

The effect of an extraneous added memory set on item recognition: A test of parallel-dependent vs. serial-comparison models

LESTER E. KRUEGER

Ohio State University, Columbus, Ohio 43210

The linear increase in response time (RT) in item recognition as a function of memory set size may be attributed either to increased memory search, as in Sternberg's serial-comparison model, or to increased interference and competition for a limited processing capacity, as in a parallel-dependent model. An additional memory load was imposed on the normal item recognition task. If the normal task involved letters, the additional task involved digits, or vice versa. The added set was presented before and tested after the normal core set was presented and tested. Across the four trial blocks, the subjects did come to dissociate the two memory sets, but this could be due as much to directed rehearsal as to directed search when other results, including the subjects' postsession reports, are considered. Other findings, such as the higher error rate and shorter RT for positive than negative trials, seem to reflect differences in effective trace strength, and thus are consistent with the parallel-dependent model.

Sternberg (1966, 1969a, b) found response time (RT) in an item recognition task to increase linearly with the size of the memory set, and at the same rate or slope for positive trials, where the test item belonged to the memory set, as for negative trials. He concluded that the memory set is scanned in a serial and exhaustive manner for the test item. Sternberg's results have been confirmed by some studies, whereas others have found negatively accelerated, more nearly logarithmic functions, or have found different slopes for positive and negative trials (see Nickerson's 1972 review). The set-size function nearly disappears when the sets greatly exceed the capacity of short-term memory: RT increased at the rate of 38 msec per item for Sternberg's (1966) small sets, but only at about 5 msec per item for Juola, Fischler, Wood, and Atkinson's (1971) sets of 10 to 26 items and Atkinson and Juola's (1973) sets of 16 to 54 items. Juola et al. and Atkinson and Juola proposed that the subject directly accesses the test item in long-term memory, responding "yes" immediately if its subjective familiarity is quite high and "no" if quite low. If the familiarity is intermediate, the subject then searches through the memory set, either in serial or parallel fashion. The number of items searched ought to increase as set size increases, thus accounting for the small slope found.

A uniprocess model, however, can account for the results, both for small and large sets. The model

proposed here involves not serial comparison, but rather direct access to the test item in memory or parallel comparison of the test item with members of the positive set. The increase in RT with set size is attributed to the increased processing load thereby imposed. According to the model, the items in the positive set share the same limited processing capacity, so that once the capacity is fully committed, any processing demanded by an additional item must come at the expense of the other items or other processes (cf. Corcoran, 1971; Moray, 1967; Shulman & Greenberg, 1971; Townsend, 1971). Corcoran (1971, Chap. 3) attributed the increase in RT with set size to the sharing of "limited power" in a parallel comparison process, i.e., the more items to be matched, the less efficient each match. The processing demands may increase with set size, not only in terms of memory comparisons, as Corcoran proposed, but also of memory load. Retaining items in short-term memory requires active processing (rehearsal, organization, etc.), whose momentary demands may exceed those of seemingly more complex tasks (Kahneman, 1973, Chap. 2) and may interfere with the identification and processing of new items or information (Aaronson, 1974; Shulman & Greenberg, 1971). If a larger set receives the greater capacity needed for its retention, then less capacity will be available for encoding the test item and comparing it with the memory items. On the other hand, if no additional capacity were assigned to the retention of the larger set, such as might occur if there were fixed rates of rehearsal (Cavanaugh, 1972, Hypothesis 2), the average trace strength of the memory items would be reduced (cf. Krueger, 1970). In either case, the result would be to introduce

This study was supported by NIH Grant HD-04869. The author is grateful to Harry Rosenberg, Madelin Weiss, and Richard A. Cohen for assistance in collecting and tabulating the data.

Requests for reprints should be sent to Lester E. Krueger, Human Performance Center, Ohio State University, 404B West 17th Avenue, Columbus, Ohio 43210.

"noise" into the process, reducing the effective trace strength of the items in memory and thereby increasing RT. Further, there may be specific (structural) as well as nonspecific (capacity) interference among items (Kahneman, Chap. 1, 10). Although the present model allows for either direct access or parallel comparison and for either specific or nonspecific interference, it will be termed "parallel-dependent" for the sake of simplicity.

One test between the parallel-dependent and serial-comparison models would involve imposing an additional memory load on the subject during the presentation and testing of the regular memory set. If the additional memory load required basically the same type of encoding, storage, etc., as the regular memory set, yet were easily dissociable from the regular set, then the two models would make different predictions. The parallel-dependent model, taking into account the larger demands on processing capacity caused by the additional memory load, would predict an increase in RT for the regular set. The serial-comparison model, taking into account only the number of items actually searched, would predict no increase in RT for the regular set.

Some evidence does indeed suggest that subjects can restrict their memory comparisons to particular sets of items in some situations. In fact, the subjects may respond as quickly in some cases as if the excluded items had not been presented. This might seem to answer the present question in favor of the serial-comparison model over the parallel-dependent model, but a closer examination reveals problems in interpreting the evidence.

Under some conditions, subjects seem to organize their access to long-term memory in a hierarchical fashion, determining first the category to which the test item belongs and then examining only members of that category in the memory set (e.g., Homa, 1973; Okada & Burrows, 1973; Seamon, 1973). Homa (1973), in fact, attributed the very small slopes on set size for long lists to a short-circuited search process in which the subject first determines the appropriate category in a subjectively categorized list and then restricts his serial search to that category. The present question, however, concerns primarily short-term memory, which Sternberg studied. Several studies have found evidence for hierarchically organized search in short-term memory (Crain & DeRosa, 1974; Kaminsky & DeRosa, 1972; Naus, Glucksberg, & Ornstein, 1972). Naus et al. found lower RT when a list was composed of, say, eight words from two taxonomic categories (four animal names, four girls names) rather than from only one category. The subjects did not restrict search only to the relevant category, but, according to Naus et al., randomly entered either category and responded overtly only after having exhaustively searched the relevant category, whether entered first or second. Thus, the

subjects, who were not precued on which category would be tested, made only incomplete use of the category structure. Furthermore, what effect was found may not be due to search processes. Dividing the short four- to eight-item list into two categories may have aided input (i.e., storage, retention) rather than output (i.e., restricted search). Sanders and Schroots (1969) found better recall, scored by items in the correct serial position, when a list contained 6 digits followed by 6 consonants, or the reverse, than when it contained 12 digits or 12 consonants.

