Journal of Experimental Psychology:
Learning, Memory, and Cognition
1982, Vol. 8, No. 1, 66–72


0096-1515/82/0801-0066$00.75

# Notes, Comments, and New Findings

## Hypermnesia: The Role of Repeated Testing

Henry L. Roediger III and David G. Payne
Purdue University

The present study was designed to determine whether the increased recall of pictures across repeated tests (hypermnesia) is due to increasing strength of imaginal traces during the retention interval or to increased retrieval practice from prior tests. Subjects studied 60 pictures and then recalled them after various delays that were filled with instructions and, in two cases, reading a passage. Recall on a first test showed no change with retention interval. With retention interval held constant, however, the number of pictures recalled varied directly with the number of prior tests subjects had been given. This finding points up the critical nature of retrieval factors in producing hypermnesia.

When people are given repeated memory tests following the presentation of some material, a typical finding is that material not recalled on an earlier test will be recalled on later tests (e.g., Brown, 1923; Tulving, 1967). This recovery is usually offset by the forgetting of other material between tests, so that the net effect is typically one of no change in overall recall, or slight forgetting. However, in an important series of studies Erdelyi and his colleagues (Erdelyi & Becker, 1974; Erdelyi, Finkelstein, Herrell, Miller, & Thomas, 1976) demonstrated that under certain conditions people actually show increases in recall with repeated tests. They referred to this phenomenon as hypermnesia. The conditions that most easily elicit the phenomenon are those in which people are given either a long series of pictures to recall or a list of concrete words with instructions to form images of the referents of the words. Repeated tests of words studied without imagery instructions did not lead to improved recall in the original studies (Erdelyi & Becker, 1974). It may be the case that hypermnesia is somehow caused by information being represented in an imaginal code, as is argued by Erdelyi and his associates, but demonstrations of the phenomenon under conditions in which use of imagery seems unlikely have lessened the force of this conclusion (Belmore, 1981; Erdelyi, Buschke, & Finkelstein, 1977; Roediger & Thorpe, 1978).

The reasons for hypermnesia are not yet clear, either theoretically or in terms of the experimental variables that affect the increases in recall on later tests. The present study addresses the role of repeated testing in producing hypermnesia. One possibility is that hypermnesia reflects the growth in strength or availability of some underlying memory trace over time. Recall may thus improve because of some continuing consolidation of traces, with such a process possibly occurring more extensively for imaginal than nonimaginal traces. A second hypothesis is that hypermnesia depends on retrieval practice from repeated testing. Accessibility of information on a later test may be greater than that on an earlier test because the earlier test permitted more efficient organization and retrieval of recalled material. On the later test the already recalled material can be retrieved again more quickly, with time remaining for the recall of new material (Erdelyi & Becker, 1974; Madigan & Lawrence, 1980). By this hypothesis, better recall after long rather than short retention intervals depends on the interval being filled with overt tests of retention or with internal review of the material.

These ideas about improvements in recall with repeated tests have been aired previously in the literature on reminiscence, a term that also refers to improvements in performance across tests (see McGeoch & Irion, 1952, chap. 5). The experimental design suggested by McGeoch and Irion (1952, p. 141) and Ammons and Irion (1954) to control for practice effects of the first test is simply to present two groups of people with the same material and then test each group once. One group is tested soon after learning and the other group after a delay. (This design was originally advocated by Ward in 1937.) True reminiscence was considered to have been demonstrated only if performance was better after the delay than immediately. Put in the terms used previously, such a

Requests for reprints should be sent to Henry L. Roediger III, Department of Psychology, Purdue University, West Lafayette, Indiana 47907.

finding would indicate that the reminiscence or hypermnesia found with repeated tests is probably due to increases in availability of information rather than an increase in accessibility due to "practice" from repeated testing. Using this type of design, Ammons and Irion found no improvements in recall on a single test of passages of poetry learned by school children over delays of 2 and 7 days. They concluded that the reminiscence phenomenon originally discovered by Ballard (1913) with a repeated testing procedure was probably due to practice on the first test aiding performance on later tests.

