Title
The Effects of Display Size on Short-Term Memory

Authors
Phillips, James L.
Atkinson, Richard C.

Publication Date
1965-08-31

Peer reviewed

THE EFFECTS OF DISPLAY SIZE ON SHORT-TERM MEMORY

by

James L. Phillips and Richard C. Atkinson

TECHNICAL REPORT NO. 78

August 31, 1965

PSYCHOLOGY SERIES

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

The Effects of Display Size on Short-Term Memory[1]

by

James L. Phillips[2] and Richard C. Atkinson

Stanford University

Abstract

This paper presents the results of an empirical study of short-term memory. Ss were presented with sequential displays which varied in size (i.e., number of stimulus items) from 3 to 14 items. After the display had been presented Ss were asked to recall one of the items of the display. Confidence ratings were then obtained for the response made by S.

Serial position curves for each of the various display sizes are presented. These curves show the classical recency and primacy effects, with the curves taking on a pronounced S-shape over the most recent serial positions. Items in the middle part of the display and items presented at the beginning of the display are most affected by display size. In general, the effect of increasing the display size is to decrease the proportion of correct responses for these positions.

A register model is presented which describes memory in terms of an information processing schema. The model is used to generate theoretical serial position curves which fit the observed data quite adequately. Estimates of the parameters of the model are given interpretations that appear to be reasonable, considering the role these parameters play in the model.

---

[1]Support for this research was provided by the National Aeronautics and Space Administration.

[2]Now at Michigan State University.

The Effects of Display Size on Short-Term Memory

Workers in the area of learning theory have become increasingly interested in the problem of short-term memory as an important and natural extension of the study of verbal behavior. There have been two somewhat different theoretical approaches to this problem. Some investigators (Atkinson and Crothers, 1964; Bernbach, 1965; Calfee and Atkinson, 1965; Greeno, 1965) have developed and tested a number of models for paired-associate learning that postulate a partition of memory into a short-term component and a long-term component. By and large, these models have been extremely accurate in their ability to predict and describe the acquisition of paired-associate responses. A second approach involves the direct investigation of short-term retention and the formulation of models for this specific task (e.g., Bower, 1964; Broadbent, 1963; Peterson, 1963; and Waugh and Norman, 1965). In this regard, the work of Broadbent has been particularly significant. He has postulated that retention may be understood in terms of the flow of information through a limited capacity channel into a short-term store. This information processing conception of short-term memory, examined in conjunction with the various models for paired-associate learning, suggested a very simple empirical study which is reported in this paper.

The present experiment was designed to obtain short-term memory data under conditions where the amount of information which S was required to process before retrieval could be systematically varied over a wide range. It was also desired to collect a large amount of data on individual Ss. Other requirements of the study were that the probability of being correct

by guessing be clearly determined, and that response interference be minimized. It was decided that the response interference problem could be solved by using responses with which S would be quite familiar, and by requiring only one response for each separate display of information.

The following procedure was considered to meet the above requirements. On each trial of the experiment a display of d items was presented sequentially to S. A display consisted of a random sequence of playing cards. The cards varied only in the color of a small patch on one side; four different colors were used. Following the presentation of the display, S was asked to recall the color of one of the cards. The S then gave a confidence rating, and the trial terminated with the experimenter informing S of the correct answer. The experiment involved a long series of such trials, and over trials the length of the display and the test position were systematically varied. This procedure is similar to that reported by Atkinson, Hansen, and Bernbach (1964).

Method

The Ss in this study were 20 females. They were drawn from a pool of Stanford University students who had expressed an interest in participating in psychological experiments, and were paid for their services. Each S participated in five sessions, each session lasting approximately one hour. The first session was a practice session, designed to familiarize S with the procedure and to eliminate practice effects. Three display sizes ($d = 8, 11, 14$) were used in session 1; the next three sessions also were restricted to these three display sizes. The last session for each S employed five different display sizes ($d = 3, 4, 5, 6, 7$).

The experiment involved a long series of discrete trials. On each trial a display of d items was presented. A display consisted of a series of 2 × 3 1/2 in. cards containing a 3/4 × 1 1/2 in. colored patch in the center. Four colors were used: black, white, blue, and green. The cards were presented to the S at a rate of one card every two seconds. The S named the color of each card as it was presented. A metronome was used to maintain a constant rate of presentation for each display. Once the color of the card had been named by S it was turned 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 Ss left and the most recently presented card to her right. The trial terminated when the experimenter pointed to one of the cards on the display board, and S attempted to recall the color of that card. The S was instructed to guess the color if uncertain and to qualify her response with a confidence rating. The confidence ratings were the numerals 1, 2, 3, and 4. The Ss were told to say 1 if they were positive; 2 if they were able to eliminate two of the four possible colors as correct; 3 if one of the four colors could be eliminated as correct; and 4 if she had no idea at all as to the correct response. These confidence ratings will be designated R_1, R_2, R_3, and R_4. Each display, regardless of size, ended at the same place on the display board. That is, displays began at different places on the display board and hence Ss knew, from the position of the first card, how long each display was to be.

