

Short-Term Memory for Words with a Perceptual-Motor Interpolated Activity¹

ROBERT G. CROWDER²

University of Michigan, Ann Arbor, Michigan

It was maintained that the Brown-Peterson short-term memory (STM) task bears important similarities to procedures used in the study of division of attention, since during STM retention intervals there is competition between the tendencies to rehearse the memory item and to execute the filler activity. It follows that when both retention and filler-task performance are scored, either or both should be sensitive to variables known to affect the division of attention, such as task-complexity and practice. In two experiments recall for 5-word stimuli was tested after 24-sec intervals filled with a self-paced keypressing task as the interpolated activity. Evidence for the presence of rehearsal was the finding that keypressing scores were lower on memory trials than on control trials where no words were recalled. Furthermore, retention scores were found to be a complex joint function of the S-R compatibility, coherence (repetitiveness), and prior practice on the keypressing activity. The conclusions were that rehearsal is objectively demonstrable in STM, that its extent may be controlled by variation of the filler task, and that the relationships revealed by such variation are consistent with the experimental literature on divided attention.

A distinguishing feature of the short-term memory (STM) procedures introduced by Brown (1958) and by Peterson and Peterson (1959) is the interpolation of a filler task between presentation and recall. It is often stated that the function of this inter-

polated activity is to prevent rehearsal, yet a systematic dependence of STM scores on filler-task difficulty has not yet been established for the Brown-Peterson technique. Presumably the complexity of the filler task would correlate negatively with rehearsal and hence with recall. This kind of reasoning serves to place STM experiments in the context of research on the division of attention, where S must perform two tasks at once with excellence on one purchased usually at the expense of the other. The general purpose of the present research was the exploitation of this analogy between the Brown-Peterson method and divided-attention studies, with two analytic strategies: (a) control of retention performance by manipulations of filler-task difficulty, and (b) examination of the effects of memory load upon performance of the interpolated activity. The former approach is

¹ This article is based on a dissertation submitted to the graduate school of the University of Michigan in partial fulfillment of the requirements for a doctoral degree. The author is greatly indebted to Arthur W. Melton, who directed the research, as well as to Ronald S. Tikofsky, Irwin Pollack, and the late Paul M. Fitts, for their advice and encouragement. J. Kirby Thomas deserves thanks for his careful assistance in the data collection of Exp. II.

This research was supported by the Advanced Research Projects Agency and was monitored by the Air Force Office of Scientific Research, under Contract Number AF 49 (638)-1235 with the Human Performance Center of the Department of Psychology, University of Michigan.

² Now at Yale University.

similar to that of Posner and Rossman (1965) while the latter is particularly consistent with experiments on the division of attention (Barrick, Noble, and Fitts, 1954). An important aspect of research of this kind is that it requires accurate performance measures on the filler task.

Posner and Rossman (1965) varied the difficulty of numerical transforms their Ss carried out while remembering digit pairs, and they found an orderly dependence of recall upon the complexity of the transforms. Their conclusion was that simple filler operations leave more information-processing capacity free for rehearsal than difficult operations. Since, however, both their memory task and interpolated activity involved digits, it could be argued that the amount of specific retroactive interference (RI) increased with the complexity of their filler task. In the present study, the danger of such confounding was minimized by the use of a perceptual-motor filler task.

The experiment of Barrick *et al.*, (1954) is typical of research on division of attention in human performance. Their Ss performed a serial keypressing task in conjunction with continuous mental arithmetic. As expected, keypressing scores were shown to suffer when arithmetic was carried out at the same time. A more interesting finding was that pretraining on a redundant version of the perceptual-motor task led to improved arithmetic scores while pretraining on a nonredundant version did not. The authors concluded that "automatization" of one task can release spare mental capacity for investment in a second task only if the first task is repetitively structured.

In the research to be reported below, S worked on a serial keypressing task while remembering (for later recall) strings of words. All independent variation was intended to affect the difficulty of the keypressing task directly and thus indirectly control the level of recall. In Exp. I

the independent variables were stimulus-response compatibility (difficulty of light-key pairings), coherence (sequential redundancy of keypressing stimuli), and pretraining (amount of practice on keypressing before it was used as the STM filler task). Experiment II was an effort to verify a tentative finding of Exp. I.

