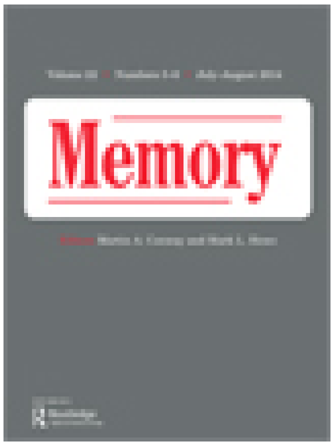

Memory
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pmem20

Different rates of forgetting following study versus test trials

Mark Wheeler [a], Michael Ewers [a] & Joseph Buonanno [a]
[a] Temple University, Philadelphia, PA, USA
Published online: 22 Sep 2010.

MEMORY, 2003, *11* (6), 571–580

Different rates of forgetting following study versus test trials

Mark A. Wheeler, Michael Ewers, and Joseph F. Buonanno
Temple University, Philadelphia, PA, USA

Two experiments investigated recall following two study conditions, (1) repeated test: a study trial followed by multiple recall trials, and (2) repeated study: multiple study trials with no tests. At a retention interval of 5 minutes, repeated study produced a higher level of recall than repeated test. When the retention interval was extended, forgetting was much more rapid in the study condition, with the repeated test manipulation leading to higher recall at an interval of 7 days. We conclude that study and test trials have different effects upon memory, with study trials promoting memory acquisition, and test trials enhancing the retrieval process itself, which protects against subsequent forgetting.

Although there is much cognitive research describing the types of encoding variables that enhance performance on subsequent recall and recognition tests, most experiments have utilised relatively short retention intervals, usually on the order of minutes. It is possible that, in some cases, different conclusions might have been drawn if the retention interval had been extended. In other words, different encoding conditions might produce different rates of forgetting. Perhaps surprisingly, not much is known about the way that forgetting interacts with other variables.

Historically, when the issue of forgetting has been addressed in cognitive psychology, it has typically been with the goal of mathematically describing the rate of forgetting over time. Of course this was the focus of the pioneering work of Ebbinghaus (1885/1964) who learned lists of nonsense syllables to a preset criterion, then gave himself a variety of recall tests for the lists at varying retention intervals. From these data, Ebbinghaus derived the first retention function, describing a monotonic function with a sharp loss in the first several minutes immediately following the learning trials, with a graduate levelling.

More than a century later, the field has made impressive progress in some aspects of forgetting research, with a dismal lack of progress in other aspects. On (perhaps) the bright side, there are now a lot of other proposed mathematical functions (or mathematical formulas, or theories) describing the rate of forgetting. In a review paper, Rubin and Wenzel (1996) described several two-parameter forgetting functions that had been suggested by others. After applying 105 different formulas to a large number of published sets of data, the authors suggested a few functions that appeared to best fit the available data.

While such an analysis (Rubin & Wenzel, 1996) surely makes a strong contribution in forgetting researcy by *describing* the way material is often forgetten across time, it is less satisfying in other ways. One question that has never been decisively resolved is whether there is a single retention function that can accurately depict the rate of forgetting across many different kinds of learning, or even across the different varieties of verbal learning that are commonly employed in psychology experiments. Put differently, is there even the possibility of a correct answer to the

Requests for reprints should be sent to Mark A. Wheeler, Department of Psychology, Weiss Hall, Temple University, Philadelphia, PA, 19122, USA. Email: mwheeler@nimbus.temple.edu
This article benefited from the comments of Robert Weisberg and Alan Sockloff.

http://www.tandf.co.uk/journals/pp/09658211.html DOI:10.1080/09658210244000414

question: what is the mathematical rate of forgetting?

When this question has been addressed, it has often been with respect to the possibility of an interaction between rate of forgetting and amount of original learning. Some researchers have claimed that the degree of learning does not affect the slope of the retention function (Anderson & Schooler, 1991; Slamecka & McElree, 1983), whereas others have argued that forgetting proceeds more slowly with relatively higher amounts of original learning (Loftus, 1985). Although the debate produced a few sharply worded exchanges, it never grabbed the attention of most memory researchers. Reasons for the relative lack of interest are unclear, but may stem from the difficulty in determining how to best measure the rate of forgetting (Wixted, 1990).

