{c} L \\ S \\ F \\ U \end{array} & \begin{bmatrix} 1 & 0 & 0 & 0 \\ a & (1-a)(1-f) & (1-a)f & 0 \\ a & (1-a)(1-f) & (1-a)f & 0 \\ ca & c(1-a)(1-f) & c(1-a)f & 1-c \end{bmatrix}
\end{array}
\quad
\begin{array}{c} Pr(Correct) \\ \begin{bmatrix} 1 \\ 1 \\ g \\ g \end{bmatrix} \end{array}
\qquad (4)
$$

where $g = 1/r$ represents the guessing probability. The probability that encoding occurs on trial n given that it has not occured on previous trial is c. If an encoded item is presented, then with probability a it goes into state L and with probability (1-a) it goes into state s. It is assumed that with each intervening event between one presentation of an item and its next presentation, there is a probability f that the item will move into state F.

Two different versions of the long-short model have been examined. The first version is the general model where all three parameters a, f, and c were taken into account. In the second case where c is set equal to 1, only the parameters a and f were considered. In the literature these two different versions are referred to as LS-3 and LS-2, respectively (the 3 and 2 designate the number of free parameters to be estimated). When c = 1 and f = 1 the model reduces to the simple one-element model.

Implications

Although the parameters of the various models discussed above usually appear under different labels, they can be classified in terms of their correspondance to changes in learning performance. For instance, the conditioning parameter c in the one-element model, θ in the single-operator linear model, the "effective reinforcement" parameter c in the RTI model, and the parameter a in the long-short model can all be classified as "learning rate parameters". The question is how do the observable individual differences in performance relate to individual differences in learning rate parameters of a given model. In one parameter case individual differences in performance are expected to reflect the differences in the learning rate parameter alone. If a model contains more than one parameter, then a certain combination of these parameters might account for the observed individual differences in performance. For this reason the RTI and the long-short models are of special interest, since they contain more than one parameter. Furhermore, the assumed states in the long-short model can be taken as an example of various attempts to incorparate psychological processes into a formal model: in the long-short model, besides the forgetting state, there is a distinction between short-term memory and long-term memory; similarly, Atkinson and Shiffrin (1968), and Bower (1967) make a distiction between short-term memory (sometimes called a buffer) and long-term memory; Waugh and Norman (1965) distinguish these states as primary memory and secondary memory; Peterson (1966) distinguishes between an active trace and a structural trace.

The proposed forgetting state in the long-short model requires further consideration. It can be shown that when the guessing parameter g is taken as a free parameter and estimated from the data there is a one-to-one correspondence between the parameter g in the one-element model and the parameter f in the LS-2 model. In other words, the one-element model with a free parameter g gives exactly the same fit as the LS-2 model. Although there seems to be no convincing interpretation of the parameter g when estimated from the data, this close relationship between the parameters g and f suggests that the parameter f in the long-short model operates in such a way to compansate for the loss which is caused by fixing the parameter g as 1/r, where r is the number of available response alternatives. If this is the case, then the incorporation of the parameter f into the long-short model could be interpreted in more than one way; it may just be the parameter of an "intelligent guesssing" process.

Another drawback of the long-short model is that the transition from state U to state F is difficult to interpret. In both the LS-3 and LS-2 models, if the state U can be considered as an unlearned state, then the transition from state U to state F can only be interpreted as forgetting an item before actually learning it.

A more general criticism of a Markovian model with a row state guessing vector is that in such a model the guessing parameter g plays no role whatsoever during the transitions from one state to another. In the forgetting state, for instance, besides other factors a correct guess might help the item to move to other states.

Keeping all these considerations in mind a modified version of a general four-stage Markovian model has been developed. The theoretical states and assumptions of this model are given below:

The model assumes four stages of learning:

State S1: The model excludes the possibility of a direct transition from the starting state S1 to the terminal state S4.

Learning is assumed to take place by forming an association between the stimulus and the response member of a given pair. It is also assumed that at each presentation an association is formed between the stimulus and the response member of a given pair. Between the presentation of an item and its test trial the association is either forgotten with a probability f, or not forgotten with a probability (1-f). If the association is not forgotten then the subject responds corretly and the item moves into state S3. If the association is forgotten then two things might happen: either the subject guesses correctly and the item moves into state S2 or the subject fails to guess correctly and the item stays in state S1. The transition probabilities from state S1 to state S2 and S3 can be summed, indicating that an item can only leave state S1 by moving into one particular state. In the present model the transition probabilities from state S1 to state S2 and S3 have been summed and assigned to the transition probability from state S1 to state S2 such that an item can only leave state S1 by moving into state S2. This restriction might prove useful if it is assumed that this particular state represents a psychological process such as short-term storage, etc., and each item is assumed to go through this state before reaching the terminal state S4.

State S2: An item can only move into state S2 from state S3. Each transition from state S3 to state S2 occurs with an error. Conversely, a correct response indicates a transition from state S2 to state S3.

State S3: If an item is in state S3, either it moves into the terminal state S4 with a probability (a) or fails to do so with a probability (1-a). Failure to move into state S4 results in either forgetting the association with a probability (f) or not forgetting it with a probability (1-f). Unforgotten items stay in state S3. If an item fails to move into state S4 and also becomes forgotten then two things might be expected to happen: either the subject guesses correctly or fails to do so. Hence, with a probability $(1-a)(1-f)+(1-a)fg$ an item stays in state S3 and the subject responds correctly. After an item has failed to move into state S4, if the association is forgotten, and the subject makes an incorrect guess then the item moves into state S2.

State S4: This state represents a permanant association between the stimulus and the response members of an item pair. Once the item moves into state S4 the subject makes no error on subsequent presentations of the item.

This model is described by its transition matrix, as follows:

$$
\begin{array}{c c}
 & \begin{array}{cccc} S4 & S3 & S2 & S1 \end{array} \\
\begin{array}{c} S4 \\ S3 \\ S2 \\ S1 \end{array} &
\left[\begin{array}{cccc}
1 & 0 & 0 & 0 \\
a & (1-a)(1-f)+(1-a)fg & (1-a)(1-g)f & 0 \\
a & (1-a)(1-f)+(1-a)fg & (1-a)(1-g)f & 0 \\
0 & (1-f)+fg & 0 & (1-g)f
\end{array}\right]
\end{array}
\qquad (5)
$$

where $g = 1/r$ represents the guessing probability, r being the number of available response alternatives. Eq. 5 can be written more compactly as follows:

$$
\begin{array}{c c}
 & \begin{array}{cccc} S4 & S3 & S2 & S1 \end{array} \\
\begin{array}{c} S4 \\ S3 \\ S2 \\ S1 \end{array} &
\left[\begin{array}{cccc}
1 & 0 & 0 & 0 \\
a & (1-a)(1-f(1-g)) & (1-a)(1-g)f & 0 \\
a & (1-a)(1-f(1-g)) & (1-a)(1-g)f & 0 \\
0 & 1-(1-g)f & 0 & (1-g)f
\end{array}\right]
\end{array}
\qquad (5/a)
$$

In Eq. 5 there are two parameters to be estimated from the data. In order to compare the model given in Eq. 5 with the other models described above, the parameter estimation was based on the predicted probabilities of 16 possible response sequences over trials 2 to 5. A detailed discussion of this estimation procedure is given on page 57. A minimum X^2 estimation procedure was carried out searching all possible parameter values until a minimum X^2 is obtained that accurate to three decimal places. Using the predictions of the model given in Eq. 5, the obtained X^2 values for the data from the eight experiments described in Atkinson and Crothers (1964) are given in Table 15. For the sake of comparison, the minimum X^2 values based on the predictions of other models have been reproduced from Atkinson and Crothers (1964, p.300). For brevity, the model described in Eq. 5 will be referred to as the LOS model. The LOS-2 model refers to the general version where the two parameters a and f are estimated. In LOS-3 model, a different forgetting parameter has been assumed in the initial state S1 and the three parameters a, f, and f_1 are estimated. Table 16 gives the parameter estimates for the LOS-2 and the LOS-3 models together with the parameter estimates for the LS-2 and the LS-3 models which are reproduced from Atkinson and Crothers (1964, p.299).

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
Ia	30.30	50.92	9.74*	6.75*	5.67*	8.50*	6.74*
Ib	39.31	95.86	13.09*	19.69*	12.42*	18.69*	18.58*
II	62.13	251.30	29.11	3.73*	3.73*	13.29*	3.23*
III	150.66	296.30	51.12	33.02	33.02	21.30*	17.98*
IV	44.48	146.95	10.66*	12.32*	10.77*	16.52*	16.13*
Va	102.02	201.98	40.17	24.41*	24.41*	24.15*	24.13*
Vc	246.96	236.15	46.43	27.12*	27.12	20.33*	13.42*
Ve	161.03	262.56	84.07	20.12*	20.12*	25.22*	16.74
TotalX^2	836.89	1542.02	281.39	147.16	137.26	148.00	116.95

* Not significant at .01 level.

Table 15: Minimum X^2 values.

Model	Parameter	Experiment							
		Ia	Ib	II	III	IV	Va	Vc	Ve
LS-2	a	.352	.305	.250	.188	.266	.109	.156	.258
	f	.719	.805	.805	.789	.836	.844	.727	.680
LS-3	a	.367	.352	.250	.188	.289	.109	.156	.266
	f	.648	.375	.850	.789	.789	.844	.727	.688
	c	.844	.500	1.000	1.000	.789	1.000	1.000	.992
LOS-2	a	.418	.397	.308	.282	.369	.168	.224	.315
	f	.619	.709	.720	.734	.733	.812	.694	.609
LOS-3	a	.355	.391	.254	.268	.375	.167	.202	.288
	f	.726	.675	.822	.652	.776	.808	.591	.742
	f_1	.001	.723	.051	.768	.715	.814	.750	.162

Table 16: Parameter estimates for the LS and the LOS models.

An inspection of Table 15 reveals that in terms of the chosen significance level the LOS-2 model does a better job than the LS-2 and even the LS-3 models. For each seperate experiment the X^2 values obtained from the LOS-2 model are below the chosen significance level whereas the X^2 values obtained from the LS-2 and the LS-3 models exceed the required significance level especially for the Experiments III and Vc.

Identification of the parameter f in the LOS model is, of course, as difficult as the parameter f in the long-short model. In particular, introducing a new forgetting parameter f_1 in the initial state S1 is questionable. The main argument which motivated the introduction of a new model was that the parameter f in the long-short model was operating in such a way to compansate for the loss which is caused by fixing the parameter g. In the LOS model, by introducing the parameter f over and above a fixed parameter g, it is hoped that the obtained values of the parameter f would reflect the assumed forgetting process better than the corresponding parameter values obtained with the long-short model.

Problems related to the identification of parameters have been deferred to a later section. The aim of this present section is to investigate the changes in individual parameter values as a function of the observable individual differences in learning performance. In the experiments which follow, the average number of trials to the last error will be taken as an index of subject's overall performance. Experimental variables such as different list lengths and guessing situations which are assumed to effect the subject's overall performance will be employed.

EXPERIMENTS II AND III

Method

The stimulus material was the same list of two-digit numbers as used in Experiment I. By utilizing the obtained association value scales, for each subject various tasks were constructed with the same difficulty level. In Experiment II the second digits were chosen without replication whereas in Experiment III the second digits were chosen with replication. These two types of experiment imply two different guessing situations. Examples of the two different types of task are as follows:

In Experiment II	In Experiment III
62	28
58	85
91	98
47	73
35	56
89	38
26	67
74	47

It was thought that the type of task used in Experiment II would allow subjects to guess "intelligently" by restricting their guesses only to the unlearned response alternatives. In Experiment III, however, the guessing probability would be the same for every unlearned item, since each possible response alternative is equally likely to occur.