Each S was given two complete blocks of displays in each of the first four sessions. A block consisted of one display for each serial position in

each display size. Thus there were (8 + 11 + 14) = 33 displays per block, and a complete session involved the presentation of 66 displays. During the fifth day each S was given five complete blocks of displays. A complete block in the final session consisted of (3 + 4 + 5 + 6 + 7) = 25 displays, and hence the total session involved the presentation of 125 displays. Each serial position of each display size was selected as the test position exactly once per block. The order of presentation of displays (display size and test position) was randomized within each block; further, the order of the cards was independently randomized for each display.

At the beginning of the second session Ss were told the proportion of correct responses that they had achieved for each of the four confidence ratings. They were reminded at that time that the "ideal" proportion correct was 100% for a confidence rating of R_1, 50% for R_2, 33% for R_3, and 25% for R_4. No further information feedback was given concerning the confidence ratings during subsequent sessions.

Results

The overall proportion of correct responses for the first four sessions is presented in Fig. 1. Each of the points in Fig. 1 is based on 7920 observations. Most of the improvement in performance occurred during the training session and the curve is reasonably stationary for session 2 through 4. Since the purpose of the first session was to eliminate practice effects, only the data from sessions 2, 3, 4, and 5 are presented in subsequent analyses. All display sizes used in session 5 were shorter than those in the first four sessions, so the overall proportion correctly recalled in session 5 cannot be meaningfully compared with those of the first four sessions, and hence was not plotted in Fig. 1.

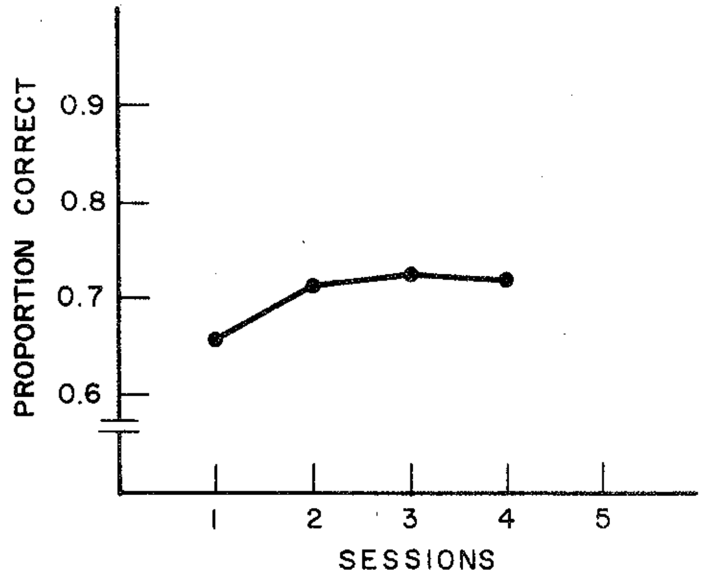

Figure 1. Proportion of items correctly recalled for each of the first four experimental sessions.

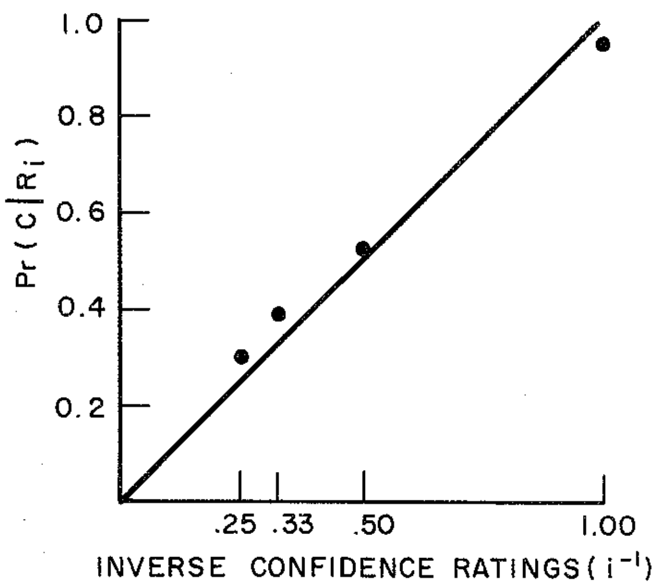

Figure 2. Proportion of items correctly recalled given confidence rating R_i plotted against the reciprocal of the confidence rating.