The possibility that there are different forgetting rates for different study, or encoding, conditions remains relevant to cognitive psychology, if mysteriously ignored. A few bits of relevant evidence exist here and there. Nelson and Vining (1978) addressed the possibility that there are different rates of forgetting for deep (or conceptual) versus shallow (or orthographic) processing. After manipulating encoding to ensure that the level of immediate learning was similar, the rate of forgetting was virtually identical between the two conditions. Hockley (1992), however, compared the forgetting rates for single words compared to that of associations between words, and concluded that item information is forgotten more quickly than associative information. Also, there appears to be an accelerated rate of forgetting in both young children (Howe & Hunter, 1986) and healthy older adults (Wheeler, 2000). Therefore, it is worth looking for further situations in which different kinds of learning lead to varying rates of forgetting across time.

One situation that may fit the foregoing description involves the use of retrieval practice during a learning, or study, phase. Numerous studies have shown that when participants are given a practice test shortly following encoding, there is a higher level of retention on a subsequent retrieval test, compared to participants who did not take an intermediate test (Gates, 1917; Runquist, 1983; Spitzer, 1939; Wheeler & Roediger, 1992). Additional research has shown that interpolated recall tests have a stronger effect on later retention than do recognition tests (Glover, 1989), and also that effects are stronger whe interpolated retrieval tests are spaced in time, as opposed to being massed (Landauer & Bjork, 1978; Rea & Modigliani, 1985).

The beneficial effects of testing on later retention are well documented and uncontroversial (if rarely applied outside the laboratory), yet other aspects of the phenomenon are puzzling, and also difficult to explain. It has been occasionally suggested that successful recall of an item also serves as an effective re-presentation of that item (Spitzer, 1939). By this logic, recalling a word on an interpolated test is similar to studying it an additional time, and follows the additional study will boost subsequent recall. Although this hypothesis is reasonable, data exist that suggest the situation is more complex.

A few experiments have addressed the idea that a test trial is equivalent, or at least very similar, to another study trial. Researchers have compared recall following two different encoding conditions, (1) repeated test: one study trial followed by multiple recall trials; and (2) repeated study: multiple study trials, with no test opportunities. After a short retention interval (on the order of minutes), repeated study commonly leads to higher levels of recall (Hogan & Kintsch, 1971; Thompson, Wenger, & Bartling, 1978, Experiments 1 and 2; but see Carrier & Pashler, 1992). After a retention interval of one or two days, however, there are often no retention differences between the two groups (Hogan & Kintsch, 1971), or an advantage for repeated test (Allen, Mahler, & Estes, 1969; Thompson et al., 1978, Experiment 3). The results imply that study and test trials are not interchangeable, and also suggest an interesting interaction between types of encoding and retention interval. There may be different rates of forgetting between the two encoding conditions.

If confirmed, the finding of differential forgetting rates following repeated study versus repeated test trials would represent a unique and important discovery. This research project is an experimental investigation of the idea that inclusion of recall trials into a study, or encoding, phase will result in a slower rate of forgetting than a similar encoding phase consisting solely of study trials. Participants studied a word list, and were then instructed to either study the same list three additional times (repeated study), or take three recall tests for the words (repeated test), before taking a final recall test at a short (5 min) or long (48 hr) retention interval. Because it is well known that elaborative study is beneficial for subsequent recall, we predicted that repeated study would lead to higher recall, at the 5-min interval. Based on previous comparisons of repeated study and repeated test trials, we also predicted that this advantage would disappear, or even be reversed, at the longer interval.

EXPERIMENT 1

Method

Participants and design. Participants were 60 Temple University students enrolled in undergraduate psychology courses, participating in exchange for partial course credit. The experiment used a 2 × 2 between-subjects design, with encoding condition (repeated test vs repeated study) and retention interval (5 min vs 48 hr) as independent variables. Fifteen participants were randomly assigned to each of the four conditions.

Materials. The stimulus materials comprised 40 common nouns. The words were selected so that they could not be easily grouped into semantic categories (i.e., animals, foods). The 40-item list was recorded onto a cassette tape four separate times, in four different random sequences. All participants listened to the first sequence, whereas those in the repeated-study conditions listened to all four sequences. The words used can be obtained from the authors.