EXPERIMENT I

Method

Experimental Design. Two levels of S-R compatibility, two levels of coherence, and three levels of pretraining were combined in a $2 \times 2 \times 3$ factorial, between-Ss design. Three Ss served in each of the 12 treatment combinations, yielding a total N of 36.

Subjects. The Ss were male University of Michigan undergraduates who had volunteered, at the time of registration, to serve in psychological experiments for pay.

Memory Materials. Three standard blocks of 10 memory items were developed in pilot investigations. Individual Ss in Exp. I saw only two of these blocks, but across the experiment as a whole all were used equally and were balanced across experimental conditions. Each stimulus was photographed on a 2×2 in.-slide and contained five four-letter English nouns which occurred with frequency greater than 13/million in the Thorndike-Lorge (1944) count. Formal and meaningful similarity among words within a stimulus were minimized.

Apparatus. The soundproof cubicle in which the experiment was conducted had been arranged so that S saw only (a) a translucent screen upon which the memory stimuli were back-projected, (b) a display of 10 jewel lights, and (c) a keyboard with 10 piano-type keys. The spatial locations of the 10 elements in the keyboard and light display were identical and suggested (by two semicircles) the positions of fingertips in two relaxed hands. A Kodak Carousel projector was modified to change slides on reception of an electric pulse; its operation, the timing of warning and recall lights, and the timing of keypressing periods were all governed by a series of four variable-interval timers. The E, however, had control over the critical interval during which S perceived the stimulus words and immediately after which the first keypressing stimulus was initiated. The keypressing unit was

automatically self-paced by a pair of stepping switches such that if and only if *S* pressed the correct key in response to a stimulus light would that light go out and another appear (with a machine cycle time of .125 sec). Keypressing scores could thus be expressed as the number of lights extinguished in a fixed interval of time.

Independent Variables. The high- and low-compatibility conditions may be described by considering the lights and keys both numbered 1-10 from left to right. With high compatibility key 1 extinguished light 1, 2 extinguished 2, etc. With low compatibility, keys and lights were paired, respectively, as follows: 1-8, 2-6, 3-2, 4-1, 5-10, 6-3, 7-9, 8-5, 9-4, and 10-7. Coherence was varied by presenting repetitive light sequences to some *Ss* and non-repetitive sequences to others. High-coherence groups saw the following cycle of 11 stimuli throughout the entire experiment (lights numbered as above, from left to right): 8, 5, 3, 9, 7, 1, 6, 10, 2, 7, and 4. On any given trial this cycle began with the position following the last correct response on the previous trial. Low-coherence groups received unpredictable sequences, although the apparatus did not permit programming of technically-random series. Three different amounts of keypress practice were allowed before collection of the critical data on the combined STM-keypressing task; Group 10 had ten 24-sec trials and Groups 60 and 110 had 60 and 110 such trials, respectively.

Dependent Variables. The important index of interpolated-task performance was not the absolute level of keypressing speed, but rather the decrement in keypressing occasioned by the simultaneous storage of words. Thus, blocks of trials on the combined task were always surrounded by blocks of undivided keypressing so as to permit calculation of decrements unconfounded by practice effects. The index of recall was based on Kendall's *P* (Kendall, 1955), a nonparametric measure of rank correlation. *P* is calculated by counting the number or proportion of ordered pairs of elements in the stimulus which are preserved in the response, thus providing a joint summary of recall for the order and identity of stimulus elements.

Procedure. Each *S* participated in three tasks during the experiment, (a) STM with backward-counting as filler (pretest used later to match *Ss* and as a baseline for evaluating recall with keypressing as filler), (b) STM with keypressing filler, and (c) keypressing alone. In task (c) each trial was initiated by a brief flash from a warning light, then, after 2 sec the illumination

of one of the 10 stimulus lights. For the next 24 sec *S* extinguished the lights as rapidly as possible under self-pacing; he had available at all times a large diagram illustrating the light-key relationships under his compatibility condition. After the 24-sec keypressing period, a continuous light lit the screen for 8 sec and then went off for 12 sec more. During the latter interval, *E* gave feedback (number of lights extinguished) on keypressing performance. The warning light then began another trial. When keypressing was used as an interpolated activity (task b) the same sequence of events obtained, with two exceptions. Two sec after the warning flash the memory stimulus was exposed on the screen. The *S* was instructed to read off the words aloud, whereupon the slide was removed and keypressing initiated. The 8-sec period immediately after keypressing was used by *S* to recall the stimulus words. The retention interval was thus always 24 sec. Work on all tasks was divided into blocks of 10 trials, with about 60-sec rest periods between blocks. No more than five such blocks were given in a single session; therefore, after the pretest, Group 10 served one day, Group 60 served two days, and Group 110 served three days.