An interesting aspect of this study was S's use of the confidence ratings. The instructions asked S to match the confidence ratings to her estimate of the number of possible correct response alternatives. That is, if S felt there was only one possible correct alternative (i.e., she was absolutely certain) she was to use R_1. If she felt there were two possible correct alternatives and was, in effect, guessing from this set of two, then her appropriate confidence rating was R_2. Likewise for R_3 and R_4. If Ss were able to follow these instructions, then the observed proportion of items correctly recalled conditionalized upon the confidence rating should be equal to the reciprocal of the confidence rating. That is, over all items to which S responded R_1, the observed proportion correct should be 1.00. For responses rated as R_2, the observed proportion correct should be 0.50. For items with R_3 and R_4, the proportions correct should be 0.33 and 0.25 respectively. Figure 2 presents the relationship between the proportion correct given R_i [denoted $Pr(C|R_i)$], and the reciprocal of the confidence rating. It is clear from these data that Ss were able to use the confidence rating reasonably accurately, in the sense implied by the instructions.

Figure 3 presents the proportion of correct responses as a function of the test position in the display. There is a separate curve for each of the display sizes used in the study. Points on the curves 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 at an increasing rate to some minimum value and then increases. Thus, there is a very powerful recency

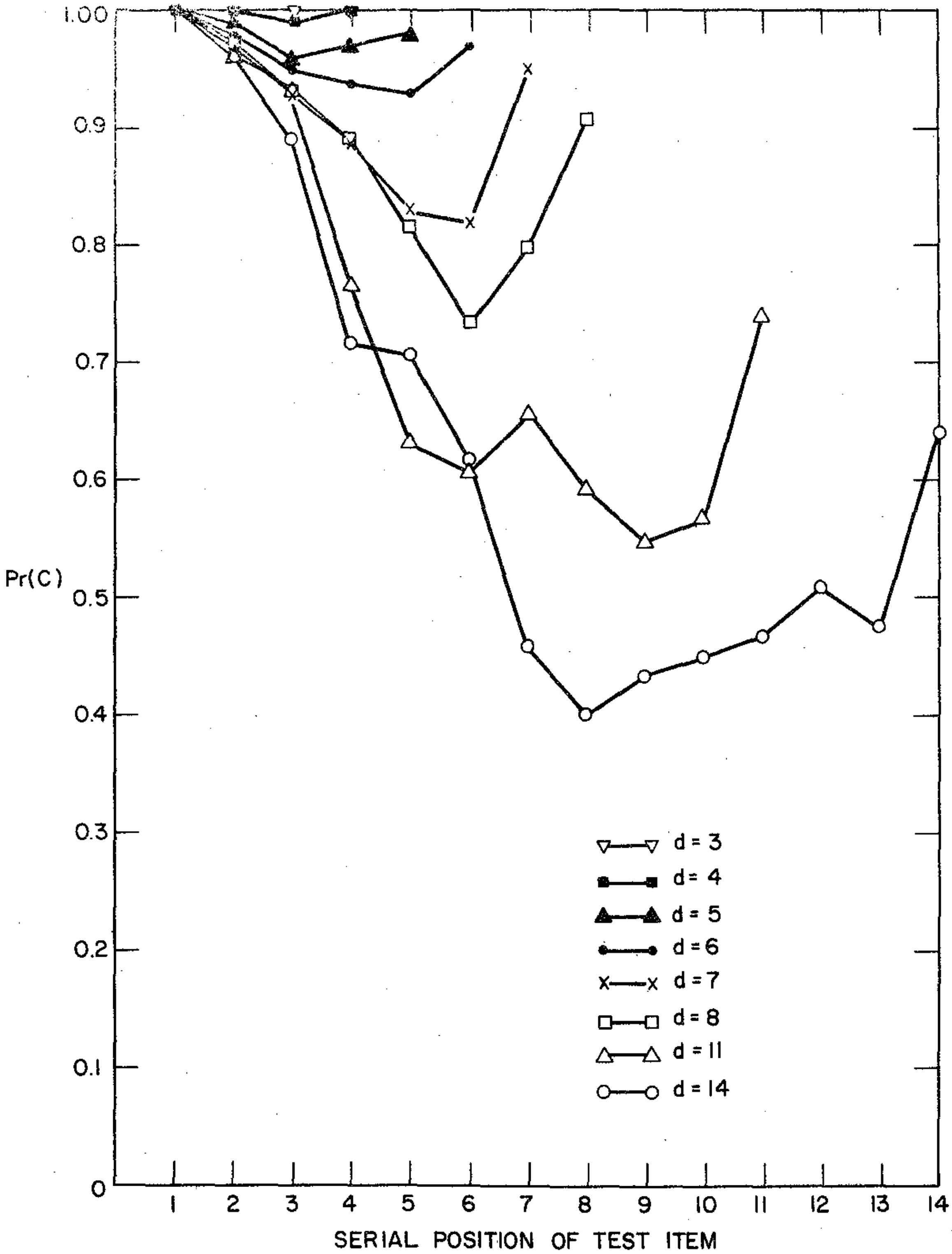

Figure 3. Proportion of items correctly recalled at each serial position for the various display sizes.

effect as well as a strong primacy effect over a wide range of display sizes. Reference to Fig. 3 also indicates that the overall proportion correct is a decreasing function of display size.

