Subsequent recognition of items subjected to proactive interference in short-term memory*

M. T. TURVEY†, DONALD, L. MOSHER, and LEONARD KATZ
University of Connecticut, Storrs, Conn. 06268

Each of 90 Ss was given five successive short-term memory (STM) tests with a triad of country names as the memory material on each test. Recall performance deteriorated across the series of STM tests. Immediately subsequent to the tests, each S was given a recognition test consisting of the 15 country names presented in the STM test series and 30 "new" country names. Each S was required to indicate whether a word was "old" or "new" and to give a confidence rating of his judgment. Operating characteristics for each of the five STM tests were computed for recalled and nonrecalled words. The data suggest that decline in recall performance across a series of STM tests (proactive interference) is due to several factors: an increasing inability to discriminate among items in storage and an increasing failure of items to gain access to storage or to stay in storage.

Recall declines across successive short-term memory (STM) tests of the Brown-Peterson variety (Brown, 1958; Peterson & Peterson, 1959). Although this observation (e.g., Keppel & Underwood, 1962) of recall decreasing as a function of prior items suggests the operation of proactive interference (PI) in STM situations, the mechanism by which PI operates is as yet poorly understood (see Posner, 1967). One source of difficulty lies in the STM test itself. Current theorizing has pointed to the fact that data obtained from STM test situations are never pure indicators of the processes of the hypothesized short-term storage (STS) mechanism (Waugh & Norman, 1965; Atkinson & Shiffrin, 1968). The probability of recalling an item in an STM test is given by the probability that an item can be retrieved from primary or STS, secondary or long-term storage (LTS), or both. Atkinson and Shiffrin have even suggested that STS may play a negligible role in STM performance, with the prime source of recall being a rehearsal buffer.

The present experiment sought to assess by means of a recognition procedure the information S has about items recently presented to him in a STM test series. A recognition test given after the series of STM tests would, it was assumed, provide data on the registration of STM test items in LTS. Of interest was the comparison between the recognition probabilities for those items which were recalled and those items which were not recalled in the STM test series. Essentially, the present experiment looks at the question: Why is an item not recalled on an STM test? Is it because the trace is no longer available in memory, STS or LTS? Or are such traces intact but temporarily inaccessible because of interference from other traces? If the former, then recognition of nonrecalled items should be minimal. If the latter, then presumably recognition of nonrecalled items should be near perfect. In any event, it was hoped that the present experiment would shed some light on the fate of items "forgotten" in an STM test series.

SUBJECTS

The Ss were 90 undergraduate males and females from the University of Connecticut who were enrolled in introductory psychology and who participated in the experiment to meet a course requirement. The Ss were seen individually in the experiment, which employed a distractor STM paradigm similar to that of Peterson & Peterson (1959) and a recognition task.

MATERIALS

The names of 45 countries were selected from Cohen, Bousfield, & Whitmarsh's (1957) cultural norms. Fifteen of these countries were presented in the five STM tests in word triads. The triads were as follows: Japan, Spain, Turkey; China, Ireland, Italy; France, Britain, Pakistan; India, Chile, Germany; Tibet, Canada, Poland.

PROCEDURE

Word triads were presented for five STM tests. The triads were projected onto a screen, using a Kodak Carousel projector with a tape timer to control presentation rates. The time parameters were as follows: an asterisk which served as a ready signal was presented for 2 sec; the triad of words was presented for 2 sec, and S was required to read the triad aloud; the retention interval lasted 18 sec and was filled by presenting nine two-digit numbers at 2-sec intervals. S was required to add the two digits and to classify their sum as odd or even; a question mark served as the cue for a 10-sec recall period; the ready signal then announced the beginning of the next trial.

Immediately following the fifth STM test, Ss were given a recognition task with the following instructions: "Listed below are 45 countries which include the 15 countries presented in the short-term memory task. In the blank space by the name of the country, write Y for "Yes" to indicate if that country was included in the STM task and write N for "No" if the country was not included in the STM task. In addition, please indicate your confidence in your decision on a scale from 1 to 5 where 1 means CERTAIN and 5 means GUESS. Rate every country."

The word triads were counterbalanced across the five STM tests. That counterbalancing was integrated with the counterbalancing of three different random lists of the 45 countries used for the recognition task. Approximately 160 sec elapsed between the ready signal of STM Test 1 and the recognition test.