List length was one of the experimental variables, two different list lengths being used. The tasks in experiments IIA and IIIA contained eight items whereas the tasks in Experiments IIB and IIIB contained seven items. It was thougth that the two different list lengths used would have a direct effect on subjects' overall performance and that this would be reflected in their parameter values.

The first group of subjects were the same eight who had completed Experiment I in five sessions. This time they completed 104 paired-associate tasks in six sessions, 52 tasks under each of the two experiments (26 tasks under both condition A and condition B). Because of the extended nature of previous experiments these eight subjects showed a substantial practice effect in working with two-digit numbers. Four of them, in particular, became so good that they could learn practically all the items in a given list after a single presentation. Therefore, the data from these four subjects had to be excluded from the analysis, since they did not provide sufficiently long response sequences.

Another group of 36 subjects who took part in these experiments as an undergraduate course requirement were equally divided between the two conditions. A group of 18 subjects were allocated to Experiments IIA and IIB and completed six paired-associate tasks under each of the two conditions. The remaining group of 18 subjects were similarly allocated to Experiments IIIA and IIIB. For these subjects, the obtained group association value scale was utilized in constructing various tasks with the same difficulty level. The data obtained from these subjects was analyzed seperately.

The apparatus used was the Wang 700B programmable electronic calculating machine with two number displays. After two practice trials subjects completed the required number of paired-associate tasks with list lengths seven and eight in an alternating order. The two-digit numbers in a given list were presented at a rate of one a second. After the presentation of all the items, the single digits representing the different number categories (twenties, thirties, etc.) were given according to a prearranged testing order. The subjects were instructed to respond with the digit associated with that number category. At the end of each test trial the subjects were informed of their incorrect responses by the presentation of single digits representing the number categories to which they responded incorrectly. This was followed by new presentation and test trials until all the items were correctly responded to twice in succession. During each task correct and incorrect responses were recorded for each stimulus item.

Results and Discussion

The parameter values of the seven models described earlier were estimated by using a minimum X^2 estimation procedure. Apart from having several desirable properties such as consistency and efficiency of the resulting estimates, the minimum X^2 also provides a measure of the adequacy of any single model and, if the degrees of freedom are equal, a method for directly comparimg the fit of various models. If several models are being analyzed, each involving a different number of free parameters, then the probabilty levels of the X^2's may be compared (Cf. Atkinson and Crothers, 1964). For each model, the parameter estimates were based on the 16 possible outcome sequence over trials 2 to 5. To illustrate the minimum X^2 method let $Pr(O_i; \theta_1, \dots \theta_k)$ denote the probability of the event O_i, where $i = 1, \dots n$ and $(\theta_1, \dots \theta_k)$ are the parameters to be estimated from the data. Further, let $N(O_i)$ denote the observed frequency of the stimulus items with outcome O_i over trials 2 to 5,and let

$$T = N(O_1) + N(O_2) + \dots + N(O_{16}).$$

Then define the function

$$X^2(\theta_1, \ldots \theta_k) = \sum_{i=1}^{16} \frac{\left[T\ Pr(O_i;\ \theta_1, \ldots \theta_k) - N(O_i)\right]^2}{T\ Pr(O_i;\ \theta_1, \ldots \theta_k)} \qquad (6)$$

and select the estimates of $(\theta_1, \ldots \theta_k)$ so that they jointly minimize the X^2 function. Because of the problems involved in carring out the minimization analytically a numerical minimization procedure was used on a computer. Given the assumption that all the stimulus items are stochastically independent and identical, then under the null hypothesis it can be shown that the degrees of freedom associated with a model that requires k parameters to be estimated from the data are

$$df = n - k - 1$$

where n is the number of possible outcomes.

Analysis of the response tuples for trials 2 to 5 is of particular importance because a major portion of the learning occurred during the first five trials. For Experiments IIA, IIB, IIIA, and IIIB the proportion of subjects who reached a correct response level were .89, .93, .86, .92, respectively.

For each experiment described above, the X^2 minimization procedure given by Eq. 6 was applied to the observed response frequencies of the O_i events obtained from the four experiments. Table 17 presents the parameter estimates and the associated X^2 values for the group data. For experiments II and III the group data was based on the initial performance of two groups of 18 subjects. There were 144 response sequences for both of the Experiments IIA and IIIA, and 126 for both of the Experiments IIB and IIIB. Tables 18 to 21 present the parameter estimates and the associated X^2 values obtained for the individual subjects (Subject 2, 5, 7, and 8) who completed 52 tasks under each of

Table 17: Parameter estimates and the X^2 values for the group data.

Model	Parameter	Experiment			
		IIA	IIB	IIIA	IIIB
One-Element	c	.245	.348	.244	.292
Linear	θ	.325	.386	.366	.434
RTI	c	.429	.522	.579	.607
	θ	.862	.905	.653	.828
LS-2	a	.824	.396	.282	.387
	f	.765	.731	.694	.611
LS-3	a	.121	.396	.287	.387
	f	.094	.731	.661	.611
	c	.420	1.000	.902	1.000
LOS-2	a	.448	.523	.354	.483
	f	.658	.564	.587	.493
LOS-3	a	.408	.512	.291	.467
	f	.481	.530	.714	.444
	f_1	.694	.570	.041	.507

Experiment	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIA	136.74	167.71	13.96*	16.88*	11.39*	12.38*	8.86*
IIB	97.83	176.11	21.31*	15.63*	15.63*	14.13*	14.02*
IIIA	142.84	77.10	33.60	15.62*	15.24*	21.93*	15.45*
IIIB	157.13	108.97	24.78*	11.42*	11.42*	9.60*	9.38*
Total X^2	534.54	529.89	93.65	59.55	53.68	58.04	47.71
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Table 18: Parameter estimates and the X^2 values for Subject 2.

Model	Parameter	Experiment IIA	IIB	IIIA	IIIB
One-Element	c	.519	.525	.597	.578
Linear	θ	.582	.653	.618	.668
RTI	c	.875	.808	.707	.758
	θ	.813	.843	.911	.889
LS-2	a	.510	.521	.544	.574
	f	.564	.393	.515	.622
LS-3	a	.510	.521	.544	.574
	f	.564	.393	.515	.622
	c	1.000	1.000	1.000	1.000
LOS-2	a	.574	.574	.640	.724
	f	.406	.313	.385	.450
LOS-3	a	.609	.601	.734	.731
	f	.534	.308	.666	.481
	f_1	.365	.283	.350	.446

Experiment	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIA	90.85	116.60	71.31	29.93	29.93	29.38	28.61
IIB	115.13	53.13	31.01	8.01*	8.01*	6.87*	6.57*
IIIA	72.21	102.57	55.66	26.34*	26.34	22.23*	19.89*
IIIB	54.88	49.04	27.88	21.14*	21.14*	18.00*	17.98*
Total X^2	333.07	321.34	185.86	85.42	85.42	76.48	73.05
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Table 19: Parameter estimates and the X^2 values for Subject 5.

Model	Parameter	Experiment			
		IIA	IIB	IIIA	IIIB
One-Element	c	.573	.644	.502	.771
Linear	θ	.512	.578	.499	.816
RTI	c	.698	.713	.661	.928
	θ	.943	.935	.852	.893
LS-2	a	.580	.608	.530	.719
	f	.560	.606	.755	.486
LS-3	a	.580	.608	.530	.701
	f	.560	.606	.755	.321
	c	1.000	1.000	1.000	.952
LOS-2	a	.613	.605	.594	.767
	f	.376	.383	.490	.300
LOS-3	a	.581	.730	.531	.882
	f	.561	.932	.756	.999
	f_1	.001	.192	.001	.222

Experiment	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIA	81.50	158.33	36.57	15.88*	15.88*	18.19*	15.88*
IIB	33.42	91.32	22.56*	7.99*	7.99*	14.82*	5.73*
IIIA	47.81	135.62	38.16	15.15*	15.15*	28.14	15.56*
IIIB	19.46*	9.21*	5.35*	3.49*	3.08*	4.45*	2.31*
Total X^2	182.19	394.48	102.64	42.91	42.50	65.60	39.48
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Table 20: Parameter estimates and the X^2 values for Subject 7.

Model	Parameter	IIA	IIB	IIIA	IIIB
		Experiment			
One-Element	c	.830	.717	.591	.963
Linear	θ	.897	.832	.846	.964
RTI	c	.987	.935	.959	1.000
	θ	.913	.906	.886	.964
LS-2	a	.697	.648	.669	.963
	f	.185	.220	.215	1.000
LS-3	a	.697	.629	.669	.980
	f	.185	.182	.215	.809
	c	1.000	.988	1.000	.978
LOS-2	a	.709	.711	.697	.941
	f	.145	.190	.168	.180
LOS-3	a	.743	.711	.424	.962
	f	.204	.190	.025	.999
	f_1	.101	.190	.235	.001

Experiment	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIA	46.57	12.80*	11.60*	2.51*	2.51*	2.40*	2.34*
IIB	59.89	22.50*	17.06*	6.06*	6.01*	5.02*	5.02*
IIIA	200.29	36.72	32.16	9.23*	9.23*	7.70*	2.93*
IIIB	.23*	.23*	.23*	.23*	.18*	1.20*	.24*
Total X^2	306.98	72.25	61.05	18.03	17.93	16.32	10.53
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Table 21: Parameter estimates and the X^2 values for Subject 8.

Model	Parameter	Experiment IIA	IIB	IIIA	IIIB
One-Element	c	.828	.821	.881	.999
Linear	θ	.782	.833	.734	.975
RTI	c	.828	.915	.881	.995
	θ	.999	.936	.999	.980
LS-2	a	.749	.717	.881	.970
	f	.374	.308	1.000	.843
LS-3	a	.749	.717	.881	.987
	f	.374	.308	1.000	.796
	c	1.000	1.000	1.000	.985
LOS-2	a	.690	.837	.968	.953
	f	.201	.245	.309	.147
LOS-3	a	.801	.949	.930	.970
	f	.482	.999	.350	.370
	f_1	.074	.228	.012	.310

Experiment	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIA	20.71*	44.67	20.85*	6.76*	6.76*	9.19*	6.45*
IIB	23.40*	24.41*	15.27*	7.60*	7.60*	5.15*	4.40*
IIIA	6.55*	68.87	6.63*	6.55*	6.55*	4.45*	4.44*
IIIB	3.76*	.10*	.10*	.12*	.08*	.63*	.33*
Total X^2	54.42	138.05	42.85	21.03	20.99	19.42	15.62
df	14	14	13	13	12	13	12

* Not significant at .01 level.

the two experiments (26 tasks under both condotion A and condotion B). The results obtained from 36 subjects who completed 12 tasks in one session are given in the Appendix II. It must be pointed out that the small sample of response sequences obtained from the subjects who participated in the experiment only one session might have an undesirable effect on the resulting estimates.

Individual differences in the performance of different subjects are clearly reflected in the estimated values of the parameters. The learning rate parameters, in particular, were closely related with the observed mean trial number of last error. These correlations are given on page 64/a. The correlations obtained between the various forgetting parameters and the subjects' overall performance were relatively low and in some cases insignificant, indicating that each subject's overall performance was more strongly related to the magnitude of the learning rate parameters rather than the forgetting parameters. The role of learning rate parameters in relation to the forgetting parameters seems to be rather crucial. It was assumed that for each individual subject the learning rate parameter would be about the same under different experimental conditions, since all the experiments were comparable except that they differed in the list lengths used. It was also assumed that only the forgetting parameters would be influenced by the list length variable (Cf. Atkinson and Crothers, 1964; Calfee and Atkinson, 1965). An inspection of Tables 18 to 21 reveal that both the learning rate parameter and the forgetting parameter change depending on various experimental conditions. Although the direction of change in the learning rate parameter closely related with subjects' overall performance, the variations in the forgetting parameters were less tractable.