Discussion

There are a number of models that can be tested against the data reported in this paper. Such a comparison of models, while useful, is beyond the scope of this paper. Instead, one model will be presented as an aid in the interpretation of the data; the model is somewhat like those proposed by Bower (1964) and Broadbent (1963).

If the organism is considered to be an information processing system, then the memory of that system may be conceptualized as a number of distinct components. First of all, it is postulated that there is a long-term memory which is relatively permanent and which has a practically unlimited capacity. Secondly, a short-term memory component or register is postulated. The capacity of the register is considered to be limited; more specifically, there are exactly r positions in the register. One item of information can be stored in each position.

When the first item of a display (item number d) is presented to S it is placed in the r^{th} position of the register. With some probability θ, item d is stored in long-term memory. Item $d-1$ is then presented and it displaces item d in the r^{th} position of the register, pushing item d down to position $r-1$. Either of these items may now be stored in long-term memory. Since storage occurs with probability θ and the items are equally likely to be stored, the effective probability of storing each item is $\theta/2$. This process continues until the register is filled. The probability of an item in the register being stored in long-term memory on

any presentation, is θ/i where i is the number of items currently in the register.

Since the register has a capacity of r items, it is filled on the presentation of the $(d - r + 1)^{st}$ item. That is, items $d, d-1, \ldots, d-r+1$ occupy positions $1, 2, \ldots, r$ of the register. Thus, on the presentation of item $d-r$, one of the first r items must be dropped from the register. If this item has not already been stored in long-term memory then it is forgotten.

Instead of developing the model further at this time, let us examine the assumptions which have been made in light of the data reported here. Since we wish to postulate that an item in the register can be recalled without error, there must be some critical display size (equal to the capacity of the register) on which S makes no errors. Reference to Fig. 3 indicates that there is such a display size. When $d = 3$, Ss in this study exhibited perfect performance. However, Fig. 3 also shows that performance is nearly perfect for $d = 4$ and $d = 5$. In fact, errors occur so infrequently for these display sizes that we are willing to assume they represent some type of attention failure, rather than memory loss *per se*. This being the case the value of $r = 5$ has been chosen as an estimate of register size for this study. The model will be developed subsequently on the assumption of $r = 5$. Although the equations derived are thereby limited in their applicability, the mathematical procedures are straightforward and sufficiently general to be used in the elaboration of the register model for any value of r (for a general statement of the model see Atkinson and Shiffrin, 1965).

It is assumed that the likelihood of an item dropping out of the

register is dependent on its position in the register. Since there is a correspondence between the time an item has been in the register and its position, this is equivalent to the statement: the probability of dropping an item depends on how long that item has been retained in the register. The process which determines the dropping probability for each position of the register will now be described. The oldest item, which is in position 1 of the register, is dropped with probability δ. If that item is <u>not</u> dropped, then the item in position 2 of the register is dropped with probability δ. This process continues until an item is dropped. If the item in the fifth position is passed over, then the process recycles through the register again. This iterative process continues until an item is dropped. As before, the new incoming item pushes the remaining items down and occupies position 5 in the register. If we define δ_i as the probability that an item in position i of the register is dropped when the register is full and the next item is presented, then

$$\delta_1 = [\delta + \delta(1-\delta)^5 + \delta(1-\delta)^{10} + \cdots]$$

$$\delta_2 = [\delta(1-\delta) + \delta(1-\delta)^6 + \delta(1-\delta)^{11} + \cdots]$$

$$\vdots$$

$$\delta_5 = [\delta(1-\delta)^4 + \delta(1-\delta)^9 + \delta(1-\delta)^{14} + \cdots] .$$

Or, more generally

$$\delta_i = \frac{\delta(1-\delta)^{i-1}}{1 - (1-\delta)^5} .$$

In order to facilitate the derivation of expressions for this model, we define an additional quantity, $\Phi(i,j)$. Given that there are j remaining items in the display to be presented, $\Phi(i,j)$ is the probability that an item currently stored in the i^{th} position of the register will be forgotten