RESULTS AND DISCUSSION

Each S was scored on each test for the number of words correctly recalled without regard for order. Figure 1 represents the mean number of words recalled as a function of STM test. Inspection of Fig. 1 clearly indicates that the PI function was obtained.

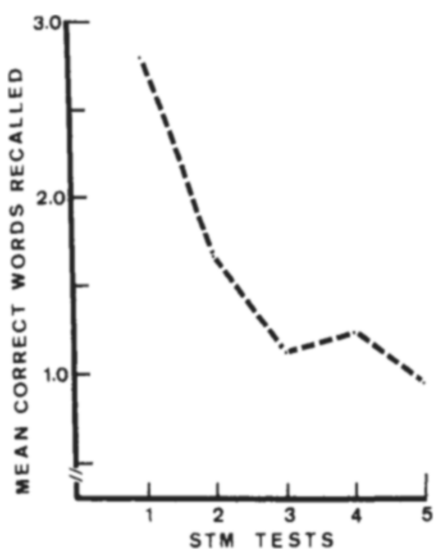

Fig. 1. Mean number of words correctly recalled as a function of STM test.

*This research was supported in part by Grants 102 and 114 from the University of Connecticut Research Foundation to the first and second authors, respectively. The authors would like to thank Mrs. Edie Sullivan for a careful job of data collection.

†Reprint requests should be sent to: M. T. Turvey, Department of Psychology, University of Connecticut, Storrs, Conn. 06268.

Table 1
Probability of Recognition for Recalled and Nonrecalled Items as a Function of STM Test

	STM Test				
	1	2	3	4	5
$P(r/R)$	0.92	0.94	0.96	0.97	0.98
$P(r/\bar{R})$	0.68	0.66	0.58	0.72	0.76

The conditional probabilities for recognizing (r) a recalled (R) and a nonrecalled ($\bar{R}$) item as a function of STM test were computed from the recognition test data. These probabilities are given in Table 1. The probability of a false positive was .09.

The recognition data were analyzed within the context of the theory of signal detection (TSD). Frequencies in each response-certainty category, 1, 2, . . . , 10 (i.e., from high certainty that an item was old to high certainty that an item was new), were aggregated across Ss for each of three types of items: (1) old items that had been recalled in a previous STM test, (2) old items that had not been recalled on a previous STM test, and (3) new items. The two old-item distributions were further partitioned into response-certainty frequencies for each STM test (T1, T2, . . . , T5). Some cells appeared to have insufficient data, so frequencies for all response-certainty categories were pooled as follows: 1 + 2, 3 + 4 + 5, 6 + 7 + 8, 9 + 10, to produce four (transformed) categories.

Following the techniques of Egan, Schulman, & Greenberg (1959), an operating characteristic for each of the five STM tests was computed for recalled and nonrecalled old items, all relative to the same "false-alarm" distribution. Figure 2 presents the 10 operating characteristics. Inspection of Table 1 and Fig. 2 suggests that a difference, or differences, existed between the trace of a recalled item and the trace of a nonrecalled item. This conclusion is based on two features of the recognition data. First, although $P(r/R)$ was nearly equal to one, $P(r/\bar{R})$ was considerably less than one. Averaged across the five STM tests, $P(r/\bar{R}) = .68$. There is a suggestion, derived from inspection of the operating characteristics in Fig. 2, that this .68 value for $P(r/\bar{R})$ is somewhat inflated. Clearly, Fig. 2 shows that items which were recalled had greater strength in memory at the time of the recognition test, i.e., greater strength in LTS, than items which had not been recalled. However, most of the items which were not recalled appeared to have positive strength, with the possible exception of items from T3. Thus, for the nonrecalled items, some information was retained, even though it was insufficient for recall in the STM test series.

Second, and perhaps of greater interest, are the relative strengths of the two types of items, recalled and nonrecalled, as a function of STM test position. In each of the two sets of curves displayed in Fig. 2, items from T4 and T5 are generally stronger than items from T1 and T2. However, items from T3 have different relative strengths in the two sets. In the set of recalled items, the strength of an item appears to increase with recency from T1 to T5. However, in the nonrecalled set of items, T3 possesses near-zero strength and is clearly the most weakly recognized of all the nonrecalled items. This suggests that subsequent items are the important source of forgetting for items which have been recalled but that traces of items which have not been recalled are susceptible to strong effects of both prior and subsequent items.