Comparisons of individual results with group results revealed that in some cases the one-element model does a better job for the group data than it does for the individual data, Correlations between the average number of last error and the learning rate parameters in various models:

Model	Experiment			
	II	IIA	III	IIIA
One-Element	.838	.771	.733	.818
Linear	.854	.822	.831	.843
RTI	.794	.763	.748	.582
LS-2	.894	.757	.805	.859
LS-3	.464	.725	.614	.743
LOS-2	.936	.669	.653	.864
LOS-3	.476	.753	.733	.820

whereas the goodness-of-fit of the linear model to the individual data was consistently better than the goodness-of-fit to the group data. Tables 22 and 23 present the total X^2 values for the group and the individual data, respectively. The total X^2 values for the groups have been summed over the X^2 values which were obtained seperately for the six different tasks under each experiment. The total X^2 values for the individual data are the summation of the X^2 values obtained for the one-session subjects. An inspection of Table 22 revealed that the one-element model gives a better fit to the group data obtained from the experiments IIB and IIIB, whereas the goodness-of-fit of the linear model is better than the goodness-of-fit of the one-element model for the experiments IIA and IIIA. The grand totals summed over the total X^2 values for each experiment indicated that there was little difference between the goodness-of-fit of the two models. The total X^2 values in Table 23 showed that the goodness-of-fit of the linear model to the individual data was consistently better than the goodness-of-fit of the one-element model. The differences between the goodness-of-fit of the two models to the individual and the group data cast some doubt on the credibility of the one-element model. It appears that for the group data the better fit obtained under the one-element model is mainly due to the effect of pooling the data from different subjects. According to the results, the response sequences obtained from individual subjects can be more adequately described by the linear model.

	Experiment				Grand Total
	IIA	IIB	IIIA	IIIB	
One-Element	575.50	614.04	507.66	678.07	2375.27
Linear	591.38	506.91	741.91	481.71	2321.91

Table 22: Total X^2 values obtained for the group data under the one-element and the linear models.

	Experiment				Grand Total
	IIA	IIB	IIIA	IIIB	
One-Element	836.61	1096.30	855.10	962.80	3750.81
Linear	552.43	455.34	615.13	355.41	1978.31

Table 23: The total X^2 values obtained for the individual data under the one-element and the linear models.

Of the two-parameter models, the random trial increments (RTI) model is less accurate than both of the LS-2 and LOS-2 models. But the variations in the goodness-of-fit of these models to the individual data brings up the possibility that the data obtained from an individual subject can sometimes be more efficiently analyzed with the RTI model rather than the LS-2 or the LOS-2 models. Of the three-parameter models, the LOS-3 model is consistently more accurate than the LS-3 model in all applications to the group data. With only a few exception this is also true for the data obtained from individual subjects. In the analysis of group data, although the obtained X^2 values might provide a measure of the adequacy of any single model, the choice between two models which differ in their fundemental assumptions seems to be rather difficult when the goodness-of-fit to the individual and the group data show considerable discrepancies.

IDENTIFICATION OF PARAMETERS

A theory can be considered as a description of a system which generates sequences of events. In the application of a theory to the results of an experiment it is generally assumed that the physical conditions of the experiment produce sequences according to the rules specified in the theory. A theory provides a set of variables which can be used to describe a system at each stage of an experiment. In some cases it happens that some of the variables or states specified by a theory cannot be identified completely in an experiment. In such cases important questions arise as to whether the experiment is relevant to certain assumptions of the theory. In learning theory, with the development of Markovian models these questions became the focus of a number of studies related to the identification of theoretical states and the parameters of the all-or-none models (e.g., Greeno and Steiner, 1964; Greeno, 1967; Greeno, 1968; Steiner and Greeno, 1969; Polson, 1970). The method used in these studies is that of constructing a second theory with all of its states identifiable in the outcome-space of the experiment and showing that these states are observable in the data-space generated by the original model.

Most of the learning models, including the two models described in the previous section (the LS and the LOS models), can be based on a simple extention of the all-or-none learning assumption. The extended theory with its initial and transition probabilities has been stated (Greeno, 1968) in general form as follows:

$$\Pr(\,L(1),\ E(1),\ C(1),\ O(1)\,) = (\,t,\ (1-s-t)r,\ (1-s-t)(1-r),\ s\,)$$

$$\begin{array}{c|cccc|}
 & L(n+1) & E(n+1) & C(n+1) & O(n+1) \\
\hline
L(n) & 1 & 0 & 0 & 0 \\
E(n) & d & (1-d)q & (1-d)p & 0 \\
C(n) & c & (1-c)q & (1-c)p & 0 \\
O(n) & ab & a(1-b)e & a(1-b)(1-e) & 1-a
\end{array} \qquad (7)$$

where in States O and E only error occur and in State C only correct responses occur. In the long-short model it has been assumed that $s = 1$, $e = q$, and $c = d$. Similarly, when $s = 1$, $a = q$, $e = 0$, and $b = 0$ the LOS model is obtained.

Consider the following Markov chain whose initial vector and the transition matrix are

$$\Pr(\,Q(1),\ R(1),\ S(1),\ I_1(1),\ I_2(1)\ \ldots\,) =$$
$$(\,\pi,\ 0,\ (1-\pi)(1-\theta),\ (1-\pi)\theta,\ 0\ \ldots\,)$$

$$\begin{array}{c|ccccc|}
 & Q(n+1) & R(n+1) & S(n+1) & \cdots & I_{j+1}(n+1) \\
\hline
Q(n) & 1 & 0 & 0 & \cdots & 0 \\
R(n) & u & (1-u)v & (1-u)(1-v) & \cdots & 0 \\
S(n) & 0 & v & 1-v & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & & \vdots \\
I_j(n) & \alpha_j & 0 & \beta_j & & \gamma_j
\end{array} \qquad (8)$$

where Q is an absorbing state entered on the trial following the last error, or the state on trial 1 if there are no errors; State R represents the occurence of an error on any trial; State S represents the occurence of a success on any trial before the last error; I_1, I_2, ..., I_j, ... : I_j applies on trial j if there are errors on all of the first j trials. Greeno and Steiner (1964), Greeno (1967) have shown that Eq.8 describes a theory with observable states. It has also been shown that Eq. 8 is equivalent to the theory given by Eq. 7. Furthermore, it has been proved that the values of the parameters of Eq. 8 can be calculated as functions of Eq. 7. In other words, any probability measure on data that can be generated by Eq. 7 corresponds to some (at least one) set of parameter values for Eq. 7. And for any such set of parameter values, it is possible to calculate a set of parameter values for Eq. 8 that would generate exactly the same probability measure on the data.

The identifiable parameters in Eq. 8 can easily be estimated from the data. But the values of these parameters will not be of much interest unless the theoretical parameters of Eq. 7 can be mapped to the identifiable parameters of Eq. 8. The derivation of the function that maps the theoretical parameters to the identifiable parameters has been given by Greeno (1968) and the result is repeated here for convenience:

$$
\begin{aligned}
\pi &= t + (1-s-t)(1-r)\left(\frac{c}{q + pc}\right), \\
(1-\pi)\theta &= s + (1-s-t)r, \\
(1-u)v &= (1-d)q, \qquad\qquad (9) \\
v &= q + pc, \\
(1-\pi)\theta(\, w-(1-u)v \,)x &= s\,(\, w-(1-u)v+(1-w)(1-b)c \,),
\end{aligned}
$$

$$(1-\pi)\theta(\ w-(1-u)v\ )xy = s\ (\ u-(1-w)\ (b+(1-b)(1-e)\ \frac{c}{q+pc}\ )\ )$$

where under the assumption that $(1-a) \neq (1-d)q$

$$w = 1-a,$$

$$x = \left(\frac{s}{s+(1-s-t)r}\right)\left(1+\frac{a(1-b)e}{(1-a)-(1-d)q}\right),$$

$$y = \frac{\left(\frac{pc+qd-a(1-b)(1-e)c}{q+pc} - ab\right)}{a(1-b)e+(1-a)-(1-d)q}$$

Following the general notational scheme developed in earlier papers (Greeno and Steiner, 1964; Greeno, 1967; Greeno, 1968), the likelihood function can be written

$$L = (\pi)^{N(A_o)}\ (\ (1-\pi)(1-\theta)\ )^{N(B_o)} \prod\ (P(A_j))^{N(A_j)} \prod\ (P(B_j))^{N(B_j)}\ (1-u)^{n(R)-n(RQ)}\ u^{n(RQ)}\ (1-v)^{n(S)}\ v^{n(R)} \qquad (10)$$

where $P(A_j)$ is the probability of an initial run of j errors and no errors after the first correct response; $P(B_j)$ is the probability of an initial run of j errors and at least one error after the first correct response; $N(A_o)$ is the number of sequences with zero errors; $N(B_o)$ is the number of sequences that start in State S; $N(A_j)$ is the number of sequences with j errors before the first correct response and no errors thereafter; $N(B_j)$ is the number of sequences with j errors before the first correct response and at least one error after the first correct response; $n(R)$ is the total number of errors after the first correct response; $n(S)$ is the total number of correct responses before the last error; $n(RQ)$ is the number of transitions from State R to S.

In the estimation procedure an iterative search was utilized by the aid of a computer to find the values of π, θ, u, v, w, x, y which minimize the value of $-2\log(L)$. Since Eq. 9 can be considered as a set of six equations with eight unknowns, some identifying restrictions are needed on the unknown theoretical parameters such that these equations can have a unique solution. Any testable restriction onthe theoretical parameters of Eq. 7 reduces the number of parameters to be estimated in Eq. 9. For notational convenience if we subscript the value of $-2\log(L)$ by the number of estimated parameters, then the seven parameters of the general model given in Eq. 8 can be estimated by minimizing $-2\log(L_7)$.

In the comparisons of different models to different cases the likelihood ratio test was used. As stated in Greeno (1968), "this is appropriate when the more general version of the model has k parameters, and when the restricted version has a j-dimentional parameter space contained in the k-dimentional parameters space of the more general version. Then we test the null hypothesis that the true values of the parameters lie on the j-dimentional surface, against the alternative that they are somewhere else in the k-dimentional space. The test statistic is $-2\log(L_j/L_k)$, which is the difference between the two minimum values of $-2\log(L)$ obtained in estimating the two sets of parameters, and the asymptotic distribution of the test statistic is chi-square with (k-j) degrees of freedom".

Because of theoretical or procedural considerations identifying restrictions can be imposed on the parameters of Eq. 7.

Two such cases will be considered:

Case 1:

In case 1 a free initial vector will be assumed. In this case if it is also assumed that $e = q$ and $b = d$, then from Eq. 9

$$y = \frac{u}{1 - (1-u)v}$$

is a testable hypothesis.