Results

In order to show that the independent variables were effective in producing differences in between-group performance on keypressing alone, keypressing scores were analyzed from the last block of pretraining trials. Displayed in Fig. 1 are the mean number of correct presses during the 24-sec periods, as a function of pretraining with compatibility and coherence as parameters. Each point is based on the mean for three *Ss* over 10 trials. A $2 \times 2 \times 3$ factorial analysis of variance showed all main effects and interactions to be statistically significant at beyond the .01 level of confidence, except for the Compatibility \times Pretraining interaction (.05 level). The one aspect of these data which merits comment is that under low compatibility, the effects of coherence were not apparent until after 110 trials of pretraining.

Keypressing during the Combined Task.

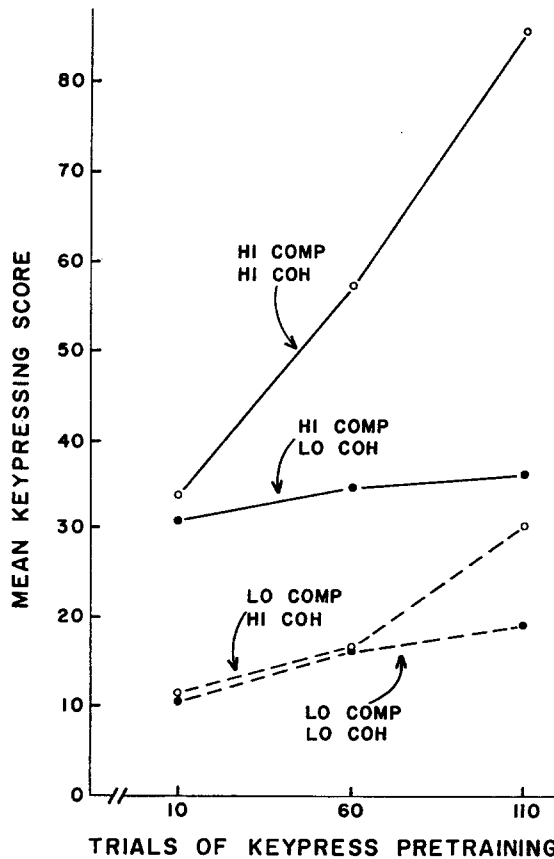


FIG. 1. Relationship between keypress pretraining and mean keypressing score prior to combined-task performance. The parameter is treatment-combination of compatibility and coherence. Each point is based on performance over 10 trials for independent groups of 3 Ss (Exp. I).

Table 1 summarizes the effect of a memory load upon keypressing speed. Mean percent keypress decrement was computed for each S by subtracting scores during the combined task from scores on keypressing alone and expressing the absolute difference as a percentage of the latter. For 34 of the 36 Ss, there was a positive difference which indicates that a memory load reduced keypressing rates. The analysis of variance showed only the main effect of compatibility to be statistically significant, $F(1, 24) = 9.34, p < .01$; the memory task disrupted keypressing more under low than high compatibility.

Recall during the Combined Task. No S failed to achieve higher recall with the perceptual-motor filler task than with the counting filler task used in the pretest. Therefore, difference scores were used to evaluate recall during the combined task; this procedure offered much the same controls on inter-S variation as covariance analysis. Figure 2 shows the mean corrected recall scores as functions of the various treatment combinations. Factorial analysis of variance revealed Compatibility to be the only significant main effect, $F(1, 24) = 6.87, p < .025$, with better recall under high than under low compatibility. The

TABLE 1
KEYPRESSING ON FINAL DAY

Compatibility	Trials of keypress pretraining	Low coherence			High coherence		
		Keypress alone	Keypress + STM	Percent decrement	Keypress alone	Keypress + STM	Percent decrement
Low	10	11.8	10.5	7.2	12.8	11.3	11.5
	60	17.9	15.1	13.7	19.6	18.5	9.8
	110	20.6	18.8	8.9	31.6	28.3	10.2
High	10	31.8	30.5	4.3	38.2	35.8	6.3
	60	35.1	34.2	2.7	62.1	61.0	1.7
	110	36.0	34.7	3.5	85.4	73.7	14.0

only other statistically significant effect was the three-way (Compatibility \times Coherence \times Pretraining) interaction, $F(2, 24) = 3.60, p < .05$. Inspection of Fig. 2 permits the following description of this interaction: With low coherence, high-compatibility Ss gave consistently high recall, whereas

low-compatibility Ss were initially low in recall but improved markedly at the highest level of pretraining. With high coherence, there were decrements in recall at the intermediate and advanced levels of keypress pretraining for high and low compatibility, respectively. In the case of the