Procedure. Participants were tested in groups of one to three. They were told that they were to listen to a list of words from a cassette player, and that they should attempt to memorise the words. The study list was presented at a rate of one word every 3 seconds.

Following the first presentation, instructions for the repeated study and the repeated test conditions diverged. Participants in the repeated test conditions were given a blank sheet of paper and told that they should write down as many of the words as they could recall from the list, in any order. They were also told that they would have 2 minutes, and they should keep trying to remember additional words throughout the test, even though they would reach a point at which further recall became difficult. The recall test began 1 minute following the end of the study list. Following the test, papers were collected and another blank sheet was given to each participant. The experimenter engaged the participants in some brief conversation, before instructing them to take another recall test for the list. They were told to write down as many words as they could remember, including words written on the previous test. The experimenter reminded them that they would have 2 minutes, and that they should keep trying to remember additional words throughout the entire interval. The recall test began 1 minute following the end of the previous test. Following this second test, the process was repeated, and participants took a third recall test, lasting for 2 minutes, beginning 1 minute after the end of the second test. After the third test, participants completed a personality questionnaire, which was not scored; its purpose was to prevent further rehearsal of the words, and also to fill the short retention interval.

Five minutes after the end of the third recall test, participants in the 5-min delay conditions were asked to take another recall test for the study list, under identical instructions as the previous tests. Participants in the 48-hr delay conditions were asked to write down the instructions that were read to them before they studied the words, and also before each of the test trials. The purpose of this task was to create a plausible cover story for the participants who were not tested immediately, so they would be less likely to expect a test on the material upon their return. After these data were collected, all participants were dismissed, and also reminded to return after 48 hours.

Following the first presentation of the words, participants in the repeated study conditions were told that they would hear the same 40 words again. They were instructed to continue trying to memorise the words. One minute after the first list had ended, participants heard the second list; again at a rate of one word per 3 seconds, in a different random order than the first presentation. Following this presentation, the experimenter informed subjects that they would hear the words yet another time, and that they should continue trying to memorise the words, even though they were already familiar with all of the words. Participants heard the words a third time, and then the words were repeated a fourth time. There was a 1-minute interval between the randomised sequences. Note that the instructions for repeated testing and repeated study conditions were synchronised: all participants spent four 2-minute periods considering the target words (either via study or test trials), with each period separated by 1 minute.

Following the final study period, instructions were identical to those in the repeated test conditions. Participants in the 5-min delay conditions filled out the personality questionnaire, then took a free recall test for the target words, under instructions identical to those for the repeated test groups. Participants in the 48-hr delay conditions recalled the instructions that were given to them before the study trials. After these data were collected, all participants were reminded that they must return 48 hours later for the second session.

When participants returned 2 days later, participants in the 48-hr delay conditions were given a blank piece of paper, and instructed to recall as many of the previously studied words as possible in 2 minutes. For participants in the 5-min delay conditions, it was simply noted that they had returned. Any participants who did not return for the second session, regardless of condition, had their data omitted from the experiment. All participants were thanked, debriefed, and then dismissed.

Results and discussion

The mean number of words recalled on the final recall test for each condition is shown in Figure 1. These data were analysed in a 2 × 2 between-subjects ANOVA, with encoding condition (repeated study vs repeated test) and retention interval (5 min vs 48 hr) as independent variables. For each statistically significant result, an effect size was computed. The partial omega squared statistic (ω_p^2) reflects the proportion of variance accounted for by the factor or interaction, after removing the variance accounted for by all other factors and interactions.

Analyses showed a main effect of encoding condition, $F(1, 56) = 4.96$, $MSe = 17.5$, $p < .05$, $\omega_p^2 = .06$, demonstrating that recall was generally higher following multiple study trials than multiple test trials. The factor of retention interval was also significant, $F(1, 56) = 11.56$, $p < .001$, $\omega_p^2 = .15$, confirming that forgetting occurred over the 48-hr retention interval. The primary question of interest in this experiment was whether the rate of forgetting differed between repeated study and repeated test conditions. The answer appears to be yes, as reflected in the significant interaction between encoding condition and retention interval, $F(1, 56) = 4.16$, $p < .05$, $\omega_p^2 = .05$. There was a steeper rate of forgetting for participants who had numerous study trials than for those who were repeatedly tested.

