Title

NO. 107 Some Two-Process Models for Memory

Authors

Atkinson, Richard C.

Shiffrin, R. M.

Publication Date

1966-09-30

Peer reviewed

SOME TWO-PROCESS MODELS FOR MEMORY

by

R. C. Atkinson and R. M. Shiffrin

TECHNICAL REPORT NO. 107

September 30, 1966

PSYCHOLOGY SERIES

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

Abstract

A general theoretical framework is developed in which to view memory and learning. The basic model is presented in terms of a memory system having two central components: a transient-memory buffer and a long-term store. Each stimulus item is postulated to enter a constant-sized push-down memory buffer, stay a variable amount of time and leave on a probabilistic basis when displaced by succeeding inputs. During the period that each item resides in the buffer, copies of the item are placed in the long-term store. The remaining feature of the model is concerned with the recovery of items from the memory system at the time of test. If at this time an item is still present in the buffer, it is perfectly retrieved. If an item is not present in the buffer, a search of the long-term store is made. This search is imperfect and the greater the number of items in the long-term store, the smaller the probability that any particular one will be retrieved. The model is applied to a set of experiments on paired-associate memory with good success.

Some Two-Process Models for Memory[1]

R. C. Atkinson and R. M. Shiffrin

Stanford University

A model for memory will be outlined in this paper. The experimental framework for which the model was constructed is that in which a series of items is presented to the subject who is then required to recall one or more of them. A familiar example is the digit span test in which the subject is required to repeat a series of digits read to him. A typical finding in digit span studies is that performance is error free until a critically large number of digits is reached. Thus a short-term memory system, called the "buffer," is proposed which may hold a fixed number of digits and allows perfect retrieval of those digits currently held. Errors are made only when the number of digits presented exceeds the capacity of the buffer, at which time the previous digits are forced out of the buffer. We propose, in addition, a long-term memory system (abbreviated LTS for long-term store) which allows items not present in the buffer to be recalled with some probability between 0 and 1. This two-process model will be presented in the first part of the paper and then applied to data from an experiment in paired-associate memory in the second part of the paper.

Insert Figure 1 about here

Figure 1 shows the overall conception. An incoming stimulus item first enters the sensory buffer where it will reside for only a brief period of time and then is transferred to the memory buffer. The sensory buffer characterizes the initial input of the stimulus item into the