Case 2:

In case 2, since in the experiments described above it was possible to be correct on the first test, it can be assumed that all sequences were in State 0 at the beginning, and that the first study trial was equal to all other study trials. In this case, the initial probabilities would be equal to the transition probabilities from State 0. That is

$$\begin{aligned} r &= e \\ s &= 1-a \\ t &= ab \end{aligned} \qquad (11)$$

The assumption imposes two restrictions on the identifiable parameters. Using Eq. 9 and 11 it can be shown that these restrictions are

$$\pi = 1 - (1-u)\left(\frac{w - x(w-(1-u)v)(1-vy)}{w - x(w-(1-u)v)}\right),$$

$$\theta = \frac{vw}{w - x\,(\,w-(1-u)v\,)(1-vy)}$$

In the present section the two Cases given above were compered with the general case both for the individual subjects and the group data. Table 24 presents these comparisons together with the likelihood ratio test results for the group data. Similarly, Table 25 presents the results obtained from the five-session subjects. For the one-session subjects, the results appear in Tables 26, 27, 28, and 29 for the experiments IIA, IIB, IIIA, and IIIB, respectively.

The likelihood ratio test results given in Tables 24 and 25 did not allow us to accept the hypothesis that all the items were in State 0 at the beginning of the experiment (Case 2). The results were strongly in favour of Case 1 where it was assumed that at the beginning of an experiment an item can be in any one of the states with a probability specified by a free initial vector. The results obtained from the one-session subjects were rather contradictory. Although Case 1 was acceptable nearly for all the subjects, the obtained likelihood ratio test results showed that Case 2 was also acceptable for some of them.

A choice between Case 1 and Case 2 poses theoretical as well as methodological problems. In the development of mathematical learning theory the aim has been the representation of psychological processes in a formal model. The examples of this can be seen in various attempts to incorparate psychological processes into a mathematical model: As mentioned earlier, in the long-short model, besides the forgetting state, there is a distinction between short-term memory and long-term memory; similary, Atkinson and Shifrin (1968), and Bower (1967) make a distinction short-term memory (sometimes called a buffer) and long-term memory; Waugh and Norman distinguish these states as primary and secondary memory; Peterson (1966) distinguishes between an active trace and a structural trace. Since in all these theoretical models an item usually starts in an unlearned state, then goes

		General Case	Case 1	Case 2	Likelihood ratio test	
Subj.	Exp.	$-2\log L_7$	$-2\log L_6$	$-2\log L_5$	$-2\log(L_6/L_7)$	$-2\log(L_5/L_7)$
S2	IIA	307.85	323.04	394.61	15.19	86.76
	IIB	283.56	297.97	343.73	14.41	60.17
	IIIA	307.07	308.91	363.58	1.84*	56.51
	IIIB	289.25	290.10	357.61	.85*	68.36
S5	IIA	230.58	253.57	253.57	22.99	22.99
	IIB	251.98	269.36	297.51	17.38	45.53
	IIIA	280.36	286.34	329.63	5.98	49.27
	IIIB	199.81	199.81	232.06	.00*	32.25
S7	IIA	257.42	257.42	288.59	.00*	31.17
	IIB	158.35	158.40	177.11	.05*	18.76
	IIIA	241.99	241.99	277.89	.00*	35.90
	IIIB	64.82	64.82	77.65	.00*	12.83
S8	IIA	170.80	170.84	203.91	.00*	33.07
	IIB	111.66	111.66	145.73	.00*	34.07
	IIIA	157.58	157.58	187.76	.00*	30.18
	IIIB	42.79	42.79	52.61	.00*	9.82

* Not significant at .05 level.

Table 25: Comparisons between the minimized values of $-2 \log L$'s and the likelihood ratio test results for the individual subjects who partioipated for five sessions.

	General Case	Case 1	Case 2	Likelihood ratio test	
Subject	$-2\log L_7$	$-2\log L_6$	$-2\log L_5$	$-2\log(L_6/L_7)$	$-2\log(L_5/L_7)$
1	88.36	88.52	97.69	.16*	9.33
2	164.10	164.14	178.37	.04*	14.27
3	197.44	197.48	203.40	.04*	5.96*
4	268.76	268.88	282.25	.12*	13.49
5	29.28	30.20	31.07	.92*	1.79*
6	16.25	20.09	21.40	3.84*	5.15*
7	81.79	82.47	87.25	.68	5.46*
8	37.85	37.94	40.18	.09*	2.33*
9	260.87	260.97	266.45	.10*	5.58
10	182.01	182.07	186.09	.06*	4.08*
11	230.88	242.00	245.68	11.12	14.80
12	121.57	121.60	125.02	.03*	3.45*
13	12.91	14.07	17.10	1.16*	4.19*
14	72.29	72.86	83.21	.57*	10.92
15	112.58	112.60	119.12	.02*	6.54
16	21.01	22.56	29.23	1.55*	8.22
17	108.39	108.83	118.64	.44*	10.25
18	97.19	98.41	101.84	1.22*	4.65*

* Not significant at .05 level.

Table 26: (Experiment IIA). Comparisons between the minimized values of −2 log L's and the likelihood ratio test results for the individual subjects who participated only for one session.

	General Case	Case 1	Case 2	Likelihood ratio test	
Subject	$-2\log L_7$	$-2\log L_6$	$-2\log L_5$	$-2\log(L_6/L_7)$	$-2\log(L_5/L_7)$
1	46.59	48.52	55.84	1.93*	9.25
2	153.61	153.68	157.59	.07*	3.98*
3	132.41	133.19	150.09	.78*	17.68
4	187.70	187.71	198.62	.01*	10.92
5	56.61	57.52	64.96	.91*	8.35
6	11.63	14.00	17.80	2.37*	6.17
7	11.10	12.93	16.80	1.83*	5.70*
8	19.67	21.64	28.81	1.97*	9.14
9	273.48	273.60	285.90	.12*	12.42
10	105.53	106.50	118.91	.97*	13.38
11	154.60	154.77	165.77	.17*	11.17
12	171.11	171.20	186.60	.09*	15.49
13	87.45	87.76	98.45	.31*	11.00
14	73.62	74.22	80.70	.60*	7.08
15	124.82	125.05	137.88	.23*	13.06
16	17.74	18.36	25.81	.62*	8.07
17	135.26	135.98	142.59	.72*	7.33
18	93.38	94.44	104.78	1.06*	11.40

* Not significant at .05 level.

Table 27: (Experiment IIB). Comparisons between the minimized values of −2 log L's and the likelihood ratio test results for the individual subjects who participated only for one session.

	General Case	Case 1	Case 2	Likelihood ratio test	
Subject	$-2\log L_7$	$-2\log L_6$	$-2\log L_5$	$-2\log(L_6/L_7)$	$-2\log(L_5/L_7)$
1	59.00	59.57	61.89	.57*	2.89*
2	134.45	134.47	139.74	.02*	5.29*
3	73.05	73.33	75.55	.28*	2.50*
4	221.23	221.29	229.49	.06*	8.26
5	38.88	39.12	40.85	.24	1.97*
6	207.81	207.89	217.80	.08*	9.99
7	283.82	285.46	394.94	1.64*	11.12
8	83.06	84.57	95.86	1.51*	12.08
9	58.59	60.48	69.10	1.89*	10.51
10	143.40	143.44	149.58	.04*	6.18
11	118.15	118.28	129.03	.13*	10.88
12	33.45	35.65	46.97	2.20*	13.58
13	85.08	85.26	90.61	.18*	5.53*
14	233.27	233.34	245.08	.07*	11.81
15	210.75	210.83	218.38	.08*	7.63
16	184.02	184.30	192.62	.28*	8.06
17	93.64	93.71	102.63	.07*	8.99
18	160.74	160.81	171.94	.07*	11.20

* Not significant at .05 level.

Table 28: (Experiment IIIA). Comparisons between the minimized values of -2 log L's and the likelihood ratio test results for the individual subjects who participated only for one session.

	General Case	Case 1	Case 2	Likelihood ratio test	
Subject	$-2\log L_7$	$-2\log L_6$	$-2\log L_5$	$-2\log(L_6/L_7)$	$-2\log(L_5/L_7)$
1	67.62	69.51	76.82	1.89*	9.20
2	77.82	78.13	81.48	.31*	3.66*
3	48.64	48.91	51.47	.27*	2.83*
4	58.40	58.92	61.09	.52*	2.69*
5	11.49	13.69	17.54	2.20*	6.05
6	250.35	250.40	259.62	.05*	9.27
7	222.41	222.47	232.07	.06*	9.66
8	134.00	134.06	137.49	.06*	3.49*
9	39.22	39.75	44.69	.53*	5.47*
10	176.92	177.03	183.54	.11*	6.62
11	114.52	115.24	122.33	.72*	7.81
12	31.94	32.18	41.50	.24*	9.56
13	57.38	59.37	77.79	1.99*	20.41
14	207.05	218.35	218.69	11.30	11.64
15	83.12	85.08	93.19	1.96*	10.07
16	106.64	106.71	117.97	.07*	11.33
17	158.71	158.90	163.63	.19*	4.92*
18	51.51	52.01	55.84	.50*	4.33*

* Not significant at .05 level.

Table 29: (Experiment IIIB). Comparisons between the minimized values of −2 log L's and the likelihood ratio test results for the individual subjects who participated only for one session.

through the other assumed states and finally becomes absorbed in a learned state, the data regarding the identification of parameters and theoretical states can be more meaningfully handled within the structure of Case 2. Although the results obtained from the group data do not support Case 2, there is some evidence for the acceptance of Case 2 only for the one-session subjects. On the other hand, if Case 1 is accepted the task facing the learning theorists would be very hard, indeed. There are numerous questions involved in why an unlearned item starts from any one of the assumed theoretical states at the first presentation. One of the simplest but unsatisfactory explanation might be the acceptance of labels such as "correct state" and "error state" which bear no psychological meaning at all. Therefore, this section remains inconclusive, since parameter identification could not be achieved without the use of a model which assumed presolution "correct" and "error" states. If this model had been employed our study would have been restricted to its parameters which bear no relationship to the psychological processes with which we want to work.

GENERAL CONCLUSIONS AND DISCUSSION

It was argued that item homogeneity as a confounding variable constitutes a major methodological problem in any study on individual differences in verbal learning, since it is difficult to determine whether the observed variance is due to the differences between subjects or the differences between the test items. In Part I an attempt was made to study individual differences in association value scales of the test items used. The correlations obtained between the individual association value scales and the group association value scale were quite low, indicating that such a scale based on group results was insufficient in controlling the item homogeneity for individual subjects. Similarly, low correlations obtained between the individual association value scales (which were obtained in an actual learning situation) and the group rated association value scale (which was obtained from the ratings of the test items) showed that the predictions from the rated association scale values to the performance on such materials were not very accurate. On the other hand, the rated association value scales obtained for individual subjects were highly correlated with their own individual association value scales obtained in an actual learning situation. These results indicated that instead of using a group association value scale, subjects' own ratings of items can be used to yield more accurate predictions of their performance in verbal learning tasks.

The comparisons between the group serial position curve and the individual serial position curves showed that there were considerable differences between the individual and the group results. Although for all the obtained serial position curves there were strong recency effects, the primacy parts of the curves varied from subject to subject. Some introspective reports obtained from subjects indicated that the different strategies used in the process of acquisition of the list might have caused the observed differences in their serial position curves.

In Part II it was argued that research on individual differences in learning must be guided by contemporary theories of human learning and performance. In an attempt to study individual differences in paired-associate learning the emphasis was placed upon various mathematical models with different theoretical assumptions. For each subject repeated measurements were taken by running them on a number of parallel tasks which were based on their association value scales. It was found that learning rate parameters, in particular, were closely related with the observed individual differences in performance. The correlations between the various forgetting parameters and the subjects' overall performance were relatively low and in some cases insignificant.