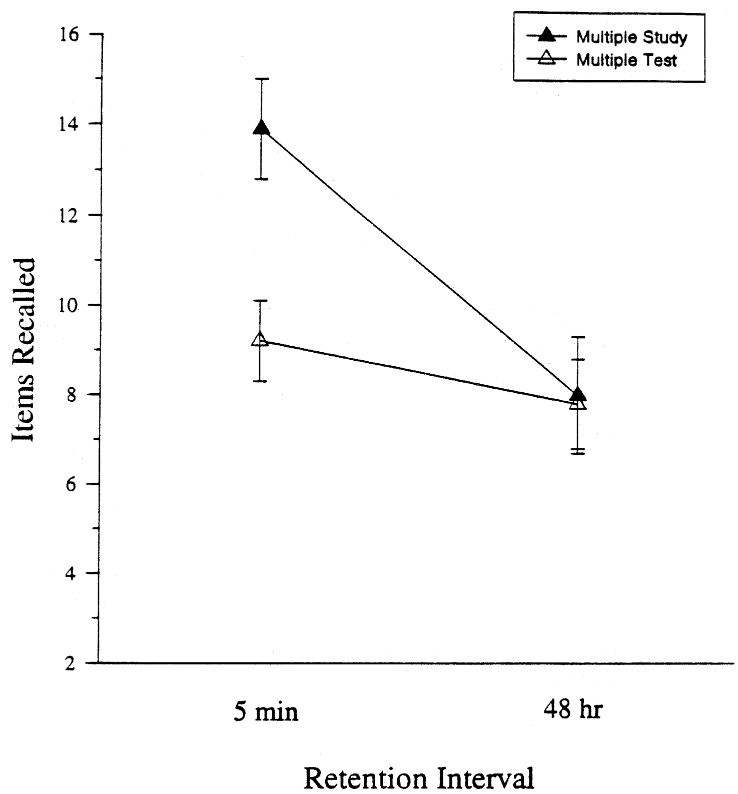

Figure 1. Mean number of words recalled as a function of encoding condition and retention interval in Experiment 1.

Simple main effects were conducted to further examine the interaction. The analyses compared the effect of encoding condition at both of the retention intervals. After the 5-min interval, participants in repeated study conditions recalled more words than those in repeated test conditions, $F(1, 56) = 9.09$, $MSe = 17.5$, $p < .005$ $\omega_p^2 = .21$. Recall at this interval is consistent with common-sense notions about memory, especially the idea that the best way to memorise is to study it repeatedly. Following the 48-hr interval, however, there was no difference between the encoding conditions, $F(1, 56) < 1$. It appears that the beneficial effect of multiple study trials does not persist over long retention intervals.

We also analysed the data from the three consecutive recall trials, collected from participants in repeated test conditions. Participants in the 5-min retention interval condition recalled an average of 9.1, 9.0, and 9.3 target words across the three tests, whereas the corresponding scores from the 48-hr retention condition were 8.3, 7.9, and 8.3. These data were entered into a 2 × 3 mixed-design ANOVA with group (5 min vs 48 hr) as a between-subjects factor, and test number (Test 1, 2, and 3) as a repeated measure. The main effect of group was not significant, $F(1, 28) < 1$, $MSe = 32.9$, confirming that the two groups, who had been treated identically, showed roughly equivalent levels of recall during this ‘‘encoding’’ stage. The factor of test number was not significant, $F(2, 56) = 1.41$, $MSe = 0.77$, $p > .25$, suggesting that there was neither hypermnesia, or net gain in recall, nor forgetting across the consecutive tests. The two factors did not interact, $F(2, 56) < 1$, $MSe = 0.77$.