How do the present data speak to the question of why an item is not recalled on a given test in a series of STM tests? The observation that most of the nonrecalled items had some strength in LTS, as determined by the recognition procedure, implicates retrieval failure as a major source of forgetting in a STM test series. Retrieval failure that occurs due to prior items can be viewed as a decreased capacity to differentiate among stored items on one basis or another. In a series of STM tests of the sort evident in the present experiment, discrimination depends on recency alone, and thus with tests presented in fairly rapid succession, performance deteriorates.

Other sources of forgetting in addition to inaccessibility, or confusion at recall, are implied by the fate of the T3 items. Decay induced by time and/or the action of other traces might be considered as such a source (see Posner, 1967). However, if one can assume that an item which is recalled probably enters LTS, or perhaps was already in LTS at the moment of recall, and that an item which was not recalled is probably still in STS, then the present data can be interpreted as follows: Entry into the large- or infinite-capacity LTS from the limited-capacity STS is determined by rehearsal. Once rehearsal of an item is stopped, however, the trace of that item decreases with each succeeding item (see Atkinson & Shiffrin, 1968). This assumption accounts for the superiority of items from the later tests over items from the earlier tests in the two sets of curves of Fig. 2. Moreoever, the similarity between the two sets of curves, i.e., T5 and T4 items are better recognized than T2 and T1 items, when viewed in this

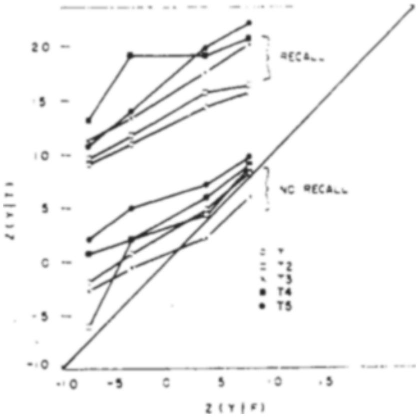

Fig. 2. Operating characteristics for the recalled and nonrecalled words.

light, asserts that the nonrecalled items like the recalled items received some rehearsal. Why then are the nonrecalled set of T3 items so poorly registered in LTS?

Given that STS is of limited capacity, later items may fail to gain access to storage because of items already there, or if they do enter storage, they are likely to be dislodged by subsequent items (Phillips, Atkinson, & Shiffrin, 1967; Waugh & Norman, 1965). Thus, the nonrecalled set of T3 items were susceptible to both prior and subsequent inputs and consequently were less likely to enter the rehearsal loop.

In short, the present data imply that the deterioration in recall performance across a series of STM tests, known as PI buildup, is due to increasing difficulty in discerning and retrieving the correct items from STS or LTS and failure of later items either to secure space in STS, or having secured space, failure of such items to remain in STS long enough to benefit substantially from rehearsal.

REFERENCES

ATKINSON, R. C., & SHIFFRIN, R. M. Human memory: A proposed system and its control processes. In K. W. Spence and J. T. Spence (Eds.), *Psychology of learning and motivation: Advances in research and theory.* Vol. 2. New York: Academic Press, 1968.

BROWN, J. Some tests of the decay theory of immediate memory. Quarterly Journal of Experimental Psychology, 1958, 10, 12-21.

COHEN, B. H., BOUSFIELD, W. A., & WHITMARSH, G. A. Cultural norms for verbal items in 43 categories. Technical Report No. 22, University of Connecticut, 1957.

EGAN, J. P., SCHULMAN, A. I., & GREENBERG, G. Z. Operating characteristics determined by binary decisions and by ratings. Journal of the Acoustical Society of America, 1959, 31, 768-773.

KEPPEL, G. K., & UNDERWOOD, B. J. Proactive inhibition in short-term retention of single items. Journal of Verbal Learning & Verbal Behavior, 1962, 1, 153-161.

PETERSON, L. R., & PETERSON, M. J. Short-term retention of individual verbal items. Journal of Experimental Psychology, 1959, 58, 193-198.
PHILLIPS, J. L., SHIFFRIN, R. M., & ATKINSON, R. C. Effects of list length on short-term memory. Journal of Verbal Learning & Verbal Behavior, 1967, 6, 303-311.
POSNER, M. I. Short-term memory systems in human information processing. Acta Psychologica, 1967, 27, 267-284.
WAUGH, N. C., & NORMAN, D. A. Primary memory. Psychological Review, 1965, 72, 89-104.