Kaminsky and DeRosa presented six-item lists which contained both letters and digits. Some lists contained more digits than letters, others the reverse. For most subjects, the size of the relevant category had little effect on RT, if the category type (letters, digits) was not precued. When category type was precued, however, RT was much shorter and did depend on the size of the relevant category. In fact, the set-size function was virtually identical, both on slope and intercept, to that obtained when only items from the relevant category were presented (control condition). Thus, the subjects made complete use of the category structure when precued. Since 2 sec intervened between cue and test item, however, the findings could reflect directed forgetting rather than directed search. Once precued, all efforts to retain items in the nonrelevant category may have ceased, thereby reducing the memory load at the time the test item appeared.

Crain and DeRosa (1974) associated every digit in the six-item memory set with a particular background color, and produced a large decrease in RT by precueing the background color of the test item. The cue preceded the test item by 2 sec, so that here, too, the results could reflect directed forgetting rather than directed search. With the precue, RT depended on the number of items in the subset associated with the particular color which was precued, yet remained higher than on the same number of items presented alone (control condition). That the subjects did not take full advantage of the advance information is understandable, considering that the color associations may not have been well learned.

Darley, Klatzky, and Atkinson (1972) precued subjects as to exactly which item in the one- to five-item set would appear if the test item was positive. Their subjects took full advantage of the advance information, responding as fast on precued trials in which the memory set contained two, three, four, or five items as on regular trials in which the memory set contained only one item. To insure that the noncued items remained as a memory load, the subjects were required to recall the memory set at the end of a trial. Darley et al. reported obtaining nearly perfect recall on all trials, but according to their Figure 1, the error rate on the recognition tests was nearly 10% for five-item noncued memory sets.

According to one explanation offered by Darley et al., memory load had no effect on RT in the precued conditions because it was held in a different memory system than that involved in the recognition decision. Since precueing occurred during presentation of the memory set, however, the subjects might have directed not their search but their prior rehearsal to the precued item.

Several other studies have imposed memory loads, but the evidence regarding the present question remains inconclusive. Wattenbarger and Pachella (1972) found no effect of the size of a memory load on choice RT. Memory load may well have had no effect, however, simply because of the much different type of processing involved. The choice RT task involved responding with the hand, either right or left, which agreed with the direction in which an arrow pointed, and the memory load consisted of letters. Spatial and verbal inputs seemingly do not compete for the same short-term store capacity (Henderson, 1972), and might well compete to only a limited extent for the same general processing capacity.

Sternberg (1969a, Experiment 5) imposed an extraneous memory load involving a similar type of item on an item recognition task and found a considerable increase in RT and in slope. He used a fixed-set procedure, with the subjects memorizing one, three, or five digits at the beginning of a series of trials. For the added memory load, he presented sequentially a new list of seven letters immediately before the test signal, which was either a test digit or a signal to recall the seven letters. The presentation of the seven letters may have prevented the subject from rehearsing and maintaining the digits in active short-term memory, as was indeed Sternberg's intention, and this factor, rather than the memory load per se, may account for the effect of the added letters. Shiffrin (1970) found no effect on free recall performance of the length of an intervening list, but he studied recall from long-term rather than from short-term memory and used an accuracy rather than RT measure. Forrin and Morin (1969) found an increase in RT for a test item pertaining to a set in short-term store due to a concurrent long-term memory load. The increase in RT, though, held constant, regardless of the size of the long-term memory load. Adding one, two, or three items to long-term store increased RT for the short-term test item about as much as adding one item to the short-term store itself. What effect Forrin and Morin did obtain may be due to the fact that their subjects never knew beforehand whether the test item would pertain to the long-term set or the short-term set.

The present experiment attempted to provide a more conclusive test of the serial-comparison vs. parallel-dependent models. The primary concern was whether RT on a regular core set would be increased by that of an added set. On each trial, first the

added-set items were presented sequentially, then the core-set items, followed immediately by a test item pertaining to the core set, and, finally, to maintain credibility, a test item pertaining to the added set. The presence of the regular core set in active memory was guaranteed by presenting it after the added set and by limiting it to a maximum of four items. For half of the subjects, the added set contained letters and the core set contained digits, and for the other half, the added set contained digits and the core set contained letters. Letters and digits both require similar verbal processing, so that any effect of memory load ought to be evident, yet letters and digits also are readily dissociable, so that any restriction on memory search ought to be evident. The subjects were in effect precued, since they knew beforehand precisely when each set would be tested, and thus were aided further in restricting search. Also, since the subjects knew that the items in the added set would be tested after the core set and, therefore, would have to be maintained in memory, the results should reveal the effects of directed search rather than directed forgetting. Finally, differential rehearsal was discouraged by giving equal stress in the instructions to the performance on both memory sets, by requiring the subject to say aloud each item as it appeared, and by having only a brief pause between the core set and its test items.

METHOD

Subjects

Thirty-two City College of New York undergraduates served as paid volunteers.

Apparatus

An Industrial Electronic Engineers Bina-View self-decoding readout display, Model 55579-1886, provided illuminated characters that were .5 cm thick, 3.5 cm high, and, depending on the character, about 2.5 cm wide. The characters were shown at 78.7 cm from the subject, with all other portions of the visual field blocked from view. The sequencing of display characters was controlled by an Ohrtronics paper tape reader, with the timing controlled by a Scientific Prototype program timer. The subjects' response times were punched out on paper tape for subsequent computer analysis.

Stimulus Materials

The letter stimuli were drawn exclusively from the set A, B, C, D, E, F, G, H, K and digit stimuli from the set 1, 2, 3, 4, 5, 6, 7, 8, 9. Each letter and digit appeared with approximately equal frequency in the stimulus strings and as a test item. A master list, containing a random ordering of the strings for the individual trials, was rendered in one of four modes to produce the specific stimuli for a given subject: The items in a particular mode were assigned as either letters or digits, in either a normal or reversed fashion. For instance, if the master list specified a "2," then the actual character presented would be either a B (normal letter), 2 (normal digit), H (reversed letter), or 8 (reversed digit). Each of the four modes was presented to 8 of the 32 subjects.

