**This dissertation has been
microfilmed exactly as received 67-6384**

**WING, John Fiske, 1931-
A STIMULUS-TRACE MODEL OF SERIAL ACQUISITION
AND RETENTION.**

**The Ohio State University, Ph.D., 1966
Psychology, experimental**

University Microfilms, Inc., Ann Arbor, Michigan

A STIMULUS-TRACE MODEL OF SERIAL ACQUISITION AND RETENTION

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

John F. ^{Fiske}Wing, A.B., M.A.

The Ohio State University

1966

Approved by

William H. Wickens

Adviser

Department of Psychology

ACKNOWLEDGMENTS

The writer wishes to express his gratitude to Dr. R. C. Miles, Dr. Delos D. Wickens and Dr. Donald R. Meyer for their helpful criticisms during the preparation of this dissertation. Appreciation is also due Dr. William B. Houston, Jr., Assistant Professor of Mathematics, Antioch College, who made mathematical derivations which are employed in the model developed here.

The writer is also very grateful to his wife, Shelley, for her help in the preparation of the many figures used in this manuscript.

CONTENTS

	Page
ACKNOWLEDGMENTS	ii
FIGURES	v
TABLES	vii

PART I INTRODUCTION

DEVELOPMENT OF THE STIMULUS-TRACE HYPOTHESIS	3
POSTULATED TRACE FUNCTIONS	4
EXPERIMENTAL EVIDENCE RELEVANT TO THE STIMULUS-TRACE HYPOTHESIS	4
Retrograde Amnesia Studies (FS-ECS and SD-ECS Intervals)	6
Delay of Response Studies (CS-CR Interval)	11
Delay of Reward Studies (CR-UCS and CR-CS Intervals)	16
Interstimulus Interval Studies (CS-UCS Interval)	20
Delay of Judgment Studies in Human Psychophysics (S-V Interval)	28
Delay of Stimulus-Reproduction by Humans (S ₁ -R ₁ Interval)	36
Summary of Evidence	46

PART II STIMULUS-TRACE INTERACTION MODEL

PROBLEM OF BACKWARD ASSOCIATIONS AND ITS RESOLUTION	52
PRINCIPLE OF ASSOCIATIVE ASYMMETRY	55
COMPUTATIONAL FORMULAS FOR FORWARD AND BACKWARD ASSOCIATIONS	57
PROPERTIES OF ONE TRIAL ASSOCIATION MATRIX	61
PROPERTIES OF MULTI-TRIAL ASSOCIATION MATRIX	63
POINT OF ENTRY AND DIRECTION OF RECALL IN A SERIES	66

PART III APPLICATIONS OF THE MODEL

REMOTE ASSOCIATION GRADIENTS	69
SERIAL-POSITION CURVES IN SERIAL RECALL	77

SERIAL-POSITION CURVES IN FREE RECALL	80
Stimulus Materials	83
Procedure	83
Results and Discussion	84
SERIAL-POSITION CURVES IN MIXED STRATEGY RECALL	90

PART IV. EVALUATION OF THE MODEL

STRENGTHS OF THE MODEL	99
WEAKNESSES OF THE MODEL	101
POSSIBLE RE-FORMULATION OF THE MODEL ALONG STOCHASTIC LINES	101

PART V. APPENDIXES

APPENDIX

A. CALCULATIONS FOR PETERSON STUDY	103
B. DERIVATION OF FORMULAS	105
C. MATRICES AND CALCULATIONS FOR REMOTE ASSOCIATION GRADIENTS	107
D. STIMULUS MATERIALS FOR FREE RECALL	111
E. MATRICES AND CALCULATIONS FOR FREE RECALL	113
F. CALCULATIONS FOR MIXED STRATEGY RECALL	116
REFERENCES	117
AUTOBIOGRAPHY	125

FIGURES

Figure	Page
1. Postulated Stimulus-trace Function	5
2. Stepdown as a Function of ECS Delay	9
3. Performance by Various Species as a Function of Response Delay	12
4. Rate of Learning as a Function of Response Delay	15
5. Performance as a Function of Primary or Secondary Reward Delay	17
6. Approach to a Positive, Secondary Reinforcer (CS) as a Function of CS-UCS Interval During Instrumental Training	21
7. Avoidance of a Negative, Secondary Reinforcer (CS) as a Function of CS-UCS Interval During Instru- mental Training	23
8. Level of Classical Conditioning as a Function of CS- UCS Interval.	27
9. Kohler's Fading Stimulus Trace as Inferred from the Point of Subjective Equality in Psychophysics	30
10. Effects of Interpolated Stimuli on Time Order Error . .	32
11. Effects of Preceding Trials on Time Order Error	33
12. Effects of Preceding Experiments on Time Order Error . .	35
13. Short-term Memory Functions: Early Studies.	37
14. Short-term Memory Functions: Later Studies	41
15. Comparison Between Cumulative Correct Recall and Log Effective Trace-Strengths at Successively Longer Latencies of Recall	43
16. Postulated Events in the Peterson, Short-term Memory Study	45
17. Stimulus Traces for a Series of Ten Stimuli	51

18.	Diagrammatic Representation of Forward and Backward Associations	54
19.	Theoretical Remote Association Gradients for Single-Trial vs. Multi-Trials	62
20.	Comparison between Theoretical and Empirical Remote Association Gradients for a 10-item List	72
21.	Comparison between Theoretical and Empirical Remote Association Gradients for an 8-item List	74
22.	Theoretical vs. Empirical Serial-position Curves for 5, 6, 7, 8 and 9 Item Lists Presented at 1 Item per Second	80
23.	Theoretical vs. Empirical Serial-position Curves for First-Presented Lists of 10, 15 and 20 Items (1-sec. Rate)	85
24.	Theoretical vs. Empirical Serial-position Curves for First-Presented Lists of 10, 15 and 20 Items (2-sec. Rate)	86
25.	Theoretical vs. Empirical Serial-position Curves for Second-Presented Lists of 10, 15 and 20 Items (1-sec. Rate)	87
26.	Theoretical vs. Empirical Serial-position Curves for Second-Presented Lists of 10, 15 and 20 Items (2-sec. Rate)	88
27.	Theoretical Serial-position Curves for Five-item Lists Presented at 2, 4, 8 and 10 Items per Second	93
28.	Theoretical Serial-position Curves for Seven-item Lists Presented at 2, 4, 8 and 10 Items per Second	94
29.	Theoretical Serial-position Curves for Nine-item Lists Presented at 2, 4, 8 and 10 Items per Second	95
30.	Empirical Serial-position Curves for Five-item Lists Presented at 2, 4, 8 and 10 Items per Second	96
31.	Empirical Serial-position Curves for Seven-item Lists Presented at 2, 4, 8 and 10 Items per Second	97
32.	Empirical Serial-position Curves for Nine-item Lists Presented at 2, 4, 8 and 10 Items per Second	98

TABLES

Table	Page
1. Summary of Evidence Relating to the Stimulus Trace	47
2. One-trial Association Matrix for a 10-item List	60
3. Multi-trial Association Matrix for a 10-item List	64
4. Multi-trial, Free-association Matrix for a 10-item List . .	71
5. One-trial Association Matrices for 5, 6, 7, 8 and 9-item Lists	78
6. Recall Order of the First Three Responses on First and Second Presented Lists	89
7. McNemar Test for Significance Changes	91

PART I

INTRODUCTION

One of the problems which challenged the earliest psychologists was how a stimulus series becomes associatively connected. The earliest explanation was possibly by Aristotle (ca. 350 B.C.), but a detailed formulation of the probable events in serial acquisition was not made until 1749 by Hartley and subsequently by Herbart (1824-25). It was Herbart's model which provided the framework for Ebbinghaus' classic experimental work. In Ebbinghaus' description of the model (1913, p. 94), "... associative threads, which hold together the remembered series, are spun not merely between each member and its immediate successor, but beyond intervening members to every member which stands to it in any close temporal relation. The strength of the threads varies with the distance of the members, but even the weaker of them must be considered as relatively of considerable significance." These associative threads which bridged intervening members were termed remote associations. Backward associations (both adjacent and remote) were also postulated, but they were conceived as being of lesser strength.

But what process spins these associative threads? The conception held by many psychologists was that each member of a series left a stimulus trace which fused with succeeding traces as long as it was active. Herbart's view of the fusing of stimulus traces, as quoted by Ebbinghaus (1913, p. 93), was

as follows:

Let a series, a, b, c, d, ... be given in perception, then a, from the very first moment of the perception and during its continuance, is subjected to inhibition by other ideas present in consciousness. While a, already partially withdrawn from full consciousness, is more and more inhibited, b comes up. The latter, at first uninhibited, fuses with the retiring a. c follows and, itself uninhibited, is united with the fast dimming b and the still more obscured a. In a similar fashion d follows and unites itself in varying degrees with a, b, and c. Thus there originates for each of these ideas a law according to which, after the whole series has been forced out of consciousness for some time, in its own way on its renewed appearance each idea struggles to call up every other idea of the same series.

Thus, at the very beginning of experimental psychology, a definite model was provided for relating association formation to the fusion or interaction of stimulus traces. Yet with the partial exception of work by Hull et al. (1940) and Hull (1943, 1951, 1952), this specific model of Herbart's has lain dormant for over half a century.¹ In the present paper Herbart's basic concepts are once again revived, and an analysis is presented which shows how interactions among a temporal series of stimulus traces will generate the adjacent and remote, forward and backward associations which are characteristic of serial learning.

¹ Hull utilized the concept of trace interaction to account for anticipatory errors based on forward associations but did not attempt to derive perseverative errors based on backward associations, although he early foresaw the need for doing so (cf. Hull et al., 1940, p. 58) and later adopted a theorem to rectify the omission (cf. Hull, 1952, pp. 107-119).

Development of the Stimulus-trace Hypothesis

Gomuliki (1953) has shown that some form of stimulus-trace theory has been in existence since the Greeks. However, active speculation about just how perseveration and consolidation of the stimulus trace could account for experimental, laboratory findings did not begin until the turn of the century. At that time, Müller and Pilzecker (1900) utilized a trace consolidation hypothesis to account for retroactive inhibition, and McDougall (1901) and Burnham (1903) utilized it to explain retrograde amnesia. Pavlov evidently used the trace concept as early as 1906 in his explanation of conditioned reflexes (Pavlov, 1928, p. 93). Later Köhler (1923) and Hull (1937) speculated about the physical properties of the stimulus trace. Finally, as part of a general theory of rote learning, Hull and his associates (1940) supplied mathematical functions to represent the significant properties of the trace. They proposed that the incremental segment of the trace follows a positive growth function of the form

$$(1) \quad X = A - Ae^{-bt}$$

and that the perseverative segment follows a negative growth function of the form

$$(2) \quad X = Ae^{-bt}$$

These particular equations were never actually used by Hull and his associates, but it was their belief that they would "be required in subsequent developments of the theory of rote learning" (Hull et al., 1940, p. 43).

It is these general formulas which have been adopted in the present model.²

Postulated Trace Functions

The dotted curve in Figure 1 shows the specific incremental and perseverative stimulus-trace functions used in this study. The trace starts at zero and rises exponentially, reaching unit height in one-half second; then it undergoes a first-order, exponential rate of decay. In the perseverative function shown here the trace drops to two-thirds of its preceding value every two seconds. This rate of change brings the trace to five per cent of its initial value in 15 sec., to one per cent in 23 sec., and to one hundredths of a percent in about 34 sec. Hull generally assumed that 30 sec. would be the maximum length of time the trace would be active (Bugelski, 1950, p. 337). Although there is no definitive, empirical support for this particular perseverative function, there is considerable experimental support for a stimulus trace of this approximate shape and duration. Evidence for this statement will be examined before considering the problem of trace interaction.

Experimental Evidence Relevant to the Stimulus-trace Hypothesis

The purpose of this study is to develop a stimulus-trace theory which will account for data on human verbal learning. Nevertheless, the trace

² Hull went on to develop a different, stimulus-trace function based upon the empirical CS-UCS data from classical eyeblink conditioning; however, we shall have occasion to question the usefulness of CS-UCS data for drawing conclusions about the stimulus trace (cf. pp. 19-28).

Fig. 1. Postulated Stimulus-trace Function

hypothesis itself derives most of its support from animal experiments, and it will be advantageous to review quite a few of these. The first set of studies concerns the ability of electroconvulsive shock (ECS) to produce in rats a retrograde amnesia (RA) for a one-trial, passive avoidance response. The problem is how long after a foot-shock will ECS erase a rat's memory for that shock and for his reflex withdrawal from it? The second set of studies concerns how long various species can remember which of two (or three) stimuli they observed so that after this enforced delay they can act differentially according to that knowledge. A third, closely related set of studies concerns a comparable delay-of-reward: How long after a response has been made will rats be able to associate the response with its positive or negative outcome? A fourth set of animal studies deals with CS-UCS delay: How much delay can intervene between a neutral stimulus and a reinforcer before the rat fails to respond to the neutral stimulus as a reward-cue or secondary reinforcer? A fifth set of experiments concerns the ability of humans to make psychophysical judgments as a function of the time lag between the standard and variable stimuli. A sixth line of research deals with human Ss' ability to recall single (or paired) verbal items after selected delays.

Retrograde amnesia studies (FS-ECS and SD-ECS intervals)

The memory trace hypothesis assumes that the fixation of an impression depends upon a physiological process which takes a certain interval of time to occur. If the hypothesis is correct, it should be possible to find agents (drugs, shock, etc.) which would interfere with the fixation process. Fur-

thermore, the amount of interference produced by an agent should depend upon how much time has elapsed between initiation of the trace and the point at which the trace process is interrupted. Glickman (1961) reviewed the results of numerous clinical and experimental studies dealing with the effects of various agents on fixation of experience. These agents included psychological trauma, electroconvulsive shock (ECS), anoxia, hyper- and hypothermia, anesthesia, and cortical stimulation. He concluded that the overall weight of the evidence favored the existence of perseverative neural processes. However, in a more analytical and detailed review, Lewis and Maher (1965) concluded that there is little, if any, support for the memory consolidation hypothesis, and they offer as an alternative explanation that ECS produces a Pavlovian inhibition to the stimuli of the ECS situation. One of their most important criticisms of the many studies they reviewed concerned the length of the interval between learning and administration of ECS. The intervals employed by most investigators were so long (often hours and days) that Lewis and Maher were forced to conclude that:

It seems quite possible that the consolidation notion has not yet been tested at all. If engram formation takes the brief span of time that usually defines immediate memory, then most of the learning-ECS intervals used by investigators are too long to disrupt it. Furthermore, it seems difficult to regard any study in which the subject has had more than one trial as appropriate to the problem, as here again the engram formed by the first trial and other early trials must surely be consolidated by the time the investigator gets to the ECS stage of investigation (p. 237).

Indeed, one-trial, passive, avoidance studies even more recent than their review add support to Lewis and Maher's contention that if ECS does produce amnesia for an event, it does so only if introduced within a very

brief period following the event. In these studies rats are placed on a raised island in the middle of a charged grid. As soon as they step down, they receive a foot-shock (FS). At selected intervals following FS various sub-groups are given ECS. The measure of RA is either the proportion of Ss who forget the FS and step down on the next occasion or the median latency of such step-down responses. In one of the most carefully controlled studies to date (Quartermain, Paolino, & Miller, 1965), ECS was found to produce significant RA for an event only if it was introduced within 30 sec. or less following the event. The closer in time it followed the event, the greater the percentage of subjects showing RA. This is shown in Figure 2A. Approximately 90 per cent showed amnesia when ECS occurred within one sec. This percentage dropped steadily to about 20 per cent when ECS was delayed for 30 sec. When ECS was delayed by 60 sec., the percentage of Ss who stepped down on the second trial did not differ significantly from that of a control group.

One problem, however, with this and many other one-trial, avoidance studies is that a detailed analysis has not been made of the aversiveness of ECS itself. Chorover and Schiller (1965) attempted to solve this problem; that is, they tried to discover any aversive effects of the ECS and to separate them from amnesic effects. They did this by testing "ECS only" groups at selected delays following stepdown. They found evidence that ECS does produce avoidance. But examination shows that the ECS avoidance gradient is also compatible with a stimulus-trace hypothesis.

