JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR 10, 522-527 (1971)

# Memory Scanning for Words Versus Categories[1]

JAMES F. JUOLA AND R. C. ATKINSON

*Stanford University, Stanford, California 94305*

Two groups of *S*s compared a target word with a memory set consisting of from one to four words (Group W) or from one to four semantic categories (Group C). The *S*s made a positive or a negative response to indicate whether or not a target word matched one of the words in the memory set for Group W, or whether or not the target word was an exemplar of one of the categories in the memory set for Group C. Reaction times for negative responses were linear functions of the memory set size for both groups, but the slope of the function for Group C was about four times the slope for Group W. The results were discussed in terms of alternative memory search mechanisms and the possible serial and parallel scanning models that were consistent with the data.

Landauer and Freedman (1968) studied information retrieval from long-term memory in an experiment designed to test the effects of category size on classification time. The *S*s were shown single words and identified them as belonging (positive response) or not belonging (negative response) to well-known semantic categories. It was shown that latencies for both positive and negative responses were greater for large categories than for small ones.

Two of the explanations offered for the data are presented in terms of the possible memory search strategies employed in making a decision. One possibility is that *S*s search through the items in a stored list of words that correspond to members of the presented category. The decision time is then dependent on the time it takes to search through all the words, in the case of a negative decision, or at least until the tested word is found, in the case of a positive decision. A second possibility is that each stimulus word has stored with it a list of attributes which includes the names of categories to which it belongs. The decision time then would include the time involved in arriving at an appropriate superordinate category name for the stimulus word, with the names being searched hierarchically from smallest to largest. Collins and Quillian (1970) have argued in favor of this latter hypothesis in that they generally failed to replicate the Landauer and Freedman results unless the categories were nested (i.e., the smaller categories were subsets of the larger categories).

It is possible to test the implications of these two processing systems in a memory search paradigm. Sternberg (1966) has shown that when a small number of digits (the memory set) is presented for *S* to remember and then a single digit (the target item) is presented as a probe, the time for *S* to decide whether or not the target item is a member of the memory set is a linear, increasing function of the number of items in memory. Further, the slopes of the reaction-time (RT) functions for positive and negative responses are equal, indicating that the extent of the memory search is the same for both types of trials (an exhaustive search), rather than ending with the location of the target item on positive trials (a self-terminating search). In terms of the model proposed by Sternberg, the slope of the RT function represents the time involved in making a single comparison between the target

[1] The writers acknowledge the assistance of Christine Wood and W. Weeks in collecting and analyzing data. Thanks are also due Bennet B. Murdock, Jr., Edward Smith, and Roberta L. Klatzky for comments on an earlier version of this paper.

item and any item in the memory set. The intercept of the RT function includes the time necessary to encode the test item and arrive at an internal representation that allows comparison with the memory items, as well as the time necessary to make a response decision and output that response.

In the present study, the memory sets consisted of from one to four words (Group W) or categories (Group C), and the target items were always words. On positive trials the target word was the same as one of the words in the memory set for Group W, or it was an exemplar of one of the categories in the memory set for Group C. If *S*x search through the exemplars of the memory set categories in seeking a match for the target word, then the slope of the RT function for Group C will be greater than that for Group W, but the intercepts should be about the same. If, however, *S*s retrieve the name of the category that includes the target word and compare that category name with those in the memory set, then the slope of the RT functions for both groups should be about the same, but the intercept for Group C should be greater than that for Group W.

## METHOD

The *S*s were 30 right-handed female students from Stanford University who participated to satisfy part of their requirements for a course in introductory psychology. Half of the *S*s were randomly assigned to each group.

Ten categories were selected from the Battig and Montague (1969) norms. The categories were: Metals, Mammals, Colors, Body Parts, Tools, Clothing, Birds, Vehicles, Insects, and Trees. These were among the largest categories (in terms of the numbers of exemplars listed) and were not obviously related or nested in any way. For each of these categories, 12 words were selected from the words most frequently given as responses to the category name. These words were all three to eight letters in length, with seven one-syllable and five two-syllable words used in every category. Each word was typed in capital letters on a 5 × 8-in. index card with an IBM Registry typewriter. The target words were shown singly using an Iconix tachistoscope and exposure box (System 153).