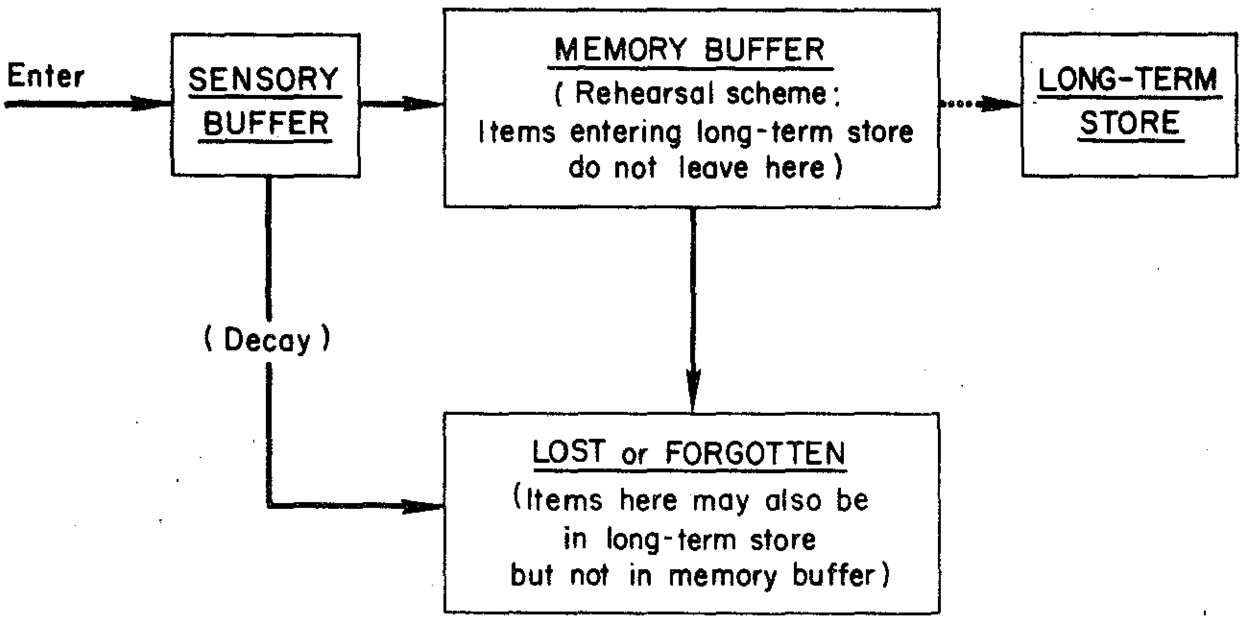

Fig. 1. Flow chart for the general system.

nervous system, and the amount of information transmitted from the sensory buffer to the memory buffer is assumed to be a function of the exposure time of the stimulus and related variables. Much work has been done on the encoding of short-duration stimuli (e.g., Estes and Taylor, 1964; Mackworth, 1963; Sperling, 1960), but the experiments considered in this paper are concerned with stimulus exposures of fairly long duration (one second or more). Hence we will assume that all items pass usccessfully through the sensory buffer and into the memory buffer; that is, all items are assumed to be attended to and entered correctly into the memory buffer. Throughout this paper, then, it will be understood that the term buffer refers to the memory buffer and not the sensory buffer. Furthermore, we will not become involved here in an analysis of what is meant by an "item." If the word "cat" is presented visually, we will simply assume that whatever is stored in the memory buffer (be it the visual image of the word, the auditory sound, or some vector of information about cats) is sufficient to permit the subject to report back the word "cat" if we immediately ask for it. This question will be returned to later. Referring back to Fig. 1, we see that a dotted line runs from the buffer to the "long-term store" and a solid line from the buffer to the "lost or forgotten" state. This is to emphasize that items are copied into LTS without affecting in any way their status in the buffer. Thus items can be simultaneously in the buffer and in LTS. The solid line indicates that eventually the item will leave the buffer and be lost. The lost state is used here in a very special way: as soon as an item leaves the buffer it is said to be lost, regardless of whether it is in LTS or not. The buffer, it should be noted, is a close correlate of what others have called a "short-term store"

(Bower, 1964; Broadbent, 1963; Brown, 1964; Peterson, 1963) and "primary memory" (Waugh and Norman, 1956). We prefer the term buffer because of the wide range of applications for which the term short-term store has been used.

Insert Figure 2 about here

Figure 2 illustrates the workings of the memory buffer. The properties of the buffer will be examined successively.

1. Constant size. The buffer can contain exactly r items and no more. This statement holds within any experimental situation. The buffer size will change when the type of items change. For example, if the items are single digits, the buffer size might be five, but if the items are five-digit numbers the buffer size would correspondingly be one. We should like eventually to be able to permanently fix the buffer size on a more molecular basis than "items": for example, on some such basis as the amount of information transmitted, or the length of the auditory code for the items. This is still an open question and at present the buffer size must be estimated separately for each experiment.

A second important point concerns what we mean by an item. In the experiments that the model is designed to handle there is a clearly separated series of inputs and a clearly defined response. In these cases, the "item" that is placed in the buffer may be considered to be an amount of information which is sufficient to allow emission of the correct response.