The gradient for ECS avoidance is shown in Figure 2B, alongside the gradient for foot-shock avoidance. The gradients are based upon the mean

Fig. 2. Stepdown as a Function of ECS Delay

latency of the stepdown, avoidance response. The FS-ECS gradient shows that when the delay between FS and ECS is very brief the animal has apparently "forgotten" that the stepdown produced a foot-shock, and he steps down with a short latency. The latency of avoidance was smallest at the .5-sec. interval and increased significantly ($p < .001$) up to 10-sec. For the 10-sec. and 30-sec. groups the mean latencies did not differ significantly from the control group's latency of 24-sec. (Maximum latency was 30-sec., after which any animal was arbitrarily removed from the test chamber.) For these long delay groups, it is assumed that the association between the step-down response and the foot-shock has already been consolidated so that ECS cannot interfere with the process.

The SD-ECS avoidance gradient similarly shows that with very brief delays the animal has forgotten that the step-down response produced ECS after-effects (presumably because ECS itself has wiped-out the trace for the step-down response), and so Ss make short latency stepdowns. These become increasingly longer, however, up to a 5-sec. delay as the probability of remembering step-down increases. After a point between 5 and 10-sec., however, Ss no longer are able to associate the step-down response with ECS even though they remember it, and so they begin again to make short latency step-down responses. In Chorover's opinion both these gradients, then, support a very brief trace consolidation hypothesis. He concludes that:

If the present findings of a rapid consolidation process dissociable from aversive effects can be confirmed and extended to other behavioral situations, it might be possible to modify that portion of the consolidation hypothesis which implies that stimulus-specific neuroelectrical activity out-lasts a stimulus by many minutes or hours. A modification of this sort would be particularly desirable since there seems to be no independent electrophysiological evidence for the existence of such perseverative activity after learning trials. Furthermore, it is difficult to

conceive of a parsimonious way in which the nervous system might handle such activity. A possible but cumbersome way for the system to do so would be to have a filtering, shunting network maintain the individual integrity of many independently coexisting traces. Such a system would probably, in addition, have to monitor the activity in each circuit in order to inhibit mutual interference among certain traces while inducing mutual facilitation among others. A system possessing rapid consolidation would require a somewhat less complex channeling network and would circumvent most of these requirements. Thus, if the initial neuroelectrical events related to learning were able to initiate the subsequent structural encoding process within a few seconds, it would be possible for stimulus-specific perseverative activity to decline rapidly without an accompanying loss of information. In addition, rapid consolidation would tend to minimize the occurrence of mutual interference among contemporary active traces during the assimilation of successively presented stimuli. (Chorover, 1965, pp. 261-262).

The most recent evidence from studies of ECS-produced RA suggests that the period of consolidation of the stimulus trace lasts between 5 sec. to 10 sec. In the following section, this finding will be compared with the length of time animals can remember a cue-stimulus upon which a differential delayed-response must be based.

Delay of response studies (CS-CR interval)

The second source of experimental data comes from classical delayed-response experiments. In these experiments the experimenter presents to the animal all the information necessary for making a choice response, but he enforces a delay in actually making the response. This CS-CR delay is usually of the order of from 0 sec. to 30 sec.

The effects of enforced delay on the performance of various species of mammals and primates is summarized in Figure 3, which is adapted from Fletcher (1965, p. 133). These curves indicate a relationship between

Fig. 3. Performance by Various Species as a Function of Response Delay

level of performance and phylogenetic status, although Fletcher has cautioned that experimental methods differed in testing these species and, therefore, comparisons of the absolute levels of performance should not be made. However, for our purposes, the significance of Figure 3 does not lie in differences in absolute level of performance, but rather in the fact that performance falls off over such a narrow range of delay intervals, regardless of the species. That is, all species appear to reach their asymptotic performance between about 15 sec. and 35 sec. This fact would seem to provide further evidence of the fixation or consolidation of the memory trace within a maximum of 30 sec. However, Fletcher has argued that the delayed-response problem cannot be viewed primarily as a simple test of recent memory because these curves would indicate that "a beast as bright as the rhesus monkey has obvious trouble remembering for as little as 20 sec., and the cat has virtually forgotten during the same delay period" (1965, p. 132). These conclusions defy common sense, and Fletcher rightly asserts that these observed decrements in performance must be due to variables which operate during the interval of delay and not to the fact of temporal delay itself. He argues that the critical feature in the delayed-response task is the maintenance during the delay of an orienting response-chain leading to the correct solution, and that the opportunities for distractibility and therefore of interference with the orienting response are cumulative with time. These cumulative distractions produce the observed decline in performance. Fletcher's analysis is heuristic and directs attention to those experimental variables which might operate during delay to produce interference. At the level of theory, however, his analysis overlooks a very pertinent question: Why do distractions only seem

to impair performance over a brief period of approximately 30 sec.? For the cat and raccoon, of course, a "chance" level of performance is reached at about this time and no further impairment could be obtained. But for the rhesus, gibbon and marmoset, performance levels off at well above chance level, and this indicates that distractibility does not affect performance for delays greater than 30 sec. Although there may be many promising hypotheses to account for this fact, certainly one of the most plausible hypotheses is the stimulus-trace hypothesis. It would account for these data by asserting that the correct orienting response chain is fixated in some way within 30 sec., and therefore it is no longer susceptible to interference.

The kinds of studies which provided the data for Figure 3 were generally ones in which the goal was in sight during the delay interval. Numerous cases of postural adjustment (and other forms of mediating responses) were reported and must have resulted in an overestimation of the delayed-response function. However, an experiment by Smith (1951) with rats effectively eliminated postural cues as an aid to delayed-response performance. He gave five, independent delay groups of rats a successive, black-white discrimination task. The rats were exposed to either a black or a white stimulus alley, then were delayed in a neutral (gray) alley, and finally were allowed to enter either a right or a left goal box. The right goal was correct if the stimulus alley was white, and the left goal was correct if the stimulus alley was black. Position habits of all five delay groups were equalized prior to the experimental phase. Smith then ran all groups until a criterion of 18 correct out of 20 trials was reached. Figure 4 compares the rate of learning for these delay-or-response groups. These data, in

Fig. 4. Rate of Learning as a Function of Response Delay

which postural attitudes have been minimized as differential cues-for-goal-pointing, show a marked foreshortening of the delay-of-response gradient. Smith concluded that the stimulus traces initiated by the black or white stimulus alleys decreased so rapidly in time that they became almost indiscriminable within 5 sec. This, of course, is about the same duration and same sort of function reported by ECS studies. In the next section we shall have occasion to compare both these stimulus-trace gradients with those for various delayed-reward studies.

Delay of reward studies (CR-UCS and CR-CS intervals)

In instrumental conditioning paradigms we can also control the time interval between an instrumental act and its outcome (positive or negative), and we can determine from a group's performance what proportion of the animals at each delay-of-reward interval associate the act with its outcome. When immediate secondary reinforcement has been eliminated, such studies also typically show an exponential curve with a maximum at zero sec. and an asymptote at between 5-15 sec. First, let us review delay of positive, primary reinforcement. Figure 5A shows Grice's data on white rats for trials to criterion in a simultaneous (black-white), discrimination-learning task as a function of delay of reward. This function shows that the animals learned most rapidly with zero delay, but encountered great difficulty with a delay as great as 5 sec. Most of his animals failed to learn at all with a 10-sec. delay even when given as many as 1,400 trials (data not shown in this graph). Riesen (1940) reports very similar gradients for individual chimpanzees who were tested at 0, 1, 2, 4, and 8 sec. Keesey (1964) also

Rate of learning as a function of delay of reward. The experimental values are represented by black dots, and the smooth curve is fitted to these data. *Adapted from Grice (1948)*

Rate of learning as a function of the delay of reward. The curve has been visually fitted to the data points. *Adapted from Keesey (1964)*

Rate of extinction as a function of the delay of punishment. Curve visually fitted to data which were transformed from data obtained by Kamin (1959) as reported in Kimble (1961).

Delay-of-reinforcement gradient obtained when the delay period is the interval between the response and the cessation of a stimulus previously associated with pain (Kamin, 1957a, 1957b).

Fig. 5. Performance as a Function of Primary or Secondary Reward Delay

obtained very similar results using a discriminative bar-press task with intracranial self-stimulation (ICS) as the reinforcer. Figure 5B shows the mean number of errors in discrimination learning as a function of the delay of reward. Again, the rate of learning is maximal at zero, falls off rapidly, and then reaches asymptote at a 5-sec. delay or slightly longer. Of course, inclusion of these data on intracranial self-stimulation is only appropriate if ICS acts in all respects like other primary reinforcers.³

Second, let us review studies on delay of negative, primary reinforcement. Kamin (1959) trained rats to avoid shock in a shuttlebox by crossing-over to the opposite compartment; but then, during an extinction session, five sub-groups were each punished with electric shock for making the previously correct, avoidance response. The shock for the sub-groups was delayed by either 0, 10, 20, 30, or 40 seconds following the avoidance response. Punishment was found to reduce markedly trials to extinction. Median number of responses in extinction for these groups were 3.0, 5.5, 10.5, 9.5, and 9.5. Although these medians confirm the presence of a delay-of-punishment gradient, the data need to be transformed to produce a function comparable to those in Figures 5A and 5B. Properly, the raw data scores should be transformed ($\text{reciprocal} \times 100$). To approximate the function which would be obtained by this procedure, I have transformed the medians themselves and visually-fitted a curve to them as shown in Figure 5C.

³ Although this delay-of-reward gradient for intracranial self-stimulation agrees with the gradient of Grice as well as with gradients obtained with other kinds of rewards, the temporal properties of self-stimulation are sometimes variable and "atypical" (cf. Valenstein, 1964).

Again, this curve shows the greatest effect of punishment at zero delay, an abrupt exponential decline, and then an asymptote at 10 sec. or greater.

Taken together, the curves of Figure 5(A-C) are grossly similar.

In all cases the delay interval has been from the occurrence of an instrumental response (CR) to the occurrence of some response-contingent outcome or reinforcement (UCS); the latter was either a positive primary reinforcer (food, intracranial stimulation) or else a primary negative reinforcer (shock).

Third, let us examine the delay of positive, secondary reinforcement. Kamin (1957a, 1957b) has studied the positive secondary reinforcement gradient. In these studies he investigated the temporal delay between an instrumental response (CR) and the cessation of a secondary, noxious stimulus (CS). In these studies the response was crossing to the other side of a shuttle box, and the secondary reinforcement consisted of the cessation of a buzzer which had been previously paired with shock using a fixed CS-UCS interval. The results for both experiments are summarized in Figure 5-D. The rate of avoidance learning as a function of delay intervals declines exponentially. Learning was much better when the instrumental response immediately terminated the (noxious) buzzer. If a 5-sec. or 10-sec. delay occurred, animals were seldom able to associate their response with the cessation of the cue stimulus. Thus, for secondary as well as primary reinforcement, the maximum delay which will produce any learning appears to be approximately 10 sec. to 15 sec. These results on delay of reward fully confirm the ECS and delay of response results.

Interstimulus interval studies (CS-UCS interval)

A fourth set of studies concerns the interval between a neutral stimulus and a primary reinforcer: the CS-UCS interval. How much delay can intervene before the animal fails to associate the CS with the UCS, i.e., fails to treat the former as a reward-cue or secondary reinforcer?

First, let us examine the CS-UCS interval in instrumental conditioning where the UCS is a positive, primary reinforcer. A study by Jenkins (1950) provides data on the interstimulus interval (CS-UCS) in this situation. Without any bar present, Jenkins paired a 3-sec. buzzer as a CS with food pellets as UCS in an operant test chamber. He assigned rats to each of five CS-UCS intervals (1, 3, 9, 27 or 81 sec.) or to a control group that did not receive the CS for food. The CS-UCS intervals were measured from time of offset of the buzzer to dropping of the food pellet. Since the buzzer did not stay on during delay intervals this represents trace conditioning. Following training, the bar was introduced. The test for how much secondary reinforcing power the CS had acquired was bar-press rate during a 6-hour period in which pressing the bar produced only the buzzer (CS). Figure 6 shows the acquired reinforcement value of the CS as a function of CS-UCS delay in terms of total number of bar-presses in the 6-hour test period. Kimble (1961, p. 184) says of these data that "... the amount of responding decreases as a function of the interval; but again, too, long intervals appear to have some effect." This latter statement, however, is not really justified by Jenkins' own report (1950, p. 241). He found large variances and specifically reports that the 27-sec. and 81-sec. delay

Fig. 6. Approach to a Positive, Secondary Reinforcer (CS) as a Function of CS-UCS Interval During Instrumental Training

groups did not differ significantly from the control group (t-tests, 5% level). The longest delay interval which differed significantly ($p < .04$) from the control group was the 9-sec. group. Thus, over the intervals examined, these results are compatible with the other data already examined.

A second study (Bersh, 1951) also paired a light (CS) with food (UCS) in an operant situation with the bar temporarily removed. However, Bersh used a delayed conditioning technique, and under delayed conditioning the CS stays on during the interstimulus interval so that there is not necessarily any differential conditioning as a function of CS trace decay. Bersh's data, therefore, are not pertinent to the trace hypothesis.

Now let us examine the CS-UCS interval in instrumental conditioning where the UCS is a negative, primary reinforcer. Kamin (1954) studied this problem. He required dogs to jump a barrier in a shuttle box. The CS was a 2-sec. buzzer and the UCS was a shock delayed for independent groups either 5, 10, 20 or 40 sec. This was a trace conditioning technique. Kamin required Ss to make five successive avoidances, and then he extinguished them to a criterion of five successive failures to respond to the CS. Both acquisition and extinction measures were significantly affected by delay of shock. The extinction data are shown in Figure 7-A. (Very similar results have also been reported by Church, Brush and Solomon, 1956.) Kamin reported a statistically significant trend based upon these data for the 5, 10 and 20-sec. groups. The 40-sec. group data were confounded by six of the ten dogs showing "spurious" avoidances (i.e., jumping responses unrelated to the CS). In Kamin's own words (1954, p. 70)

Fig. 7. Avoidance of a Negative, Secondary Reinforcer (CS) as a Function of CS-US Interval During Instrumental Conditioning

With the 40-sec. animals, the very high rate of spontaneous jumping and concomitant diffuse emotionality obscured all other aspects of behavior. The most general statement that can be made about these data is this: as CS-US interval increased, the quality of the emotional behavior passed from a focalized (about the CS) to a diffuse nature.

These observations suggest that the majority of the dogs in the 40-sec. delay group had experienced such a long delay that they had not associated the buzz with the shock. However, the fact that the remaining four of his six dogs did learn to avoid with 40-sec. delays requires either that a 40-sec. stimulus trace be postulated, or else that these animals somehow bridged this time span with some kind of mediating response chain. It is known, of course, that rather long spans can be bridged by the development of rigid, idiosyncratic, or superstitious response chains (Skinner, 1948; Kellogg, 1949; Morse and Skinner, 1957; Bower, 1961). Kamin's observations, however, do not include any references to such behavior on the part of the four dogs that learned with a 40-sec. delay.

Davitz et al. (1957) also studied the effect of CS-US interval on the acquisition of a trace conditioned, emotional response. They paired a blinking light with 3-sec. electric shock for 20 trials using CS-US onset delays of -300, 0, 2, 5, 10, 30, 120, and 600 sec. with independent groups of rats. (The 2-sec. group, of course, was actually a delayed conditioning group because the CS lasted 3 sec.) Then a test was given to determine whether the CER differed in strength for the different delay groups. The strength of the CER was considered to be adequately indexed by inhibited motor activity. Activity was measured by counting the number of spaces traversed on the grid-floor by the rats' right front paws during the five-

minute test. A one minute pre-test was used as a measure of operant level, free exploration. Figure 7-B compares the pretest vs. test activity measures for the various CS-UCS groups. Pre-test activity did not differ significantly between groups; however, the authors report that the test differences between the 600-sec. (control) group and the 0, 2, 5, 10, and 30-sec. groups were all significant. Then they say that

Although the difference between the 600-sec. and the 120-sec. groups was not statistically significant, the direction of the results suggests the possibility of some conditioning even at the 120-sec. interval. Four of the 7 animals in the 120-sec. group exhibited less activity than any of the 14 animals in the 600-sec. and 300-sec. control groups. (Davitz et al., 1957, p. 73.)

They concluded that the stimulus trace may actually be as long as 120 sec. In drawing such a conclusion, however, they have overlooked the problem (raised by Kamin) of distinguishing between diffuse emotional responses to shock and focused emotional responses which are actually conditioned to the CS. Without a basis for separating the two, it is hazardous to draw a firm conclusion regarding the actual length of the stimulus trace. Thus, these two studies by Kamin (1954) and Davitz et al. (1957) do show possible evidence that the stimulus trace is of somewhat longer duration than is suggested by the ECS, delay-of-response, and delay-of-reward studies; however, it is impossible to draw any firm conclusion because of the failure to separate diffuse from focal responsiveness.