On each trial, the subject received two input strings, first the added set and then the core set. For half of the subjects, the added set always contained letters and the core set always digits, and for the other half, the reverse assignment. The added and core sets

each could contain zero, one, two, or four items. Taking all possible combinations of number of added-set items (zero, one, two, four) and core-set items (zero, one, two, four), omitting the nonexistent 0, 0 combination, gave a basic set of 15 trials. Each of the 15 was presented once each in four different types of trials: Both test items were positive, only the added-set item was positive, only the core-set item was positive, and both test items were negative. The 60 trials thus obtained were randomly permuted and formed one block on the master list. Having four blocks, or 240 trials in all, allowed each serial position in the two- and four-item sets to be probed equally often. There were not enough trials, however, to probe equally often every combination of added-set serial position and core-set serial position. In addition, there were 30 practice trials, which were balanced as much as possible on the various types of trials.

Procedure

Upon pressing the footpedal to initiate a trial, the subject first received the added set: One # character appeared for 1.6 sec, then each letter or digit, also for 1.6 sec. The core set followed immediately: again an # and then each letter or digit, all for 1.6 sec each. Immediately after the core set had been presented, a large homogeneously white field was shown on the Bina-View and a buzzer was sounded concurrently for 1.6 sec, followed by a test item which pertained to the core set (i.e., if the core set contained digits,

the test item was a digit) and which remained on until the subject responded. Immediately after the subject's response, an # appeared for 1.6 sec, a white-field-and-buzzer combination for 1.6 sec, and then a test item pertaining to the added set. On trials where there was only one set (i.e., the other set contained zero items), the presentation and testing were as described above for the core set.

In order to equalize attention to various items in a string, subjects were required to pronounce aloud each letter and digit as it appeared. The subjects were offered a 50-cent bonus if their average response time fell in the top quartile for the study, while their error rate was kept below 5%. Half of the subjects made a positive response with their left hand and a negative response with their right hand, and for the other half, the reverse assignment.

RESULTS

"Target set" will refer here to the set, either core or added, to which the test item pertains, and "extraneous set" to the other set. For example, if the subject first received a two-letter added set and then a four-digit core set, the first test item, a digit, would have a four-item target set (core set) and a two-item

Table 1
Mean Reaction Time (in Milliseconds) and Error Rate (Percentage), in Parentheses, by Trial Block, Positive vs. Negative, Size of Target Set, and Size of Extraneous Set for Core-Set Test Items

Size of Target Set (Core Set)		Size of Extraneous Set (Added Set)							
		0		1		2		4	
Block 1									
Positive	1	764	(0.0)	716	(1.6)	777	(1.6)	845	(1.6)
	2	707	(0.0)	778	(1.6)	842	(3.1)	929	(3.1)
	4	878	(0.0)	951	(6.3)	881	(0.0)	1105	(12.5)
Negative	1	742	(0.0)	770	(0.0)	824	(1.6)	899	(0.0)
	2	865	(0.0)	875	(0.0)	905	(4.7)	949	(1.6)
	4	977	(6.3)	969	(0.0)	977	(7.8)	945	(4.7)
Block 2									
Positive	1	608	(0.0)	766	(1.6)	754	(3.1)	844	(1.6)
	2	744	(0.0)	738	(0.0)	878	(3.1)	815	(9.4)
	4	828	(4.7)	865	(9.4)	962	(1.6)	883	(3.1)
Negative	1	759	(3.1)	775	(0.0)	869	(1.6)	817	(0.0)
	2	770	(1.6)	796	(0.0)	805	(0.0)	918	(4.7)
	4	885	(0.0)	933	(4.7)	899	(0.0)	938	(1.6)
Block 3									
Positive	1	664	(1.6)	688	(0.0)	716	(0.0)	783	(4.7)
	2	682	(3.1)	781	(3.1)	782	(1.6)	873	(3.1)
	4	783	(1.6)	866	(3.1)	877	(4.7)	977	(14.1)
Negative	1	766	(3.1)	788	(0.0)	922	(1.6)	863	(1.6)
	2	909	(1.6)	829	(0.0)	800	(3.1)	829	(3.1)
	4	877	(0.0)	932	(3.1)	916	(3.1)	937	(6.3)
Block 4									
Positive	1	661	(3.1)	715	(6.3)	696	(3.1)	716	(1.6)
	2	725	(3.1)	700	(3.1)	786	(1.6)	801	(4.7)
	4	803	(6.3)	858	(4.7)	927	(1.6)	845	(4.7)
Negative	1	734	(0.0)	899	(0.0)	823	(3.1)	782	(0.0)
	2	828	(0.0)	832	(0.0)	823	(1.6)	820	(3.1)
	4	815	(0.0)	929	(0.0)	885	(4.7)	934	(0.0)
All Blocks									
Positive	1	674	(1.2)	721	(2.3)	735	(2.0)	797	(2.3)
	2	715	(1.6)	749	(2.0)	822	(2.4)	854	(5.1)
	4	823	(3.1)	885	(5.9)	912	(2.0)	953	(8.6)
Negative	1	750	(1.6)	808	(0.0)	859	(2.0)	840	(0.4)
	2	843	(0.8)	833	(0.0)	833	(2.3)	879	(3.1)
	4	888	(1.6)	941	(2.0)	919	(3.9)	939	(3.1)

extraneous set (added set); whereas the second test item, a letter, would have a two-item target set (added set) and a four-item extraneous set (core set). The data reported below are collapsed over the letters vs. digits variable, because, although performance was generally better on digits, there were few significant interactions with the other, more important variables. Since the core-set test items provide the most important information, they will be presented first.

Core-Set Test Items

The data are presented for every treatment combination in Table 1. The RT decreased significantly across the four trial blocks, $F(3,90) = 4.63, p < .01$, but error rate did not, $F < 1$. The subjects responded significantly faster on positive than negative test items, $F(1,30) = 29.06, p < .001$, but made more errors on the positive items, 3.2% vs. 1.7%, $F(1,30) = 8.69, p < .01$. Increasing the size of the target set (core set) increased significantly both RT, $F(2,60) = 93.25, p < .001$, and error rate, $F(2,60) = 9.85, p < .001$. Increasing the size of the extraneous set (added set) also increased significantly both RT, $F(3,90) = 24.61, p < .001$, and error rate, $F(3,90) = 7.66, p < .001$.