Comparisons of individual resuls with group resuls revealed that the linear model does a better job for the individual data than it does for the group data. Although, in the application of the one-element and the linear models to the group data the obtained total X^2's were about the same, the goodness-of-fit of the linear model to the individual data was consistently better than the goodness-of-fit of the one-element model. The goodness-of-fit of the two-parameter models (RTI, LS-2, LOS-2) varied between subjects, indicating that a particular model with specific assumptions might not be the best one for the data obtained from every subject. Of the three-parameter models, the LOS-3 model was consistently more accurate than the LS-3 model in all the applications to the group data. With only a few exceptions this was also true for the data obtained from individual subjects.

Since a study on individual differences in the parameter values of the learning models would not be complete without the identification of the theoretical parameters and the states of the models used, an attempt was made to examine the problems involved in the identification of these theoretical parameters and the states. Unfortunately, the results indicate that the data obtained from most of the subjects can be best explained with a three-state Markov chain which is equivalent to many of the current formalizations of all-or-none learning theories. Such a general model assumes an initial transition vector whose transition probabilities determine the three starting states which can be best labeled as "correct state", "error state", and "absorbing state". In the face of this evidence, the generally accepted labels of the present models such as forgetting state, short-term and long-term memory states, etc., do not seem to be very well founded as far as the data regarding identification of theoretical parameters and states is concerned. It was thought that a model with presolution "correct" and "error" states would not be satisfactory in the study of individual parameter values unless the parameters involved in this model bore some relationship to the psychological processes with which it was intended to work.

It appears that the need for a theory which affords a source of testable hypothesis for the psychological processes which exist within individuals still pervades the research on individual differences in verbal learning. When the present models or theories of learning are tested against the averaged characteristics of a group of subjects, the obtained goodness-of-fit does not necessarily indicate that such models or theories of learning really represent the state of affairs that exists within individuals. When these models are tested against data obtained from individual subjects, the goodness-of-fit seems to vary from subject to subject. If these variations were due only to differences in the parameter values of the models used, then it would be possible to account for the variance by assuming a suitable distribution for the parameter values involved. However, when fitting models to data from individual subjects we find that one particular model may best fit the behaviour of one group of subjects whereas the behaviour of other subjects may be better represented by other models. If one particular model gives the best fit to the behaviour of a particular subject (and this particular model might be different from subject to subject), then there is a possibility that different psychological mechanisims take place in different subjects. These mechanisms are, we hope, reflected in the assumptions of best fitting models. It is also possible that the behaviour of a particular subject may be fitted equally well by more than one type of model. In this case, in order to make a choice between these models we have to test them further against data obtained with different experimental variables which are assumed to effect the subjects' overall performance. In the present study only two such experimental variables were manipulated: length of task (or amount of material to be learned) and different guessing situations. We could also manipulate such variables as association value or meaningfulness (in our study we kept this variable constant), task complexity, the amount of practice on a given task or a particular class of tasks, original learning and relearning after some interpolated activity, etc.. As suggested by Jensen (1967), besides the ones listed above, some other procedural variables such as stimulus duration, distribution of practice and instructional variables such as differentially motivating sets constitute one of the most important "dimensions" in the study of individual differences in learning.

Individual differences approach can be used as a further source of information for the development and testing of psychological models. It is our contention that by developing new models which postulate different psychological mechanisims for certain group of subjects, and testing these models against the data obtained from individual subjects, some progress can be made towards more general law and theories. This requires an extensive study of all available data from individual subjects, not gross generalizations based on group results.

In evaluating the findings of the present research it would be useful to make a distinction between intrinsic and extrinsic individual differences (Cf. Jensen, 1967). The essence of the difference is exemplified by the two expressions: a) individual differences <u>in</u> learning, and b) the effects of individual differences <u>on</u> learning.

Certain attitudes and personality traits, chronological age, mental age, IQ, sex, and other personal characteristics are included in the extrinsic individual differences category. All these variables exert some influence on the subject's performance in a learning situation. These variables operate in a way to exert some influence upon the functioning of more basic processes of perception and learning. Much of the research work in the past was aimed at eliminating or controlling this kind of individual differences, assuming that subjects would all perform alike in the laboratory learning tasks when these controls have been applied. However, very little success was achieved in reducing the variance among subject performance.

On the other hand, as Jensen (1967) puts it, the term intrinsic individual differences refer to those individual differences which are inherent in learning and which do not exist independently of the learning process. That is, intrinsic individual differences consist of intersubject variability in the learning process itself.

The present research was aimed at discovering some of the experimental conditions under which these intrinsic differences show their effect. Intrinsic individual differences were demonstrated by our findings on variations in association values, serial position curves, and paired-associate learning. If we consider the differences in association values as a product of subjects' past experiences and some sort of ability to attach meanings to the tests items used, then we would be dealing with an intrinsic variable, probably a recognition process which is assumed to play an important role in the learning process itself. Introspective reports obtained from different subjects indicate that different strategies used by the subjects cause the observed differences in serial position curves. In other words, if we can talk of an ability to develop new strategies, we might correlate this ability with chronological age, mental age, IQ, and certain other personal characteristics. Similarly, when fitting models to data from individual subjects if we find that one particular model may best fit the behaviour of one group of subjects whereas the behaviour of other subjects may be better explained by other models, then it is interesting to know what sorts of extrinsic individual differences cause these different groups of subjects to show variations in the learning process itself.

This line of research might prove to be useful if we assume that the development of some extrinsic individual differences such as certain attitudes and personality traits are based on essential variables in the learning domain. In this case certain extrinsic individual differences and the learning performance both would have the same source. Future research on individual differences would be more beneficial if we could discover the relationship between the intersubject variability in the learning process and the extrinsic individual differences which are based on essential variables in the learning domain.

APPENDIX I

Mean rated association values and SD's for numbers 0-100 (Battig and Spera, 1962).

No	M	SD	No	M	SD	No	M	SD	No	M	SD
31	0.72	1.07	54	1.15	1.22	44	1.73	1.25	8	2.21	1.34
53	0.79	1.02	29	1.18	1.26	80	1.74	1.30	45	2.21	1.38
83	0.79	1.13	91	1.22	1.29	55	1.78	1.32	6	2.23	1.40
71	0.83	1.13	38	1.23	1.33	17	1.85	1.38	15	2.25	1.36
57	0.84	1.23	86	1.23	1.21	48	1.86	1.30	9	2.29	1.46
46	0.85	0.96	93	1.23	1.39	96	1.91	1.39	69	2.32	1.33
73	0.86	1.13	62	1.27	1.36	30	1.96	1.33	98	2.32	1.38
59	0.88	1.11	52	1.33	1.18	27	1.98	1.34	20	2.36	1.38
74	0.90	0.99	85	1.33	1.21	40	1.98	1.28	12	2.43	1.43
37	0.93	1.14	39	1.34	1.36	77	1.98	1.37	75	2.43	1.33
47	0.95	1.15	63	1.34	1.32	95	1.99	1.29	3	2.55	1.51
87	0.95	1.10	23	1.35	1.23	66	2.00	1.45	5	2.57	1.45
58	0.98	1.18	72	1.36	1.28	19	2.01	1.45	7	2.62	1.50
41	1.00	1.16	89	1.38	1.33	65	2.03	1.34	18	2.67	1.31
51	1.02	1.22	94	1.38	1.28	90	2.05	1.36	50	2.69	1.40
79	1.02	1.20	42	1.47	1.24	14	2.06	1.43	4	2.71	1.35
43	1.03	1.17	35	1.54	1.30	24	2.06	1.28	16	2.72	1.28
34	1.07	1.06	26	1.55	1.25	49	2.07	1.37	99	2.80	1.30
78	1.09	1.19	56	1.55	1.26	64	2.07	1.45	25	2.85	1.26
84	1.09	1.11	81	1.61	1.34	11	2.08	1.48	10	2.96	1.37
61	1.12	1.32	97	1.61	1.41	36	2.13	1.34	21	3.03	1.17
67	1.12	1.16	70	1.65	1.30	60	2.14	1.36	13	3.06	1.28
68	1.14	1.22	28	1.69	1.32	33	2.15	1.39	2	3.11	1.28
82	1.14	1.14	32	1.69	1.35	88	2.15	1.37	0	3.31	1.16
92	1.14	1.19	76	1.69	1.45	22	2.18	1.30	1	3.38	1.20
									100	3.56	0.93

APPENDIX II

Parameter estimates and the X^2 values for the data obtained from individual subjects who completed 12 tasks in one session.

Subject II-1

Model	Parameter	Experiment IIA	IIB
One-Element	c	.416	.506
Linear	θ	.450	.425
RTI	c	.542	.510
	θ	.896	.994
LS-2	a	.438	.506
	f	.838	.999
LS-3	a	.438	.734
	f	.838	.999
	c	1.000	.633
LOS-2	a	.641	.864
	f	.632	.691
LOS-3	a	.498	.506
	f	.999	.999
	f_1	.149	.002

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIA	12.70*	32.20	8.93*	8.10*	8.10*	8.48*	6.46*
IIB	3.97*	32.34	3.96*	3.97*	2.43*	8.69*	3.97*
Total X^2	16.67	64.54	12.89	12.07	10.53	17.17	10.43
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject II-1

Model	Parameter	Experiment	
		IIA	IIB
One-Element	c	.416	.506
Linear	θ	.450	.425
RTI	c	.542	.510
	θ	.896	.994
LS-2	a	.438	.506
	f	.838	.999
LS-3	a	.438	.734
	f	.838	.999
	c	1.000	.633
LOS-2	a	.641	.864
	f	.632	.691
LOS-3	a	.498	.506
	f	.999	.999
	f_1	.149	.002

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIA	12.70*	32.20	8.93*	8.10*	8.10*	8.48*	6.46*
IIB	3.97*	32.34	3.96*	3.97*	2.43*	8.69*	3.97*
Total X^2	16.67	64.54	12.89	12.07	10.53	17.17	10.43
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject II-2

Model	Parameter	Experiment	
		IIA	IIB
One-Element	c	.283	.288
Linear	θ	.437	.421
RTI	c	.639	.721
	θ	.668	.568
LS-2	a	.314	.283
	f	.718	.681
LS-3	a	.336	.296
	f	.494	.594
	c	.684	.795
LOS-2	a	.426	.376
	f	.598	.604
LOS-3	a	.314	.323
	f	.719	.771
	f_1	.001	.159

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIA	39.79	17.74*	12.31*	12.33*	9.34*	13.98*	12.35*
IIB	57.03	22.11*	19.94*	17.90*	17.45*	18.70*	17.24*
Total X^2	96.82	39.85	32.25	30.23	26.79	32.68	29.59
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject II-3

Model	Parameter	Experiment	
		IIA	IIB
One-Element	c	.218	.355
Linear	θ	.368	.464
RTI	c	.517	.655
	θ	.776	.722
LS-2	a	.263	.334
	f	.684	.639
LS-3	a	.234	.346
	f	.223	.582
	c	.477	.878
LOS-2	a	.417	.467
	f	.621	.564
LOS-3	a	.382	.479
	f	.433	.631
	f_1	.671	.544

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIA	86.88	52.79	17.02*	17.65*	17.22*	14.02*	12.56*
IIB	33.40	15.37*	7.56*	6.86*	6.71*	5.67*	5.61*
Total X^2	120.28	68.06	24.58	24.51	23.93	19.69	18.17
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject II-4

Model	Parameter	Experiment	
		IIA	IIB
One-Element	c	.137	.190
Linear	θ	.214	.275
RTI	c	.269	.436
	θ	.744	.703
LS-2	a	.138	.171
	f	.826	.675
LS-3	a	.148	.024
	f	.810	.165
	c	.837	.398
LOS-2	a	.276	.287
	f	.785	.643
LOS-3	a	.273	.188
	f	.750	.372
	f_1	.796	.711