Shapiro and Erdelyi (1974) used a similar design to question whether hypermnesia in picture recall could occur independently of potential practice effects arising from use of multiple recall tests. They varied type of material and retention interval by presenting people with either 60 pictures or 60 words and then testing them once, either after a delay of 30 sec or 5 min. Both variables were manipulated between subjects. During the retention interval subjects were instructed to review the material covertly. Shapiro and Erdely reported that people who studied pictures performed better after 5 min. than after 30 sec, whereas there was no reliable difference between the two tests for people who studied words. Taken at face value, these results indicate that recall increases over a retention interval without practice due to tests, but of course just such practice was encouraged by Shapiro and Erdelyi during the retention interval. The question remains as to whether hypermnesia will occur in recall of pictures when attempts are made to minimize any attempt to rehearse the material, either by overt tests or internal review, during the retention interval.

The purpose of the present study was to examine simultaneously the effects of increased retention interval and repeated testing on recall of pictures under conditions that minimized the possibility of covert review. Retention interval and the number of prior tests were naturally confounded in all previous studies using multiple tests. They can be disentangled by providing subjects with repeated tests but by beginning the tests at various delays from the time the material is presented. Performance can then be compared across groups that have received varying numbers of recall tests but are still tested after a constant delay.

We used three groups of subjects who all studied the same set of 60 pictures and were then tested three times. In the immediate test condition, the first test was given after a brief set of instructions intended to reduce recency effects, whereas the remaining groups had their first tests delayed by various periods of time. So that subjects in these delay conditions would not rehearse the material during the retention interval, they were required to read a prose passage. We selected this interpolated task because we thought a verbal task would be unlikely to interfere much with memory for pictorial materials but at the same time would prevent conscious rehearsal of the pictorial information. One of the delay groups (short delay) read the passage for 9 min., was given instructions, and began taking the first test after the same time interval following presentation of the pictures that the immediate test group had prior to taking its second test. Subjects in the long delay group read for 18 min., were given instructions, and received their first test when the immediate group began their third test and the short delay group began their second test. This design (illustrated in Table 1) allows examination of changes in recall over time with the number of prior memory tests held constant, as did the design of the Shapiro and Erdelyi (1974) study. In this case, however, the time was filled with an interpolated task. The other novel feature of this design is that it permits examination of the effects of number of prior tests with retention interval held constant.

## Method

### *Subjects*

Subjects were 120 undergraduates, participating to fulfill part of a course requirement in introductory psychology.

### *Design*

Three groups of 40 subjects were presented with drawings of common objects. Following list presentation, one group was given instructions and tested immediately (immediate test condition), the second group read for 9 min. before being tested (short delay condition), and the third group read for 18 min. before being tested (long delay condition).

### *Materials*

Sketches of 60 easily recognizable objects (e.g., airplane, arrow, lamp, rifle) were prepared as slides.

### *Procedure*

Subjects were randomly assigned in groups to the immediate, short delay, and long delay conditions and were tested in groups of 4 to 18. Subjects were instructed that they would see a series of 60 slides containing pictures of objects and that they should study each slide carefully because following presentation of the slides their memory for the pictures would be tested. They were told that they would be required to recall the names of the pictured objects. Following these instructions the slides were presented at a 5-sec rate.

After the slides were presented, the subjects were given instructions appropriate to their condition. The subjects in the immediate test condition were told that their task would be to write the names of the objects on their test sheets. They were also told that after each minute of the test had passed they would be asked to draw a line below the last item they had recalled. Subjects were informed that they would have 7 min. for the recall test. They were also told that although the recall period might seem quite long, they should keep trying to recall items throughout the entire time. It took 2 min. to read these instructions, answer questions, and distribute the test sheets. This delay was intended to reduce any recency effects. The first recall test began immediately following these instructions.

The subjects in the short delay and long delay conditions were given an article from the *American Scientist* (Premack, 1976) and asked to read the article carefully, as if they were to be tested on it later. Subjects in the short delay condition read for 9 min., and subjects in the long delay condition read for 18 min. Following the reading interval the subjects from both the short and long delay conditions were treated in exactly the same fashion as the subjects in the immediate test condition were treated following the presentation of the slides. They were given 2 min. of recall instructions and then the first recall test.

Thus, for the subjects in the immediate test condition, testing began 2 min. after the last slide had been presented. Subjects in the short and long delay conditions began their first test 11 and 20 min., respectively, after the presentation of the final slide, at the same times that the immediate test group received their second and third tests (see Table 1).

When the first 7-min. test period had expired, there was a 2-min. period during which the recall sheets were collected and new blank sheets distributed. The subjects were also given a second set of instructions similar to those for the first test. Subjects were again explicitly encouraged to keep trying even when they thought they had recalled all that they were able to. They were told that they should try to recall both previously recalled pictures and any new items they could remember.