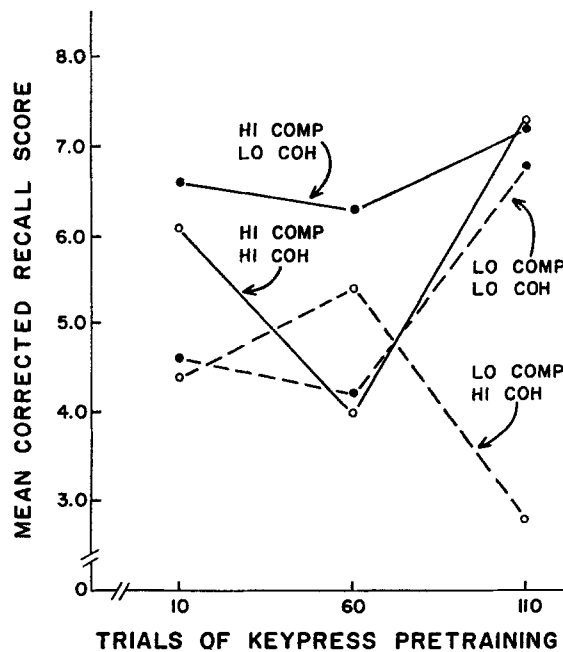


FIG. 2. Relationship between mean corrected recall scores and trials of keypress pretraining. The parameter is treatment-combination of compatibility and coherence. Each point is based on performance over 10 trials for independent groups of 3 Ss (Exp. I).

high-compatibility, high-coherence conditions, there was a recovery in recall at the advanced level of pretraining.

EXPERIMENT II

Although the effect of stimulus-response compatibility upon the experimental tasks was unequivocal, in Exp. I, the relationships involving coherence and pretraining suggested by Fig. 2 were statistically reliable only through their participation in the three-way interaction. The second study was designed to verify the apparent nonmonotonic effect upon recall, in Exp. I, of pretraining on a redundant keypressing task. The tentative interpretation placed on recall performance for high-coherence groups in the first study was that with high compatibility recall was related by a *U*-shaped function to pretraining and would have been also, for low-compatibility Ss, except that insufficient trials were given to reveal the right arm of the *U*. In light of the data of Fig. 1, it is entirely reasonable that the hypothesized *U*-shaped function should appear later for low- than for high-compatibility groups, since those data show that the effects of coherence were not felt until after 60 trials by low-compatibility Ss. In other words, the imposition of a difficult light-key pairing had, it was assumed, delayed the formation of serial associations for groups with a redundant task; consequently, the functional relationship between redundancy and recall was also delayed. To be tested in the present study was the implication that if high-coherence, low-compatibility groups were run with a wider range of pretraining levels, retention performance would display the full *U*-shaped function.

Experimental Design. All Ss received the same low-compatibility, high-coherence task used in Exp. I. Three groups of eight Ss each received 10, 110, or 210 trials of keypress pretraining before being tested on the combined task.

Procedure. The three groups were matched on the basis of a pretest employing 39 Ss. The final day, for all groups, consisted of 10 trials of keypressing alone, 10 trials on the combined task, and then a final block of five keypressing-alone trials. The extra pretraining given Groups 110 and 210 was accomplished over two or four additional days, respectively. In all other respects the details of Exp. II were identical with those of Exp. I.

Results

The 10 keypressing trials surrounding the block of combined-task trials were examined to verify choice of appropriate practice levels. The mean scores over 24-sec intervals for the three groups were 9.5, 24.5, and 49.2, with virtually no overlap of distributions about these values. Uncorrected *P* scores showed the mean levels of recall to be 6.4, 5.8, and 7.5 for Groups 10, 110, and 210, respectively. A one-way, repeated-measures analysis of variance showed these differences to fall just short of statistical significance, $F(2, 14) = 3.65$, with a value of 3.74 required at the .05 level. However, since exactly the obtained form of nonmonotonic relation had been predicted, the results of Exp. II were not considered unsupportative of the hypothesized *U*-shaped function.