(i.e., will be neither in long-term memory nor in the register) at the termination of the display. We note that for the first position of the register ($i = 1$) these expressions are first order difference equations of the form

$$\Phi(1,j) = \delta_1 + (1-\frac{\theta}{5})(1-\delta_1)\Phi(1,j-1) \ .$$

For $i \geq 2$ the expressions are somewhat more formidable:

$$\Phi(2,j) = \delta_2 + (1-\frac{\theta}{5})\Big[\delta_1\Phi(1,j-1) + (\delta_3+\delta_4+\delta_5)\Phi(2,j-1)\Big]$$

$$\Phi(3,j) = \delta_3 + (1-\frac{\theta}{5})\Big[(\delta_1+\delta_2)\Phi(2,j-1) + (\delta_4+\delta_5)\Phi(3,j-1)\Big]$$

$$\Phi(4,j) = \delta_4 + (1-\frac{\theta}{5})\Big[(\delta_1+\delta_3)\Phi(3,j-1) + \delta_5\Phi(4,j-1)\Big]$$

$$\Phi(5,j) = \delta_5 + (1-\frac{\theta}{5})(1-\delta_5)\Phi(4,j-1) \ .$$

The initial condition for each of these expressions is $\Phi(i,0) = 0$.

The probability that the i^{th} item in a display of size d has been forgotten at the end of the display is given by the following expressions:

$$F_i^{(d)} = \begin{cases} \left[\prod_{j=d-i+1}^{5}(1-\frac{\theta}{j})\right]\Phi(d-i+1,d-5) & , \text{ for } i \leq d-4 \\ (1-\frac{\theta}{5})\Phi(5,i-1) & , \text{ for } i > d-4 \ . \end{cases}$$

The probability of a correct response is a function of $F_i^{(d)}$ appropriately corrected for guessing. A fairly natural assumption is that the probability of guessing correctly is just the reciprocal of the number of alternative responses available, which in this case is four. That is, if S fails to remember a particular item, then she responds by guessing one of the four colors used in the study. Thus, the probability of correctly

identifying the i^{th} item in a display of length d is

$$\Pr[C_i^{(d)}] = (\tfrac{1}{4})F_i^{(d)} + [1 - F_i^{(d)}] \ .$$

Since the register size r was assumed to be equal to 5, only the parameters θ and δ need be estimated from the data. The estimates of these parameters were obtained by computing $\Pr[C_i^{(d)}]$ for a grid of values on both θ and δ [θ = (00.00, 0.01, 0.02, ..., 1.00) and δ = (0.00, 0.01, 0.02, ..., 1.00)]. The values of θ and δ that minimized the sum of squared deviations between observed and predicted values of $\Pr[C_i^{(d)}]$ were selected as the parameter estimates. The theoretical curves for the various display sizes, along with the corresponding observed values, are presented in Fig. 4. It is obvious from this figure that the model provides an accurate account of the data of this study.

The estimates of the parameters δ and θ, and the sums of squared deviations are listed in Fig. 4 for each display size. There are a number of comments to make concerning these estimates. To begin with, $\hat{\delta}$ does not exhibit a great deal of variability over the various display sizes. It is clear that a single estimate of δ could have been made over all display sizes without seriously disturbing the correspondence between the theoretical curves and the data.[3] Furthermore, within the framework of the theory presented here, there is no reason to expect that δ should vary in any systematic fashion with display size. It might, of course, be

---

[3]Carrying out a minimization where a single δ is estimated (and separate θ's for each display) yields fits about as good as those presented in Fig. 4.