Following this 2-min. period, there was a second 7-min. test. When this recall period was over, there was another 2-min. period during which the second test sheets were collected, new test sheets distributed, and instructions similar to those for the second test read. Subjects were also told that this was the last test. This 2-min. period was followed by the third and final 7-min. test.

## Results

The mean number of items recalled on each test for the three conditions is presented in Table 1. There are three main findings of interest. First, hypermnesia was obtained in each of the conditions. There was an increase in recall across every pair of successive tests. The amount of hypermnesia obtained was similar in the three conditions, with subjects recalling an average of about five more items on the third test than on the first test. Second, there was no increase in recall with delay of recall from initial presentation. This can be seen by examining recall in the upper left to lower right diagonals of Table 1. On the first test mean recall was 25.6, 25.1, and 25.6 items in the immediate, short delay, and long delay conditions, respectively. On the second test the comparable means for the three conditions were 27.9, 27.5, and 28.9, and on the third test they were 30.1, 29.8, and 31.3. Thus, so long as the number of prior tests is held constant, there is no evidence that recall increases as the retention interval is increased.

The third main point to be gleaned from Table 1 is the critical role of retrieval practice in producing hypermnesia. When the retention interval is held constant but the number of prior tests increases, recall increases. This can be seen by looking up the columns of Table 1, from bottom to top. The clearest case is that with three observations. With the retention interval held constant at 20 min., subjects taking their first test (long delay) recalled an average of 25.6 items, subjects taking their second test (short delay) recalled 27.5 items, and subjects taking their third test (immediate) recalled 30.1 items. The same pattern can also be seen in the columns to the left and right of the center column.

The results of a 3 × 3 (Delay Condition × Test) analysis of variance supported the first two of these observations. (All effects are reliable at the .01 level of confidence unless otherwise noted.) There was a significant main effect only for test, $F(2, 234) = 208.4$, $MS_e = 3.53$, indicating that recall increased across the three tests. There was no effect of increasing the retention interval on recall, as neither the main effect of delay ($F < 1$, $MS_e = 105.2$) nor the Delay × Test interaction, $F(4, 234) = 1.34$, $MS_e = 3.53$, approached significance. A one-way analysis of variance on the

Table 1
*Mean Recall on the Three Tests for Each Delay Condition*

| Condition | Items recalled | | | | |
|---|---|---|---|---|---|
| Immediate | Test 1<br>25.6 | Test 2<br>27.9 | Test 3<br>30.1 | | |
| Short delay | | Test 1<br>25.1 | Test 2<br>27.5 | Test 3<br>29.8 | |
| Long delay | | | Test 1<br>25.6 | Test 2<br>28.9 | Test 3<br>31.3 |

means in the middle column of Table 1 (Test 1, long delay; Test 2, short delay; Test 3, immediate) indicated reliable variation among these conditions, $F(2, 117) = 5.38$, $MS_e = 37.9$.

Presented in Figure 1 are cumulative recall curves for subjects in each of the three conditions. The results presented in Table 1 represent the total (7-min.) recall. The curves in Figure 1 allow examination of changes in the pattern seen in Table 1 across the 7-min. recall period. Several conclusions can be drawn from Figure 1. First, analyses of recall at each minute of the recall period suggest the same conclusion as that drawn from recall at the end of 7 min.: The increased recall shown from successive tests is due to repeated testing and not to the increasing retention interval per se. Second, increases in recall with repeated testing are somewhat greater at intermediate times (3–5 min.) than at the end of the entire 7-min. period, replicating an observation of Roediger and Thorpe (1978, Experiment 1). Third, even on the third test, recall had not yet reached asymptote at the end of the 7-min. recall period, indicating that subjects would likely have recalled more items if given additional tests or simply more time (Erdelyi & Kleinbard, 1978; Roediger & Thorpe, 1978).

A $3 \times 3 \times 7$ (Delay Condition $\times$ Test $\times$ Minute) analysis of variance revealed that the pattern of recall did not differ across the three delay conditions. There were reliable main effects of Test, $F(2, 234) = 208.4$, $MS_e = .51$, and minute, $F(6, 702) = 1478.3$, $MS_e = 5.64$, but no reliable effect of delay condition ($F < 1$, $MS_e = 15.0$). The only other reliable effect was the Test $\times$ Minute interaction, $F(12, 1404) = 40.5$, $MS_e = 2.36$.