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIA	41.34	48.34	11.16*	9.86*	9.66*	7.71*	7.60*
IIB	201.81	70.86	25.19*	25.05*	22.85*	21.72*	16.23*
Total X^2	243.15	119.20	36.35	34.91	32.51	28.43	23.83
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject II-5

Model	Parameter	Experiment	
		IIA	IIB
One-Element	c	.455	.458
Linear	θ	.525	.721
RTI	c	.645	.983
	θ	.906	.733
LS-2	a	.491	.513
	f	.739	.456
LS-3	a	.055	.421
	f	.024	.173
	c	.573	.840
LOS-2	a	.659	.561
	f	.543	.357
LOS-3	a	.416	.514
	f	.101	.457
	f_1	.604	.001

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIA	23.10*	22.12*	3.68*	5.83*	4.66*	6.43*	5.10*
IIB	35.18	1.19*	1.17*	3.03*	1.92*	3.46*	3.03*
Total X^2	58.28	23.31	4.85	8.86	6.58	9.89	8.13
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject II-6

Model	Parameter	Experiment IIA	Experiment IIB
One-Element	c	.499	.764
Linear	θ	.688	.612
RTI	c	.800	.765
	θ	.876	.999
LS-2	a	.564	.764
	f	.564	.999
LS-3	a	.001	.778
	f	.016	.999
	c	.748	.778
LOS-2	a	.658	.999
	f	.410	.424
LOS-3	a	.088	.999
	f	.022	.001
	f_1	.500	.424

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIA	33.35	10.20*	5.53*	5.92*	3.37*	6.14*	4.20*
IIB	4.84*	31.09	4.87*	4.84*	3.76*	1.02*	1.01*
Total X^2	38.19	41.29	10.40	10.76	7.13	7.16	5.21
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject II-7

Model	Parameter	Experiment	
		IIA	IIB
One-Element	c	.443	.867
Linear	θ	.451	.879
RTI	c	.630	.999
	θ	.832	.880
LS-2	a	.413	.867
	f	.657	.999
LS-3	a	.413	.938
	f	.657	.999
	c	1.000	.923
LOS-2	a	.601	.868
	f	.545	.336
LOS-3	a	.732	.867
	f	.999	.999
	f_1	.535	.001

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIA	30.25	54.57	18.43*	13.50*	13.50*	9.79*	9.03*
IIB	.77*	.70*	.70*	.77*	.55*	2.30*	.78*
Total X^2	31.02	55.27	19.13	14.27	14.05	12.09	9.81
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject II-8

Model	Parameter	Experiment IIA	Experiment IIB
One-Element	c	.774	.851
Linear	θ	.853	.797
RTI	c	.999	.852
	θ	.854	.999
LS-2	a	.732	.851
	f	.532	.999
LS-3	a	.710	.863
	f	.294	.724
	c	.934	.863
LOS-2	a	.752	.998
	f	.306	.388
LOS-3	a	.833	.998
	f	.863	.166
	f_1	.146	.389

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIA	5.41*	.33*	.33*	1.38*	1.16*	2.14*	1.24*
IIB	.57*	5.48*	.58*	.57*	.27*	.77*	.77*
Total X^2	5.98	5.81	.91	1.95	1.43	2.91	2.01
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject II-9

Model	Parameter	Experiment	
		IIA	IIB
One-Element	c	.181	.217
Linear	θ	.314	.338
RTI	c	.480	.392
	θ	.677	.875
LS-2	a	.208	.281
	f	.651	.710
LS-3	a	.208	.281
	f	.651	.710
	c	1.000	1.000
LOS-2	a	.295	.363
	f	.599	.598
LOS-3	a	.241	.290
	f	.433	.730
	f_1	.652	.065

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIA	148.51	61.98	25.52*	12.45*	12.45*	11.26*	9.36*
IIB	68.43	83.10	25.68*	14.63*	14.63*	16.36*	14.42*
Total X^2	216.94	145.08	51.20	27.08	27.08	27.62	23.78
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject II-10

Model	Parameter	Experiment	
		IIA	IIB
One-Element	c	.373	.168
Linear	θ	.357	.444
RTI	c	.390	.589
	θ	.969	.762
LS-2	a	.366	.253
	f	.870	.413
LS-3	a	.366	.253
	f	.870	.413
	c	1.000	1.000
LOS-2	a	.435	.282
	f	.641	.383
LOS-3	a	.418	.261
	f	.999	.424
	f_1	.188	.054

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIA	12.03*	57.64	11.53*	8.52*	8.52*	16.15*	7.56*
IIB	205.81	40.80	25.49*	11.10*	11.10*	11.28*	11.02*
Total X^2	217.84	98.44	37.02	19.62	19.62	27.43	18.58
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject II-11

Model	Parameter	Experiment	
		IIA	IIB
One-Element	c	.215	.284
Linear	θ	.337	.398
RTI	c	.615	.763
	θ	.568	.518
LS-2	a	.189	.252
	f	.617	.619
LS-3	a	.192	.251
	f	.509	.562
	c	.717	.861
LOS-2	a	.285	.345
	f	.594	.576
LOS-3	a	.278	.393
	f	.567	.779
	f_1	.608	.493

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIA	127.44	22.14*	12.12*	12.85*	11.48*	11.02*	10.97*
IIB	55.18	17.73*	15.81*	14.68*	14.03*	13.92*	12.16*
Total X^2	182.62	39.87	27.93	27.53	25.51	24.94	23.13
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject II-12

Model	Parameter	Experiment IIA	Experiment IIB
One-Element	c	.260	.194
Linear	θ	.450	.362
RTI	c	.552	.432
	θ	.846	.788
LS-2	a	.352	.261
	f	.594	.630
LS-3	a	.355	.262
	f	.581	.588
	c	.972	.886
LOS-2	a	.430	.300
	f	.504	.549
LOS-3	a	.443	.261
	f	.571	.630
	f_1	.463	.001

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIA	79.10	39.95	17.10*	11.93*	11.89*	12.25*	11.98*
IIB	106.76	39.27	22.61*	14.87*	14.14*	17.98*	14.88*
Total X^2	185.86	79.22	39.71	26.80	26.33	30.23	26.86
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject II-13

Model	Parameter	Experiment	
		IIA	IIB
One-Element	c	.491	.394
Linear	θ	.549	.487
RTI	c	.699	.649
	θ	.915	.736
LS-2	a	.535	.392
	f	.675	.508
LS-3	a	.535	.392
	f	.675	.508
	c	1.000	1.000
LOS-2	a	.680	.416
	f	.472	.410
LOS-3	a	.250	.448
	f	.039	.581
	f_1	.539	.210

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIA	33.25	34.95	9.61*	10.53*	10.53*	9.46*	7.61*
IIB	41.68	43.55	32.40	10.86*	10.86*	11.93*	10.52*
Total X^2	74.93	78.50	42.01	21.39	21.39	21.39	18.13
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject II-14

Model	Parameter	Experiment	
		IIA	IIB
One-Element	c	.336	.432
Linear	θ	.550	.598
RTI	c	.689	.915
	θ	.786	.653
LS-2	a	.393	.375
	f	.648	.496
LS-3	a	.220	.375
	f	.113	.496
	c	.618	1.000
LOS-2	a	.515	.480
	f	.539	.459
LOS-3	a	.329	.537
	f	.163	.719
	f_1	.630	.346

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIA	65.65	12.56*	8.17*	12.66*	7.63*	13.41*	11.37*
IIB	51.55	10.91*	10.56*	13.00*	13.00*	11.70*	10.69*
Total X^2	117.20	23.47	18.73	25.66	20.63	25.11	22.06
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject II-15

Model	Parameter	Experiment IIA	IIB
One-Element	c	.377	.285
Linear	θ	.514	.502
RTI	c	.682	.708
	θ	.777	.703
LS-2	a	.399	.319
	f	.650	.544
LS-3	a	.413	.302
	f	.396	.353
	c	.729	.772
LOS-2	a	.536	.408
	f	.535	.485
LOS-3	a	.541	.397
	f	.570	.437
	f_1	.524	.512

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIA	27.39*	14.35*	5.90*	6.47*	5.12*	6.05*	6.04*
IIB	79.52	17.05*	12.32*	10.46*	9.10*	9.96*	9.88*
Total X^2	106.91	31.40	18.22	16.93	14.22	16.01	15.92
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject II-16

Model	Parameter	Experiment	
		IIA	IIB
One-Element	c	.774	.787
Linear	θ	.853	.888
RTI	c	.999	.999
	θ	.853	.888
LS-2	a	.730	.621
	f	.530	.156
LS-3	a	.710	.621
	f	.294	.156
	c	.934	1.000
LOS-2	a	.752	.624
	f	.306	.131
LOS-3	a	.833	.659
	f	.863	.173
	f_1	.146	.075

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIA	5.41*	.33*	.33*	1.38*	1.16*	2.14*	1.24*
IIB	9.36*	2.04*	2.04*	.69*	.69*	.71*	.65*
Total X^2	14.77	2.37	2.37	2.07	1.85	2.85	1.89
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject II-17

Model	Parameter	Experiment	
		IIA	IIB
One-Element	c	.379	.222
Linear	θ	.518	.455
RTI	c	.721	.797
	θ	.747	.572
LS-2	a	.393	.251
	f	.609	.528
LS-3	a	.398	.243
	f	.408	.441
	c	.771	.834
LOS-2	a	.514	.331
	f	.511	.503
LOS-3	a	.527	.319
	f	.599	.463
	f_1	.480	.524

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIA	34.51	14.69*	7.34*	8.32*	7.05*	7.75*	7.62*
IIB	113.33	11.15*	9.14*	10.65*	10.21*	9.21*	9.16*
Total X^2	147.84	25.84	16.48	18.97	17.26	16.96	16.78
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject II-18

Model	Parameter	Experiment	
		IIA	IIB
One-Element	c	.396	.416
Linear	θ	.410	.552
RTI	c	.500	.721
	θ	.909	.781
LS-2	a	.388	.457
	f	.781	.636
LS-3	a	.388	.468
	f	.781	.548
	c	1.000	.886
LOS-2	a	.587	.543
	f	.630	.486
LOS-3	a	.580	.457
	f	.599	.637
	f_1	.634	.001

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIA	30.50	55.55	14.26*	13.16*	13.16*	11.30*	11.28*
IIB	27.11*	10.60*	5.53*	4.11*	3.59*	5.26*	4.12*
Total X^2	57.61	66.15	19.79	17.27	16.75	16.56	15.40
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject III-1

Model	Parameter	Experiment	
		IIIA	IIIB
One-Element	c	.483	.643
Linear	θ	.745	.752
RTI	c	.999	.937
	θ	.746	.807
LS-2	a	.540	.609
	f	.481	.539
LS-3	a	.460	.591
	f	.180	.339
	c	.842	.903
LOS-2	a	.583	.668
	f	.361	.374
LOS-3	a	.541	.745
	f	.483	.850
	f_1	.001	.222

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIIA	35.97	1.45*	1.46*	5.15*	3.58*	6.03*	5.15*
IIIB	10.28*	1.61*	1.20*	2.08*	1.83*	2.60*	1.63*
Total X^2	46.25	3.06	2.66	7.23	5.41	8.63	6.78
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject III-2

Model	Parameter	Experiment	
		IIIA	IIIB
One-Element	c	.296	.372
Linear	θ	.398	.513
RTI	c	.626	.669
	θ	.662	.757
LS-2	a	.274	.394
	f	.626	.699
LS-3	a	.279	.424
	f	.568	.506
	c	.866	.745
LOS-2	a	.400	.535
	f	.578	.572
LOS-3	a	.421	.394
	f	.682	.699
	f_1	.544	.001