Next, let us consider the CS-UCS interval in classical conditioning. There have been at least a dozen experiments which have shown that the optimal CS-UCS interval for the establishment of a classical, conditioned response is at .50 sec. ($\pm .25$ sec.) rather than at a zero delay interval

(cf. Kimble, 1961, pp. 156-157, and Hall, 1966, pp. 198-199). Careful examination shows that most of these studies employed delayed conditioning techniques. Only a very few studies used trace conditioning techniques. One might expect different interstimulus functions for the delayed vs. trace techniques in classical conditioning, since different functions for the two have been obtained in instrumental avoidance conditioning. However, the functions appear to be very similar, as Figure 8 shows. This figure is adapted from Riesen (1958) and compares trace and delayed techniques in both eyelid and GSR conditioning. The eyelid data are taken from Reynolds (1945) and Kimble (1947), and the GSR data are taken from Moeller (1954) and White and Schlosberg (1952). The eyelid and GSR data represent different stages in conditioning and, therefore, their slopes differ. Nevertheless, the most effective interval for both trace and delayed conditioning is the same: .50 sec.

In all the previous temporal gradients which we have examined (ECS, delayed-responses and delayed-reward) the peak or optimum has been at zero delay. The question arises, then, as to why the optimal CS-UCS interval in classical trace conditioning is at some value other than zero. Jones (1962) has offered the explanation that the CS-UCS interval actually consists of two intervals: (1) the CS-UCR interval and (2) the CR-UCS interval. The first represents the contiguity gradient, and the second the reinforcement gradient. Simultaneity of events is the optimum condition for both CS-UCR and CR-UCS. However, no single CS-UCS interval can be selected which will produce simultaneity for both. In any experiment the response has some latency, and thus the CS-UCS interval which maximizes CS-UCR is backward conditioning, but backward conditioning minimizes CR-UCS

Fig. 8. Level of Classical Conditioning as a Function of CS-UCS Interval

since the reinforcement precedes the CR. Thus, the optimum CS-UCS interval must be a compromise function.

Studies by Boneau (1958), Jones (1962), Ebel and Prokasy (1963), and Prokasy, Ebel and Thompson (1963) all have supported the conception of classical conditioning as composed of two processes with special emphasis upon the CR-UCS process as a kind of instrumental shaping. This conception must cause us to consider carefully the advantages and disadvantages of using CS-UCS interval effects in deducing stimulus-trace properties. Obviously it will tell us directly very little about the real peak or maximum of the stimulus trace. On the other hand, it should be a real challenge to trace theory to account for the optimum CS-UCS interval in terms of the overlapping contiguity and reinforcement gradients.

We may conclude that the CS-UCS interval data are the first data we have encountered which do not show a maximum at zero. However, the finding that the CS-UCS interval actually consists of two intervals removes these data as a direct contradiction to all the other sets of data we have already examined.

Delay of judgment studies in human psychophysics (S-V interval)

Fechner (1860) himself noted that the observer's judgments in successive comparisons are affected by the elapsed time between the standard (S) and variable (V) stimuli. The manner in which judgments are affected can be of two sorts: in negative time order error (TOE) the impression of S is weakened (and an objectively weaker V is judged equal to S). This is the error Fechner observed. In positive time order error (TOE) the impression of S is augmented (and an objectively stronger V is judged equal to S).

Fechner interpreted negative TOE on the basis of a fading image. Later the Gestalt theorists (Köhler, 1923; Koffka, 1935) postulated fading traces (rather than images) which served to link the two stimuli in time. These traces were not necessarily available in consciousness for introspection: they were "silent traces" (Figure 9). A fading trace for S would, of course, provide a temporally weakened standard against which to compare the objectively weaker V (thus producing a negative TOE); but how could this same fading trace account for cases where positive TOE's are obtained?

The positive TOE which occurs with a delay of about 1.5 sec. between S and V is well enough established to require attention. It may be due to fatigue, adaptation or the refractory period. For example, when the S and V are two spots of light presented successively to the same retinal area, the V is underestimated (Fullerton & Cattell, 1892; Laurenstein, 1933). Such a positive TOE may be directly attributed to retinal brightness adaptation. However, such explanations may not cover all cases in which a 1.5-sec. positive TOE is obtained.

In other circumstances in which positive TOE's are obtained, it appears to be due to (1) the effects of the "context" in which the S and V are presented, (2) the effects of stimuli interpolated during the S-V delay interval, and (3) the proactive effects of stimuli from previous trials and experimental sessions. McClelland (1943) has shown that with standard and variable lines drawn on white cards the contours of the cards themselves will reliably shift time error in the positive direction. That is, their "traces" exert an effect. This may be contrasted with conditions in which no context is present to affect line judgments, such as Tresselt's experi-

Fig. 9. Kohler's Fading Stimulus Trace as Inferred from the Point of Subjective Equality in Psychophysics

ment. (1944) in which bright lines were projected in a dark room. Under this latter condition, no positive TOE is obtained.

Concerning the second cause, Guilford and Park (1931) and Laurenstein (1933) have shown that positive TOE may be produced by interpolating high intensity background stimuli during the S-V interval. Interpolating low intensity background stimuli produced negative TOE. These results are shown in Figure 10. They are considered favorable to trace theory because "... a stimulus interpolated between S and V should displace the level of the trace, i.e., of PSE: if strong, it should raise the level and make V appear unduly weak; if weak, it should depress the level and make V unduly strong" (Woodworth, 1938, p. 443).

Concerning the third source of positive TOE (i.e., proactive effects), it should be noted that hundreds of successive trials are run in psychophysical studies. On any one trial, Ss judgment is not restricted to the information contained in that trial alone, because of the operation of long-term memory. There is evidence that memory for stimuli from earlier trials affects the perception of both the standard and variable on any given trial (and vice versa). Therefore, the variable on any given trial is not judged solely against the fading trace of an "absolute" standard. Rather it is judged against the fading trace of a standard which has itself been shifted by the experiment as a whole (or even over a series of experiments). Evidence for such shifting has been provided by Needham (1935) and his results on loudness judgments are reproduced here as Figure 11. Note that in all cases the standard has shifted in the direction of the "general standard" for the experiment as a whole, but that the amount of shift declines as a function of the S-V interval. It is as if the general standard can

Fig. 10. Effects of Interpolated Stimuli on Time Order Error

Fig. 11. Effects of Preceding Trials on Time Order Error

shift the specific standard of a given trial (and probably vice versa) only during the period in which the latter's trace is active. As Woodworth remarks: "Each of the curves is a rudimentary curve of forgetting. We can think of 0 as adjusted to the general level of tonal intensity used in the experiment, 40 db, and of his adjustment being displaced by each standard and gradually reverting toward the general level" (1938, p. 447). The question is: how long is it before the specific standard for a given trial is "forgotten"? That is, over how many seconds is the trace of the specific standard still operating? At what point in time is it replaced by the general standard? The straight lines in Figure 11 all converge somewhere between 11 sec. and 15 sec. This suggests that the memory for the specific standard used on a given trial lasts but a brief period. A truly valid estimate of this duration, of course, cannot be obtained by projecting these lines because an infinite number of functions could pass through the pairs of points upon which each line is based. The fact of their convergence, however, is of the utmost theoretical importance and supports the stimulus-trace hypothesis.

There are proactive effects not only from trials within the same experiment but also from preceding experimental days as shown in Figure 12. Again the effect is to gradually shift a negative TOE on the first day to a positive TOE on the ninth day. These data are from Needham (1934).

These various data on time order errors do not provide unambiguous support for a decaying memory trace. But the fact that these complex TOE effects are restricted to about 4-sec. to 20-sec. duration is certainly consistent with the trace hypothesis. Furthermore, there seems to be no

Fig. 12. Effects of Preceding Experiments on Time Order Error

necessity for taking the position held by Osgood (1953, p. 55) that "in any case, the conception of a fading physiological trace will not account for cases where positive time errors are obtained." On the contrary, if the traces of proactive stimuli, interpolated stimuli and background stimuli are taken into account (as well as the traces for S and V), then in principle it should be possible to account for both negative and positive time errors.

Delay of stimulus-reproduction by
humans ($S_1 - R_1$ interval)

A final source of data pertinent to the concept of a stimulus trace comes from recent experiments on ability to recall single (or paired) verbal items after enforced delay. The procedure for testing recall was devised by Peterson and Peterson (1959). They presented S a single consonant syllable (e.g., CHJ) and immediately followed it by a three-digit number. The S had to repeat the number and then count backwards from it by three's (or four's) in time to a 1-sec. beat of a metronome. Counting was employed to minimize if not eliminate rehearsal of the syllable. The S's interpolated counting was interrupted after either 3, 6, 9, 12, 15, or 18 sec., and he was asked to recall the original syllable. Any responses which occurred within 15 sec. were recorded.

Peterson and Peterson found that the proportion of correct recalls declined exponentially as a function of the recall interval (i.e., the amount of time Ss recall was delayed by interpolated counting). Their data are shown in Figure 13 as open circles. A curve fitted to their data points was offered by the authors as a "forgetting curve" for a single item; it is shown in Figure 13 as a solid line. As a forgetting curve it bears a similarity to the decaying memory trace postulated here. For comparison

Fig. 13. Short-term Memory Functions: Early Studies

the decremental portion of the memory trace of Figure 1 is shown as a dashed line paralleling their solid curve.

Murdock (1960) and Hellyer (1962) have both replicated the Peterson and Peterson study with a few differences in procedure. Murdock's results are shown as the solid circles in Figure 13. The main difference in Murdock's study was that he measured the recall interval from the beginning of the stimulus exposure period, whereas Peterson and Peterson reckoned it from the end of the exposure period. Hellyer used a visual rather than an auditory presentation technique, and he had Ss read aloud random groups of numbers as an interpolated task. His results, based on recall intervals of 3, 9, 18, and 27 sec. are shown as solid triangles in Figure 13.

If it were true that Peterson and Peterson's interpolated counting procedure or Hellyer's reading aloud procedure actually prevented rehearsal without at the same time interfering with the acquisition process, then these data in Figure 13 would certainly provide empirical support for the notion of a decaying stimulus trace. Unfortunately, there is now every reason to assume that the "forgetting curve" shown here is as much a product of the interpolated activity as it is of any stimulus decay process. Loess and McBurney (1965) have recently shown that the proportion of correct recalls after a given interval of delay (9 sec.) varies markedly with the type of interpolated task. When their Ss were given an active, interfering task such as reading random letters aloud, the proportion of correctly recalled trigrams after 9 sec. was about .30; but when their Ss were given a passive, non-interfering task such as reciting digits (beginning with the digit one), the proportion of correctly recalled trigrams after 9 sec. jumped to above .90. Thus, in order to choose the stimulus trace function

shown in Figure 1 on the basis of its apparent fit to the Peterson and Peterson data, one would have to argue that Peterson and Peterson's interpolated counting task was somehow a unique task, at once absolutely free of any interfering effects and yet completely effective in preventing rehearsal. Such an argument would be hard to develop. At present it must be concluded that the slope of their forgetting curve does not offer direct empirical support for the probable slope of the stimulus trace function.

However, the type of performance curve which the Petersons' study yields does offer some additional support for a stimulus trace of brief duration. In developing this point it is instructive to compare their experiment with humans to the delayed-response experiment given to animals. In both cases a stimulus is supplied and there is an enforced delay in the response to the stimulus. In the animal experiments the response is a choice response; in the human experiment it is a "reproduction" of the stimulus itself. There are other differences as well. In most classical, delayed-response experiments, the animal is not prevented from "rehearsing" the position of the baited food-well by maintaining an orienting response chain, but the maintenance of this response chain is presumably interrupted by accidental, distracting events during the delay interval. In the Petersons' experiment the human subject is discouraged from rehearsing by the introduction of an interpolated activity during the enforced delay period; however, the effectiveness of the interpolated activity in discouraging rehearsal seems to vary directly with the amount of retroactive interference it produces.⁴

⁴ Any estimate of interference effects must also take into account proactive interference (PI) from preceding trials as a result of the within-Ss design commonly employed. Evidence of a rapid build-up of PI is contained in Shepard and Teghtsoonian (1961), Murdock (1961) and Keppel and Underwood (1962). Melton (1963) has reviewed this evidence.

However, in both experimental paradigms it might be presumed that the distracting or interfering stimuli (whether accidentally present or experimentally introduced) would only produce their effects during the period of trace consolidation. As we have already noted, this was apparently true in the delayed-response experiments: all the performance curves reach asymptote in a maximum of 30-35 sec. (cf. Figure 3). Is this also true in the Petersons' delayed-recall paradigm?

The evidence on this question is not definitive, but it certainly is suggestive. The curves in Figure 13, for example, do appear to reach asymptote within the presumed period of trace consolidation. More recent data also tend to corroborate this proposition. Figure 14 shows the delayed-recall performance functions obtained by Melton (1963) and by Bruce and Cofer (1965). Melton's performance function has not definitely levelled off even by 30 sec., but the functions reported by Bruce and Cofer have clearly reached asymptote well within that time. This is true both for a recall measure of retention and for a recognition measure (where the item had to be correctly identified out of a group of five high-similarity items (HS) or else out of a group of five low-similarity items (LS)). These curves, like the delayed-response curves of Figure 3, at least suggest that interference with short-term memory is only effective if the interference comes within a brief span of time.

So far the data reviewed here have dealt only with the proportion of subjects who correctly recall a stimulus. With the Peterson and Peterson study, however, it is also possible to examine the latency of correct recalls. In their analysis, Peterson and Peterson did not simply record the number of

Fig. 14. Short-term Memory Functions: Later Studies

correct responses, but they kept a separate tally of how many of them occurred in each successive, 2-sec. latency period. Then they plotted the proportion of correct recalls as a cumulative function of latency. This was done separately for each of the six delay of recall intervals (i.e., 3, 6, 9, 12, 15, and 18 sec.). These cumulative functions are shown as the solid circles in Figure 15, where proportion of correct recalls is plotted on the left-hand ordinate and latency is plotted on the abscissa. Concerning these functions, Peterson and Peterson say:

Note that the number correct with latency less than 2 secs. does not constitute a majority of the total correct. These responses would not seem appropriately described as identification of the gradually weakening trace of a stimulus. There is a suggestion of an oscillatory characteristic in the events determining them (1959, p. 195).

Their statement indicates they believe probability of recall is at any given time a function of the instantaneous value of the gradually weakening trace. However, it might be just as reasonable to assume that probability of recall is a function of the integrated trace-strength; that is, a function of how long the trace has had a chance to accumulate in the brain. This alternative was not investigated. Their statement, then, was a rejection of the trace concept without a complete investigation of its potential for explaining their data.

In the present stimulus-response model the events occurring in the Peterson and Peterson experiment might be viewed as follows. At time zero minus .5 sec., a subject is presented a single consonant-syllable and a stimulus trace is initiated in the brain. A half-second later (when the trace reaches unity) the three-digit number which is the stimulus for interpolated counting is given. This stimulus and the subsequent interpolated counting may be presumed to directly interfere with the consolidation or storage of information from the fading trace. However, at the end

Fig. 15. Comparison Between Cumulative Correct Recall and Log Effective Trace-Strengths at Successively Longer Latencies of Recall

of the interpolated counting period (i.e., delay of recall period), there will be more or less trace still available to S on which to base his recall of the item (depending on the length of this interpolation period). For example, Figure 16 shows the amount of trace left for the specific case in which S does interpolated counting for only .3 sec. The portion of the trace subjected to interpolated interference is cross-hatched; the amount of trace not interfered with is not cross-hatched.

Under this application of trace theory, the proportion of correct recalls at any given latency ought to be some function of the integral of the trace up to the end of the latency period minus the integral of that portion of the trace during which S did interpolated counting. This residual integral might be termed the "effective" trace. Following exactly this procedure, the integrated area of the "effective" trace may be found for all intervals of delay for each of the six interpolation durations. (The exact computations are supplied in Appendix A.) Then, if a log relationship is assumed between trace strength and latency of recall, the log values of these integrated "effective" trace strengths may be calculated and plotted against the right-hand ordinate (as has been done in Figure 15). These log, "effective" trace strengths are shown as open circles in Figure 15. Despite some differences in shape, the family of cumulative functions generated by the stimulus-trace model clearly parallels the cumulative latency functions reported by Peterson and Peterson. In the light of this, it is difficult to accept their statement that the latency functions "would not seem appropriately described as identification of the gradually weakening trace of a stimulus." Their study simply shows that the shape of the stimulus-trace function required to fit their latency data may be slightly

Fig. 16. Postulated Events in the Peterson, Short-term Memory Study

different than the function postulated here. Their data, then, are also compatible with the stimulus-trace concept.