A final analysis examined the fate of specific items over the three successive tests. Presented in Table 2 are the mean values for intertest forgetting and recovery. Intertest forgetting refers to the situation in which items are correctly recalled on a first test and not recalled on a second test, or the CN (correctly recalled–not recalled) component of recall. Recovery, of the NC component,

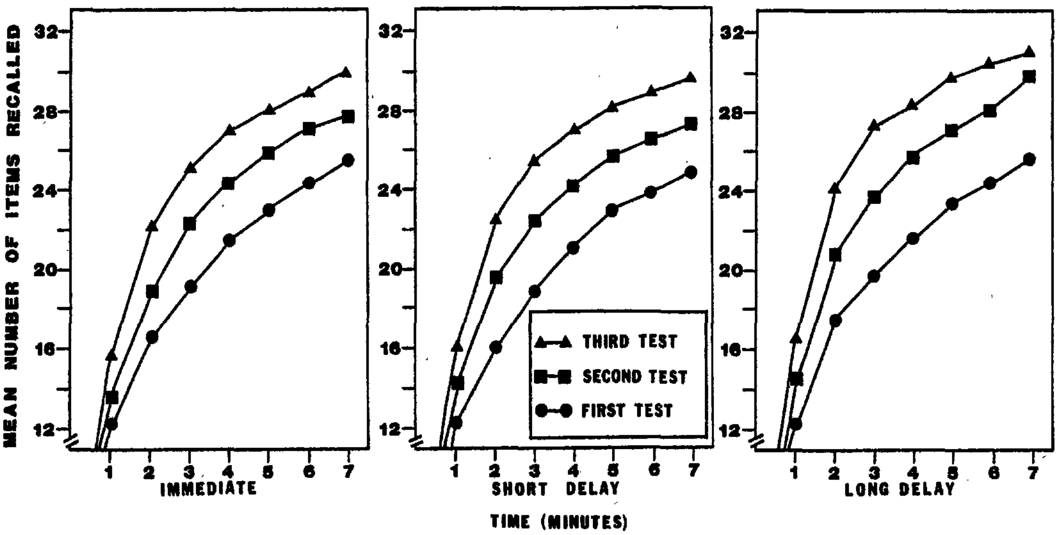


*Figure 1.* Mean cumulative recall in the three tests for all conditions.

Table 2
*Mean Number of Items Forgotten (CN) and Recovered (NC) Between Successive Tests*

| Condition | Intertest forgetting (CN) and recovery (NC) | Test trials | |
|---|---|---|---|
| | | 1–2 | 2–3 |
| Immediate | CN | 2.7 | 2.2 |
| | NC | 4.9 | 4.5 |
| Short delay | CN | 2.1 | 1.6 |
| | NC | 4.6 | 3.8 |
| Long delay | CN | 1.6 | 1.6 |
| | NC | 5.0 | 3.7 |

refers to the situation in which an item is not recalled on a first test and then is correctly recalled on a second test (Tulving, 1964).

As shown in Table 1, there was a larger fluctuation in the specific items recalled between Tests 1 and 2 than between Tests 2 and 3, $F(1, 78) = 17.26$, $MS_e = 2.37$. This is shown by the fact that both intertest forgetting (CN component) and recovery (NC component) were larger between Tests 1 and 2 than between Tests 2 and 3 in all three delay conditions. There was also a consistent trend in the amount of fluctuation in specific items recalled across delay conditions. Item fluctuation was greatest in the immediate test condition and least in the long delay condition. The results of the analysis of variance revealed reliable differences in item fluctuation among the three delay conditions, $F(2, 156) = 4.29$, $MS_e = 3.88$; $p < .05$. None of the interactions was reliable.

## Discussion

The purpose of the present study was to determine if the increased recall of pictures across repeated tests was due to some aspect of repeated practice from taking the tests or to an increasing strength of traces that occurred over time independently of overt or covert recall. The results provide an unambiguous answer: Increased recall across multiple tests in this situation is due to the effect of the earlier tests on the later tests. When recall of pictures was delayed by a relatively neutral reading task, performance did not improve across delays of 2 to 20 min. On the other hand, with the retention interval held constant at 20 min., recall on a 7-min. test was greatest if subjects had been given two prior tests, reduced if subjects had only been given one prior test, and poorest if the test was the subjects' first.