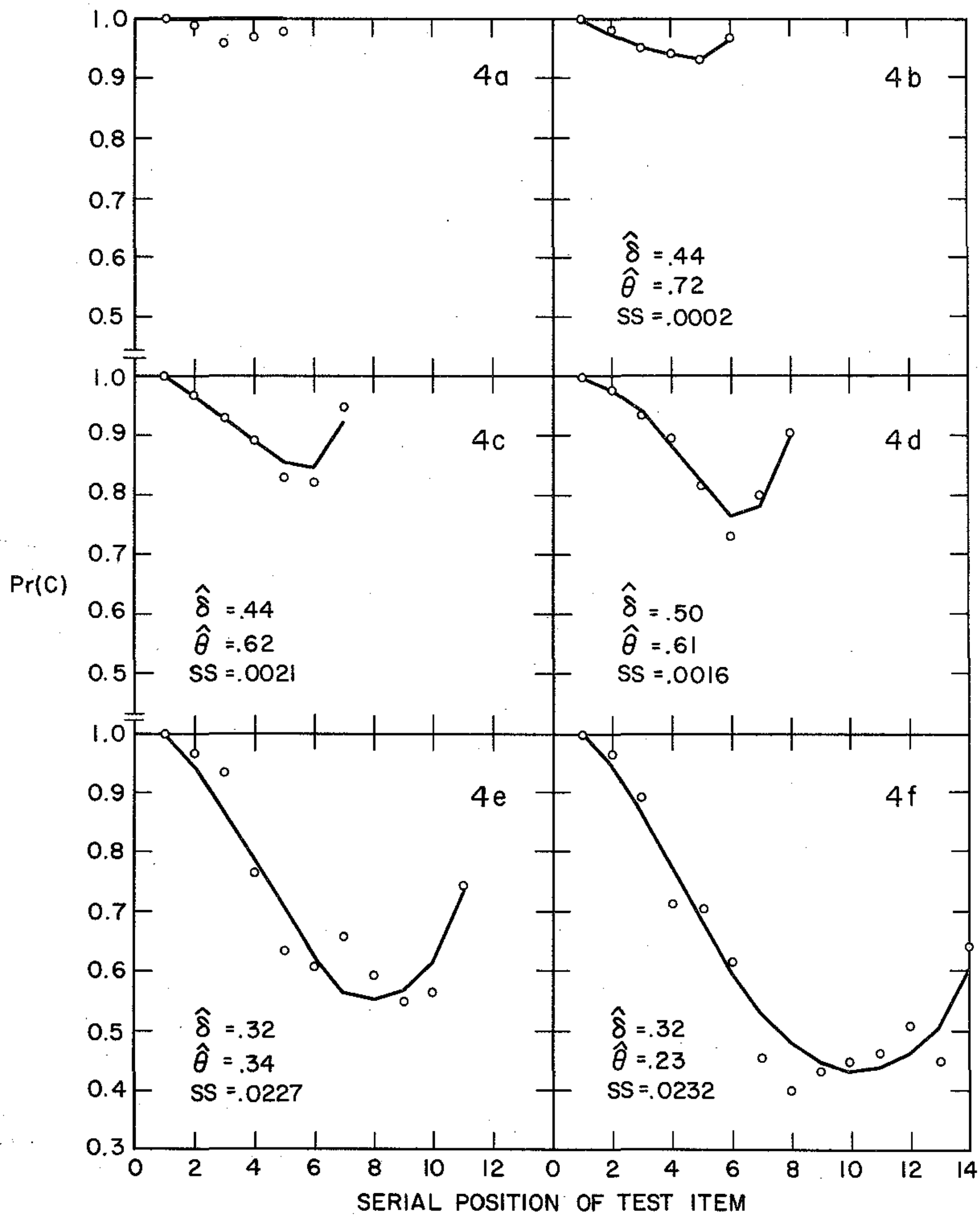

Figure 4. A comparison of theoretical and observed serial position curves. (Observed values are indicated by open circles, and theoretical values by connected line segments.)

affected by the nature of the experimental material. One outstanding feature of the data of this study was the S-shaped curve over the most recent positions of the display. Other experiments using different experimental materials (c.f. Atkinson, Hansen and Bernbach, 1964) have observed a more exponentially shaped curve for the more recent items. An interesting feature of the register model is that it is capable of generating either the exponential or the S-shaped curve. As δ approaches zero, the curve over the recent positions of the display changes from S-shaped to exponential.

As indicated in Fig. 4, the various estimates of the parameter θ are clearly related to display size. As d goes from 6 to 14 the estimates of θ monotonically decrease from .72 to .23. It should be noted in this regard that S cannot attain perfect retention for display size in excess of r even if θ is adjusted to its maximum. This is the case because the effective probability of storing an item in long-term memory when the register is filled is θ/r. Also, with a long list S would undoubtedly perform near a chance level on the middle items and hence it would be necessary to postulate that θ/r was quite small. In view of these observations it seems clear that the model needs to be generalized to specify a function relating display size to θ. Several such generalizations are examined and tested in a paper by Atkinson and Shiffrin (1965), but they are too complex to be considered here.

A finding that contradicts the interpretation given to the data by the register model is the observed relation between confidence ratings and the likelihood of a correct response. The register model describes the recall process as an all-or-none event. An item is either retained (i.e.,

in long-term memory or in the register) or it has been forgotten. However, the data in Fig. 2 clearly contradicts this assumption. The Ss were able to order their confidence ratings in a manner consistent with a response-strength or response-elimination schema. It may be sufficient to argue that a single item of information in this experimental task consists of more than a simple association between a color and a position in the display. Further analysis of this discrepancy between the model and the data is beyond the scope of this paper. A detailed discussion of this problem is given in the previously cited paper by Atkinson and Shiffrin.

References

Atkinson, R. C., and Shiffrin, R. M. Mathematical models for memory and learning. Tech. Rep. No. 81, Institute for Mathematical Studies in the Social Sciences, Stanford University, 1965.

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

Atkinson, R. C., Hansen, D. N., and Bernbach, H. A. Short-term memory with young children. Psychon. Sci., 1964, 1, 255-256.