On a table to the right of *S* three telegraph keys were arranged along an arc with each key separated from the adjacent key by 3 cm. The *S* could comfortably rest her right forefinger on the center key between trials and could make a short, natural movement to the right or left to strike either of the two response keys. The *S*s were randomly assigned to depress either the right or the left key to indicate a match between the target word and one of the words or categories in the memory set (positive response), and to depress the other key if no match occurred (negative response).

At the start of the experimental session *S* was given ten 3 × 5-in. index cards. Each card had the name of one of the categories typed at the top, with the 12 exemplars typed in a column below the name. The *S* was instructed to read through each of the categories to be certain to be familiar with the words used.

The *S* was then seated at the tachistoscope and was told that the following sequence of events would occur on each trial: (*a*) The *E* would place an index card with a single word typed on it into the tachistoscope. (*b*) The *E* would read a list of from one to four items (consisting of words for Group W and names for Group C). (*c*) The *S* would repeat the memory set items aloud and then press a button in her left hand. (*d*) One-half sec later the display would be illuminated for 0.5 sec. (*e*) The *S* would then make a positive or negative response by depressing the appropriate key. (*f*) After an interval of approximately 10 sec the next trial would begin.

The test session consisted of 120 trials, divided into five consecutive 24-trial blocks. The first block of trials (Trials 1–24) was regarded as a warm-up block, and the data are not included in the analyses. Every word was presented as a test item exactly once. Each block contained six presentations of each of the four memory set sizes. The items in the memory set were selected randomly from the set of category names for *S*s in Group C and were selected randomly from the pool of 120 words for *S*s in Group W, with the constraints that half the trials would be positive and half negative for each memory set size within each block. The order of the memory set was further constrained for positive trials to insure that the target item was represented approximately equally often at each serial position within the memory set.

The *S*s were instructed to respond as rapidly as possible on every trial while at the same time being careful to be correct. No feedback was given for correct responses, but *S*s were informed of all errors. The test session averaged 40 min.

## RESULTS

Mean latencies for correct positive and negative responses for the four memory set

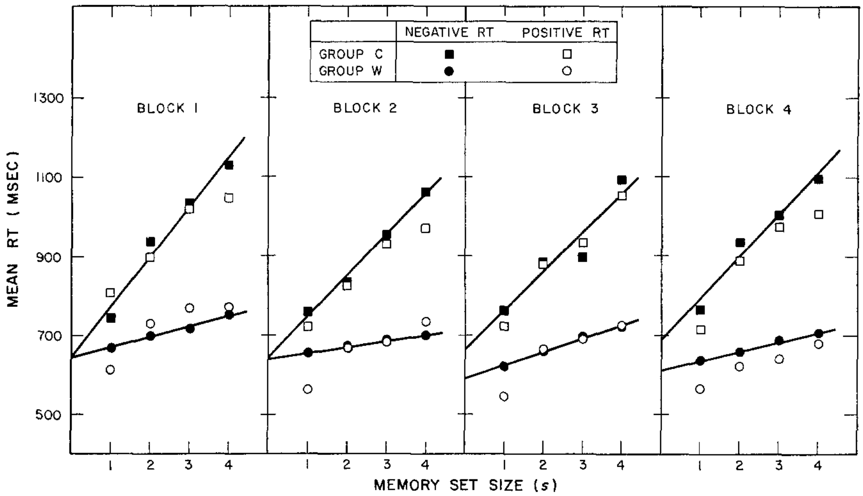


FIG. 1. Mean RT for positive and negative responses as functions of memory set size ($s$) for Group C (categories) and Group W (words) for four consecutive trial blocks.

sizes were obtained for each *S*. The means across *S*s for trial blocks 1–4 (excluding the warm-up block) are presented in Figure 1. The straight-line functions are least-squares fits to the data for negative responses. Note that the slopes and intercepts of the RT functions remain fairly constant across trial blocks.