Another finding of Exp. II deserves mention, although it pertains more to the research problem as a whole than to the specific purpose of the study. For each of 39 Ss pretested on the combined STM-keypressing task, the regression coefficient was computed for predicting memory scores from keypressing scores on individual trials. Two Ss had coefficients of zero by dint of perfect recall on each trial, 16 had positive coefficients, and 21 had negative coefficients; this division between negative and positive regression was well within chance expectations. Thus, knowledge of keypressing performance on a given trial was of no predictive value with respect to retention on that trial.

DISCUSSION

STM and the Division of Attention

The present research confirmed the similarity between the Brown-Peterson STM technique and the traditional methods of studying division of attention. The most fundamental generalization these latter methods have yielded is that a given task suffers a decrement when it must be executed at the same time as another, unrelated, task (Bahrick *et al.*, 1954; Binet, 1890; Brown, 1962; McQueen, 1917; Shouten, Kalsbeek, and Leopold, 1962). In Exps. I and II, 53 out of the 60 Ss made fewer correct keypressing responses during STM trials than during trials on keypressing alone. This result may be accepted as direct evidence for covert (and in the present case illegal by the terms of the instructions) rehearsal during filled retention intervals. There was, however, no consistent negative correlation between STM and keypressing on combined-task trials. This dependence of one task on the *presence or absence* of another, but independence of their relative levels of performance was unexpected for it had been hoped to demonstrate a reciprocity between rehearsal and the filler task on individual trials. It is of substantial interest that McQueen (1917) reported the same situation in an early experiment on divided attention. In the case of the present studies, *E* noticed a tendency on the part of some Ss to fuse the keypressing and memory tasks by whispering words in time with keypresses. It was not possible to determine how widespread this strategy was, but future research may show it to be most prevalent among Ss who show positive regression between the two tasks.

A second major empirical generalization about divided attention is that performance decrements increase with the complexity of the simultaneous tasks (Brown

and Poulton, 1964; Hylan, 1903; Jastrow, 1891-92; Shouten, Kalsbeek, and Leopold, 1962). This generalization was supported in Exp. I by the significantly greater keypress decrements under low than under high compatibility. Finally, certain experiments have been concerned with reductions in intertask disruption through practice or pretraining on one of the tasks (Bahrick *et al.*, 1954; Bahrick and Shelley, 1958; Brown and Poulton, 1962). There was no evidence that pretraining affected keypress decrements in the present experiments, although portions of the retention data showed effects of keypress pretraining.

Retention as a Function of the Filler Task

The perceptual-motor activity used in the present research turned out to be a generally poor method of controlling rehearsal when compared with those verbal tasks typically used in STM studies (counting backwards, digit reading, color naming). Though the memory stimuli were more difficult and the retention interval longer in the present study than those used by Peterson and Peterson (1959), the Petersons showed substantially greater forgetting in their study. Even the highly-compatible digit-naming task used by Hellyer (1962) was apparently more effective in discouraging rehearsal than the keypressing activity used here. It may be that retroactive interference from any *verbal* filler task plays a more important role in short-term forgetting than has previously been supposed.

Permissive as the interpolated activity was, the objective of controlling recall through filler-task difficulty was nonetheless realized unequivocally in Exp. I where recall was significantly better under high than under low compatibility keypressing. This finding directly confirms those of Posner and Rossman (1965) and Pillsbury

and Sylvester (1940). Moreover, since a nonverbal filler task was used, (a) there was little chance for variation in RI from the interpolated activity, and (b) the result implies a single information-processing channel serving both verbal and nonverbal operations.

The effects of coherence and pretraining upon recall were internally consistent but of only marginal significance statistically; therefore only tentative interpretations may be entertained. There seemed to be two important departures from the results of Bahrck *et al.*, (1954). In the first place Ss with low-compatibility, *non-repetitive* tasks showed improved recall at 110 trials of pretraining as compared with Ss given less pretraining. Although Bahrck *et al.*, (1954) found no such indirect practice effects with nonredundant tasks, their keypressing task was, for all Ss, maximally compatible and therefore probably placed minimal demands on the central capacity.