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIIA	62.14	35.33	20.71*	18.10*	17.80*	15.35*	14.96*
IIIB	28.50*	10.80*	6.45*	7.21*	5.64*	7.83*	7.22*
Total X^2	90.64	46.13	27.16	25.31	23.44	23.18	22.18
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject III-3

Model	Parameter	Experiment	
		IIIA	IIIB
One-Element	c	.357	.643
Linear	θ	.569	.752
RTI	c	.753	.937
	θ	.773	.807
LS-2	a	.427	.609
	f	.526	.539
LS-3	a	.428	.591
	f	.458	.339
	c	.916	.903
LOS-2	a	.495	.668
	f	.429	.374
LOS-3	a	.428	.745
	f	.527	.850
	f_1	.001	.222

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIIA	68.54	14.48*	8.55*	6.22*	6.02*	6.68*	6.22*
IIIB	10.28*	1.61*	1.20*	2.08*	1.83*	2.60*	1.63*
Total X^2	78.82	16.09	9.75	8.30	7.85	9.28	7.85
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject III-4

Model	Parameter	Experiment	
		IIIA	IIIB
One-Element	c	.250	.398
Linear	θ	.360	.572
RTI	c	.532	.846
	θ	.684	.662
LS-2	a	.269	.417
	f	.729	.472
LS-3	a	.269	.417
	f	.729	.472
	c	1.000	1.000
LOS-2	a	.366	.430
	f	.615	.370
LOS-3	a	.269	.418
	f	.730	.473
	f_1	.003	.001

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIIA	41.66	54.59	28.41	14.56*	14.56*	15.55*	14.56*
IIIB	33.56	16.90*	16.10*	6.03*	6.03*	6.97*	6.03*
Total X^2	75.22	71.49	44.51	20.59	20.59	22.52	20.59
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject III-5

Model	Parameter	Experiment	
		IIIA	IIIB
One-Element	c	.660	.919
Linear	θ	.725	.924
RTI	c	.858	.999
	θ	.864	.925
LS-2	a	.628	.919
	f	.603	.999
LS-3	a	.628	.961
	f	.603	.999
	c	1.000	.956
LOS-2	a	.720	.906
	f	.404	.263
LOS-3	a	.767	.919
	f	.999	.999
	f_1	.217	.001

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIIA	9.27*	4.88*	2.28*	2.19*	2.19*	2.48*	1.70*
IIB	.26*	.25*	.25*	.26	.19*	.97*	.27*
Total X^2	10.53	5.13	2.53	2.45	2.38	3.45	1.97
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject III-6

Model	Parameter	Experiment	
		IIIA	IIIB
One-Element	c	.208	.109
Linear	θ	.288	.269
RTI	c	.407	.669
	θ	.723	.393
LS-2	a	.217	.106
	f	.778	.627
LS-3	a	.217	.106
	f	.778	.627
	c	1.000	1.000
LOS-2	a	.333	.135
	f	.690	.597
LOS-3	a	.217	.123
	f	.778	.669
	f_1	.001	.135

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIIA	40.31	57.04	16.84*	9.20*	9.20*	9.49*	9.21*
IIB	145.08	31.42	27.67*	13.54*	13.54*	14.24*	12.81*
Total X^2	185.39	88.46	44.51	22.74	22.74	23.73	22.02
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject III-7

Model	Parameter	Experiment	
		IIIA	IIIB
One-Element	c	.133	.158
Linear	θ	.199	.310
RTI	c	.255	.676
	θ	.700	.438
LS-2	a	.114	.166
	f	.812	.675
LS-3	a	.114	.166
	f	.812	.631
	c	1.000	.858
LOS-2	a	.216	.212
	f	.780	.614
LOS-3	a	.209	.166
	f	.726	.676
	f_1	.799	.001

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIIA	56.02	50.70	14.19*	10.90*	10.90*	9.88*	9.58*
IIIB	65.17	25.72*	23.34*	12.68*	12.49*	14.13*	12.68*
Total X^2	121.19	76.42	37.53	23.58	23.39	24.01	22.26
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject III-8

Model	Parameter	Experiment IIIA	Experiment IIIB
One-Element	c	.446	.233
Linear	θ	.402	.505
RTI	c	.496	.825
	θ	.930	.604
LS-2	a	.434	.300
	f	.935	.485
LS-3	a	.496	.299
	f	.915	.474
	c	.784	.982
LOS-2	a	.729	.354
	f	.686	.439
LOS-3	a	.435	.363
	f	.937	.474
	f_1	.002	.410

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIIA	10.20*	41.38	8.39*	9.33*	8.56*	11.19*	9.33*
IIIB	90.61	13.39*	11.97*	6.38*	6.37*	6.07*	6.03*
Total X^2	100.81	54.77	20.36	15.71	14.93	17.26	15.36
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject III-9

Model	Parameter	Experiment	
		IIIA	IIB
One-Element	c	.285	.587
Linear	θ	.569	.456
RTI	c	.793	.587
	θ	.715	.999
LS-2	a	.404	.559
	f	.421	.787
LS-3	a	.404	.559
	f	.412	.787
	c	1.000	1.000
LOS-2	a	.428	.642
	f	.345	.475
LOS-3	a	.404	.636
	f	.421	.999
	f_1	.001	.144

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIIA	98.76	28.08*	23.12*	9.64*	9.64*	10.07*	9.64*
IIB	9.46*	56.16	9.46*	6.00*	6.00*	6.02*	4.96*
Total X^2	108.22	84.24	32.58	15.64	15.64	16.09	14.60
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject III-10

Model	Parameter	Experiment	
		IIIA	IIIB
One-Element	c	.323	.274
Linear	θ	.380	.366
RTI	c	.457	.425
	θ	.868	.798
LS-2	a	.334	.264
	f	.873	.612
LS-3	a	.410	.264
	f	.266	.609
	c	.410	.992
LOS-2	a	.577	.300
	f	.712	.533
LOS-3	a	.528	.357
	f	.412	.718
	f_1	.745	.397

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIIA	19.78*	34.04	10.28*	13.60*	9.14*	16.26*	15.55*
IIIB	53.56	55.40	34.23	16.92*	16.92*	19.30*	15.25*
Total X^2	73.34	89.44	44.51	30.52	26.06	35.56	30.80
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject III-11

Model	Parameter	Experiment IIIA	Experiment IIIB
One-Element	c	.245	.239
Linear	θ	.394	.520
RTI	c	.541	.832
	θ	.859	.622
LS-2	a	.343	.343
	f	.641	.513
LS-3	a	.001	.331
	f	.044	.408
	c	.516	.861
LOS-2	a	.459	.389
	f	.533	.439
LOS-3	a	.406	.348
	f	.352	.522
	f_1	.587	.020

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIIA	104.64	61.15	13.21*	12.78*	9.65*	11.67*	9.96*
IIIB	88.33	12.11*	10.52*	7.95*	7.44*	8.39*	7.95*
Total X^2	192.97	73.26	23.73	20.73	17.09	20.06	17.91
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject III-12

Model	Parameter	Experiment	
		IIIA	IIIB
One-Element	c	.725	.780
Linear	θ	.744	.879
RTI	c	.823	.999
	θ	.900	.879
LS-2	a	.725	.678
	f	.999	.263
LS-3	a	.867	.671
	f	.999	.244
	c	.823	.995
LOS-2	a	.788	.685
	f	.485	.193
LOS-3	a	.725	.739
	f	.999	.322
	f_1	.001	.106

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIIA	3.08*	3.67*	3.10*	3.08*	1.73*	8.62*	3.10*
IIIB	7.86*	1.15*	1.15*	.75*	.75*	.85*	.68*
Total X^2	10.94	4.82	4.25	3.83	2.48	9.47	3.78
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject III-13

Model	Parameter	Experiment IIIA	Experiment IIIB
One-Element	c	.388	.507
Linear	θ	.442	.459
RTI	c	.661	.530
	θ	.728	.966
LS-2	a	.407	.488
	f	.628	.742
LS-3	a	.407	.488
	f	.628	.742
	c	1.000	1.000
LOS-2	a	.475	.482
	f	.485	.495
LOS-3	a	.442	.594
	f	.699	.999
	f_1	.124	.206

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIIA	29.02*	40.13	17.79*	5.02*	5.02*	5.40*	4.68*
IIIB	12.75*	43.78	11.95*	7.14*	7.14*	11.99*	5.37*
Total X^2	41.77	83.91	29.74	12.16	12.16	17.39	10.05
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject III-14

Model	Parameter	Experiment	
		IIIA	IIIB
One-Element	c	.200	.200
Linear	θ	.339	.376
RTI	c	.422	.525
	θ	.802	.704
LS-2	a	.258	.248
	f	.752	.620
LS-3	a	.301	.248
	f	.570	.620
	c	.626	1.000
LOS-2	a	.380	.331
	f	.650	.562
LOS-3	a	.258	.335
	f	.752	.584
	f_1	.001	.549

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIIA	49.93	30.10	8.60*	9.29*	7.10*	10.68*	9.31*
IIIB	78.60	30.58	15.39*	6.30*	6.30*	5.80*	5.78*
Total X^2	128.53	60.68	23.99	15.59	13.40	16.48	15.09
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject III-15

Model	Parameter	Experiment IIIA	IIIB
One-Element	c	.228	.332
Linear	θ	.403	.630
RTI	c	.339	.944
	θ	.476	.665
LS-2	a	.227	.396
	f	.515	.456
LS-3	a	.201	.268
	f	.412	.160
	c	.822	.774
LOS-2	a	.265	.455
	f	.477	.393
LOS-3	a	.290	.397
	f	.561	.458
	f_1	.393	.001

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIIA	99.83	17.55*	16.90*	14.98*	12.77*	15.30*	14.70*
IIIB	75.86	3.42*	3.29*	5.63*	3.31*	5.70*	5.64*
Total X^2	175.69	20.97	20.19	20.61	16.08	21.00	20.34
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject III-16

Model	Parameter	Experiment	
		IIIA	IIIB
One-Element	c	.275	.314
Linear	θ	.405	.516
RTI	c	.724	.737
	θ	.563	.680
LS-2	a	.252	.341
	f	.552	.536
LS-3	a	.252	.341
	f	.552	.536
	c	1.000	1.000
LOS-2	a	.300	.404
	f	.500	.460
LOS-3	a	.330	.401
	f	.660	.450
	f_1	.317	.465

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIIA	69.67	33.63	26.77*	10.58*	10.58*	11.04*	8.93*
IIIB	55.32	23.13*	19.38*	10.44*	10.44*	10.51*	10.51
Total X^2	124.99	56.76	46.15	21.02	21.02	21.55	19.44
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject III-17

Model	Parameter	Experiment	
		IIIA	IIIB
One-Element	c	.423	.242
Linear	θ	.405	.407
RTI	c	.437	.691
	θ	.976	.583
LS-2	a	.408	.202
	f	.724	.435
LS-3	a	.408	.202
	f	.723	.435
	c	1.000	1.000
LOS-2	a	.390	.237
	f	.506	.422
LOS-3	a	.543	.259
	f	.999	.489
	f_1	.296	.292

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIIA	19.90*	62.18	19.52*	10.75*	10.75*	17.92*	7.71*
IIIB	163.00	22.67*	19.00*	8.71*	8.71*	8.59*	7.93*
Total X^2	179.90	84.85	38.52	19.46	19.46	26.51	15.64
df	14	14	13	13	12	13	12

* Not significant at .01 level.