The left panel of Figure 2 presents mean latency for the data from trial blocks 1–4 combined. Again straight lines were fit to the data for negative responses by the method of least squares. The slope of the RT function for Group C was 110.7 msec, and the slope for Group W was 26.3 msec. This difference be-

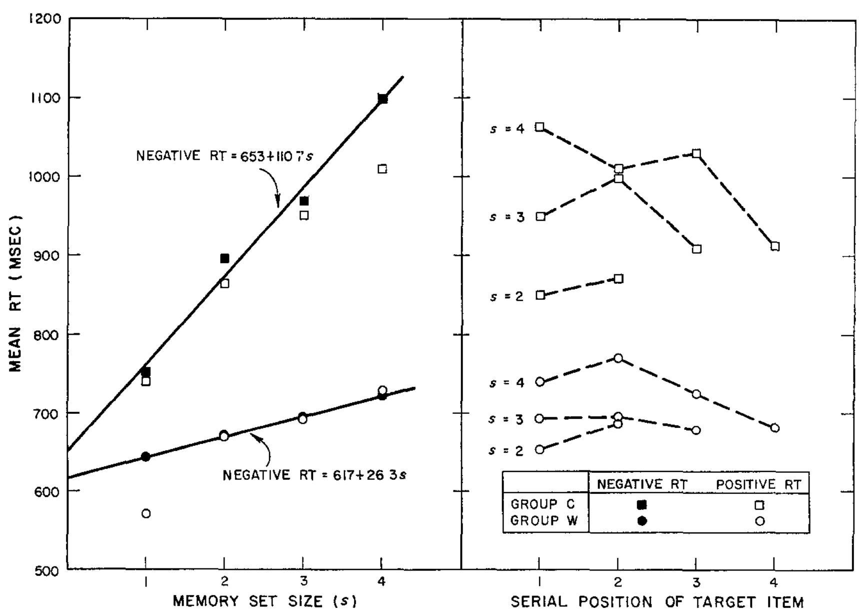


FIG. 2. Mean RT for positive and negative responses as functions of memory set size ($s$) (left panel). Mean RT for positive responses as a function of the serial position of the target item in the memory set (right panel).

tween slopes for the two groups was highly significant, $t(13) = 6.12$, $p < 0.001$. The intercept of the RT function for Group C was 653 msec and the intercept for Group W was 617 msec, but this difference was not significant, $t(13) = 0.96$. Straight lines were not fit to the data for positive responses since there were marked departures from linearity in both groups. The right panel of Figure 2 presents mean positive response latency as a function of the serial position of the target item in the memory set. The mean percentage errors for Group C and Group W as a function of set size are shown in Table 1.

TABLE 1

MEAN PERCENTAGE ERRORS FOR POSITIVE AND NEGATIVE RESPONSES AS A FUNCTION OF MEMORY SET SIZE FOR GROUP C AND GROUP W

| | Memory set size | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| Group C | | | | |
| Positives | 1.1 | 7.9 | 7.5 | 9.4 |
| Negatives | 0.6 | 4.6 | 2.9 | 6.3 |
| Group W | | | | |
| Positives | 0.0 | 1.1 | 1.7 | 0.0 |
| Negatives | 0.6 | 0.0 | 0.6 | 0.6 |

## DISCUSSION

The results presented in the left panel of Figure 2 generally conform to the predictions of the serial and exhaustive scanning model proposed by Sternberg (1966). However, there are some aspects of the data which are inconsistent with this model, that is, the nonlinearities in the data for positive responses and the marked recency effects in the serial position curves (right panel of Figure 2). It is necessary to make some additional assumptions about the memory-scanning processes involved in order to account for the finer details of the data. (A discussion of this problem is available from the authors on request.)

It is possible to examine the results in light of the processing systems proposed by Landauer and Freedman (1968). If, on negative trials, *S*s in Group W compare the target word with all of the words in the memory set, then the slope of the RT function (26.3 msec) represents the average time necessary to make a single comparison. If *S*s in Group C convert the test word to its appropriate superordinate category name and then compare this name with all of the names in the memory set, the slope of the RT function (110.7 msec) again represents the average time for a single comparison. Since both groups would then be comparing an internal representation of a target word against those contained in the memory set, the expectation is that the slopes for the two groups should be similar, if not equal.[2] Further, since *S*s in Group C would have to convert the target word into a superordinate category name before they could begin the comparison process, the intercept of the RT function for Group C should be higher than that of Group W [at least about 75 msec higher, which is the amount of time Collins and Quillian (1969) estimate is necessary to arrive at a superset name given a word as an exemplar]. These predictions clearly are not upheld by the data. The slope for Group C is more than four times the slope of the negative RT function for Group W, and although the intercept of the Group C function is 36 msec higher than that of Group W, this difference was not significant.