The size of the extraneous set had an effect, but it was not as large as that of the target set. As Table 1 indicates, each additional item increased RT by about 45 msec when added to the target set and about 25 msec when added to the extraneous set. In making statistical comparisons, however, the trials having an extraneous set of Size 0 were excluded, because these involved not only a decrease in the size of the extraneous set but also the elimination of the 1.6-sec delay between sets. Further, on positive trials, serial position of the test item in the overall presentation was controlled. Target set size had a significantly larger

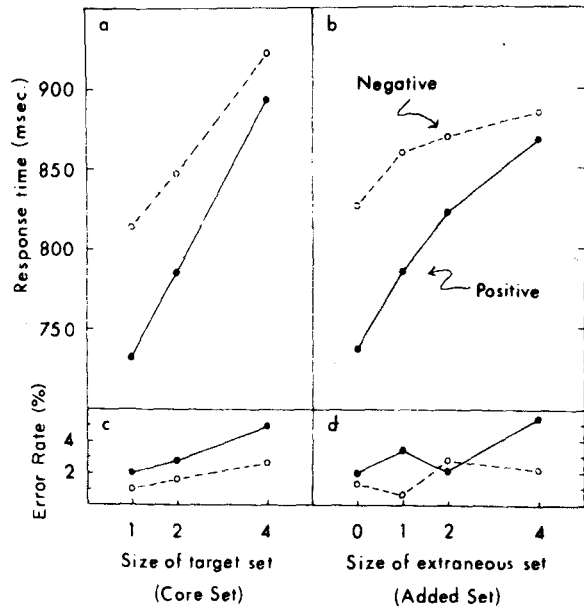


Figure 2. Mean response time, in milliseconds, and mean error rate (percentage) for core-set test items by positive vs. negative trials and by size of target set (core set) and size of extraneous set (added set).

effect than extraneous set size only in the case of negative RT ($p < .01$).

Trial block did not interact significantly with target set size but did interact significantly with extraneous set size on RT, $F(9,270) = 2.26, p < .025$, though not on error rate. As Figure 1 shows, the effect of extraneous set size decreases fairly steadily from Block 1 to Block 4, but the effect of target set size shows virtually no such change. Separate analyses of variance on each block revealed significant effects of extraneous set size ($p < .025$ or better) and target set size ($p < .001$) on RT for each of the four blocks. Thus, the effect of extraneous set size decreased, but was not eliminated, as the subjects became more practiced. On errors, significant effects ($p < .05$ or better) for both extraneous and target set size were found for Blocks 1 and 3 but not for Blocks 2 and 4.

The Blocks by Target Set by Extraneous Set interaction was significant both for RT, $F(18,540) = 1.79, p < .05$, and error rate, $F(18,540) = 2.70, p < .001$, but is largely uninterpretable, reflecting no meaningful pattern of change over the four blocks. Since each block remained the same for all subjects, except for changes in the assignment of particular digits and letters, peculiarities in the random ordering of trials within particular blocks may account for the different patterns of results in different blocks. Other complex and uninterpretable interactions were obtained on RT: Blocks by Positive-Negative by Extraneous Set, $F(9,270) = 2.49, p < .025$; Blocks by Positive-Negative by Target Set by Extraneous Set, $F(18,540) = 2.59, p < .01$.

As Figure 2 shows, set size had a consistently larger effect on RT for positive than negative trials. The Positive-Negative by Target Set interaction was

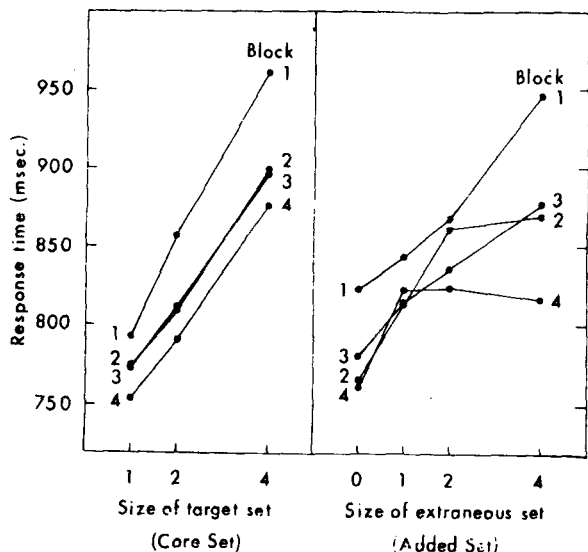


Figure 1. Mean response time, in milliseconds, for core-set test items by trial block and by size of target set (core set) and size of extraneous set (added set).

Table 2
 Mean Reaction Time (in Milliseconds) and Error Rate (Percentage), in Parentheses, by Trial Block, Positive vs. Negative, Size of Target Set, and Size of Extraneous Set for Added-Set Test Items

Size of Target Set (Added Set)		Size of Extraneous Set (Core Set)							
		0		1		2		4	
Block 1									
Positive	1	695	(0.0)	711	(3.1)	736	(1.6)	793	(7.8)
	2	768	(1.6)	804	(3.1)	814	(0.0)	794	(4.7)
	4	842	(0.0)	781	(3.1)	924	(18.8)	923	(4.7)
Negative	1	833	(1.6)	765	(3.1)	834	(0.0)	905	(4.7)
	2	866	(1.6)	887	(0.0)	816	(0.0)	963	(12.5)
	4	794	(0.0)	923	(4.7)	919	(6.3)	1037	(26.6)
Block 2									
Positive	1	826	(0.0)	705	(1.6)	642	(4.7)	707	(0.0)
	2	752	(0.0)	749	(0.0)	798	(7.8)	724	(1.6)
	4	811	(1.6)	809	(7.8)	799	(7.8)	878	(14.1)
Negative	1	810	(0.0)	716	(0.0)	931	(0.0)	842	(1.6)
	2	767	(3.1)	799	(0.0)	845	(1.6)	839	(6.3)
	4	915	(0.0)	826	(0.0)	921	(3.1)	1017	(4.7)
Block 3									
Positive	1	655	(1.6)	685	(1.6)	692	(1.6)	745	(6.3)
	2	737	(3.1)	726	(0.0)	778	(0.0)	816	(17.2)
	4	795	(7.8)	775	(4.7)	869	(15.6)	855	(14.1)
Negative	1	814	(0.0)	774	(1.6)	811	(3.1)	811	(1.6)
	2	780	(7.8)	848	(6.3)	750	(1.6)	856	(1.6)
	4	807	(0.0)	839	(4.7)	830	(3.1)	992	(17.2)
Block 4									
Positive	1	726	(1.6)	672	(4.7)	664	(1.6)	685	(1.6)
	2	768	(0.0)	684	(3.1)	752	(1.6)	734	(4.7)
	4	789	(1.6)	919	(17.2)	776	(7.8)	822	(18.8)
Negative	1	764	(0.0)	823	(3.1)	739	(1.6)	801	(4.7)
	2	792	(1.6)	817	(1.6)	844	(1.6)	794	(9.4)
	4	873	(6.3)	856	(7.8)	884	(7.8)	837	(20.3)
All Blocks									
Positive	1	726	(0.8)	693	(2.7)	683	(2.3)	732	(3.9)
	2	756	(1.2)	741	(1.6)	785	(2.3)	767	(7.0)
	4	809	(2.7)	821	(8.2)	842	(12.5)	870	(12.9)
Negative	1	806	(0.4)	769	(2.0)	829	(1.2)	840	(3.1)
	2	801	(3.5)	838	(2.0)	814	(1.2)	863	(7.4)
	4	848	(1.6)	861	(4.3)	888	(5.1)	971	(17.2)