The overall results are consistent with the ideas that motivated the experiment. At a relatively short retention interval of 5 minutes, recall clearly benefited by the opportunity to repeatedly study the target words. Yet when the retention interval was extended, participants in the multiple study conditions rapidly forgot the target words, at a steeper rate than participants who had studied the words only a single time before taking multiple recall tests. Two days after the initial study phase, the beneficial effects of repeated study were no longer evident. Experiment 1 supports the hypotheses that additional elaborative study is beneficial for encoding into episodic memory, while the free recall of previously studied material may buffer against rapid forgetting.

Although the results were suggestive, they were not overwhelming. The critical interaction between encoding condition and retention interval explained only 5% of the variance, once the effects of the other independent variables were factored out. This effect size can be classified as small-to-moderate (Cohen, 1988). Because these results could have implications for applied pursuits such as education and employee training, it is important to determine whether the effect could be made more robust. Experiment 2 replicated the first experiment, and also attempted to create the potential for a more powerful effect.

EXPERIMENT 2

The second experiment utilised a design and procedure that were very similar to Experiment 1 in order to replicate the basic finding, with a few changes that were intended to bolster and effect. First, the final recall test was taken after a delay of 1 week. If there are differences in rates of forgetting between the encoding conditions, it seems reasonable that the effect should be more pronounced following 1 week than 48 hours. Also, a fourth repeated study or repeated test trial was added to the encoding stage of the experiment. If it is the case that additional study opportunities serve to increase the amount that is successfully encoded, and test opportunities decrease the rate of forgetting, then the inclusion of yet another study or test trial should lead to additional gains in recall at the shorter retention interval for the repeated study group, and reduce the rate of forgetting in the repeated test group. Following initial presentation of the target list, participants in Experiment 2 took either four free-recall tests, or had four additional study opportunities. We predicted that results would fully replicate Experiment 1, and that the critical interaction between encoding group and retention interval would be substantially stronger than in the previous experiment, as indicated by the effect size.

Method

Participants, design, and materials. Participants were 48 Temple University students enrolled in undergraduate psychology courses, participating in exchange for partial course credit. The experiment used a 2 × 2 between-subjects design, with encoding condition (repeated test vs repeated study) and retention interval (5 min vs 1 week) as independent variables. Twelve participants were randomly assigned to each of the four conditions. Materials were the same 40 target words used in Experiment 1.

Procedure. The procedure was very similar to the first experiment, with two changes. There was an additional study trial, or test trial, in the repeated study or repeated test conditions, respectively. Therefore, the encoding phase of the repeated study conditions involved the presentation of the same target list a total of five times, with the words in five different random orders. Participants in repeated test conditions studied a single presentation of the target list, then took four free-recall trials. The other critical difference was the length of the retention interval. All participants were required to return 7 days after the initial session.

Results and discussion

The mean number of words recalled on the final recall test for each condition is shown in Figure 2. These data were analysed in a 2 × 2 between-subjects ANOVA, with encoding condition (repeated study vs repeated test) and retention interval (5 min vs 7 days) as independent variables.

Levene's test for homogeneity of variance was statistically significant in the overall ANOVA, $F(3, 44) = 3.13$, $p < .05$. According to Box (1954), ominibus *F*-tests are still robust with respect to moderate violations of the assumptions of equal variance, provided that cell sizes are equal, as they are in the present experiment.

Analyses showed a marginal effect of encoding condition $F(1, 44) = 3.88$, $MSe = 23.4$, $p < .06$, $\omega_p^2 = .06$, demonstrating that recall was generally higher following multiple study trials than multiple test trials. The factor of retention interval was significant, $F(1, 44) = 44.65$, $p < .001$, $\omega_p^2 = .48$, confirming that forgetting occurred over the 7-day retention interval. Again, the primary question involved the rate of forgetting between repeated study and repeated test conditions. There was a significant interaction between encoding condition and retention interval, $F(1, 44) = 16.95$, $p < .001$, $\omega_p^2 = .25$. There was a steeper rate of forgetting for participants who had numerous study trials than for those who were repeatedly tested