2. Push-down buffer: temporal ordering. These two properties are equivalent. As it is shown in the diagram the spaces in the buffer (henceforth referred to as "slots") are numbered in such a way that when an item first enters the buffer it occupies the r^{th} slot. When the next item is

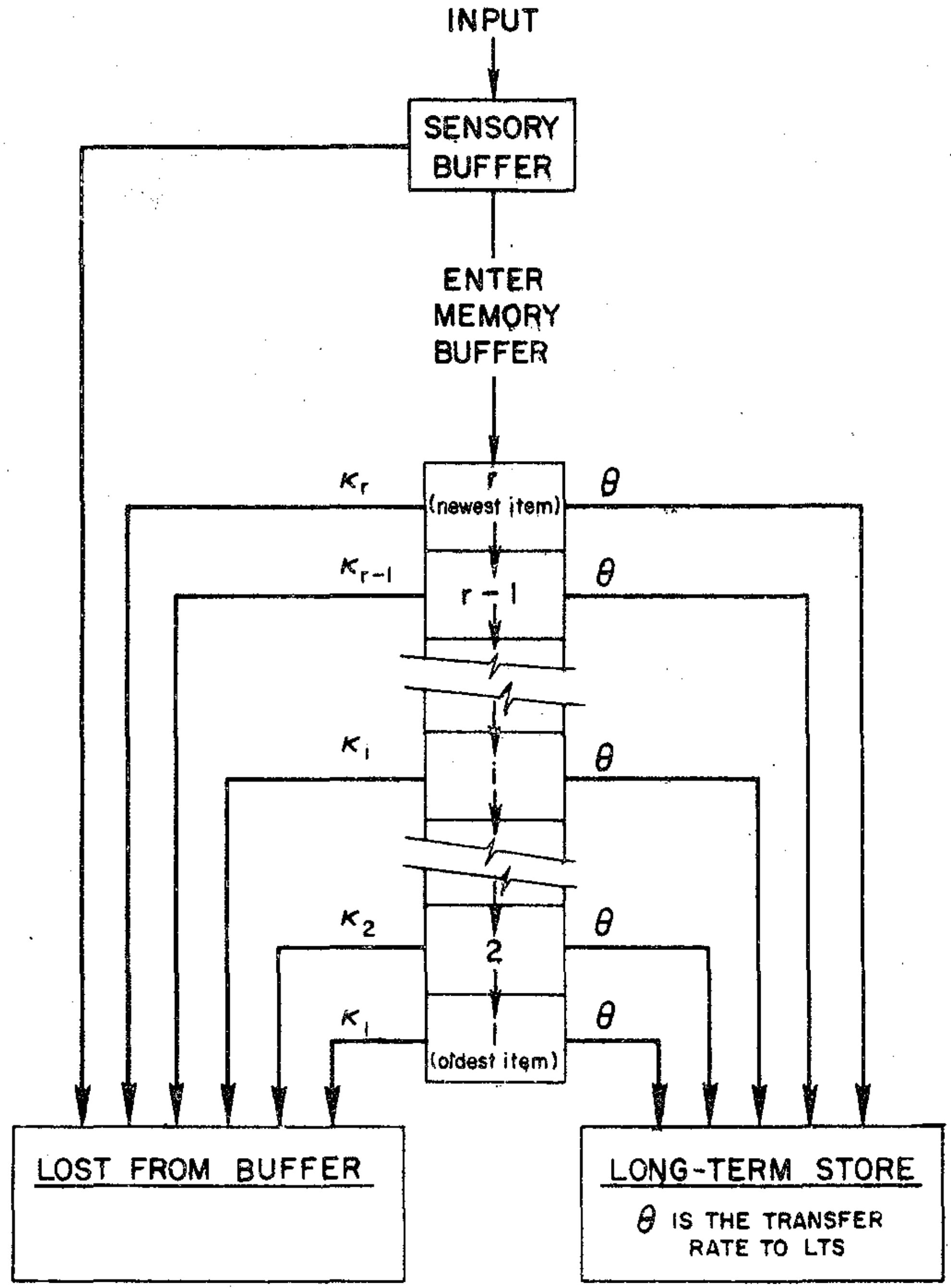

Fig. 2. A flow chart characterizing inputs to the memory system.

presented it enters the r^{th} slot and pushes the preceding item down to the $r-1^{st}$ slot. The process continues in this manner until the buffer is filled; after this occurs each new item pushes an old one out on a basis to be described shortly. The one that is pushed out is lost. Items stored in slots above the one that is lost move down one slot each and the incoming item is placed in the r^{th} slot. Hence items in the buffer at any point in time are temporally ordered: the oldest is in slot number 1 and the newest in slot r. It should be noted that the lost state refers only to the fact that an item has left the buffer and says nothing regarding the item's presence in LTS.