significant, $F(2,60) = 5.80$, $p < .01$, as was the Positive-Negative by Extraneous Set interaction, $F(3,90) = 6.91$, $p < .001$. The Positive-Negative by Target Set by Extraneous Set interaction also was significant, $F(6,180) = 2.65$, $p < .025$, but cannot be readily interpreted. The target set size functions were quite linear, the least-squares slope being 53.7 msec per item for positive trials, 36.2 msec per item for negative trials, and 45.0 msec per item overall. On error rate the pattern was more complex, but again set size generally had a larger effect on positive than negative trials. The Positive-Negative by Extraneous Set interaction was significant, $F(3,90) = 5.28$, $p < .01$, though the Positive-Negative by Target Set interaction was not.

Added-Set Test Items

The data are presented for every treatment combination in Table 2. Trial block had a significant effect on both RT, $F(3,90) = 4.00$, $p < .025$, and

error rate, $F(3,90) = 6.52$, $p < .001$. The RT decreased steadily across the four blocks, but error rate decreased only from Block 1 to Block 2, then increased to Block 3 and Block 4. The subjects responded significantly faster on positive test items, $F(1,30) = 45.0$, $p < .001$, but with a higher, though not significantly so, error rate on positive items. Increasing the size of the target set (added set) increased significantly both RT, $F(2,60) = 57.76$, $p < .001$, and error rate, $F(2,60) = 35.59$, $p < .001$. Increasing the size of the intervening extraneous set (core set) increased significantly both RT, $F(3,90) = 6.92$, $p < .001$, and error rate, $F(3,90) = 26.86$, $p < .001$. Special analyses revealed no significant differences between the effect of target set size and extraneous set size.

As Figure 3 shows, the effect of extraneous set size on RT decreased steadily across the four blocks, except for one discrepant point on Block 2, but the effect of target set size showed virtually no such

change. The Blocks by Extraneous Set interaction was significant, $F(9,270) = 3.13, p < .01$, but not the Blocks by Target Set interaction, $F < 1$. Separate analyses of variance on each block revealed that the effect of target set size was significant for each block ($p < .001$), whereas extraneous set size was significant for Blocks 1, 2, and 3 ($p < .05$ or better), but not Block 4. Similar interactions were obtained on error rate, but these reflect largely uninterpretable patterns.

As Figure 4 shows, target set size had a larger effect on positive than negative RT, as attested to by the significant Positive-Negative by Target Set interaction, $F(2,60) = 3.22, p < .05$. The target set size functions were generally linear, the least-squares slope being 41.2 msec per item for positive trials, 27.6 msec per item for negative trials, and 34.4 msec per item overall.

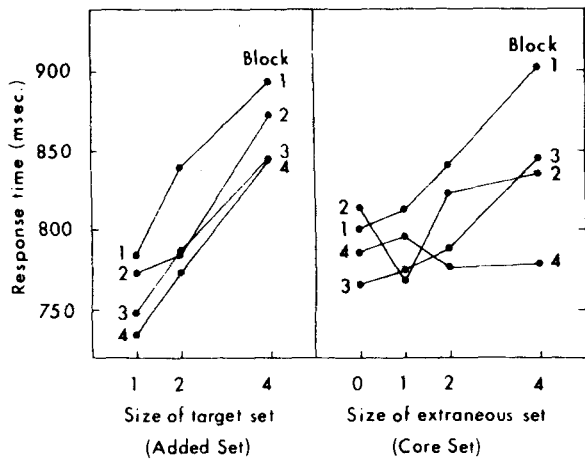


Figure 3. Mean response time, in milliseconds, for added-set test items by trial block and by size of target set (added set) and size of extraneous set (core set).

Serial Position Effects

Figure 5 shows RT and error rate on positive trials by serial position in the target set. For added-set test items, whose onset was typically delayed by the intervening core set, there is only a primacy effect: Both RT and error rate were lowest for the first item in the set. The effect of serial position was not significant in the two-member case, but was in the four-member case, on both RT, $F(3,90) = 8.28, p < .001$, and error rate, $F(3,90) = 4.32, p < .01$. The Serial Position by Extraneous Set interaction was significant both for RT on two-member added sets, $F(3,90) = 6.59, p < .001$, and error rate on four-member added sets, $F(9,270) = 2.88, p < .01$. Both interactions reflect the emergence of a recency effect, when the extraneous set (core set) had a size of 0, thus allowing the test item to appear immediately after the target set. Though not significant, the same trend also was evident for two-member error rate and four-member RT.