Shapiro and Erdelyi (1974) concluded that hypermnesia occurred independently of repeated testing because they found that recall of pictures on a single test was better after a 5-min. delay than after a 30-sec delay. In their experiment, however, subjects were encouraged to think about the material during the retention interval. This internal review may be thought of as self-testing, so that the subjects gave themselves extra practice. In the present experiment, subjects in the different conditions were prevented from covert review of the material by reading an article. Under these conditions, no improvement in recall was found. Taking these results together, we can conclude that in order to observe improvements in performance on a free-recall test by lengthening the interval between study and test, it is necessary for the material to be presented again to subjects in one way or another. Of course, performance would be improved if the experimenter presented the material again between study and test, but the improvement also occurs if the re-presentation occurs via overt or covert recall on the part of the subjects.

The current findings complement those reported by Roediger and Thorpe (1978). In Experiment 1 they presented people with either 50 pictures or words (names of the pictures). Half of the subjects receiving each type of material were given three successive 7-min. free-recall tests, and the other half were given a single 21-min. test. People who were given the three tests showed increased recall across tests (hypermnesia), though the effect was larger for people who studied pictures than for those who studied words. The finding of central interest was that when subjects given three tests were compared with subjects given one long test of equivalent duration, there was no difference in the total number of unique items recalled for either pictures or words. Thus the procedure of using multiple recall tests, which provides increases in recall, can be seen as simply extending the amount of available recall time (see also Madigan & Lawrence, 1980). The successive recall tests are not independent assessments of recall; rather, performance on the later tests depends on having the earlier tests, as was shown in the present experiment. Because of the similarity in recall of words over time between the groups given three tests and those given one long test, Roediger and Thorpe (1978) suggested that perhaps subjects given the long test were effectively retrieving the material in the same manner as the ones given the successive tests. That is, even the subjects given the long test were repeatedly retrieving (though not overtly recalling) already recalled items. If they were retrieving only new items, one would expect them to recall more words than subjects given three tests because they were not required to retrieve and write down items they had already recalled in a previous test.

How are we to account for hypermnesia theoretically? If hypermnesia is conceived as increasing recall with additional recall time, then in some sense almost any theory can account for it. Whatever processes are postulated to produce recall in the first place can simply be assumed to grind along inexorably with consistently diminishing returns. Roediger and Thorpe (1978) suggested that one way to think of hypermnesia was to relate it to similar increases in recall when external retrieval cues are presented. It is well known that, under appropriate conditions, providing people with external cues can greatly augment recall relative to free recall (e.g., Tulving & Pearlstone, 1966). Although hypermnesia occurs in a free-recall situation, we may assume that even here recall is guided by "cues" (information in the retrieval environment), even if these cues remain invisible to the experimenter.

A recent model of Raaijmakers and Shiffrin (1980) made a formal assumption concerning this process and satisfactorily accounted for the hypermnesia results of Roediger and Thorpe (1978). The model assumed a probabilistic search of information stored in an associative memory. Although complete discussion of the model is outside the scope of this short article, three points are worth making briefly. First, according to the model, increases in recall in the multiple test situation should only occur with repeated testing and not with the passage of time per se, which is just the finding of this experiment. Second, one of the basic mechanisms producing hypermnesia is that the "cue set" that people use to sample and recover memories over long test intervals changes depending on the items and context "sampled" as cues. Thus alternate retrieval routes are used and allow retrieval of previously unrecalled information. This may be seen as a formalization of the vague assumption given previously that changes in the invisible cues in the free-recall test environment lead to hypermnesia.

Third, the other process within the Raaijmakers and Shiffrin (1980) model that leads to hypermnesia is "incrementing," or the assumption that associations from the retrieval context to successfully recalled items are increased, as is (in effect) the strength of the item itself. This has the result of making a recalled item more likely to be retrieved again, so that across repeated tests the act of recall will make future recall of previously recalled items more probable. This factor was suggested previously as potentially important in determining hypermnesia (Erdelyi & Becker, 1974; Madigan & Lawrence, 1980; Roediger & Thorpe, 1978). Simulations varying parameters within the Raaijmakers and Shiffrin model show that either the incrementing or alternative retrieval route assumption is sufficient to produce increased recall across repeated tests. Absence of both processes in the multiple test situation produces no change in recall across repeated tests.

In the present experiment there was no evidence that recall of pictures increased over various filled retention intervals. However, these data point up the equally remarkable fact that there was no forgetting across retention intervals either. This observation reinforces the point made by Erdelyi and Kleinbard (1978) that an assumption of a single exponential function for forgetting from long-term memory is inaccurate. Some years ago Reitman (1971) addressed the issue of whether memories decay in short-term memory by asking subjects to remember small amounts of information over brief intervals while engaged in a distracting task. Her subjects remembered verbal materials and were distracted by detecting tones. She initially found no forgetting in this situation, an observation confirmed by Shiffrin (1973; but see Reitman, 1974).