Summary of evidence

A rather remarkable agreement emerges from the data of these six classes of experiments. The extent of agreement is summarized in Table I. The ECS, delayed-response and delayed-reward studies all show that (1) maximum strength of association occurs at a zero delay between the two events in question, (2) strength of association falls off very rapidly at first and then more slowly (according to a negatively-accelerated function), and (3) strength of association approaches a value of zero somewhere between 5 sec. and 20 sec. The interstimulus interval studies, on the other hand, offer discrepant results, at least in classical conditioning where the maximum strength of association between CS and UCS appears to occur at about one-half sec. However, the classical conditioning process has recently been formulated in terms of two processes which involve a CS-UCR interval and a CR-UCS interval. This development means that these data may not be discrepant after all. For it is possible to assume that maximum conditioning for these separate pairings alone would be at a zero interval, whereas under simultaneous acquisition the optimum interval for one process is incompatible with the optimum for the other, thereby producing the .5-sec. CS-UCS optimum. Finally, certain results on TOE in psychophysics are also at variance with the results of the ECS, delayed-response and delayed-reward studies. Positive TOE in the first 1.5 sec. of the TOE curve has not been completely explained. On the other hand, the shifts of the overall TOE curve from negative to positive can be attributed to effects of

Table 1

Summary of Evidence Relating to the Stimulus Trace

Studies	Estimated Peak	Estimated Asymptote	Reference Experiments
<u>Retrograde-Amnesia (FS-ECS)</u>	0.00 sec.	30 sec.	Quartermain et al (1965)
	0.00 sec.	5-10 sec.	Chorover & Schiller (196
<u>Delay of Response (CS-CR)</u>	0.00 sec.	5-10 sec.	Smith (1951)
<u>Delay of Reward (CR-UCS)</u>	0.00 sec.	5-10 sec.	Keesey (1964)
	0.00 sec.	5-10 sec.	Grice (1948)
	0.00 sec.	5-10 sec.	Kamin (1959)
<u>Interstimulus Interval (CS-UCS)</u>	----	9 sec. or greater	Jenkins (1950)
	----	10-20 sec.	Kamin (1954)
	0.45 sec.	3 sec. or greater	Reynolds (1945)
	0.40 sec.	----	Kimble (1947)
	0.50 sec.	2 sec. or greater	White & Schlosberg (1952)
	0.45 sec.	4- 6 sec.	Moeller (1954)
<u>Psychophysical Judgments (S-V)</u>	----	12 sec.	Köehler (1923)
	----	4-8 sec.	Needham (1934)
	----	10-15 sec.	Needham (1935)
<u>Short-term Memory (S_1-R_1)</u>	----	20-30 sec.	Peterson & Peterson (1959)
	0.00 sec.	30-35 sec.	Murdock (1961)
	----	30-40 sec.	Hellyer (1962)
	0.00 sec.	32 sec. or greater	Melton (1963)
	----	9-12 sec.	Bruce & Cofer (1965)
	0.00 sec.	16 sec.	Peterson (1966)

preceding experiments or trials, to effects of interpolated stimuli, or to effects of background stimuli. They can be theoretically handled by considering the accumulated interaction of many traces other than just the S and V traces for a particular trial.

We may conclude from this brief review that some evidence which is contradictory to the stimulus-trace hypothesis exists, but that it all comes from complicated, experimental paradigms and therefore constitutes very indirect, negative evidence. On the other hand, those experimental situations which support the trace hypothesis are rather direct tests of it, and they show remarkable agreement concerning the general properties of the stimulus trace.

But what about the specific properties of the trace? On what basis should a particular perseverative function be chosen? Ideally, a rational function based upon sound theoretical principles should be chosen (Lewis, 1960). Although a rationale for the shape of a stimulus trace has never been worked out in detail, Hull has implicitly employed certain basic premises which seem sound. First of all, the stimulus trace should intersect the ordinate at some value. (This value is usually arbitrarily designated as unity.) That is, by definition, the perseverative segment represents a decline from some maximum trace strength. This premise eliminates parabolic and hyperbolic functions and restricts consideration to negatively-accelerated exponential curves whose Y intercept is unity. Among this class of curves we have empirical data which justifies choosing a negative exponent which will bring the trace to asymptote somewhere between 5 sec. and 20 sec. This is based on the assumption that trace activity has effec-

tively dissipated during this interval (otherwise ECS would show interference with the trace or Ss would be able to utilize trace information in delayed responses, delayed rewards, or delayed judgments). Most of the data converge on a brief trace whose effective duration is 15 sec. or less. Thus an exponential, perseverative function has been chosen which has a negative exponent of $b = -.200$. This value of b is somewhat smaller than has been employed in earlier work (Wing, 1964). It brings the trace to five per cent of unity in 15 sec. and to one per cent of unity in 23 sec., as shown in Figure 1. (It is assumed that trace strengths any weaker than either five per cent or one per cent of unity would be ineffective in the formation of associations.) This, then, is the postulated trace; the problem still remains as to how the many traces of a stimulus series might interact to produce associations among adjacent and remote members of the series.

PART II

STIMULUS-TRACE INTERACTION MODEL

In developing a model of trace interaction only the interaction among the perseverative segments of the successive stimulus traces will be considered. This is for the sake of simplicity. Eventually, of course, the effects of the incremental portion of the trace may have to be considered. Also, at least initially, it will be desirable to assume that stimulus traces are identical for all of the percepts generated by a stimulus series. Both these assumptions are made explicit in Figure 17. Here the decremental segments of a succession of ten, identical traces are shown. They are labeled A, B, C, ... J and may be conveniently thought of as the traces for a series of randomly chosen stimuli such as adjectives or nonsense syllables which are presented at a 2-sec. rate. Figure 17 shows that the traces from earlier stimuli are still active when later traces are being established. The stimulus trace of A, for example, is still active while B, C, D, and other percepts are being established. This lingering of traces is necessary if such a series is to become associatively connected.

Problem of Backward Associations and its Resolution

Just how can a model be developed which will quantify the degree of trace-interaction? Consider the traces A and B as an example. The simplest procedure would be to sum or multiply the instantaneous values of A and B and integrate these sums or products from the instant B is presented to the instant A is no longer active. But what would the integral of these sums or products represent? Would it represent the strength with which A would tend to call up B or would it represent the strength with which B would tend to call up A? Or would it represent both associations? This is an important question for it is already well established from laboratory experiments that two kinds of associations exist (i.e., forward and backward associations), and that the forward association ($A \rightarrow B$) is stronger than the backward association ($B \rightarrow A$) (see McGeoch and Irion, 1952, 90-114). Unfortunately, neither the integration of sums nor the integration of products will generate bidirectional associations of differing strength.

The present model solves this problem in the following manner. For purposes of exposition, let us assume that the traces for all stimuli consist of physical processes occurring at separate sites in the brain. Under these circumstances, the strength of the forward association ($A \rightarrow B$) might reasonably depend upon just a fraction of A's trace activity being represented at B's site in the brain; and, conversely, the strength of the backward association might reasonably depend upon just a fraction of B's trace activity being represented at A's site. That is to say, the strength of both the forward and backward associations would be obtained by raising

the trace value of the eliciting percept to some fractional power before adding or multiplying it by the value of the percept being elicited.

Figure 18 shows this diagrammatically.

In Figure 18, sites 1 and 2 represent the separate places in the brain where A's trace activity and B's trace activity are occurring. By means of the upper pathway, a fraction of A's activity (i.e., A^k where $k < 1.00$) is represented at B's site and is available to interact with B's trace strength. A multiplicative rather than an additive function will be adopted. This is an arbitrary decision and is made in the absence of any rational criteria for deciding between the two. Using a multiplicative function, the forward association at any given instant in time becomes

$$(A)^k(B) \text{ where } k < 1.00.$$

Similarly, by means of the lower pathway a fraction of B's trace strength (B^k where $k < 1.00$) is represented at A's site and is available to interact with it. Again using a multiplicative function, the backward association at any given instant in time becomes

$$(A)(B)^k \text{ where } k < 1.00.$$

Very importantly, this simple operation always yields stronger forward than backward associations for two stimuli which were presented at different points in time. That is,

$$(A)^k(B) > (A)(B)^k$$

simply because $B > A$ at every instant of time. The amount the forward association exceeds the backward association depends both upon the time lag between the stimuli and upon the fractional exponent k . As the time lag approaches zero, the backward association approaches the forward

Fig. 18. Diagrammatic Representation of Forward and Backward Associations

association in strength; and as k approaches unity the backward association also comes closer and closer to equalling the forward association. With a time lag of zero and/or with $k = 1.00$ both associations are equal in strength.

The physical analog to the mathematical model which has been described above was given in order to help clarify the latter, but in no sense is the mathematical model dependent on it.

Principle of Associative Asymmetry

It is clear from the above description that the present model is based on a principle of "associative asymmetry" which states that the $A \rightarrow B$ association is stronger than the $B \rightarrow A$ association. Asch and Ebenholtz (1962) recently reviewed the assumptions underlying such a principle and concluded that the existence of backward associations is an embarrassment to the asymmetry principle. The main reasons are that (1) backward associations contradict our experience of learning a series as a forward-going sequence; (2) they contradict our conception of the $S \rightarrow R$ bond as a temporal, causal sequence; and (3) they contradict the law of forward conduction across the synapse. As a result they say that "given the principle of asymmetry, one can deal with the empirical fact of backward association only by denying its authenticity (p. 137)." But in the present model there has been no necessity for denying the authenticity of backward associations! Using the simplest possible assumptions the model generates weaker backward associations as readily and as automatically as it does the stronger forward associations.

Furthermore, the concept of backward associations does not contradict our experience in learning a series as our ability to perform backward digit span testifies. Nor does it violate either the notion of the $S \rightarrow R$ bond or the law of forward conduction across synapses. The fundamental requirement of an $S \rightarrow R$ bond is contiguity, and in the trace model A and B are contiguous from the time B is presented until A's trace has disappeared. This contiguity provides the occasion for each stimulus to become a response for the other. Furthermore, the principle of asymmetry would not accord with the law of forward conduction only if the forward and backward associations are assumed to travel the same pathway. However, it is not violated in the present model because separate pathways are postulated for the $A \rightarrow B$ and $B \rightarrow A$ associations. Each pathway, of course, utilizes forward conduction. Furthermore, it is not even necessary to postulate that different traces are located in different neurons. They could all be represented within a single neuron if the memory code is a molecular one. Therefore, it is really unnecessary to respond at all to this particular criticism. In short, all the embarrassing factors which Asch and Ebenholtz cite are only embarrassing to a theory of asymmetrical associations which overlooks the fact that a temporal series of discrete stimuli may be represented in the brain as spatially differentiated centers of activity simultaneously activating one another. Finally, it should be noted that all the data so far cited in support of Asch and Ebenholtz' principle of symmetry pertains to the relative strength of forward and backward associations between paired-associates. These are presented either simultaneously (e.g., Murdock, 1962) or with a maximum of 2-sec.

delay between pairs (Asch and Ebenholtz, 1962; Horowitz et al., 1964; Horowitz and White, 1965; Horowitz et al., 1966). The present model in fact predicts equally strong forward and backward associations with the simultaneous presentation, and it predicts only a relatively small difference for brief time lags. But there is a more fundamental objection to these paired-associate studies. In the paired-associate task the problem of interpreting relative strengths of directional association is extremely complicated. The stimulus member of each pair never serves as a response, and so it has been objected that it is consequently less "available" to the subject at time of recall. To correct for this, certain artificial techniques for increasing its "availability" have recently been employed. However, Ekstrand (1966), who has reviewed paired-associate experiments using such techniques, concludes "... it is not clear that equation of availability will entirely eliminate the difference between forward and backward learning under all circumstances (p. 60)." In the serial learning task, of course, objection to differences in availability cannot be made because every stimulus is also a response. Thus for purposes of testing relative strengths of forward and backward associations it would seem to be a superior task; and, as already noted, forward associations are almost universally reported to be stronger than backward associations in serial learning.

Computational Formulas for Forward and Backward Associations

The operations $(A)^k(B)$ and $(A)(B)^k$, of course, are not formulas but are simply symbolic notations. To obtain formulas for computing forward and backward associations, it is necessary to replace each of the symbols

A and B with the equation for the stimulus trace and then derive a simple expression which will actually represent the integration of the traces over time. Such expressions have been derived by calculus and are shown in Appendix B. The formula for the forward association, which will be designated ϕ , is

$$(3) \quad \phi = \frac{1}{b(k+1)e^{kb\lambda}}$$

where k is that fractional power of the one trace represented at the site of the other trace, e is the base of the natural logarithms, b is the decay constant in the exponent of the stimulus-trace equations, and λ is the time lag between any earlier stimulus and any later stimulus for which the associative strength is being computed. Similarly, it may be shown (see Appendix B) that the backward association, β , may be represented as

$$(4) \quad \beta = \frac{1}{b(k+1)e^{b\lambda}}$$

A significant property of both equations should be noted. The terms b , k , and λ are all in the denominator of the expressions. This means that an increase in any one of them (or combination of them) will decrease associative strength. For example, consider the value of λ , which represents the time lag between stimuli. If the rate of presentation is decreased, λ will increase, the values of $e^{kb\lambda}$ and $e^{b\lambda}$ will also increase, and the result will be a decrease in forward and backward associative strengths. It is to be expected that the longer the time between stimuli, the weaker will be their associative strengths.

The only difference between equations (3) and (4) is the presence of k as part of the exponent in the denominator of the formula for the forward

association. As already noted, for any given time lag between stimuli, k determines the relative strength of the backward and forward association. In the present study, a value of $k = .75$ has been used. This value of k was chosen specifically because it yielded a theoretical, average ratio of backward to forward associative strengths of 1.00 : 1.15 (when averaged over multiple trials) which compared favorably with an empirically obtained ratio of backward to forward associations of 1.00 : 1.18 reported by McGeoch and Irion (1952, p. 105). The value of this parameter, then, has been selected to fit one set of empirical data. It will be used hereafter as a constant in all further applications, however. Having chosen a value of k for equations (3) and (4), the forward and backward associations may be quickly computed between all possible pairs of stimuli for a series of any length and for any rate of presentation. These computations yield what may be called an association matrix. Table 2, for example, shows an association matrix containing all the forward and backward associations established in a single trial for the 10-item series shown in Figure 17. (However, this matrix is based upon a 1.8-sec. rate rather than a 2.0-sec. rate in order to facilitate later comparison with an empirical study.) Such a one-trial matrix can in principle be used to predict performance in forward and backward digit span, in serial recall after one presentation and in free recall after one presentation. To use such a matrix, however, it is important to understand its properties.

Table 2

One-trial Association Matrix for a 10-item List

	A	B	C	D	E	F	G	H	I	J	
A	---	218	167	127	97	74	43	33	25	19	8
B	199	---	218	167	127	97	74	43	33	25	7
C	139	199	---	218	167	127	97	74	43	33	6
D	97	139	199	---	218	167	127	97	74	43	5
E	68	97	139	199	---	218	167	127	97	74	4
F	47	68	97	139	199	---	218	167	127	97	3
G	33	47	68	97	139	199	---	218	167	127	2
H	23	33	47	68	97	139	199	---	218	167	1
I	16	23	33	47	68	97	139	199	---	218	0
J	11	16	23	33	47	68	97	139	199	---	0
	8	7	6	5	4	3	2	1	0		

Forward
DEGREE OF REMOTENESS

Backward
DEGREE OF REMOTENESS

Properties of the One-trial Association Matrix

Table 2 may be read in three different ways. Reading down the columns the tabled values show with what relative strength each stimulus heading a column will be elicited as a response to each of the stimuli heading the rows. Reading across the rows, the tabled values show with what relative strength each stimulus heading a row will elicit as a response each of the stimuli heading the columns. Finally, reading down the left to right diagonals the tabled values show the strength of association between stimuli which are immediately adjacent (0 degrees remote) or which are once, twice, or more removed from one another (1,2,3 ... 8 degrees remote). Regardless of which way the table is read, the associations above and to the right of the diagonal line of dashes always represent forward associations, while the associations below and to the left of the diagonal line of dashes always represent backward associations.

Inspection of Table 2 in terms of its diagonals reveals three very important characteristics. First, strength of association is the same at any given degree of remoteness (the values in Table 2 are identical along any given diagonal). Second, strength of association between any given stimulus and the other stimuli of the series falls off exponentially with the degree of remoteness of those stimuli. Third, at any given degree of remoteness, the forward association is stronger than the backward association. These latter two observations may be more readily drawn by inspecting the dotted curve in Figure 19 which shows the weighted average of the theoretical associative strengths following a single trial as a function of degree of remoteness.