3. Buffer stays filled. Once the first r items have arrived the buffer is filled. Each item arriving after that knocks out exactly one item already in the buffer; thus the buffer is always filled thereafter. This state of affairs is assumed to hold as long as the subject is paying attention. In this matter we tend to follow Broadbent (1963) and view the buffer as the input-output mechanism for information transmission between the subject and the environment. At the end of a trial, for example, attention ceases, the subject "thinks" of other things, and the buffer gradually empties of that trial's items.

4. Each new item bumps out an old item. This occurs only when the buffer has been filled. The item to be bumped out is selected as a function of the buffer position (which is directly related to the length of time each item has spent in the buffer). Let

κ_j = probability that an item in slot j of a full buffer is lost when a new item arrives.

Then of course $\kappa_1 + \kappa_2 + \cdots + \kappa_r = 1$, since exactly one item is lost. Various schemes can be proposed for the generation of the κ's. The simplest scheme, requiring no additional parameters, is to equalize the κ's: i.e., let $\kappa_j = 1/r$ for all j.

A useful one-parameter scheme can be derived as follows: the oldest item (in slot 1) is dropped with probability δ. If that item is not dropped, then the item in position 2 is dropped with probability δ. If the process reaches the r^{th} slot and it also is passed over, then the process recycles to the first slot. This process continues until an item is dropped. Hence

$$\kappa_j = \frac{\delta(1-\delta)^{j-1}}{1-(1-\delta)^r} \quad .$$

It is easy to see that as δ approaches 0, κ_j approaches $1/r$ for all j, which was the earlier case mentioned. On the other hand, when $\delta = 1$, the oldest item is always the one lost. Intermediate values of δ allow a bump-out process between these two extremes. We would expect that the tendency to bump out the oldest item first would depend on such factors as the serial nature of the task, the subject's instructions, and the subject's knowledge concerning the length of the list he is to remember.

5. Perfect representation of items in the buffer. Items are always encoded correctly when initially placed in the buffer. This, of course, only holds true for experiments with fairly slow inputs, such as the experiment to be considered later in this paper.

6. Perfect recovery of items from the buffer. Items still in the buffer at the time of test are recalled perfectly (subject to the "perfect representation" assumption made above). This point leads to the question,

"What is stored in the buffer?" and "What is an item?" In terms of the preceding requirement (and in accord with the mathematical structure of the model) we may be satisfied with the definition, "an item is that amount of information that allows correct performance at the time of test." Because the model does not require a more precise statement than the above, it is not necessary in the present analysis to spell out the matter in detail. Nevertheless, in view of the work of Conrad (1964), Wickelgren (1965), and others on auditory confusions in short-term memory, we would be satisfied with the view that items in the buffer are acoustic mnemonics and are kept there via rehearsal, at least for experiments of a verbal character.

7. Buffer is unchanged by the transfer process to LTS. We will say more about LTS and transfer to it in the next section, but here it may be said that whatever transfer takes place, and whenever the transfer takes place, the buffer remains unchanged. That is, if a copy of an item is placed in LTS, the item remains represented in the buffer, and the buffer remains unchanged.

This set of seven assumptions characterizes the memory buffer. Now we consider the long-term memory system. In recent years a number of mathematical models for memory and learning have made use of a state labeled "long-term store." In most of these cases, however, the term is used to denote a completely learned state. LTS in this case is used in a very different manner; information concerning each item is postulated to enter LTS during the period the item remains in the buffer. This information may or may not be sufficient to allow recall of the item, and even if sufficient information to allow recall is stored, the subject may fail to recall because he still must search LTS for the appropriate information.