It should be noted that there are alternative representations of the memory search process that could account for slope differences between Group C and Group W even though the average number of items scanned for the two groups are equal. An example of such a process is discussed by Sternberg (1969). He

[2] Additional *S*s were tested with the category names used both for the memory set and for the target items; that is, the conditions were the same as for Group W, only the category names were used instead of the exemplar words. The data for these *S*s were comparable to those for Group W in the present paper.

describes a task in which *S*s either are able to maintain the memory set in "active" memory by rehearsing it until the target item is presented, or are momentarily prevented from this rehearsal by the introduction of an intervening task between the presentation of the memory set and the target item. While *S*s are attending to the intervening task, the memory set is not rehearsed, and it passes into "inactive" memory until the target item is presented. The results showed that if the memory items are not in active memory at the time of test, both the slope and the intercept are greater than those obtained when rehearsal is allowed in the interval prior to the presentation of the target item. In the present study a similar situation might occur if *S*s in Group C convert the target word into its superordinate category name and scan this name against those in memory. If the memory set is lost from active memory during this conversion process, then, in terms of Sternberg's explanation, *S* must locate the memory set in inactive memory and serially transfer the items into active memory before the scan for the target item can begin. The addition of these stages to the memory search process for Group C could account for the types of differences observed between the RT functions for the two groups.

Alternatively, if *S*s in Group C generate lists of exemplars of each of the categories in the memory set and then compare the target word with these exemplars in seeking a match, the slope of the negative RT function should represent the time necessary for comparing the target word with the average number of exemplars generated for each category. Since this comparison process could begin as soon after the presentation of the target word as for Group W, the intercepts of the RT functions for the two groups should be about the same. With the slope ratio of the negative RT function of Group C to that of Group W being about 4 to 1 it might be concluded that *S*s in Group C do not compare the target word with all of the possible exemplars of a given category. If the exemplars must be retrieved from memory and then be compared with the target word, both of these stages would add to the slope of the RT function. If the comparison process is identical for both groups, it would appear that for each category in the memory set no more than about two or three exemplars are retrieved and then compared with the target word. In this case, an alternative decision rule must be proposed for Group C other than the rule that a positive response is made if and only if there is a direct match between the target item and one of the items in memory. One such rule could be that *S*s arrive at an index of "association value" [or semantic similarity in the Schaeffer and Wallace (1969) sense] between the target word and the category exemplars, and if this value exceeds a certain criterion a positive response is made, otherwise a negative response is emitted. Such a decision process would account for the fact that the error rate was higher in Group C than it was in Group W.

Neither of the proposed mechanisms seems to provide an adequate representation of the memory processes involved in determining whether or not a given word is a member of one of a set of specific semantic categories. However, with certain additional assumptions about the search and decision aspects of the processes involved, either of the two models may be used to provide an account for our data. That is, the slope differences between the RT functions for Group C and Group W can be predicted by a memory scanning model with *S*s in Group C either comparing the target word with words in lists of category exemplars or searching for the name of the category which includes the target word in the names of the memory set.

## References

Battig, W. F., & Montague, W. E. Category names for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monograph*, 1969, **80** (3, Part 2), 1–46.

Collins, A. M., & Quillian, M. R. Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 240–247.

Collins, A. M., & Quillian, M. R. Does category size affect categorization time? *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 432–438.

Landauer, T. K., & Freedman, J. L. Information retrieval from long-term memory: Category size and recognition time. *Journal of Verbal Learning and Verbal Behavior*, 1968, **7**, 291–295.

Schaeffer, B., & Wallace, R. Semantic similarity and the comparison of word meanings. *Journal of Experimental Psychology*, 1969, **82**, 343–346.

Sternberg, S. High-speed scanning in human memory. *Science*, 1966, **153**, 652–654.

Sternberg, S. Memory scanning: Mental processes revealed by reaction-time experiments. *American Scientist*, 1969, **57**, 421–457.

(Received March 20, 1971)