Fig. 19. Theoretical Remote Association Gradients for Single-Trial vs. Multi-Trials

Properties of the Multi-trial Association Matrix

In multi-trial acquisition of a serial list there is the problem of associations formed between trials as well as the problem of accumulated associative strengths. To illustrate this problem let us consider the theoretical, forward and backward, associative strengths which would result from the interaction of the stimulus traces of the 10-stimulus series of Figure 17 presented for 20, consecutive trials at a 1.8-sec. rate with a 5-sec. inter-trial interval. (This rate and inter-trial interval have been chosen in order to later compare the theoretical values with empirical data.) Using this rate of presentation and inter-trial interval, there is ample opportunity for traces from end items on any previous trial to interact with traces from beginning items on any subsequent trial. This means both forward and backward associations will be established across the inter-trial interval. For reasons to be discussed below, these inter-trial associations will always be lumped together with the intra-trial associations. For example, the relatively strong backward associations from stimulus A of trial N to stimulus J of trial N-1 have been added to the weak forward associations formed from A to J on all trials. The effect of lumping together the inter-trial and intra-trial associations is to increase the associative strengths at the higher degrees of remoteness. This is shown in Table 3 which gives the associative strengths between all pairs of the series averaged over 20 trials. Note the increase in the values in the upper-right and lower-left corners of Table 3 over those of Table 2. Figure 19 shows that this increase is also reflected in the weighted, average associative strengths for multi-trial presentation (dashed line).

Table 3

Multi-trial Association Matrix for a 10-item List

	A	B	C	D	E	F	G	H	I	J
A	---	224	175	139	115	99	92	94	106	130
B	215	---	224	175	139	115	99	92	94	106
C	159	215	---	224	175	139	115	99	92	94
D	124	159	215	---	224	175	139	115	99	92
E	103	124	159	215	---	224	175	139	115	99
F	93	103	124	159	215	---	224	175	139	115
G	93	93	103	124	159	215	---	224	175	139
H	102	93	93	103	124	159	215	---	224	175
I	119	102	93	93	103	124	159	215	---	224
J	146	119	102	93	93	103	124	159	215	---

Forward
DEGREE OF REMOTENESS

Backward
DEGREE OF REMOTENESS

Such a multi-trial matrix can be used to predict performance in serial anticipation learning. The matrix will generate the serial-position curve and the relative frequency of anticipatory and perseverative intrusions. It will also generate the remote-association gradients obtained in serial learning. It is this latter prediction which we will consider here in some detail.

One of the best methods of measuring the strength of remote associations for multiple presentations is to run Ss in a serial-anticipation learning experiment and then elicit free-associations to the separate stimuli of the series when these are presented in random order upon termination of learning. Despite the advantages of this method, it should be noted that it cannot distinguish between inter-trial and intra-trial associations. For example, when stimulus J is given as a free-associative response to stimulus A, it is impossible to tell with this method whether it is given on the basis of a forward, remote, intra-trial association or on the basis of a backward, adjacent, inter-trial association. It is for this reason that the inter-trial and intra-trial associative strengths have been lumped together in the multi-trial matrix.

Point of Entry and Direction of Recall in a Series

In learning or relearning of a series Ss are always provided with a starting point at the beginning of the series; but in other tasks, such as free recall, the beginning of the series is not necessarily the point of entry used by Ss. The point of entry will depend upon idiosyncratic factors such as the unique association-value of a particular item for a particular subject. Everyday observations similar to these experimental observations led Aristotle to make a number of pertinent remarks about point of entry and direction of recall:

Recollecting differs also in this respect from relearning, that one who recollects will be able, somehow, to move, solely by his own effort, to the term next after the starting-point. When one cannot do this of himself, but only by external assistance, he no longer remembers ... It often happens that, though a person cannot recollect at the moment, yet by seeking he can do so, and discovers what he seeks. This he succeeds in doing by setting up many movements, until finally he excites one of a kind which will have for its sequel the fact he wishes to recollect. For remembering ... is the existence, potentially, in the mind of a movement capable of stimulating it to the desired movement, and this, as has been said, in such a way that the person should be moved ... from within himself, i.e., in consequence of movements wholly contained within himself.

But one must get hold of a starting-point. This explains why it is that persons are supposed to recollect sometimes by starting from mnemonic loci. ... to recollect, this is what he will do: he will try to obtain a beginning of movement whose sequel shall be the movement which he desires to reawaken. This explains why attempts at recollection succeed soonest and best when they start from a beginning (of some objective series).

... It seems true in general that the middle point also among all things is a good mnemonic starting-point from which to reach any of them. For if one does not recollect before, he will do so when he has come to this, or, if not, nothing can help him; as, e.g. if one were to have in mind the numerical series denoted by the symbols 1, 2, 3, 4, 5, 6, 7, 8, 9. For if he does not remember what he wants at 5, then at 5 he remembers 9; because from 5 movement in either direction is possible, to 4 or to 6. (Aristotle, ca. 350 B.C.)

Not only the beginning and the middle, but also the end of the series is a useful point of entry. However, the usefulness of each point of entry depends upon the conditions of recall. (As we shall see, for example, the optimal time for beginning recall from the end of a series is when several previous series have been experienced and are interfering proactively.) However, whatever the starting point of the series, if recall proceeds in a forward direction, it is assumed to be by means of forward associations; and if in a backward direction, it is assumed to be by means of backward associations. This carries certain implications for use of an association matrix. An association matrix for a one-trial or multi-trial experiment is built up solely in terms of the interaction of the stimulus traces. It only represents the input for the memory-storage system. The output, however, will depend upon Ss' point of entry and direction of movement through the association matrix. If Ss are instructed to perform forward recall, it is assumed their entry will be the first item and that they will proceed through the series by means of forward, adjacent associations. All forward remote associations and all backward associations (adjacent and remote) will interfere with this process. If Ss are instructed to perform backward recall, the opposite situation will prevail.

However, if Ss are instructed to perform free recall, then different Ss may adopt different points of entry and different directions of recall (or the same S may operate differently when naive than when sophisticated). Such considerations are of major importance in the application of the stimulus-trace model.

PART III

APPLICATIONS OF THE MODEL

Remote association gradients

In Table 2 (p. 60) a single-trial association matrix was presented. This matrix represented the theoretical, associative strengths of every item in the series with every other item following just one presentation. Unfortunately, no empirical data are available which would serve as an index of the relative strength of remote associations following a single trial. Thus the strategy followed here will be to confirm the validity of the multi-trial, theoretical values and thereby (indirectly) the single-trial, theoretical values upon which the multi-trial values were based.

In Table 3 (p. 64) the multi-trial association matrix was presented. This matrix represented the theoretical strengths of every item in the series with every other item including those associations formed across the inter-trial interval. In this section we shall compare these theoretical values with an empirical measure of the strength of remote associations. The empirical data were all obtained from serial-anticipation learning experiments in which learning was interrupted and Ss were then instructed to make free-associations to the separate stimuli of the series. The assumptions underlying this method are that (1) the frequency with which an item will be given as a free-associate to another item will vary with strength of association; and (2) this strength of

association is a function of the degree of remoteness of the two items in the original series.

McGeoch (1936) used this free-association technique with several different series of verbal stimuli. One of these series consisted of ten, two-syllable adjectives, presented at a 1.8-sec. rate with a 5-sec. inter-trial interval. For this series he obtained free-associations to all but the first and last stimuli. The frequency of these free-associations is shown in Table 4. It shows that number of McGeoch's Ss who gave the stimulus listed on the top of the table as a free-associative response to the stimulus listed on the left-hand margin of the table.

Comparison of Table 4 with Table 3 shows certain striking similarities. First, the number of free-associations in Table 4 remain relatively constant (if somewhat unstable) within each degree of remoteness, just as the theoretical values of Table 3 do. Second, frequency of free-associations does fall off exponentially over the first few degrees of remoteness just as the theoretical association values do, and then there is even a subsequent upswing in free-associations at the higher degrees of remoteness. Third, the frequency of the forward free-associations is greater than the frequency of the backward free-associations (at least at the smaller degrees of remoteness).

The last two conclusions may be more readily drawn from Figure 20. This figure compares at each degree of remoteness the weighted, mean proportion of Ss who gave free-associations as obtained by McGeoch (solid lines) with the weighted, mean, theoretical associative strengths (dashed line)

Table 4

Multi-trial Free-association Matrix for a 10-item List

	A	B	C	D	E	F	G	H	I	J	
A	---	---	---	---	---	---	---	---	---	---	
B	114	---	138	42	24	20	60	15	48	72	8
C	48	126	---	132	48	42	42	18	102	60	7
D	21	24	66	---	204	42	30	66	24	108	6
E	30	30	30	162	---	66	60	54	66	80	5
F	54	24	24	60	144	---	162	96	72	30	4
G	24	18	42	72	24	138	---	150	48	49	3
H	30	30	42	36	54	42	72	---	204	72	2
I	48	48	30	24	36	60	42	80	---	210	1
J	---	---	---	---	---	---	---	---	---	---	0
	8	7	6	5	4	3	2	1	0		

Forward
DEGREE OF REMOTENESS

Backward
DEGREE OF REMOTENESS

Fig. 20. Comparison between Theoretical and Empirical Remote Association Gradients for a 10-item List

for multiple presentation of the series.⁵ This multi-trial, theoretical curve is the same one depicted in Figure 19. There are two solid lines which represent two sets of empirical data. The solid line with filled circles represents the data for the ten, two-syllable adjectives. These values are the weighted means of the diagonals of Table 4 expressed as proportions of total subject sample. The solid line with the filled triangles represents comparable weighted proportions taken from data for a series of ten, consonant-syllables also used by McGeoch (1936, p. 236). The matrices for this condition are reported in Appendix C. Two main discrepancies may be noted between the predicted and obtained functions of Figure 20. One is failure to predict the extremely large number of free-associations at zero degrees remoteness in the forward direction. The other discrepancy is that a greater upswing was predicted for backward associations than for forward associations, whereas the exact reverse was obtained by McGeoch. However, this latter discrepancy may not be due to a failure in the theory. Data reported by Raskin and Cook (1937) conform more closely to the theory than do McGeoch's data. Figure 21 shows their data for a series of eight nonsense syllables which were presented at a 2.5-sec. rate with a 6-sec. inter-trial interval. Comparison of their data with the theoretical curve generated by the present theory shows that the greater upswing predicted for backward associations was actually obtained. It will be noted, however, that once again the theory failed to predict

⁵ Using the weighted mean produces the same correction for differing number of opportunities at each degree of remoteness as does the correction traditionally applied by Raskin and Cook (1937).

Fig. 21. Comparison Between Theoretical and Empirical Remote Association Gradients for an 8-item List

the very large frequency of free-associations at zero forward, degrees of remoteness. (It is reasonable to assume that this large frequency at zero degrees forward represents the implicit and explicit rehearsal of this association during the learning process.)

No test of fit has been made between the theoretical and empirical functions of Figures 20 and 21 for two reasons. First, it seems premature to apply such tests until a more detailed rationale is established for the constants used in the equations (i.e., b and k). Secondly, the empirical data themselves suffer from certain inadequacies. McGeoch's data were obtained from experiments whose primary objectives were to study transfer of training, and they probably contain systematic sources of error. Raskin and Cook's data, while not containing such sources of error, were nevertheless obtained from a very small sample. A prerequisite to curve-fitting should be the collection of more adequate, empirical data.

Despite the existence of certain unknowns in both theory and data, it seems reasonable to conclude that the present trace theory satisfactorily generates the strength of remote associations formed in multi-trial learning. The implications of this are important. First of all, all the phenomena explicable in terms of remote associations now become explicable in terms of trace theory as well. Secondly, the mathematical description of the processes whereby remote associations are formed opens the way for rigorous mathematical deductions which are testable in the laboratory. For example, the present model predicts that no upswing in free-associations will be obtained at the higher degrees of remoteness after multi-trial learning with inter-trial intervals longer than about 20 secs.

The fairly rigorous efforts that have been made here to generate remote associations have naturally been made on the assumption that remote associations actually exist! Recently, however, Slamecka (1964, 1965) vigorously disputed the existence of remote associations. As part of his refutation of the association doctrine, Slamecka ran an experiment in which no meaningful remote associations could have been formed during learning, and yet he derived from the experiment a table of free-associations between stimuli which resembled in certain respects the tables of free-associations reported by McGeoch and Raskin and Cook (cp. Slamecka, 1964, p. 70).

As a result of the present analysis, it is possible to show that, while the totals of the columns and diagonals of Slamecka's table are reasonable, the values in the body of his table are not reasonable; that is, they depart radically from those values actually obtained in a serial learning experiment. Specifically, in Slamecka's table, the values in any given column tend to be constant, whereas in both the present theory and in McGeoch's empirical table of free-associations, it is the values in the diagonal which tend to be constant. This difference in predictions is even more pronounced in a theoretical table of free-associations subsequently reported by Slamecka (1965, p. 244) in which the column values are shown to be identical. These comparisons show that the processes postulated by the trace model generate the values in the body of the empirical free-association table, whereas the prefamiliarization process postulated by Slamecka does not. It is concluded, therefore, that the evidence suggests true remote associations are formed during serial learning, and that they are the possible result of the interaction of stimulus traces of the general form presented in this study.

The finding that for the multi-trial situation the theoretical and empirical association-matrices are equivalent in their general properties strengthens the case for the single-trial, theoretical matrix as well, because the former is based on the latter. This finding, therefore, raises one's level of confidence in using the one-trial matrix as well as the multi-trial matrix to predict various other phenomena such as the serial-position curve.

Serial-position curves in serial recall

So far we have examined the diagonals of the association matrix and corroborated that they adequately represent the relative strengths of association as a function of degree of remoteness. Now we turn our attention to the rows of the matrix. In any association matrix the values in a given row represent the strengths with which the stimulus heading that row can elicit the other stimuli of the series as responses. Thus in the first row of Table 3 (p. 64), A can elicit B with 2.24 units, C with 1.75 units, and so on. Now consider the meaning of these row values for a serial recall task where Ss are required to reproduce a series in order after a single presentation. Under these conditions, only the forward, adjacent associations will produce the correct order of the items; all the other associations will act in competition with these correct associations. That is, such tendencies as $A \rightarrow C$, $A \rightarrow D$, and $A \rightarrow E$ will act to displace B from its proper position in the sequence by eliciting C, D or E in its place. In serial recall, then, the sum of all row values (other than the forward, adjacent value) represents the intraserial interference inhibiting the item heading the row from producing the correct response. Table 5 lists such row totals for a nine-item list presented at a 1-sec. rate. (No total is

TABLE 5

One-trial Association Matrices for 5, 6, 7, 8 and 9-item Lists¹

	A	B	C	D	E	5 F	6 G	7 H	8 I	9	Serial Position	Interference Totals for 9-item list	Log
*	---	---	---	---	---	---	---	---	---		1	(-----) ²	(0.84)
A	---	246	212	182	157	135	116	100	86		2	9.880	1.00
B	234	---	246	212	182	157	135	116	100		3	11.360	1.06
C	191	234	---	246	212	182	157	135	116		4	12.274	1.09
D	157	191	234	---	246	212	182	157	135		5	12.679	1.10
E	128	157	191	234	---	246	212	182	157		6	12.613	1.10
F	105	128	157	191	234	---	246	212	182		7	12.096	1.08
G	86	105	128	157	191	234	---	246	212		8	11.133	1.05
H	71	86	105	128	157	191	234	---	246		9	9.722	0.99

¹All entries to two decimals (i.e. 246 = 2.46)²No interference can be computed for serial position no. 1 because its stimulus is not known; therefore the log value for 100% recall has been arbitrarily entered.

given for the first item because its stimulus is not known, and therefore the kind and amount of interference to which it is subjected are not known either.) These row totals show minimum interference for the first few items, maximum interference for the middle items, and an intermediate amount of interference for the final items of the series. In short, these interference values describe a typical, asymmetrical, serial-position curve.

To test the power of the stimulus-trace model a family of such serial-position curves may be generated for lists of different length and for lists of different rates of presentation, and then these curves can be compared with empirical data. Let us consider first the curves for 5, 6, 7, 8 and 9 items each of which are presented at a 1-sec. rate. Jahnke (1963) has reported serial recall data for this set of conditions. He tested different Ss on each length list each of which was presented just once. Lists were composed of English consonants (which were equated across lists for average frequency of occurrence). The Ss were instructed (1) to learn the items in order and to write them down in order, and (2) to draw a dash during the recall test for any position in which an item had been forgotten (i.e. guessing was minimized). Results were reported as proportion of total consonants recalled correctly at each position by the Ss. The solid-line curves in Figure 22 show his results.