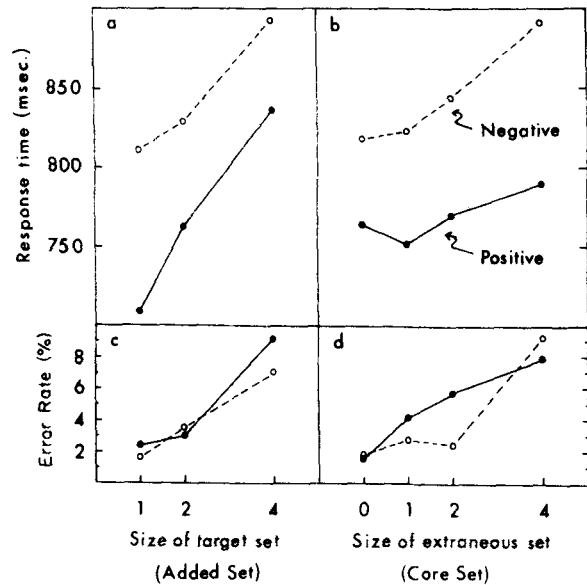


Figure 4. Mean response time, in milliseconds, and mean error rate (percentage) for added-set test items by positive vs. negative trials and by size of target set (added set) and size of extraneous set (core set).

For core-set test items, which were presented immediately after the core set, there is mainly a recency effect, but also some indication of a primacy effect for the four-member sets. Serial position effect was not significant on two-member error rate, but was significant on two-member RT, $F(1,30) = 7.45, p < .025$, four-member error rate, $F(3,90) = 6.32, p < .001$, and four-member RT, $F(3,90) = 8.65,$

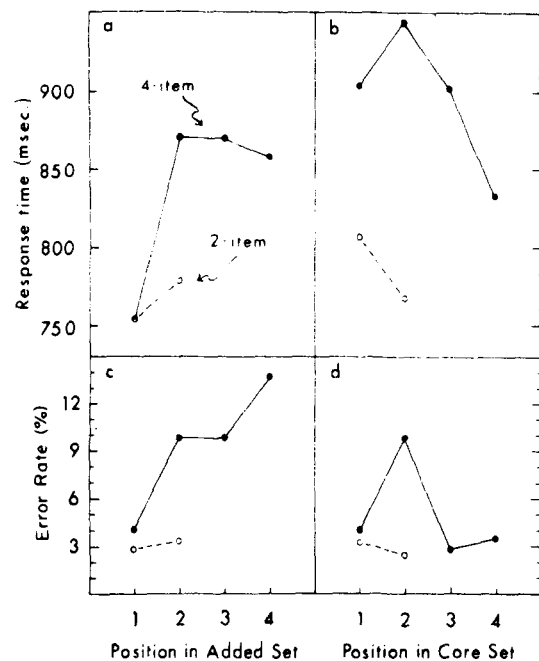


Figure 5. Mean response time, in milliseconds, and mean error rate (percentage) for test items in two-item and four-item memory sets by serial position of test item in added set and core set.

$p < .001$. The Serial Position by Extraneous Set interaction was significant on two-member error rate, $F(3,90) = 4.15$, $p < .01$, on four-member error rate, $F(9,270) = 2.59$, $p < .01$, and on four-member RT, $F(9,270) = 3.33$, $p < .01$. The error data were erratic, but on RT there was a tendency for the recency effect to increase and the primacy effect to decrease as the number of extraneous items increased.

Combined analyses of added-set and core-set items generally confirmed the difference between the sets on serial position effects and the effect of extraneous set size on serial position effects.

DISCUSSION

The present findings would seem at first glance to provide more support for the serial-comparison model than the parallel-dependent model. The size of the extraneous added set had less effect on core-set test items than did the size of the core set itself. Further, the effect of the extraneous added set on RT diminished considerably across the four trial blocks, whereas that of the target core set remained virtually unchanged (Figure 1). These findings indicate that the subjects employed different processes in dealing with the two memory sets and thus could and did dissociate them. What effect the extraneous added-set items did have might simply reflect a failure to completely dissociate added-set from core-set items, especially on early trial blocks.

Much evidence suggests, however, that the dissociation of the two sets was based more on directed rehearsal than directed search. Although the subjects had to say aloud each item as it was presented, about half (17) reported having rehearsed other items while a particular item was being shown. The 1.6-sec warning interval before onset of the test item also provided an opportunity to rehearse the target set, which 22 subjects (70%) reported having exploited. Nearly half the subjects (13) reported they had rehearsed the letters and digits separately and differently, forming words, phonetics, names, etc., with the letters, and sequences, telephone numbers, dates, etc., with the digits. The added-set, judging by its primacy effect, higher error rate, and subjects' reports, was shunted into a long-term store and only moved back into active memory after the core set had been probed. The extraneous added-set items thus may have had less effect only because the subjects concentrated their rehearsal on the core-set items.

The basic dilemma in the present type of test is that whatever conditions better enable the subject to direct his search (precueing, use of letters vs. digits, a warning interval before the test item appears, etc.) also better enable the subject to direct his rehearsal. A similar issue has arisen in the work on directed forgetting. When a subject receives a paired-associate test stimulus 3 sec after being told which portion of

the list (i.e., first half, second half, entire list) will be tested, he may use the 3-sec interval either to restrict his search or actually to change his memory by differentially rehearsing and organizing the to-be-remembered portion of the list. Spector, Laughery, and Finkelman (1973) found that filling the 3-sec interval with an arithmetic problem reduced the advantage provided by the half-list cue. This result would seem to implicate differential rehearsal and organization, except that the subjects may yet have been rehearsing during the filled 3-sec interval: 45% of Epstein and Wilder's (1972) subjects reported that they could rehearse during some filled postcue intervals, especially on easy arithmetic problems. In any case, differential rehearsal and organization cannot account for the entire cueing effect, because many of the to-be-forgotten items can be recovered when specially cued; the recovery has been complete in some cases (Shebilske & Epstein, 1973), but not in others (Epstein & Wilder, 1972).

The main variable of the present study, size of the extraneous added set, does not provide a conclusive test of the two models, because it may reflect the effects of either directed rehearsal or directed search; but what about the other variables? The data on error rate, serial position effects on RT, and the slope on RT for positive and negative trials seem generally more consistent with the parallel-dependent model than the serial-comparison model. On core-set test items, error rate was quite low, yet did increase significantly as the size of the core set increased (Figure 2). The increase in error rate cannot be attributed simply to an increase in the opportunity for a negative test item to be confused with a memory (positive) item, as the set increased in size. Such increased confusion would account for the increase in false-positive responses but not for the increase in false-negative responses. Since errors actually increased more on positive than on negative trials, it seems more likely that increasing the size of the target set simply reduced the effect trace strength of individual members of the memory set.