There are many possible representations of the transfer process to LTS. Let θ_{ij} be the transfer parameter representing the amount transferred to LTS of an item in slot i of the buffer between one item presentation and the next if there are currently j items in the buffer. In the present version θ_{ij} is the probability of copying an item into LTS during each presentation period.

For this discussion we will assume that θ does not depend on the position in the buffer, but does depend on the number of other items currently in the buffer. The justification for this is based on the amount of attention that an item will receive during each presentation period; thus an item will receive r times as much attention if it is the only item in the buffer than if all r buffer positions were filled. Hence θ_{ij} is set equal to θ/j. It is further assumed that there may be more than one copy of any item in LTS. Since one copy may be made during each presentation period, the maximum number of copies that can exist in LTS for a particular item equals the number of presentation periods that the item stayed in the buffer.

The retrieval rules are relatively simple. At the time of test any item in the buffer is recalled perfectly. If the item is not present in the buffer then a search of LTS is made. If the item is found in LTS it is recalled; if not, then the subject guesses. The search process the subject engages in is postulated to be a search made uniformly with replacement from the pool of items in LTS which are not in the buffer. (An alternative scheme is to pick from all the items in LTS, which gives very similar results to those given by the stated scheme.) In particular, the subject is said to make R random picks in LTS; if none of these picks finds the desired item, it is reported; otherwise the subject guesses.

The mathematical development of this model is presented in Atkinson and Shiffrin (1965). For present purposes, it is sufficient to note that there are four parameters available to fit the data: r, the buffer size; θ, the transfer probability; δ, the tendency to bump out the oldest item in the buffer first; and R, the number of searches into LTS.

We now turn to an experiment in human paired-associate memory (Phillips, Shiffrin, and Atkinson, 1967). The experiment involved a long series of discrete trials. On each trial a display of items was presented. A display consisted of a series of cards each containing a small colored patch on one side. Four colors were used: black, white, blue, and green. The cards were presented to the subject at a rate of one card every two seconds. The subject named the color of each card as it was presented. Once the color of the card had been named by the subject it was placed face down on a display board so that the color was no longer visible, and the next card was presented. After presentation of the last card in a display the cards were in a straight row on the display board: the card presented first was to the subject's left and the most recently presented card to the right. The trial terminated when the experimenter pointed to one of the cards on the display board, and the subject attempted to recall the color of that card. The subject was instructed to guess the color if uncertain.

Following the subject's response, the experimenter informed the subject of the correct answer. The display size (list length) will be denoted as d. The values of d used in the experiment were 3, 4, 5, 6, 7, 8, 11, and 14. Each display, regardless of size, ended at the same place on the display board, so that the subject knew at the start of each display how long that particular display would be. Twenty subjects, all females, were run for a total of five sessions, approximately 70 trials per session.

Insert Figure 3 about here

Figure 3 presents the proportion of correct responses as a function of the test position in the display. Display sizes 3 and 4 are not graphed because performance was essentially perfect for these cases. Observed points for $d = 8$, 11, and 14 are based on 120 observations, whereas all other points are based on 100 observations. Serial position 1 designates a test on the most recently presented item. These data indicate that for a fixed display size, the probability of a correct response decreases to some minimum value and then increases. Thus there is a very powerful recency effect as well as a strong primacy effect over a wide range of display sizes. Note also that the recency part of each curve is S-shaped and could not be well described by an exponential function. Reference to Fig. 3 also indicates that the overall proportion correct is a decreasing function of display size.