Jahnke's data can be compared with theoretical curves generated by the stimulus-trace model. All the values necessary for generating the theoretical curve for each length list are contained in Table 5. The row totals for the nine-item list are already listed on the right-hand margin along with their log values. The row totals for the shorter lists are

Fig. 22. Theoretical vs. Empirical Serial-position Curves for 5, 6, 7 and 9 Item Lists Presented at 1 item per Second

obtained by summing the row values within the appropriate, restricted portions of the larger matrix and converting to logs. Figure 22 shows these log associative-interference values for each position in each length list plotted as dashed lines on the inverted, right-hand ordinate. Each theoretical curve is matched by its empirical counterpart. These empirical values were plotted using the left-hand ordinate. An alignment of the left and right ordinates has been made which is favorable for the comparison of the theoretical and empirical curves. However, this particular alignment will be retained in all further comparisons. It is useful, therefore, to think of the theoretical log value of .84 units (which is opposite 100% recall) as a "threshold" intraserial interference value above which recall will be degraded.

Comparison of the theoretical and empirical curves shows that with log increases in the intraserial interference from competing associations, there is a linear decline in the proportion of correct recalls. Both theoretical and empirical curves show a greater primacy than recency effect and a maximum dip past the middle of the list.

Serial-position curves in free recall

In the free recall experiments, Ss are asked to recall as many items as they can in any order they wish following a single presentation of the series. The experimenter does not supply any stimulus for evoking specific responses as he does in the serial-learning or free-association techniques, nor does he specify the direction of recall as he does in digit span or serial recall. Under these circumstances, Ss are free to utilize either forward recall, backward recall or both. The following experiment was

designed to determine which procedure Ss used when presented three, successive lists. It was hypothesized that a forward recall procedure would be employed for the first of the three lists, but that Ss would switch to a backward recall procedure for the second and third lists in order to (1) avoid proactive interference effects from the first list and (2) capitalize on the recency of the final items. If the hypothesis of a switch in strategy is correct, then for the first list all the association values except the forward adjacent value should be summed up to yield a predictive, serial-position curve. (That is, free recall would be the same as serial recall following a single list.) For the second and third lists, however, all the association values except the backward, adjacent value should be added up. The former (forward recall process) will yield a bow-shaped curve with greater primacy than recency effect, whereas the latter (backward recall process) will yield a bow-shaped curve with a greater recency than primacy effect.

Preliminary support for the hypothesis of a switch in strategy has been provided by Dallett (1963), Wittig (1964), Jahnke (1965) and Wing and Thomson (1965). However, in the Wing and Thomson study which used three, different-length lists of nonsense syllables the switch only occurred with a 5-sec. rate of presentation. With a 2-sec. rate of presentation Ss displayed greater recency than primacy on the very first list of nonsense syllables, regardless of list length. This finding does not accord with Wittig's results for English nouns presented at a 2-sec. rate. Although there is a possibility of the difference in results being due to type of stimulus materials, it is considered more likely that some

slight variation in Wing and Thomson's procedure may have produced discrepant results at the 2-sec. rate. The present experiment repeats their 2-sec. condition. A 1-sec. rate is also incorporated into the design, so that six conditions obtain. These conditions represent all combinations of two rates of presentation (1 sec. and 2 sec.) and three list lengths (10, 15, and 20 items).

Stimulus materials

The stimuli were high association-value nonsense syllables. The materials are reproduced in Appendix D.

Procedure

A total of 288 Antioch College students were randomly assigned to the six conditions of three list lengths and two rates of presentation, so that independent groups of $N = 48$ took each condition. To control for effects of item uniqueness, each of these six groups was divided into six subgroups of $N = 8$. Three of these subgroups each received a different list of the same length; the other three received the same three lists but projected on the screen in reverse order. It was hoped that the use of three lists of each length in both a forward and reverse sequence would produce smoother, more reliable serial-position curves.

All lists were presented by automatic slide projector using either a 1-sec. or 2-sec. interstimulus interval and a .5-sec. exposure time. After seeing the list just once, Ss were allowed 1.5 minutes to write down in any order as many items as they could recall. Results were scored in terms of proportion of correct recalls at each serial-position.

Results and discussion

Figures 23 and 24 present the results for the first of the three lists. Figure 23 shows the serial-position curves for the 10, 15 and 20-item lists presented at a 1-sec. rate while Figure 24 shows the serial-position curves for the 10, 15 and 20-item lists presented at a 2-sec. rate. All six curves show greater primacy than recency. This suggests that Ss used the forward recall process just as they do in serial recall experiments. The predictive curves generated by the stimulus-trace model are also shown in Figures 23 and 24. These curves were obtained under the assumption that all associations except the forward-adjacent association introduce competition. Appendix E gives the association-matrices and row totals from which these curves were constructed. Comparison of the family of empirical curves with the family of theoretical curves shows essential agreement, although recall tends to be somewhat lower at the 2-sec. rate than is predicted by the model.

Figures 25 and 26 show the results for the second of the three lists. For both a 1-sec. rate (Figure 25) and a 2-sec. rate (Figure 26) there is a reversal in the magnitude of the primacy-recency effects. (Almost identical curves were obtained for the third list of the series so these are not shown.) This reversal of primacy and recency confirms the hypothesis. It suggests that at least some Ss have already switched to use of backward recall upon presentation of the second list. More direct evidence that this is so can be obtained by analysing the order in which Ss recall the items. Table 6 compares the recall order of the first three responses given in the first list and second list of the series.

Fig. 23. Theoretical vs. Empirical Serial-position Curves for First-Presented Lists of 10, 15, and 20 items (1-sec. Rate)

Fig. 24. Theoretical vs. Empirical Serial-position Curves for First-Presented Lists of 10, 15, and 20 Items (2-sec. Rate)

Fig. 25. Theoretical vs. Empirical Serial-position Curves for Second-Presented Lists of 10, 15 and 20 Items (1 1/2 sec. Rate)

Fig. 26. Theoretical vs. Empirical Serial-position Curves for Second-Presented Lists of 10, 15 and 20 Items (2-sec. Rate)

Table 6

RECALL ORDER AS A FUNCTION OF SERIAL POSITION

I First List

1-SEC RATE										2-SEC RATE									
Posit.	10-item			15-item			20-item			Recall Order	10-item			15-item			20-item		
	1	2	3	1	2	3	1	2	3		1	2	3	1	2	3	1	2	3
n	31	7	2	28	6	2	21	6	3		41	6	1	28	6	5	32	6	3
n+1	7	17	6	0	11	8	2	10	2		2	31	7	4	23	7	8	20	9
n+2	3	3	13	1	2	7	4	1	8		5	5	26	5	4	19	4	3	13
L-2	4	2	14	1	0	3	1	1	3		4	6	20	9	7	14	10	7	16
L-1	1	11	2	3	8	1	1	8	3		9	16	7	8	21	8	4	18	6
L	22	4	2	13	1	2	20	2	4		12	8	21	10	8	20	9	6	21
<u>II Second List</u>																			
n	8	4	4	7	4	6	5	2	7		17	8	12	17	9	8	17	7	7
n+1	4	8	5	3	8	4	1	8	4		9	16	10	6	18	5	4	18	6
n+2	3	6	6	1	5	8	2	1	7		10	10	14	8	7	14	7	5	14
L-2	4	2	11	6	4	16	4	4	18		6	4	26	3	2	29	6	9	22
L-1	3	17	6	5	26	6	6	24	6		7	22	8	9	26	7	9	27	6
L	27	4	11	32	7	8	38	2	7		28	6	12	31	9	8	33	3	10

*Where the list goes from n to L.

By the second list, the predominance of first responses already comes from the final three positions rather than the initial three positions. The McNemar test for significance of changes can be employed to determine if a change in frequency of recall at a given position has occurred from the first to second list of the series. Table 7 summarizes the results of this statistical test which generally confirm the interpretation already adopted.

The most parsimonious view of the present findings, then, is that Ss utilize different recall strategies in free recall depending upon whether the list they are recalling is the first or a later list in a series. The characteristics of the serial-position curves differ. A greater primacy effect is obtained for the first list, whereas a greater recency effect is obtained for any subsequent lists. These differential shapes are directly generated by the stimulus-trace model if it is assumed that (1) a forward recall strategy is employed for the first list, and (2) a backward recall strategy is employed for any subsequent list. The present findings, then, confirm the power of the stimulus-trace model to predict free recall performance as well as serial recall performance.

Serial-position curves in mixed-strategy recall

One study by Yntema, Wozencraft and Klem (1964) tested immediate recall of 3-item, 5-item, 7-item and 9-item lists of digits presented at .500-sec., .250-sec., .125-sec., and .100-sec. rates. All conditions were given to all Ss and following each list Ss were required to guess which digits occupied each serial position. Although they cite their study as a serial recall study, their procedure of running Ss under all

Table 7

McNEMAR TEST FOR SIGNIFICANCE OF CHANGES*

List 1 vs. List 2

(1-sec. Rate)

Position**	<u>10-item List</u>	<u>15-item List</u>	<u>20-item List</u>
N	< .001	< .001	< .001
N+1	< .001	< .001	< .001
N+2	< .01	< .001	< .001
L-2	< .01	< .001	< .001
L-1	< .01	< .001	< .001
L	< .01	< .001	< .001

List 1 vs. List 2

(2-sec. Rate)

Position	<u>10-item List</u>	<u>15-item List</u>	<u>20-item List</u>
N	< .001	< .001	< .001
N+1	< .001	< .001	< .001
N+2	< .01	< .001	< .001
L-2	< .01	< .001	< .001
L-1	< .01	< .001	< .001
L	< .001	< .001	< .001

* Based on one-tailed test as given by Siegel (1956, p. 249)

** Where list goes from N to L

conditions as well as their stricture to fill in a response at every position may have encouraged Ss to retrieve the final item(s) of each list by backward recall. At least, only under this assumption will the present model generate their serial-position data. In Appendix F are presented the matrices for all four rates of presentation. These matrices have been used to generate a mixed recall strategy, i.e. one in which Ss retrieve optimally using both forward and backward recall. The serial-position curves based on these matrices utilize whichever associative-interference value is lowest. Thus, the initial items of each list are recalled by means of forward associations while the final items are recalled by means of backward associations. The serial-position curves based on the mixed-strategy hypothesis are shown in Figures 27, 28 and 29. (No 3-item lists were generated because the authors' error data were too meagre for lists of this length and so were not reported by them.) These theoretical curves may be compared with the empirical curves shown in Figures 30, 31 and 32 which are based upon the proportion of correct recalls at each position. The overall magnitude of recall is very well predicted; however, the shapes of the theoretical curves are all flat, whereas the empirical curves show a sharper dip at the middle serial-positions. The reasons for such a discrepancy are not apparent. Although the model fails to produce the shape of these curves, it is encouraging that it does generate the magnitude of recall for different length lists presented at such very fast rates.

Fig. 27. Theoretical Serial-position Curves for Five-item Lists Presented at 2, 4, 8 and 10 Items per Second

Fig. 28. Theoretical Serial-position Curves for Seven-item Lists Presented at 2, 4, 8 and 10 Items per Second

Fig. 29. Theoretical Serial-position Curves for Nine-item Lists Presented at 2, 4, 8, and 10 items per Second

Fig. 30. Empirical Serial-position Curves for Five-item Lists Presented at 2, 4, 8 and 10 Items per second.

Fig. 31. Empirical Serial-position Curves for Seven-item Lists Presented at 2, 4, 8 and 10 Items per Second

Fig. 32. Empirical Serial-position Curves for Nine-item Lists Presented at 2, 4, 8 and 10 Items per Second

PART IV

EVALUATION OF THE MODEL

In evaluating any model there is a question of its probable validity. The stimulus-trace model has an advantage in this respect because of its long historical priority. Many of the best minds who considered the problems of serial acquisition and recall were led by extensive introspection to adopt the trace hypothesis. Thus it possesses a great deal of face validity.

Even the most respectable and ancient hypotheses, however, must be able to handle a wide range of data gathered in the psychological laboratory. Otherwise they possess only face validity. The question, then, is whether or not stimulus-trace theory has sufficiently wide-ranging explanatory power to warrant its continued use. All the present evidence suggests that it does. This evidence may be briefly summarized.

Strengths of the model

First, the stimulus-trace, interaction model can generate the data on remote associations. Thus it bridges a gap between intraserial interference theory and the simple decay theory of memory. Even today these two theories have been put up against one another as if they were incompatible. Because the model emphasizes trace interaction rather than trace decay per se, it provides a trace explanation for competing responses. It subsumes intraserial interference theory by providing a lower-level explanation of the same phenomena. All the predictive power of interference

theory thereby becomes (at least potentially) predictive power for trace-interaction theory. We have only examined a few cases in point: the serial-position curve for serial recall; the serial-position curve for free recall; and the serial-position curve for a mixed recall strategy. There is every reason to suppose, however, that the model will handle other phenomena such as the Marbè effect (Bousfield, Cohen and Silva, 1956) and the von Restorff effect (von Restorff, 1933). Wing (1962, 1964) and Thomson and Wing (1965) have shown, for example, that the von Restorff effect can be predicted using intraserial interference rating scales whose properties may be derived directly from the present trace model. However, if there is to be wider application of trace theory, it should be done methodically. The fewest possible postulates should be added. From the present vantage point, only two or three additional assumptions need to be made. One would concern the effect of frequency of rehearsal of a trace. It was noted that the forward, adjacent associations produced a disproportionately large number of free-associations (as if S_s had augmented the trace strength by rehearsing the association during the interstimulus interval). Another additional postulate would deal with the relative amounts of associative interference amongst various classes of stimulus materials (e.g. digits, consonants, CVC's, CCC's, and English nouns). More preferably, these classes would be defined phonemically since verbal memory seems to be intimately related to speech-motor patterns (Wickelgren, 1965, 1966). As it is, there is already much additional data which the model can presumably handle (such as digit span gradients); but space and time have not allowed us to deal with data from every experimental paradigm.

Weakness of the model

So far trace theory has been granted rather extraordinary virtues. It is important to ascribe these to trace theory in general rather than to the particular model developed here. There is every reason to suppose that the model presented here will succumb to a series of revisions which might make it virtually unrecognizable. There are major assumptions in it which certainly will not stand the test of time. These might include the basic formula for the trace and the type of transformation which produces the differential strength of forward and backward associations. The value of the present model lies in the fact that it works well enough (whether it is correct in detail or not) that it thereby establishes that trace-interactions will produce the kinds of associations which various serial phenomena require.

Possible reformulation of the model along stochastic lines

One of those kinds of revisions which might make the model almost unrecognizable (but substantially improve it) is to reformulate it in stochastic terms. There is a fundamental weakness in the concept of a stimulus as a single entity which has an associated trace. A more reasonable view is to treat any event (such as the presentation of a word or syllable) as a set of stimuli. The composition of stimulus elements in the set are assumed to fluctuate or change from time-to-time, and S_s are considered to be sampling the set whenever an observing response is made. From any sampled set, it may be postulated that the number of original conditioned elements in that set which are still available from one

moment to the next declines according to an exponential function. Thus the probability of two successive sets being associated also declines as the proportion of original conditioned elements in each set falls off in time. Such a restatement of trace theory has the distinct advantage of placing it within the context of general probability theory and of mathematical learning theory in particular.

APPENDIX A

CALCULATIONS FOR PETERSON STUDY

The effective trace strength, i.e. that part of the trace not subjected to interpolated interference is found by evaluating the integral of $e^{-.2t}$ from time t_1 to time t_2 and adding to it .200 which is the area under the incremental segment of the trace.