The serial position effects (see Figure 5) suggest that items in the memory set varied in effective trace strength. In the core set, RT and error rate were lowest for the last items (recency effect) and the first item (primacy effect). Similar serial position effects on RT have been found in previous studies (e.g., Corballis, Kirby, & Miller, 1972; Forrin & Cunningham, 1973; Forrin & Morin, 1969). These findings indicate that, if memory search is serial, it at least is self-terminating (cf. Theios, Smith, Haviland, Traupmann, & Moy, 1973). It might be instead that greater trace strength of the last items in a list earns them direct access to the decision process, by-passing any scanning procedure (Corballis, Kirby, & Miller, 1972; Forrin & Cunningham, 1973). The recency effects on error rate found in the present study would

seem especially to indicate a parallel-dependent process, since they seem more clearly to reflect variations in effective trace strength. If the test item is delayed by several seconds, the short-term persistence favoring the last items is dissipated and the recency effect disappears (Forrin & Cunningham, 1973). In the present study, the long delay imposed by the intervening core set eliminated the recency effect on added-set test items, leaving only a primacy effect.

The 45-msec per item slope for the core set and 36-msec per item slope for the added set match quite well the 38-msec per item slope obtained by Sternberg (1966). The present data differ from those of Sternberg, however, in that the slope was only two-thirds as large for negative as for positive trials, both for the core set (36.2 vs. 53.7 msec) and the added set (27.6 vs. 41.2 msec). Forrin and Morin (1969) similarly found about half as great a slope for negative as for positive trials. Corballis, Kirby, and Miller (1972) also found a steeper slope for positive than negative trials. These findings, though they are exceptions to the usual finding of about equal slope for positive and negative trials (see Nickerson's 1972 review), do pose a problem for the serial-comparison model. The slope ought to be equal for positive and negative if serial search is exhaustive, and steeper for negative than positive if search is self-terminating, but *never* steeper for positive than negative trials. The steeper slope for positive trials might yet be explained by a serial self-terminating model which allows both positive and negative items on the list scanned, with the positive items being more likely to be at the top of the list due to their recent presentation (Theios et al., 1973). The steeper slope for positive items, however, also was found for the added set, where there was no recency effect.

Another difference between positive and negative trials was the curious fact that, at least in the core set, RT was lower but error rate higher on positive than negative trials. These results cannot be ascribed simply to a speed-accuracy tradeoff, because the subjects never knew whether a trial would be positive or negative and, thus, whether to speed up or not. Nor can the faster RT on positive trials be ascribed to a simple response bias, because there actually was a greater tendency to respond negative than positive, i.e., there were more false-negative than false-positive responses. Krueger (1973) obtained similar results with a same-different visual comparison task, and presented a coding or effective trace-strength model, which will now be modified and applied to the present situation.

First, positive items are assumed to acquire special tags in memory, specifying recent activation or membership in the positive set, whereas negative items bear "null" tags. When the positive tags are effaced, they revert to the null state. This would account for the fact that the error rate is higher on positive than negative trials.

Second, the comparison process may be seen as comprised of a series of ever more finely detailed examinations of the test item's tag. The comparison process may improve either because it itself improves, coming into better focus, or because the representation of the test item improves, having not yet been fully encoded when the initial comparisons began (cf. Krueger, 1973). In any case, when the positive trace is strong and clear, only a single, gross examination might be necessary to determine that the tag is positive. When the trace is weaker, as when an item receives less rehearsal time in a larger memory set, a closer, longer comparison would be required, thereby increasing RT. This would account for the effect of set size on RT.

Third, since the effacement or blurring of the positive tags can change a positive item's tag into a null or "all other" tag, but not a negative item's null tag into a positive tag, the subject may safely respond as soon as he affirms a positive tag. The negative response, however, would not be made as soon as a null tag had been found, but only after further checks had insured that the tag was not simply a somewhat defaced or misread positive tag. This would account for the fact that negative RT generally is longer than positive RT.

The RT would generally be shorter for positive than negative responses, that is, because the subject could self-terminate and respond "yes" as soon as he affirmed a positive tag. The difference between positive and negative trials would lessen, however, as the target set increased in size (see Figures 2 and 4), because the further weakening of positive tags would force processing to be even more exhaustive on both positive and negative trials alike. This would account for the fact that the RT slope is steeper for positive than negative trials.

Nickerson (1972) and Briggs and Johnsen (1973) have presented models that are very similar to the present one. In Nickerson's case, positive RT is held to be inversely related to the trace strength of the test item. Positive responses are made as soon as the test item "rings a bell" (i.e., reveals it bears a positive tag), whereas negative decisions occur by default: The subject responds "no" only after he decides the bell is not going to ring. The response-by-default accounts for the typically longer RT for negative than positive trials. To account for the increase in negative RT with memory set size, Nickerson posited that the time needed to make a negative (default) decision is determined by the minimum strength of the positive items. When the positive set is made larger, the trace strength of the positive items will generally be weaker, and the subject will wait longer before deciding that the bell is not going to ring.

As in the present experiment, Briggs and Johnsen, as well as other investigators (Corballis et al., 1972; Darley et al., 1972; Forrin & Cunningham, 1973; Forrin & Morin, 1969), have found faster RT but

higher error rate for positive than negative trials. Briggs and Johnsen further found a divergence of the positive and negative functions on both RT and error rate, with an increase in set size, whereas the present study, as well as Forrin and Morin, found divergence on error rate but a convergence on RT. Briggs and Johnsen explained their findings by proposing that the subject responds "yes" as soon as he has registered a positive match but rechecks his central processing before making a negative response. The effect of rechecking is held to be a power function of an information measure, H_c , which corresponds roughly to memory set size, so that the larger the memory set, the more rechecking time is added when the response is negative. Murdock (1971) also posited the making of "second looks" on some proportion of negative trials as a way to explain the steeper slopes sometimes found for negative trials. In the present model, by contrast, additional processing and rechecking occurs for both positive and negative items. The subject must process more carefully all items which are perceived as negative, whether actually positive or negative. Further, increasing the memory set size ought to have a relatively larger effect on positive RT, because more and more positive items would be misclassified, both initially and permanently (i.e., false-negative responses). Consistent with this prediction are the present data and those of Corballis et al. (1972) and Forrin and Morin (1969), all showing a convergence rather than divergence on RT.