The model was fit to the data using a minimum chi-square technique.[2] The details are presented in Atkinson and Shiffrin (1965). It will merely be pointed out here that the value of r was set equal to 5 before the minimization because performance was essentially error free for list lengths of 5 and less. The other three parameters were fit using a grid search

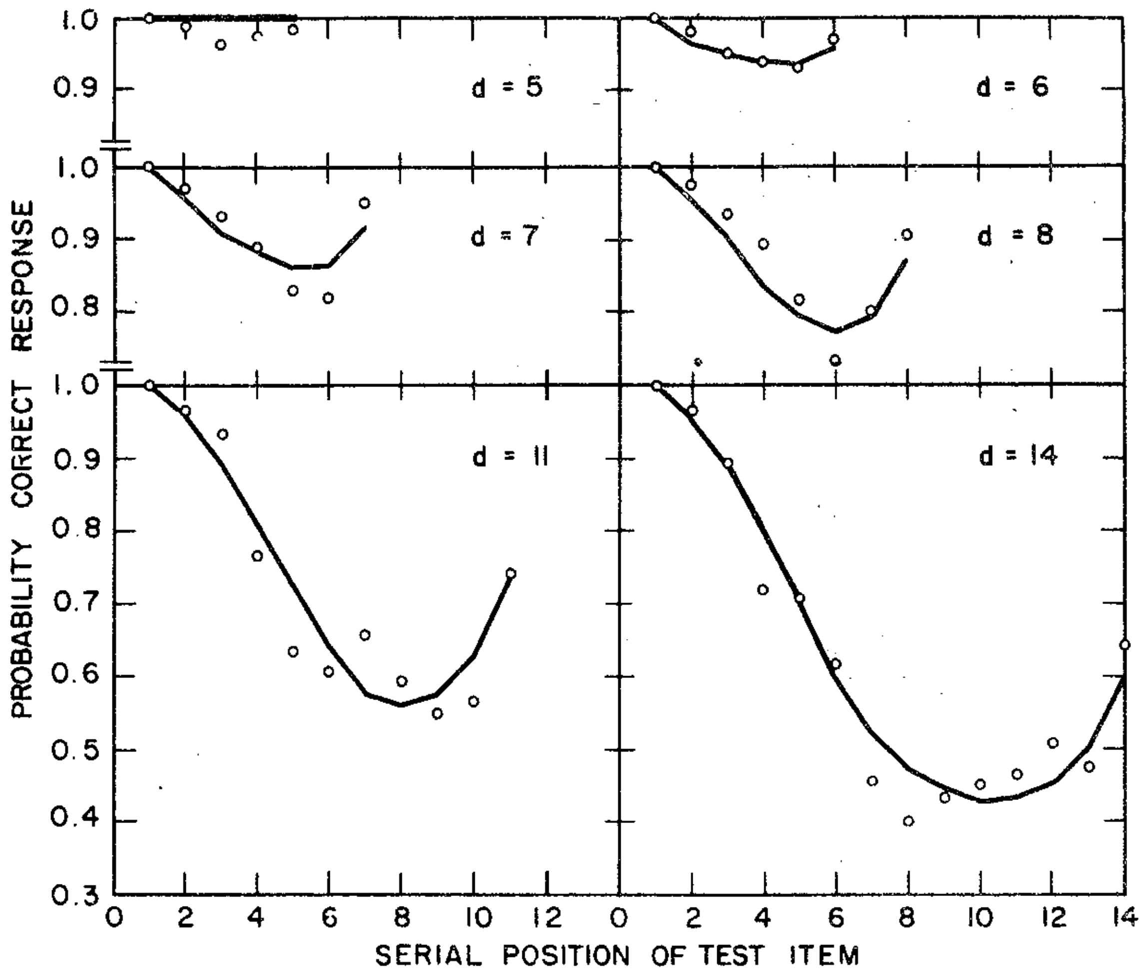

Fig. 3. Goodness-of-fit results for the paired-associate memory experiment.

procedure on a computer. The parameter estimates were as follows:

$$\hat{\delta} = .39$$

$$\hat{\theta} = .72$$

$$\hat{R} = 3.15 \ .$$

The predicted curves are given in Fig. 3. It should be emphasized that the same 4 parameters are used to fit the serial position curves for all five list lengths. It can be seen that the fit is quite good with a minimum chi-square of 46.2 based on 43 degrees of freedom.