The following expression was used for making calculations:

$$\int_{t_2}^{t_1} e^{-.2t} dt + .200 = -5 (e^{-.2t_2} - e^{-.2t_1}) + .200$$

$$= 5 (e^{-.2t_1} - e^{-.2t_2}) + .200$$

3-sec Delay		6-sec Delay	
Latency	Log		Log
2	$5 (e^{-.2(3)} - e^{-.2(5)}) + .200 = 1.105 = -.430$	$5 (e^{-.2(6)} - e^{-.2(8)}) + .200 = .695 = -.158$	
4	$5 (e^{-.2(3)} - e^{-.2(7)}) + .200 = 1.710 = -.233$	$5 (e^{-.2(6)} - e^{-.2(10)}) + .200 = 1.030 = .013$	
6	$5 (e^{-.2(3)} - e^{-.2(9)}) + .200 = 2.120 = -.326$	$5 (e^{-.2(6)} - e^{-.2(12)}) + .200 = 1.250 = .097$	
8	$5 (e^{-.2(3)} - e^{-.2(11)}) + .200 = 2.390 = -.378$	$5 (e^{-.2(6)} - e^{-.2(14)}) + .200 = 1.400 = .146$	
10	$5 (e^{-.2(3)} - e^{-.2(13)}) + .200 = 2.575 = -.411$	$5 (e^{-.2(6)} - e^{-.2(16)}) + .200 = 1.500 = .176$	
12	$5 (e^{-.2(3)} - e^{-.2(15)}) + .200 = 2.695 = -.431$	$5 (e^{-.2(6)} - e^{-.2(18)}) + .200 = 1.570 = .196$	
14	$5 (e^{-.2(3)} - e^{-.2(17)}) + .200 = 2.780 = -.444$	$5 (e^{-.2(6)} - e^{-.2(20)}) + .200 = 1.615 = .208$	

9-sec Delay			12-sec Delay		
Latency		Log			Log
2	5	$(e^{-.2(9)} - e^{-.2(11)}) + .200 = .470 = -.328$	5	$(e^{-.2(12)} - e^{-.2(14)}) + .200 = .350 = -.456$	
4	5	$(e^{-.2(9)} - e^{-.2(13)}) + .200 = .655 = -.185$	5	$(e^{-.2(12)} - e^{-.2(16)}) + .200 = .450 = -.347$	
6	5	$(e^{-.2(9)} - e^{-.2(15)}) + .200 = .775 = -.111$	5	$(e^{-.2(12)} - e^{-.2(18)}) + .200 = .520 = -.284$	
8	5	$(e^{-.2(9)} - e^{-.2(17)}) + .200 = .860 = -.066$	5	$(e^{-.2(12)} - e^{-.2(20)}) + .200 = .565 = -.248$	
10	5	$(e^{-.2(9)} - e^{-.2(19)}) + .200 = .915 = -.039$	5	$(e^{-.2(12)} - e^{-.2(22)}) + .200 = .595 = -.226$	
12	5	$(e^{-.2(9)} - e^{-.2(21)}) + .200 = .950 = -.022$	5	$(e^{-.2(12)} - e^{-.2(24)}) + .200 = .615 = -.211$	
14	5	$(e^{-.2(9)} - e^{-.2(23)}) + .200 = .975 = -.011$	5	$(e^{-.2(12)} - e^{-.2(26)}) + .200 = .625 = -.204$	
15-sec Delay			18-sec Delay		
Latency		Log			Log
2	5	$(e^{-.2(15)} - e^{-.2(17)}) + .200 = .285 = -.545$	5	$(e^{-.2(18)} - e^{-.2(20)}) + .200 = .245 = -.611$	
4	5	$(e^{-.2(15)} - e^{-.2(19)}) + .200 = .340 = -.469$	5	$(e^{-.2(18)} - e^{-.2(22)}) + .200 = .275 = -.561$	
6	5	$(e^{-.2(17)} - e^{-.2(21)}) + .200 = .375 = -.426$	5	$(e^{-.2(18)} - e^{-.2(24)}) + .200 = .295 = -.530$	
8	5	$(e^{-.2(17)} - e^{-.2(23)}) + .200 = .400 = -.398$	5	$(e^{-.2(18)} - e^{-.2(26)}) + .200 = .305 = -.516$	
10	5	$(e^{-.2(17)} - e^{-.2(25)}) + .200 = .415 = -.382$	5	$(e^{-.2(18)} - e^{-.2(28)}) + .200 = .315 = -.502$	
12	5	$(e^{-.2(17)} - e^{-.2(27)}) + .200 = .425 = -.372$	5	$(e^{-.2(18)} - e^{-.2(30)}) + .200 = .325 = -.488$	
14	5	$(e^{-.2(17)} - e^{-.2(29)}) + .200 = .435 = -.362$	5	$(e^{-.2(18)} - e^{-.2(32)}) + .200 = .325 = -.488$	

APPENDIX B*

Derivation of equations (3) and (4)

Let the "t" of equation (2) continue to represent the elapsed time since any stimulus was presented. Then let E = any earlier trace in time and let L = any later trace in time. If this is the case, then, if E was presented at time p and if L was presented at time q,

$$E = e^{-b(t-p)}$$

$$L = e^{-b(t-q)}$$

Then, according to the model presented in the paper, where the instantaneous forward association is $(E)^k L$, the total forward association at time t would be given by

$$\int_q^t e^{-kb(t-p) - b(t-q)} dt$$

where $t = \infty$

$$\begin{aligned} &= e^{kbp+bq} \int_q^t e^{-(k+1)bt} dt \\ &= \frac{e^{kbp+bq} \left[-e^{-(k+1)bt} + e^{-(k+1)bq} \right]}{(k+1)b} \\ &= \frac{e^{-kb(q-p)} \left[-e^{-b k(t-q) + k(q-p) + (t-q)} \right]}{b(k+1)} \\ &= \frac{e^{-kb(q-p)}}{b(k+1)} \left[1 - e^{-b(k+1)(t-q)} \right] \end{aligned}$$

* This derivation was made by Dr. William B. Houston, Jr., Assistant Professor of Mathematics, Antioch College. The derivation omits many intermediate steps for the sake of brevity. The author also wishes to express his appreciation to Jerry Satlow, of Behavior Research Laboratory, Antioch College, Yellow Springs, Ohio.

The limiting value for this expression of the forward association is, therefore,

$$\frac{e^{-kb(\underline{q}-p)}}{b(k+1)}$$

Since $\underline{q}-p$ is the time lag between stimuli, this may be more conveniently represented as λ , giving

$$\frac{e^{-kb\lambda}}{b(k+1)} \quad \text{or alternatively,} \quad \frac{1}{b(k+1) e^{kb\lambda}}$$

which is the equation cited as equation (3) on page 58 of the text.

Similarly, if $E(L)^k$ is the instantaneous backward association, then the total backward association between stimuli E and L may be given by

$$\int_{\underline{q}}^t e^{-b(t-p)} e^{-b(t-\underline{q})k} dt$$

where $t = \infty$.

For convenience, if we let $\lambda = \underline{q}-p$, $u = t-\underline{q}$, and $t = u+\underline{q}$, then the expression becomes

$$\begin{aligned} & \int_{\underline{q}}^{\infty} e^{-b(u+\lambda+ku)} du \\ &= e^{-b\lambda} \int_{\underline{q}}^{\infty} e^{-b(k+1)u} du \end{aligned}$$

The limiting value for this expression of the backward association is, therefore,

$$\frac{e^{-b\lambda}}{b(k+1)} \quad \text{or alternatively,} \quad \frac{1}{b(k+1) e^{b\lambda}}$$

which is the equation cited as equation (4) on page 58 of the text.

MATRICES AND CALCULATIONS FOR REMOTE ASSOCIATION GRADIENTS

Intra-trial Associations for 10-item List Presented at 1.8-sec. Rate for a Single Trial

---	218	167	127	97	74	57	43	33	25
199	---	218	167	127	97	74	57	43	33
139	199	---	218	167	127	97	74	57	43
97	139	199	---	218	167	127	97	74	57
68	97	139	199	---	218	167	127	97	74
47	68	97	139	199	---	218	167	127	97
33	47	68	97	139	199	---	218	167	127
23	33	47	68	97	139	199	---	218	167
16	23	33	47	68	97	139	199	---	218
11	16	23	33	47	68	97	139	199	---

Inter-trial Associations between Any Two Trials for 10-item List
Presented at 1.8-sec. Rate with 5-sec. Inter-trial Interval

---	6	8	12	17	25	36	51	73	105
15	---	6	8	12	17	25	36	51	73
20	15	---	6	8	12	17	25	36	51
27	20	15	---	6	8	12	17	25	36
35	27	20	15	---	6	8	12	17	25
46	35	27	20	15	---	6	8	12	17
60	46	35	27	20	15	---	6	8	12
79	60	46	35	27	20	15	---	6	8
103	79	60	46	35	27	20	15	---	6
135	103	79	60	46	35	27	20	15	---

Combined Intra-trial and Inter-trial Matrix for 10-item
List at 1.8-sec. Rate with 5-sec. Inter-trial Interval*

108

---	224	175	139	115	99	92	94	106	130
215	---	224	175	139	115	99	92	94	106
159	215	---	224	175	139	115	99	92	94
124	159	215	---	224	175	139	115	99	92
103	124	159	215	---	224	175	139	115	99
93	103	124	159	215	---	224	175	139	115
93	93	103	124	159	215	---	224	175	139
102	93	93	103	124	159	215	---	224	175
119	102	93	93	103	124	159	215	---	224
146	119	102	93	93	103	124	159	215	---

* Note: Values are for any trial but the first and last trial. Values in each matrix were computed by formulas (3) and (4) on page 58 of the text. The values entered in the matrices have been rounded from four-place values actually used; therefore the sum of the values for the intra-trial and inter-trial elements may not exactly equal the value given for the combined matrix element.

Intra-trial Associations for 8-item List Presented
at 2.5-sec. Rate for a Single Trial

---	196	135	93	44	30	21	14
173	---	196	135	93	44	30	21
105	173	---	196	135	93	44	30
64	105	173	---	196	135	93	44
39	64	105	173	---	196	135	93
23	39	64	105	173	---	196	135
14	23	39	64	105	173	---	196
9	14	23	39	64	105	173	---

Inter-trial Associations between Any Two Trials for
8-item List Presented at 2.5-sec. Rate with 6-sec.
Inter-trial Interval

---	4	7	12	19	32	52	86
12	---	4	7	12	19	32	52
18	12	---	4	7	12	19	32
26	18	12	---	4	7	12	19
38	26	18	12	---	4	7	12
55	38	26	18	12	---	4	7
80	55	38	26	18	12	---	4
116	80	55	38	26	18	12	---

Combined Intra-trial and Inter-trial Matrix for
an 8-item List Presented at 2.5-sec. Rate with
a 6-sec. Inter-trial Interval

---	201	142	105	63	62	73	100
186	---	201	142	105	63	62	73
123	186	---	201	142	105	63	62
90	123	186	---	201	142	105	63
76	90	123	186	---	201	142	105
78	76	90	123	186	---	201	142
94	78	76	90	123	186	---	201
125	94	78	76	90	123	186	---

Note: Values are for any trial but the first and last trial. Values in each matrix were computed by formulas (3) and (4) on page 58 of the text. The values entered in the matrices have been rounded from four-place values actually used; therefore the sum of the values for the intra-trial and inter-trial elements may not exactly equal the value given for the combined matrix element.

APPENDIX D

STIMULUS MATERIALS FOR FREE RECALL*

(100% Association-Value)

10-item Lists

SOF	SEM	RUF	SOF	SEM	RUF
DUC	DOV	TAM	DUC	DOV	TAM
PHIL	HAZ	BEV	PIL	HAZ	BEV
TER	GUL	SOC	TER	GUL	SOC
WOM	BEC	DUL	WOM	BEC	DUL
FAS	JIN	VIN	FAS	JIN	VIN
BIZ	ROF	WAT	BIZ	ROP	WAT
JEL	TAK	MEX	JEL	TAK	MEX
LON	CUS	HUR	LON	CUS	HUR
SUR	QIL	DOB	SUR	QIL	DOB
			FAM	KEN	LIF
			BER	WOM	VAC
			NOV	RUD	JUS
			LK	WAK	COL
			VES	VIF	WIS

* Each list served equally often as first, second and third list in the series and was given to half of each group as a forward sequence and to half as a backward sequence. For each forward sequence sub-group of eight Ss, the starting-point of the list was randomly-selected; but for the backward sequence sub-group, the mirror-image of the list was given so that a new, random, starting-point was not chosen.

20-item Lists

SOF	SEM	RUF
DUC	DOV	TAM
PIL	HAZ	BEV
TER	GUL	SOC
WOM	BEC	DUL
FAS	JIN	VIN
BIZ	ROP	WAT
JEL	TAK	MEX
LON	CUS	HUR
SUR	QIL	DOB
FAM	KEN	LIF
BER	WOM	VAC
NOV	RUD	JUS
LIK	WAK	COL
VES	VIF	WIS
PUF	HOB	NAV
RAV	TEX	REG
TIG	NAR	HON
CEN	SUT	DIS
HOL	MOD	PAC

APPENDIX E

MATRICES AND CALCULATIONS FOR FREE RECALL

The matrices are too large to conveniently reproduce here; therefore the values for every element in every matrix are reproduced instead.

Values for 1-sec. Rate Matrices

λ	$b\lambda$	<u>Backward</u>		<u>Forward</u>		
		$e^{-b\lambda}$	$e^{-b\lambda}/b(k+1)$	$bk\lambda$	$e^{-bk\lambda}$	$e^{-bk\lambda}/b(k+1)$
1	0.2	.819	2.340	.15	.861	2.460
2	0.4	.670	1.914	.30	.741	2.117
3	0.6	.549	1.568	.45	.638	1.823
4	0.8	.449	1.283	.60	.549	1.568
5	1.0	.368	1.051	.75	.472	1.349
6	1.2	.301	.860	.90	.407	1.163
7	1.4	.247	.706	1.05	.350	1.000
8	1.6	.202	.577	1.20	.301	.860
9	1.8	.165	.471	1.35	.259	.740
10	2.0	.135	.386	1.50	.223	.637
11	2.2	.111	.317	1.65	.192	.549
12	2.4	.091	.260	1.80	.165	.471
13	2.6	.074	.211	1.95	.142	.406
14	2.8	.061	.174	2.10	.122	.349
15	3.0	.050	.143	2.25	.105	.300
16	3.2	.041	.117	2.40	.091	.260
17	3.4	.033	.093	2.55	.078	.223
18	3.6	.027	.077	2.70	.067	.191
19	3.8	.022	.063	2.85	.058	.166

Values for 2-sec. Rate Matrices

λ	Backward			Forward		
	$b\lambda$	$e^{-b\lambda}$	$e^{-b\lambda}/b(k+1)$	$bk\lambda$	$e^{-bk\lambda}$	$e^{-bk\lambda}/b(k+1)$
2	.4	.670	1.914	.30	.741	2.117
4	.8	.449	1.283	.60	.549	1.568
6	1.2	.301	.860	.90	.407	1.163
8	1.6	.202	.577	1.20	.301	.860
10	2.0	.135	.386	1.50	.223	.637
12	2.4	.091	.260	1.80	.165	.471
14	2.8	.061	.174	2.10	.122	.349
16	3.2	.041	.117	2.40	.091	.260
18	3.6	.027	.077	2.70	.067	.191
20	4.0	.018	.051	3.00	.050	.143
22	4.4	.012	.034	3.30	.037	.106
24	4.8	.008	.023	3.60	.027	.077
26	5.2	.006	.017	3.90	.020	.057
28	5.6	.004	.011	4.20	.015	.043
30	6.0	.002	.006	4.50	.011	.031
32	6.4	.002	.006	4.80	.008	.023
34	6.8	.001	.003	5.10	.006	.017
36	7.2	.001	.003	5.40	.005	.014
38	7.6	.000	.000	5.70	.003	.009

Instructions for Constructing Matrices

Matrices for any length list or rate of presentation may be constructed as shown in Table 5 (p. 78). This table shows 5, 6, 7, 8 and 9-item lists using the 1-sec. rate values (rounded to hundredths). This table may be used as a guideline.