The present study can be seen as a similar demonstration of lack of decay in long-term memory. The first test in the present experiment occurred after 2 min. in the immediate condition and after 20 min. in the long delay condition. Despite this tenfold increase in the retention interval, there was no forgetting between tests. The amount of forgetting in such a situation depends on the similarity between the processes involved in learning the material and those used in performing the distractor task. If the memory and distractor tasks tap different kinds of processing, little forgetting may be expected (Roediger, Knight, & Kantowitz, 1977). The reading task used in the present experiment apparently caused no interference in the retention of pictures but presumably prevented subjects from spending much time ruminating over their memories of the pictures (but see Roediger et al., 1977). If we assume that different processes mediate picture memory and reading, then reading might be expected to have no effect on picture recall. That was just the outcome obtained here.

## References

Ammons, H., & Irion, A. L. A note on the Ballard reminiscence phenomenon. *Journal of Experimental Psychology*, 1954, *48*, 184–186.

Ballard, P. B. Oblivescence and reminiscence. *British Journal of Psychology Monograph Supplements*, 1913, *1* (2).

Belmore, S. M. Imagery and semantic elaboration in hypermnesia for words. *Journal of Experimental Psychology: Human Learning and Memory*, 1981, *7*, 191–203.

Brown, W. To what extent is memory measured by a single recall? *Journal of Experimental Psychology*, 1923, *6*, 377–382.

Erdelyi, M. H., & Becker, J. Hypermnesia for pictures: Incremental memory for pictures but not words in multiple recall trials. *Cognitive Psychology,* 1974, *6,* 159–171.

Erdelyi, M. H., Buschke, H., & Finkelstein, S. Hypermnesia for Socratic stimuli: The growth of recall for an internally generated memory list abstracted from a series of riddles. *Memory & Cognition,* 1977, *5,* 283–286.

Erdelyi, M. H., Finkelstein, S., Herrell, N., Miller, B., & Thomas, J. Coding modality vs. input modality in hypermnesia: Is a rose a rose a rose? *Cognition,* 1976, *4,* 311–319.

Erdelyi, M. H., & Kleinbard, J. Has Ebbinghaus decayed with time?: The growth of recall (hypermnesia) over days. *Journal of Experimental Psychology: Human Learning and Memory,* 1978, *4,* 275–289.

Madigan, S., & Lawrence, V. Factors affecting item recovery and hypermnesia in free recall. *American Journal of Psychology,* 1980, *93,* 489–504.

McGeoch, J. A., & Irion, A. L. *The psychology of human learning.* New York: Longmans, Green, 1952.

Premack, D. Language and intelligence in ape and man. *American Scientist,* 1976, *64,* 674–683.

Raaijmakers, J. G. W., & Shiffrin, R. M. SAM: A theory of probabilistic search of associative memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 14). New York: Academic Press, 1980.

Reitman, J. S. Mechanisms of forgetting in short term memory. *Cognitive Psychology,* 1971, *2,* 185–195.

Reitman, J. S. Without surreptitious rehearsal, information in short-term memory decays. *Journal of Verbal Learning and Verbal Behavior,* 1974, *13,* 365–377.

Roediger, H. L., Knight, J. L., & Kantowitz, B. H. Inferring decay in short term memory: The issue of capacity. *Memory & Cognition,* 1977, *5,* 167–176.

Roediger, H. L., & Thorpe, L. A. The role of recall time in producing hypermnesia. *Memory & Cognition,* 1978, *6,* 296–305.

Shapiro, S. R., & Erdelyi, M. H. Hypermnesia for pictures but not words. *Journal of Experimental Psychology,* 1974, *103,* 1218–1219.

Shiffrin, R. M. Information persistence in short-term memory. *Journal of Experimental Psychology,* 1973, *100,* 39–49.

Tulving, E. Intratrial and intertrial retention: Notes toward a theory of free recall verbal learning. *Psychological Review,* 1964, *71,* 219–237.

Tulving, E. The effects of presentation and recall in free recall learning. *Journal of Verbal Learning and Verbal Behavior,* 1967, *6,* 175–184.

Tulving, E., & Pearlstone, Z. Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior,* 1966, *5,* 381–391.

Ward, L. B. Reminiscence and rote learning. *Psychological Monographs,* 1937, *49* (4, Whole No. 220).

Received April 9, 1981 ■