We have outlined only one example of how this model can be applied to data. Other applications of the model have been made including experiments involving a continuous-presentation memory task, free-verbal recall, paired-associate learning, and serial-anticipatory learning; also, the model has been used to perdict not only response probabilities, but confidence ratings and latency data. Time does not permit us to present these developments here; for a review of such applications see Atkinson and Shiffrin (1965), Atkinson, Brelsford, and Shiffrin (1967), Brelsford and Atkinson (1967), and Phillips, Shiffrin, and Atkinson (1967). In conclusion, it should be pointed out that of all the assumptions introduced, three are crucial to the theory. First is the set of buffer assumptions; i.e., constant size, push-down list, and so on. Second is the assumption that items can be in the buffer and LTS simultaneously. Third is what was called the retrieval process--the hypothesis that the decrement in recall caused by increasing the list length occurs as the result of an imperfect search of LTS at the time of test. Within this framework, we feel that a number of the results in memory and learning can be described in quantitative detail.

References

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

Atkinson, R. C., Brelsford, J. W., Jr., and Shiffrin, R. M. Multi-process models for memory with applications to a continuous presentation task. Journal of Mathematical Psychology, 1967, 4, in press.

Atkinson, R. C., and Shiffrin, R. M. Mathematical models for memory and learning. Technical Report 79, Institute for Mathematical Studies in the Social Sciences, Stanford University, 1965. (To be published in D. P. Kimble (Ed.), Proceedings of the Third Conference on Learning, Remembering, and Forgetting. New York: The New York Academy of Sciences, 1966.)

Bower, G. H. Notes on a descriptive theory of memory. In D. P. Kimble (Ed.), Proceedings of the Second Conference on Learning, Remembering, and Forgetting. New York: The New York Academy of Sciences, 1966, in press.

Brelsford, J. W., Jr., and Atkinson, R. C. Recall of paired-associates as a function of overt and covert rehearsal procedures. Technical Report 111, Institute for Mathematical Studies in the Social Sciences, Stanford University, Stanford, California, 1966.

Broadbent, D. E. Flow of information within the organism. Journal of Verbal Learning and Verbal Behavior, 1963, 4, 34-39.

Brown, J. Short-term memory. British Medical Bulletin, 1964, 20(1), 8-11.

Conrad, R. Acoustic confusions in immediate memory. British Journal of Psychology, 1964, 55, 75-84.

Estes, W. K., and Taylor, H. A. A detection method and probabilistic models for assessing information processing from brief visual displays. *Proceedings of the National Academy of Sciences*, 1964, 52, 2, 446-454.

Mackworth, J. F. The relation between the visual image and post-perceptual immediate memory. *Journal of Verbal Learning and Verbal Behavior*, 1963, 2, 113-119.

Peterson, L. R. Immediate memory: data and theory. In C. N. Cofer (Ed.), *Verbal learning and behavior: Problems and processes*. New York: McGraw-Hill, 1963.

Phillips, J. L., Shiffrin, R. M., and Atkinson, R. C. The effects of display size on short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1967, in press.

Sperling, G. The information available in brief visual presentations. *Psychological Monographs*, 1960, 74, Whole No. 498.

Wickelgren, W. A. Acoustic similarity and retroactive interference in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1965, 4, 53-61.

Waugh, Nancy C., and Norman, D. A. Primary memory. *Psychological Review*, 1965, 72, 89-104.

Footnotes

[1] Support for this research was provided by the National Aeronautics and Space Administration, Grant NGR-05-020-036, and United States Public Health Service MH06154. This paper was read at the Symposium on Mathematical Models of Psychological Processes, XVIII International Congress of Psychology, Moscow, USSR, August, 1966.

[2] For a discussion of minimum chi-square procedures for parameter estimation, see Atkinson, Bower, and Crothers (1965).