Secondly, Exps. I and II suggested that intermediate amounts of pretraining on repetitive tasks lead to decreased recall, or, at a more fundamental level, to decreased ability to rehearse. (Joint consideration of the two studies reported permits relating the term "intermediate" legitimately to stage of mastery rather than to a specific number of pretraining trials.) In a theoretical analysis of perceptual-motor skill learning, Fitts (1964) has characterized the intermediate stage of skill acquisition as one where the basic associations to the task are being formed, in the presence of mediation, anticipation, trial-and-error, verbalization, and rehearsal. Surely such behavior would be of disastrous consequences to any simultaneous task, especially if the latter were verbal. It is therefore suggested that in the present data, the observed decrements in recall for Ss with repetitive keypressing tasks may be associated with this critical phase of

skill learning. It is from the final phase of Fitts' analysis that the principle of automatization derives its validity. At this stage, all necessary associations are formed and gradual improvements in performance are due to the streamlining of mediational processes and, eventually, to the dropping-out of such mediators (as must occur, for example, in tying shoelaces over a period of years). The ultimate improvement in recall for Ss with the largest amounts of pretraining, in the present data, may indicate entry into this final phase of learning on the keypressing task. The failure of Bahrck *et al.*, (1954) to find a nonmonotonic relation between arithmetic performance and keypress pretraining may tentatively be ascribed to their choice of only two practice levels at which to introduce the arithmetic task in conjunction with keypressing. It is interesting to speculate that if these authors had happened to choose early (and then intermediate) levels of pretraining, they would have found secondary-task performance to *decrease* with primary-task practice, and the notion of automatization might summarily have been rejected. Indeed, such a possibility may account for the failure of some investigators (Bahrck and Shelley, 1958) to find clear-cut practice effects at all in divided-attention situations.

REFERENCES

- BAHRICK, H. P., NOBLE, M., AND FITTS, P. M. Extratask performance as a measure of learning a primary task. *J. exp. Psychol.*, 1954, 48, 298-302.
- BAHRICK, H. P., AND SHELLEY, C. Time sharing as an index of automatization. *J. exp. Psychol.*, 1958, 56, 288-293.
- BINET, A. La concurrence des états psychologiques. *Rev. Phil.*, 1890, XXIX.
- BROWN, I. D. Measuring the "spare mental capacity" of car drivers by a subsidiary auditory task. *Ergonomics*, 1962, 5, 247-250.
- BROWN, I. D., AND POULTON, E. C. Measuring the "spare mental capacity" of car drivers by a subsidiary task. *Ergonomics*, 1961, 4, 35-40.

- BROWN, J. Some tests of the decay theory of immediate memory. *Quart. J. exp. Psychol.*, 1958, **10**, 12-21.
- FITTS, P. M. Skill learning. In A. W. Melton (Ed.), *Categories of human learning*. New York: Academic Press, 1964. Pp. 243-285.
- HELLYER, S. Supplementary report: Frequency of stimulus presentation and short-term decrement in recall. *J. exp. Psychol.*, 1962, **64**, 650.
- HYLAN, J. P. The distribution of attention. *Psychol. Rev.*, 1903, **10**, 313-403, 498-533.
- JASTROW, J. The interference of mental processes. *Amer. J. Psychol.*, 1891-92, **4**, 219-223.
- KENDALL, M. G. *Rank correlation methods* (2nd. Ed.), London: Griffin, 1955.
- MCQUEEN, E. N. The distribution of attention. *Brit. J. Psychol., Monog. Supp.*, 1917, 1919, **2**, No. 5.
- PETERSON, L. R., AND PETERSON, M. J. Short-term retention of individual verbal items. *J. exp. Psychol.*, 1959, **58**, 193-198.
- PILLSBURY, W. B., AND SYLVESTER, A. Retroactive and proactive inhibition in immediate memory. *J. exp. Psychol.*, 1940, **27**, 532-545.
- POSNER, M. I., AND ROSSMAN, E. The effects of size and location of informational transforms upon short-term retention. *J. exp. Psychol.*, 1965, **70**, 496-505.
- SHOUTEN, J. F., KALSBECK, J. W. H., AND LEOPOLD, F. F. On the evaluation of perceptual and mental load. *Ergonomics*, 1962, **5**, 252-260.
- THORNDIKE, E. L., AND LORGE, I. *The Teacher's Word Book of 30,000 Words*. New York: Columbia Univ. Press, 1944.

(Received February 7, 1966)