Bernbach, H. A. A forgetting model for paired-associate learning. J. math. Psychol., 1965, 2, 128-144.

Bower, G. H. Notes on a descriptive theory of memory. Paper read at Second Conference on Learning, Remembering, and Forgetting, Princeton, New Jersey, September, 1964.

Broadbent, D. E. Flow of information within the organism. J. verb. Learn. verb. Behav., 1963, 4, 34-39.

Calfee, R. C., and Atkinson, R. C. Paired-associate models and the effects on list length. J. math. Psychol., 1965, 2, 254-165.

Greeno, J. G. Mathematical analysis of paired associate learning with short term retention. 1965, unpublished manuscript.

Peterson, L. R. Associative memory over brief intervals of time. J. verb. Learn. verb. Behav., 1963, 2, 102-106.

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

Appendix

Supplementary Data

Certain aspects of the data, while not relevant to the analyses presented in this paper, are of general interest. Table 1 presents the probability of a correct response and the probability of each confidence rating as a function of the serial position for display sizes $d = 3, 4, 5, 6,$ and 7. Table 2 presents the same information for display sizes $d = 8$ and 11. These same data for display size $d = 14$ are presented in Table 3.

Figure 5 presents the serial position curves for $Pr[R_1]$ over all display sizes. The shape of these curves is quite similar to the shape of the serial position curves for the proportion of correct response data presented in Fig. 3, and can be well fitted by the $F_i^{(d)}$ function derived earlier.

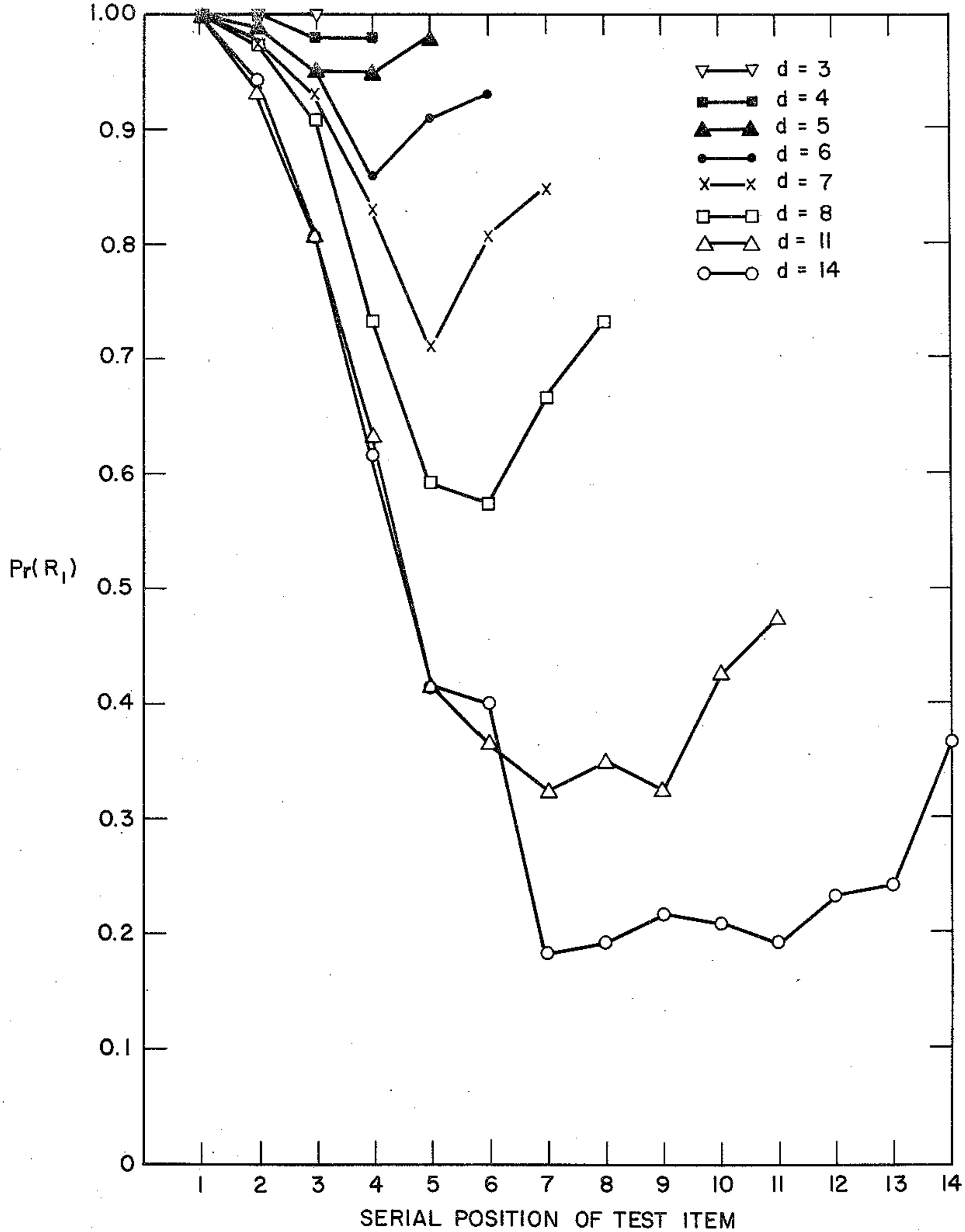

Figure 5. Proportion of R_1 responses at each serial position for the various display sizes.