Since Briggs and Johnsen found a decided increase in the error rate for positive items with set size, they, too, ought to have allowed for the rechecking of some positive items. It is perhaps easy to lose sight of the positive items because, after all, it is mainly the negative items which undergo the closer scrutiny. According to the present analysis, however, it is really the positive items, or at least those that are wont to appear disguised initially as negative or "null" items, that are the real culprits in this matter. These false-negatives force the subject to spend more time processing all perceived negatives, both false and real, in order to avoid responding "no" incorrectly to a positive item.

REFERENCES

- AARONSON, D. Stimulus factors and listening strategies in auditory memory: A theoretical analysis. *Cognitive Psychology*, 1974, 6, 108-132.
- ATKINSON, R. E., & JUOLA, J. F. Factors influencing speed and accuracy of word recognition. In Kornblum, S. (Ed.), *Attention and performance IV*. New York: Academic Press, 1973. Pp. 583-612.
- BRIGGS, G. E., & JOHNSEN, A. M. On the nature of central processing in choice reactions. *Memory & Cognition*, 1973, 1, 91-100.
- CAVANAUGH, J. P. Relation between the immediate memory span and the memory search rate. *Psychological Review*, 1972, 79, 525-530.
- CORBALLIS, M. C., KIRBY, J., & MILLER, A. Access to elements of a memorized list. *Journal of Experimental Psychology*, 1972, 94, 185-190.
- CORCORAN, D. W. J. *Pattern recognition*. Baltimore: Penguin, 1971.
- CRAIN, R. D., & DEROSA, D. V. Retrieval of information from multiple ensembles in short-term memory. *Memory & Cognition*, 1974, 2, 255-260.
- DARLEY, C. F., KLATZKY, R. L., & ATKINSON, R. C. Effects of memory load on reaction time. *Journal of Experimental Psychology*, 1972, 96, 232-234.
- EPSTEIN, W., & WILDER, L. Searching for to-be-forgotten material in a directed forgetting task. *Journal of Experimental Psychology*, 1972, 95, 349-357.
- FORRIN, B., & CUNNINGHAM, K. Recognition time and serial position of probed item in short-term memory. *Journal of Experimental Psychology*, 1973, 99, 272-279.
- FORRIN, B., & MORIN, R. E. Recognition times for items in short- and long-term memory. In W. G. Koster (Ed.), *Attention and performance II*. Acta Psychologica, 1969, 30, 126-141.
- HENDERSON, L. Spatial and verbal codes and the capacity of STM. *Quarterly Journal of Experimental Psychology*, 1972, 24, 485-495.
- HOMA, D. Organization and long-term memory search. *Memory & Cognition*, 1973, 1, 369-379.
- JUOLA, J. F., FISCHLER, I., WOOD, C. T., & ATKINSON, R. C. Recognition time for information stored in long-term memory. *Perception & Psychophysics*, 1971, 10, 8-14.
- KAHNEMAN, D. *Attention and effort*. Englewood Cliffs, N.J.: Prentice-Hall, 1973.
- KAMINSKY, C. A., & DEROSA, D. V. Influence of retrieval cues and set organization on short-term recognition memory. *Journal of Experimental Psychology*, 1972, 96, 449-454.
- KRUEGER, L. E. Effect of stimulus probability on two-choice reaction time. *Journal of Experimental Psychology*, 1970, 84, 377-379.
- KRUEGER, L. E. Effect of irrelevant surrounding material on speed of same-different judgment of two adjacent letters. *Journal of Experimental Psychology*, 1973, 98, 252-259.
- MORAY, N. Where is capacity limited? A survey and a model. *Acta Psychologica*, 1967, 27, 84-92.
- MURDOCK, B. B. A parallel-processing model for scanning. *Perception & Psychophysics*, 1971, 10, 289-291.
- NAUS, M. J., GLUCKSBERG, S., & ORNSTEIN, P. Taxonomic word categories and memory search. *Cognitive Psychology*, 1972, 3, 643-654.
- NICKERSON, R. S. Binary-classification reaction time: A review of some studies of human information-processing capabilities. *Psychonomic Monograph Supplements*, 1972, 4(17, Whole No. 65), 275-318.
- OKADA, R., & BURROWS, D. Organizational factors in high-speed scanning. *Journal of Experimental Psychology*, 1973, 101, 77-81.
- SANDERS, A. F., & SCHROOTS, J. J. F. Cognitive categories and memory span, III. Effects of similarity on recall. *Quarterly Journal of Experimental Psychology*, 1969, 21, 21-28.
- SEAMON, J. G. Retrieval processes for organized long-term storage. *Journal of Experimental Psychology*, 1973, 97, 170-176.
- SHEBILSKE, W., & EPSTEIN, W. Effect of forget instructions with and without the conditions for selective search. *Memory & Cognition*, 1973, 1, 261-267.
- SHIFFRIN, R. M. Forgetting: Trace erosion or retrieval failure? *Science*, 1970, 168, 1601-1603.
- SHULMAN, H. G., & GREENBERG, S. N. Perceptual deficit due to division of attention between memory and perception. *Journal of Experimental Psychology*, 1971, 88, 171-176.
- SPECTOR, A., LAUGHERY, K. R., & FINKELMAN, D. G. Rehearsal

- and organization in intentional forgetting. *Journal of Experimental Psychology*, 1973, **98**, 169-174.
- STERNBERG, S. High-speed scanning in human memory. *Science*, 1966, **153**, 652-654.
- STERNBERG, S. Two operations in character recognition: Some evidence from reaction-time measurements. *Perception & Psychophysics*, 1967, **2**, 45-53.
- STERNBERG, S. Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist*, 1969, **57**, 421-457. (a)
- STERNBERG, S. The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), *Attention and performance II*. *Acta Psychologica*, 1969, **30**, 276-315. (b)
- THEIOS, J., SMITH, P. G., HAVILAND, S. E., TRAUPMANN, J., & MOY, M. C. Memory scanning as a serial self-terminating process. *Journal of Experimental Psychology*, 1973, **97**, 323-336.
- TOWNSEND, J. T. A note on the identifiability of parallel and serial processes. *Perception & Psychophysics*, 1971, **10**, 161-163.
- WATTENBARGER, E. L., & PACHELLA, R. G. The effect of memory load on reaction time in character classification. *Perception & Psychophysics*, 1972, **12**, 100-102.

(Received for publication July 31, 1974;
revision accepted October 31, 1974.)