Subject III-18

Model	Parameter	Experiment IIIA	Experiment IIIB
One-Element	c	.251	.474
Linear	θ	.358	.706
RTI	c	.531	.858
	θ	.684	.830
LS-2	a	.293	.541
	f	.748	.479
LS-3	a	.293	.211
	f	.748	.046
	c	1.000	.785
LOS-2	a	.362	.608
	f	.604	.366
LOS-3	a	.299	.400
	f	.763	.096
	f_1	.028	.450

Exp.	One-Element	Linear Model	RTI	LS-2	LS-3	LOS-2	LOS-3
IIIA	36.38	44.75	22.86*	11.53*	11.53*	15.92*	11.49*
IIB	34.32	5.31*	3.37*	3.16	1.79*	3.28*	2.62*
Total X^2	70.70	50.06	26.23	14.69	13.32	19.20	14.11
df	14	14	13	13	12	13	12

* Not significant at .01 level.

REFERENCES

Anderson, N.H., "An analysis of sequential dependencies", in R.R. Bush and W.K. Estes (Eds.), Studies in mathematical learning theory, Stanford: Stanford Univer. Press., pp. 125-134, 1959.

Atkinson, R.C. and Shiffrin, R.M., "Human memory: A proposed system and its control processes", In K.W. Spence and J.T. Spence (Eds.), Psychology of learning and motivation, New York: Academic., 2, pp. 89-195, 1968.

Atkinson, R.C. and Crothers, E.J., "A comparison of paired-associate learning models having different acquisition and retention axioms", J. Math. Psychol., 1964, 1, 285-315.

Atkinson, R.C., Bower, G.H., and Crothers, E.J., "Introduction to mathematical learning theory", New York: Wiley., 1965.

Audley, R.J., "A stochastic description of the learning behaviour of an individual subject", Quart. J. Exp. Psychol., 1957, 9, 12-20.

Audley, R.J. and Jonckheere, A.R., "The statistical analysis of the learning process", Brit. J. Stat. Psychol., 1956, 9, 87-94.

Bakan, D., "A generalization of Sidman's results on group and individual functions, and a criterion", Psychol. Bull., 1954, 51, 63-64.

Battig, W.F., "Paired-associate learning under simultaneous repetition and nonrepetition conditions", J. Exp. Psychol., 1962, 64, 87-93.

Battig, W.F. and Brackett, H.R., "Comparison of anticipation and recall methods in paired-associate learning", Psychol. Rep., 1961, 9, 59-65.

Battig, W.F. and Spera, A.J., "Rated association values of numbers from 0-100", J. Verb. Learn. Verb. Behav., 1962, 1, 200-202.

Birnbaum, A., "Statistical theory of logistic mental test models with a prior distribution of ability", J. Math. Psychol., 1969, 6, 258-276.

Bower, G.H., "Application of a model to paired-associate learning", Psychometrika, 1961, 26, 255-280.

Bower, G.H., "A model for response and training variables in paired-associate learning", Psychol. Rev., 1962, 69, 34-53.

Bower, G.H., "Notes on a descriptive theory of memory", in D.P. Kimble (Ed.), Proceedings of the second conference on learning, remembering, and forgetting., New York: Academic., 1967, pp. 112-185.

Brunswik, E., "The conceptul framework of psychology", Univ. of Chicago Press., 1969.

Bush, R.R. and Mosteller, F., "Stochastic models for learning", New York: Wiley, 1955.

Bush, R.R. and Mosteller, F., "A comparison of eight models", in R.R. Bush and W.K. Estes (Eds.), Studies in mathematical learning theory., Stanford: Stanford Univer. Press, 1959, pp. 293-307.

Bush, R.R. and Sternberg, S., "A single-operator model", in R.R. Bush and W.K. Estes (Eds.), Studies in mathematical learning theory., Stanford: Stanford Univer. Press., 1959, pp. 204-214.

Bush, R.R. and Wilson, T.R., "Two choice behaviour of paradise fish", J. Exp. Psychol., 1956, 51, 315-322.

Bush, R.R., Galanter, E. and Luce, R.D., "Tests of the beta model", in R.R. Bush and W.K. Estes (Eds.), Studies in mathematical learning theory., Stanford: Stanford Univer. Press., 1959, pp. 382-399.

Calfee, R.C. and Atkinson, R.C., "Paired-associate models and the effects of list length", J. Math. Psychol., 1965, 2, 254-265.

Carroll, J.B. and Burke, M.L., "Parameters of paired-associate verbal learning: length of list, meaningfulness, rate of presentation, and ability", J. Exp. Psychol., 1965, 69, 543-553.

Cohen, J.C. and Musgrave, B.S., "Effect of meaningfulness on cue selection in verbal paired-associate learning", J. Exp. Psychol., 1964, 68, 284-291.

Crothers, E.J., "Paired-associate learning with compound response", J. Verb. Learn. Verb. Behav., 1962, 1, 66-70.

Duncanson, J.P., "Intelligence and the ability to learn", Princeton, N.J.: Educational Testing Service, 1964.

Estes, W.K., "The problem of inference from curves based on group data", Psychol. Bull., 1956, 53, 134-140.

Estes, W.K., "Learning theory and the new mental chemistry", Psychol. Rev., 1960, 67, 207-223.

Estes, W.K., "New developments in statistical behavior theory: differential tests of axioms for associative learning", Psychometrika, 1961, 26, 73-84.

Gagne', R.M. (Ed.), "Learning and individual differences", Colombus, Ohio: Merrill, 1967.

Games, P.A., "A factorial analysis of verbal learning tasks", J. Exp. Psychol., 1962, 63, 1-11.

Glaser, R., "Some implication of previous work on learning and individual differences", in R.M. Gagne' (Ed.), Learning and individual differences., Colombus, Ohio: Merrill, 1967, pp. 1-18.

Glaze, J.A., "The association value of non-sense syllables", J. Genet. Psychol., 1928, 35, 255-267.

Greeno, J.G., "Paired-associate learning with short-term retention: mathematical analysis and data regarding identification of parameters", J. Math. Psychol., 1967, 4, 430-472.

Greeno, J.G., "Identifiability and statistical properties of two-stage learning with no successes in the initial stage", Psychometrika, 1968, 33, 173-215.

Greeno, J.G. and Steiner, T.E., "Markovian processes with identifiable states: general considerations and application to all-or-none learning", Psychometrika, 1964, 29, 309-333.

Gregg, L.W. and Simon, H.A., "Process models and stochastic theories of simple concept formation", J. Math. Psychol., 1967, 4, 246-276.

Hayes, K.J., "The backward curve: a method for the study of learning", Psychol. Rev., 1953, 60, 269-275.

Hull, C.L., "The meaningfulness of 320 selected nonsense syllables", Amer. J. Psychol., 1933, 45, 730-734.

Hull, C.L., "The place of innate individual and species differences in a natural-science theory of behaviour", Psychol. Rev., 1945, 52, 55-60.

Jenkins, J.J., "Individual differences in verbal learning", in R.M. Gagne' (Ed.), Learning and individual differences., Colombus, Ohio: Merrill, 1967, pp. 45-57.

Jensen, A.R., "Varieties of individual differences in learning", in R.M. Gagne' (Ed.), Learning and individual differences., Colombus, Ohio: Merrill, 1967, 117-140.

Kelley, H.P., "A factor analysis of memory ability", Princeton: Princeton University and Education Testing Service, 1954.

Krueger, W.C.F., "The relative difficulty of nonsense syllables", J. Exp. Psychol., 1934, 17, 145-153.

Mandler, G., "Associative frequency and associative prepotency as measures of response to nonsense syllables", Amer. J. Psychol., 1956, 68, 662-665.

Melton, A.W., "Individual differences and theoretical process variables: general comments on the conference", in R.M. Gagne' (Ed.), Learning and individual differences., Colombus, Ohio: Merrill, 1967, pp. 238-252.

Merrel, M., "The relationship of individual growth to average growth", Hum. Biol., 1931, 3, 37-40.

Noble, C.E., "Measurements of association value (a), rated associations (a'), and scaled meaningfulness (m') for the 2100 CVC combinations of the English alphabet", Psychol. Rep., 1961, 8, 487-521.

Noble, C.E., "An analysis of meaning",Psychol.Rev.,1952,59,421-430.

Noble, C.E., Stockwell, F.E. and Pryer, M.W., "Meaningfulness (m') and association value (a) in paired associate syllable learning", Psychol. Rep., 1957, 3, 441-452.

Noble, C.E., Noble, C.L. and Alcock, W.T., "Prediction of individual differences in human trial-and-error learning", Percept. Mot. Skills, 1958, 8, 151-172.

Norman, M.F., "Incremental learning on random trials", J. Math. Psychol., 1964, 1, 336-350.

Offir, J.D., "Stochastic learning models with distributions of parameters", J. Math. Psychol., 1972, 9, 407-417.

Paivio, A., "Imagery and verbal processes", Holt, Rinehart and Winston, Inc., 1971.

Peterson, L.R., "Short-term verbal memory and learning", Psychol. Rev., 1966, 73, 193-207.

Polson, P.G., "Statistical methods for a general theory of all-or-none learning", Psychometrika, 1970, 35, 51-72.

Restle, F., "The significance of all-or-none learning", Psychol. Bull., 1965, 62, 313-324.

Reynolds, B. and Adams, J.A., "Psychomotor performance as a function of initial level of ability", Amer. J. Psychol., 1954, 67, 268-277.

Sidman, M., "A note on functional relations obtained from group data", Psychol. Bull., 1952, 49, 263-269.

Stake, R.E., "Learning parameters, aptitudes, and achievements", Psychometric Monogr., No.9, 1961.

Steiner, T.E. and Greeno, J.G., "An analysis of some conditions for representing N state Markov processes as general all or none models", Psychometrika, 1969, 34, 461-487.

Sternberg, S.H., "Applications of four models to sequential dependence in human learning", in R.R. Bush and W.K. Estes (Eds.), Studies in mathematical learning Theory., Stanford: Stanford Univer. Press, 1959, pp. 340-381.

Sternberg, S.H., "Stochastic learning theory", in R.D. Luce, R.R. Bush and E. Galanter (Eds.), Handbook of Mathematical Psychology, Vol. II., New York: Wiley, 1963, pp. 1-120.

Suppes, P. and Ginsberg, R., "A fundemantal property of all-or-none models", Psychol. Rev., 1963, 70, 139-161.

Travers, R.M.W., "Learning measures and psychometric variables", in R.M. Gagne' (Ed.), Learning and individual differences., Colombus, Ohio: Merrill, 1967, pp. 19-22.

Tulving, E. and Madigan, S.T., "Memory and verbal learning", Annual Rev. Psychol., 1970, 21, 437-484.

Underwood, B.J. and Schulz, R.W., "Meaningfulness and verbal learning", Chicago: Lipincott, 1960.

Underwood, B.J., Rehula, R. and Keppel, G., "Item selection in paired-associate learning", Amer. J. Psychol., 1962, 75, 353-371.

Waugh, N.C. and Norman, D.A., "Primary memory", Psychol. Rev., 1965, 72, 89-104.

Witmer, L.R., "The association value of three-place consonant syllables", J. Genet. Psychol., 1935, 47, 337-360.

Woodrow, H.A., "The ability to learn", Psychol. Rev., 1946, 53, 147-158.

Zeaman, D. and Kaufman, H., "Individual differences and theory in a motor learning task", Psychol. Monogr., 69 (Whole No. 391), 1955.