Table 1

Proportion of correct responses and proportions of each confidence rating for the five smallest display sizes.

Display Size	Serial Position	$Pr(C_i)$	$Pr(R_1)$	$Pr(R_2)$	$Pr(R_3)$	$Pr(R_4)$
3	1	1.00	1.00	0.00	0.00	0.00
	2	1.00	1.00	0.00	0.00	0.00
	3	1.00	1.00	0.00	0.00	0.00
4	1	1.00	1.00	0.00	0.00	0.00
	2	1.00	1.00	0.00	0.00	0.00
	3	0.99	0.98	0.02	0.00	0.00
	4	1.00	0.98	0.02	0.00	0.00
5	1	1.00	1.00	0.00	0.00	0.00
	2	0.99	0.99	0.01	0.00	0.00
	3	0.96	0.95	0.05	0.00	0.00
	4	0.97	0.95	0.05	0.00	0.00
	5	0.98	0.98	0.02	0.00	0.00
6	1	1.00	1.00	0.00	0.00	0.00
	2	0.98	0.99	0.01	0.00	0.00
	3	0.95	0.95	0.05	0.00	0.00
	4	0.94	0.86	0.14	0.00	0.00
	5	0.93	0.91	0.08	0.01	0.00
	6	0.97	0.93	0.06	0.01	0.00
7	1	1.00	1.00	0.00	0.00	0.00
	2	0.97	0.98	0.02	0.00	0.00
	3	0.93	0.93	0.07	0.00	0.00
	4	0.89	0.83	0.14	0.02	0.01
	5	0.83	0.71	0.27	0.02	0.00
	6	0.82	0.81	0.18	0.01	0.00
	7	0.95	0.85	0.12	0.02	0.01

Table 2

Proportion of correct responses and proportions of each confidence rating for display sizes 8 and 11.

Display Size	Serial Position	$Pr(C_i)$	$Pr(R_1)$	$Pr(R_2)$	$Pr(R_3)$	$Pr(R_4)$
	1	1.000	0.992	0.008	0.000	0.000
	2	0.975	0.975	0.025	0.000	0.000
	3	0.933	0.908	0.075	0.008	0.008
8	4	0.892	0.733	0.225	0.033	0.008
	5	0.817	0.592	0.292	0.117	0.000
	6	0.733	0.575	0.342	0.050	0.033
	7	0.800	0.667	0.258	0.067	0.008
	8	0.908	0.733	0.208	0.050	0.008
	1	1.000	1.000	0.000	0.000	0.000
	2	0.967	0.933	0.058	0.008	0.000
	3	0.933	0.808	0.183	0.008	0.000
	4	0.767	0.633	0.233	0.108	0.025
11	5	0.633	0.417	0.383	0.150	0.050
	6	0.608	0.367	0.450	0.117	0.067
	7	0.658	0.325	0.375	0.233	0.067
	8	0.592	0.350	0.425	0.133	0.092
	9	0.550	0.325	0.442	0.175	0.058
	10	0.567	0.425	0.375	0.167	0.033
	11	0.742	0.475	0.433	0.067	0.025

Table 3

Proportion of correct responses and proportions of each confidence rating for the display size of 14.

Serial Position	$Pr(C_i)$	$Pr(R_1)$	$Pr(R_2)$	$Pr(R_3)$	$Pr(R_4)$
1	1.000	1.000	0.000	0.000	0.000
2	0.967	0.942	0.058	0.000	0.000
3	0.892	0.808	0.158	0.033	0.000
4	0.717	0.617	0.342	0.033	0.008
5	0.708	0.417	0.417	0.133	0.033
6	0.617	0.400	0.367	0.158	0.075
7	0.458	0.183	0.375	0.308	0.133
8	0.400	0.192	0.425	0.258	0.125
9	0.433	0.217	0.392	0.233	0.158
10	0.450	0.208	0.433	0.267	0.092
11	0.467	0.192	0.517	0.183	0.108
12	0.508	0.233	0.467	0.267	0.033
13	0.475	0.242	0.500	0.175	0.083
14	0.642	0.392	0.475	0.100	0.033