APPENDIX F

CALCULATIONS FOR MIXED STRATEGY RECALL

Matrices for 5, 7, and 9-item lists were constructed for each rate of presentation (.100, .125, .250 and .500) exactly as was done in Appendix E. Then all values in a row except the forward adjacent were added to get forward strategy interference (F); and all values except backward adjacent were added to get backward strategy interference (B). Finally, that value which was lowest for any given serial position was chosen as the mixed-strategy value, yielding the following theoretical serial-position values:

S. P.	Strategy	<u>.100</u>		<u>.125</u>		<u>.250</u>		<u>.500</u>	
		<u>Interf.</u>	<u>Log</u>	<u>Interf.</u>	<u>Log</u>	<u>Interf.</u>	<u>Log</u>	<u>Interf.</u>	<u>Log</u>
1	F	---	(.840)	---	(.840)	---	(.840)	---	(.840)
2	F	8.193	.913	8.123	.910	7.653	.884	6.857	.836
3	B	8.251	.917	8.168	.912	7.812	.893	7.108	.852
4	B	8.074	.907	7.954	.900	7.386	.869	6.371	.804
5	B	---	(.840)	---	(.840)	---	(.840)	---	(.840)
1	F	---	(.840)	---	(.840)	---	(.840)	---	(.840)
2	F	13.455	1.129	13.259	1.123	12.242	1.088	10.589	1.025
3	F	13.644	1.135	13.505	1.130	12.773	1.106	11.352	1.055
4	B	13.659	1.135	13.520	1.131	12.803	1.107	11.482	1.060
5	B	13.474	1.129	13.277	1.123	12.378	1.093	10.756	1.020
6	B	13.194	1.120	12.937	1.112	11.729	1.069	9.673	.985
7	B	---	(.840)	---	(.840)	---	(.840)	---	(.840)
1	F	---	(.840)	---	(.840)	---	(.840)	---	(.840)
2	F	18.560	1.269	18.227	1.261	16.550	1.219	13.902	1.143
3	F	18.826	1.275	18.555	1.269	17.150	1.134	14.920	1.174
4	F	19.001	1.279	18.764	1.273	17.545	1.244	15.569	1.192
5	B	18.976	1.278	18.749	1.273	17.571	1.245	15.496	1.190
6	B	18.779	1.274	18.503	1.267	17.146	1.234	14.784	1.170
7	B	18.491	1.267	18.134	1.258	16.509	1.218	13.744	1.138
8	B	18.111	1.258	17.674	1.247	15.657	1.195	12.376	1.083
9	B	---	(.840)	---	(.840)	---	(.840)	---	(.840)

REFERENCES

- Aristotle. De memoria et reminiscentia. (Trans. by J. I. Beare in Ross, W. D. (Ed.) The works of Aristotle, Vol. III, Oxford: Clarendon Press, 1931.
- Asch, S. E., and Ebenholtz, S. M. The principle of associative symmetry. Proceed. American Philos. Soc., 1962, 106, 135-163.
- Bersh, P. J. The influence of two variables upon the establishment of a secondary reinforcer for operant responses. J. exp. Psychol. 1951, 41, 62-73.
- Boneau, C. A. The interstimulus interval and the latency of the conditioned eyelid response. J. exp. Psychol., 1958, 56, 464-472.
- Bousfield, W. A., Whitmarsh, G. A., and Esterson, J. Serial position effects and the "Marbe effect" in the free recall of meaningful words. J. gen. Psychol., 1958, 59, 255-262.
- Bower, G. H. Correlated delay of reinforcement. J. comp. physiol. Psychol., 1961, 54, 196-203.
- Bruce, D., and Cofer, C.N. A comparison of recognition and recall in short-term memory. Proceed. 73rd Ann. Conv. Amer. Psychol. Assoc., 1965, 81-82.
- Bugelski, B. R. A remote association explanation of the relative difficulty of learning nonsense syllables in a serial list. J. exp. Psychol., 1950, 40, 336-348.
- Burnham, W. H. Retroactive amnesia: Illustrative cases and a tentative explanation. Amer. J. Psychol., 1903, 14, 382-396.
- Chorover, S. L. Contributed remarks in McGaugh, J. L. Facilitation and impairment of memory storage processes. In Kimble, D. P. (Ed.) The anatomy of memory, Vol.1. Palo Alto, Calif.: Science and Behavior Books, Inc., 1965, 240-291.
- Chorover, S. L., and Schiller, P. H. Short-term retrograde amnesia in rats. J. comp. physiol. Psychol., 1965, 59, 73-78.

- Church, R. M., Brush, F. R., and Solomon, R. L. Traumatic avoidance learning: The effects of CS-UCS interval with a delayed conditioning procedure in a free-responding situation. J. comp. physiol. Psychol., 1956, 49, 301-308.
- Dallett, K. M. Practice effects in free and ordered recall. J. exp. Psychol., 1963, 66, 65-71.
- Davitz, J. R., Mason, D. J., Mowrer, O. H., and Viek, P. Conditioning of fear: A function of the delay of reinforcement. Amer. J. Psychol., 1957, 70, 69-74.
- Ebel, H. C., and Prokasy, W. F. Classical eyelid conditioning as a function of sustained and shifted interstimulus intervals. J. exp. Psychol., 1963, 65, 52-58.
- Ebbinghaus, H. Memory: a contribution to experimental psychology. (Trans. by H. A. Ruger and C. E. Bussenius.) New York: Teachers College, Columbia Univ., 1913.
- Ekstrand, B. R. Backward associations. Psychol. Bull., 1966, 65, 50-64.
- Fechner, G. T. Elemente der psychophysik. Leipzig: Breitkopf und Hartel, 1860.
- Fletcher, H. J. The delayed-response problem. Chapter 4 in A. M. Schrier, H. F. Harlow, and F. Stollnitz (Eds.), Behavior of non-human primates: Modern research trends. New York: Academic Press, 1965, 129-165.
- Fullerton, G. S., and Cattell, J. McK. On the perception of small differences. Philos. Series, No. 2. Philadelphia: Univ. of Penn., 1892.
- Glickman, S. E. Perseverative neural processes and consolidation of the memory trace. Psychol. Bull., 1961, 58, 218-233.
- Gomulicki, B. R. The development and present status of the trace theory of memory. Brit. J. Psychol., Monogr. Suppl., 1953, 29, 1-94.
- Grice, G. R. The relation of secondary reinforcement to delayed reward in visual discrimination learning. J. exp. Psychol., 1948, 38, 1-16.
- Guilford, J. P., and Park, O. G. The effect of interpolated weights upon comparative judgments. Amer. J. Psychol., 1931, 43, 589-599.

- Hall, J. F. The psychology of learning. Philadelphia: Lippincott, 1966.
- Hartley, D. Observations on man, his frame, his duty, and his expectations. London, England: W. Eyres, 1801 (First published, 1794).
- Hellyer, S. Supplementary report: Frequency of stimulus presentation and short-term decrement in recall. J. exp. Psychol., 1962, 64, 650.
- Herbart, J. F. Psychologie als Wissenschaft, neu gegründet auf Erfahrung, Metaphysik und Mathematik. Königsburg: Vol. I, 1824, Vol. II, 1825.
- Horowitz, L. M., and White, Margaret A. Producing symmetrical associations. Proceed. 73rd Ann. Conv. Amer. Psychol. Assoc., 1965, 31-32.
- Horowitz, L. M., Brown, Z. M., and Weissbluth, S. Availability and the direction of associations. J. exp. Psychol., 1964, 68, 541-549.
- Horowitz, L. M., Norman, Sandra A., and Day, Ruth S. Availability and associative symmetry. Psychol. Rev., 1966, 73, 1-15.
- Hull, C. L. Mind, mechanism, and adaptive behavior. Psychol. Rev., 1937, 44, 1-32.
- Hull, C. L. Principles of behavior. New York: Appleton-Century, 1943.
- Hull, C. L. Essentials of behavior. New Haven, Conn.: Yale Univ., 1951.
- Hull, C. L. A behavior system. New Haven, Conn.: Yale Univ., 1952.
- Hull, C. L., Hovland, C. I., Ross, R. T., Hall, M., and Perkins, D. T. Mathematico-deductive theory of rote learning: a study in scientific methodology. New Haven, Conn.; Yale Univ., 1940.
- Jahnke, J. C. Serial position effects in immediate serial recall. J. verb. Learn. verb. Behav., 1963, 2, 284-287.
- Jahnke, J. C. Supplementary report: Primacy and recency effects in serial-position curves of immediate recall. J. exp. Psychol., 1965, 70, 130-132.
- Jenkins, W. O. A temporal gradient of derived reinforcement. Amer. J. Psychol., 1950, 63, 237-243.

- Jones, J. E. Contiguity and reinforcement in relation to CS-UCS intervals in classical aversive conditioning. Psychol. Rev., 1962, 69, 176-186.
- Kamin, L. J. Traumatic avoidance learning: The effects of CS-UCS interval with a trace-conditioning procedure. J. comp. physiol. Psychol., 1954, 47, 65-72.
- Kamin, L. J. The gradient delay of secondary reward in avoidance learning. J. comp. physiol. Psychol., 1957a, 50, 445-449.
- Kamin, L. J. The delay of secondary reward gradient in avoidance learning tested on avoidance trials only. J. comp. physiol. Psychol., 1957b, 50, 450-456.
- Kamin, L. J. The delay-of-punishment gradient. J. comp. physiol. Psychol., 1959, 52, 434-437.
- Keeseey, R. Intracranial reward delay and the acquisition rate of a brightness discrimination. Science, 1964, 143, 702.
- Kellogg, W. N. "Superstitious" behavior in animals. Psychol. Rev., 1949, 56, 172-175.
- Keppel, G., and Underwood, B. J. Proactive inhibition in short-term retention of single items. J. verb. Learn. verb. Behav., 1962, 1, 153-161.
- Kimble, G. A. Conditioning as a function of the time between conditioned and unconditioned stimuli. J. exp. Psychol., 1947, 37, 1-15.
- Kimble, G. A. Hilgard and Marquis' conditioning and learning. New York: Appleton-Century-Crofts, 1961.
- Koffka, K. Principles of gestalt psychology. New York: Harcourt, Brace, 1935.
- Köhler, W. Zur Theorie des Sucesswergluchs und der Zaltfehler. Psychol. Forsch., 1923, 4, 115-175.
- Laurenstein, O. Ansatz zu einer physiologischen Theorie des Vergleichs und der Zeitfehler. Psychol. Forsch., 1933, 17, 130-177.
- Lewis, D. Quantitative methods in psychology. New York: McGraw-Hill, 1960.
- Lewis, D. J., and Maher, B. A. Neural consolidation and electroconvulsive shock. Psychol. Rev., 1965, 72, 225-239.

- Loess, H., and McBurney, Judith. Short-term memory and retention-interval activity. Proceed. 73rd Ann. Conv. Amer. Psychol. Assoc., 1965, 85-86.
- McClelland, D. C. Factors influencing the time error in judgments of visual extent. J. exp. Psychol., 1943, 33, 81-95.
- McDougall, W. Experimentelle Beitrage zur Lehre von Gedachtness: Von G., E. Muller und A. Pilzecker. Mind, 1901, 10, 388-394.
- McGeoch, J. A. The direction and extent of intra-serial associations at recall. Amer. J. Psychol., 1936, 48, 221-245.
- McGeoch, J. A., and Irion, A. L. The psychology of human learning. (2nd ed.) New York: Longmans, Green, 1952.
- McNemar, Q. Note on the sampling error of the differences between correlated proportions or percentages. Psychometrika, 1947, 12, 153-157.
- Melton, A. W. Implications of short-term memory for a general theory of memory. J. verb. Learn. verb. Behav., 1963, 2, 1-21.
- Moeller, G. The CS-UCS interval in GSR conditioning. J. exp. Psychol., 1954, 48, 162-166.
- Morse, W. H., and Skinner, B. F. A second type of superstition in the pigeon. Amer. J. Psychol., 1957, 70, 308-311.
- Muller, G. E., and Pilzecker, A. Experimentelle Beitrage zur Lehre vom Gedachtnis. Z. Psychol., 1900, 1, 1-300.
- Murdock, B. B., Jr. The retention of individual items. J. exp. Psychol., 1961, 62, 618-625.
- Murdock, B. B., Jr. Direction of recall in short-term memory. J. verb. Learn. verb. Behav., 1962, 1, 119-124.
- Needham, J. G. The time error as a function of continued experimentation. Amer. J. Psychol., 1934, 46, 558-567.
- Needham, J. G. The effect of the time interval upon the time error at different intensive levels. J. exp. Psychol., 1935, 18, 530-543.
- Pavlov, I. P. Lectures on conditioned reflexes. New York: Intern. Publ., 1928.

- Peterson, L. R. Short-term verbal memory and learning. Psychol. Rev., 1966, 73, 193-207.
- Peterson, L. R., and Peterson, M. J. Short-term retention of individual verbal items. J. exp. Psychol., 1959, 58, 193-198.
- Prokasy, W. F., Ebel, H. C., and Thompson, D. D. Response shaping at long interstimulus intervals in classical eyelid conditioning. J. exp. Psychol., 1962, 66, 138-141.
- Quartermain, D., Paolino, R. M., and Miller, N. E. A brief temporal gradient of retrograde amnesia independent of situational change. Science, 1965, 149, 1116-1118.
- Raskin, E., and Cook, S. W. The strength and direction of associations formed in the learning of nonsense syllables. J. exp. Psychol., 1937, 20, 381-395.
- Restorff, H. von. Über die Wirkung von Bereichsbildungen im Spurenfeld. Psychol. Forsch., 1933, 18, 299-342.
- Reynolds, B. The acquisition of a trace conditioned response as a function of the magnitude of the stimulus trace. J. exp. Psychol., 1945, 35, 15-30.
- Riesen, A. H. Delayed reward in discrimination learning by chimpanzees. Comp. Psychol. Monogr., 1940, 15, No.5 (whole No. 77)
- Riesen, A. H. Plasticity of behavior: Psychological aspects. In Harlow, H. F. and Woolsey, C. N.: Biological and biochemical bases of behavior. Madison: Univ. of Wisconsin, 1958, 425-450.
- Shepard, R. N., and Teghtsoonian, M. Retention of information under conditions approaching a steady state. J. exp. Psychol., 1961, 62, 302-309.
- Siegel, S. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill, 1956.
- Skinner, B. F. "Superstition" in the pigeon. J. exp. Psychol., 1948, 38, 168-172.
- Slamecka, N. J. An inquiry into the doctrine of remote associations. Psychol. Rev., 1964, 71, 61-76.
- Slamecka, N. J. In defense of a new approach to old phenomena. Psychol. Rev., 1965, 72, 242-246.

- Smith, M. P. The stimulus trace gradient in visual discrimination learning. J. comp. physiol. Psychol., 1951, 44, 154-161.
- Thomson, Barbara P., and Wing, J. F. Serial-position and isolation effects in free recall of heterogeneous lists. Paper read at Midwestern Psychological Association meetings, Chicago, Illinois, May, 1965.
- Tresselt, M. E. Time errors in successive comparison of simple visual objects. Amer. J. Psychol., 1944, 57, 555-558.
- Valenstein, E. S. Problems of measurement and interpretation with reinforcing brain stimulation. Psychol. Rev., 1964, 71, 415-437.
- White, C. T., and Schlosberg, H. Degree of conditioning of the GSR as a function of the period of delay. J. exp. Psychol., 1952, 43, 357-362.
- Wickelgren, W. A. Acoustic similarity and intrusion errors in short-term memory. J. exp. Psychol., 1965, 70, 102-108.
- Wickelgren, W. A. Short-term memory for phonemically similar lists. Amer. J. Psychol., 1965, 78, 567-574.
- Wickens, D. D., Born, D. G., and Allen, C. K. Proactive inhibition and item similarity in short-term memory. J. verb. Learn. verb. Behav., 1963, 62, 440-445.
- Wing, J. F. Prediction of isolation effects in immediate recall. Paper read at American Psychological Association meetings, St. Louis, Missouri, September, 1962.
- Wing, J. F. A memory trace derivation of remote associations. Chapter XXI in Merrifield, P. R. (Ed.) Experimental and factor analytic measurement of personality. Kent, Ohio: Kent Univ. (in press).
- Wing, J. F., and Painter, Barbara A. Intraserial interference vs. structural change explanations of isolation effects in immediate recall. Paper read at the American Psychological Association meetings, Los Angeles, California, September, 1964.
- Wing, J. F., and Thomson, Barbara P. Primacy-recency effects in free recall. Proceed. 73rd Ann. Conv. Amer. Psychol. Assoc., 1965, 57-58.
- Wittig, A. F. Immediate free recall of serially presented items. Ph.D. Dissertation, Ohio State University, 1964.

Woodworth, R. S. Experimental psychology. New York: Holt and Company, 1938.

Yntema, D. B., Wozencraft, F. T., and Klem, Laura. Immediate serial recall of digits presented at very high rates. Paper read at Psychonomic Society meetings, Niagara Falls, Ontario, 1964.

AUTOBIOGRAPHY

I, John Fiske Wing, was born in Detroit, Michigan, on September 10, 1931. I attended public schools in South Pasadena, California, and graduated from South Pasadena-San Marino Senior High School in 1949. I attended John Muir (Junior) College from 1949 to 1951 where I received my Associate of Arts degree in June, 1951. From 1951 to 1954 I attended Antioch College. I received my A. B. degree from there in June, 1954. I enrolled in the graduate school of the University of Pennsylvania in 1957 and attended for one year. In 1959 I transferred my University of Pennsylvania credits to Antioch College and enrolled in a Master of Arts program. I received my M. A. degree in June, 1961. During this period at Antioch College I also served as Instructor in Psychology and Assistant Director of the Behavior Research Laboratory. Since that date I have been enrolled in the graduate program of the Ohio State University.