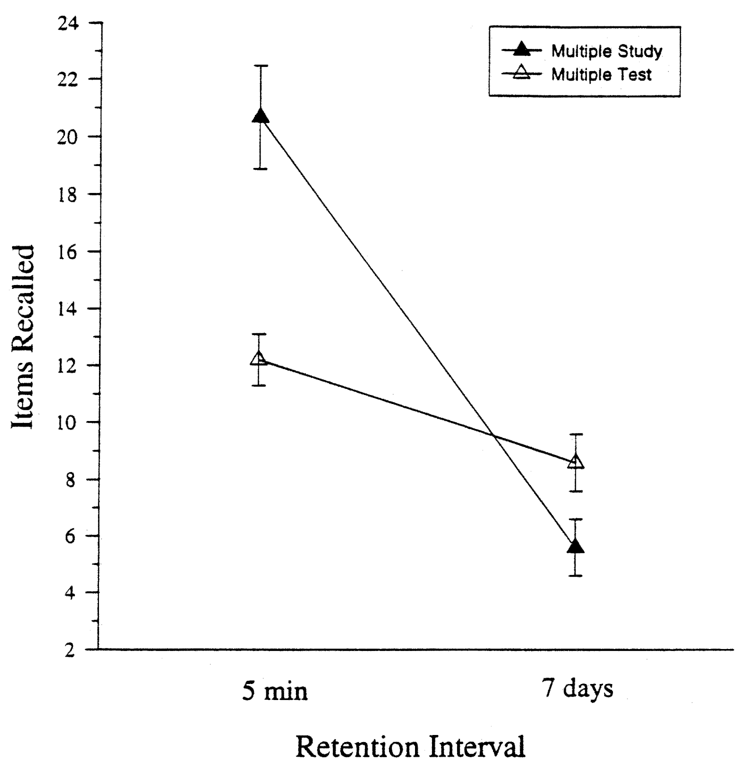

Figure 2. Mean number of words recalled as a function of encoding condition and retention interval in Experiment 2.

for the target words, and the effect was larger than in Experiment 1.

Follow-up analyses investigated the effect of encoding condition at both of the retention intervals. Because of the statistically significant Levene test, *t*-tests were performed without the assumption of equal variances. After the 5-min interval, participants in repeated study conditions recalled more words than those in repeated test conditions, $t(22) = 3.56$, $p < .005$, $\omega_p^2 = .33$. Once again, the participants who repeatedly studied the target words were able to recall more of them at the shorter retention interval. One week later, however, the effect was very different. Participants in repeated test conditions recalled more of the target words, $t(22) = -2.07$, $p = .05$, $\omega_p^2 = .12$. Delayed recall was enhanced by test trials, more than study trials.

Next, we analysed the data from the four consecutive recall trials, collected from participants in repeated testing conditions. Participants in the 5-min retention interval condition recalled an average of 11.6, 11.3, 11.7, and 11.7 target words across the four tests, whereas the corresponding scores from the 7-day retention interval condition were 10.5, 10.5, 10.7, and 10.8. These data were entered into a 2 × 4 mixed-design ANOVA with group (5 min vs 1 week) as a between-subjects factor, and test number (Test 1, 2, 3, and 4) as a repeated measure. The main effect of group was not significant, $F(1, 22) < 1$, $MSe = 71.0$, confirming that the two groups, who had been treated identically, showed roughly equivalent levels of recall during the "encoding" stage. The factor of test number was not significant, $F(3, 66) < 1$, $MSe = 0.52$, and the two factors did not interact, $F(3, 66) < 1$, $MSe = 0.52$.

The overall result of Experiment 2 is in accord with Experiment 1, and also with the hypotheses that motivated the project. Repeated study of target material is a useful method of episodic encoding, but the beneficial effects of studying are short-lived if there are no tests for the target material. Across an interval of 1 week, target material was rapidly forgotten in repeated study conditions; the rate of forgetting was greatly reduced after repeated tests, implying that the act of recalling helps to buffer those recalled items against forgetting. Further, when the study and test manipulations were made stronger, as they were from the first to the second experiment, the critical interaction also became stronger; the resulting effect size of .25 is classified as very large (Cohen, 1988).

GENERAL DISCUSSION

Results support the hypothesised differences in rates of forgetting between two different encoding conditions. Repeated study of target material results in a higher level of original learning, and also a steeper rate of forgetting, compared to only a single study trial followed by numerous recall opportunities.

Previous, similar studies comparing repeated test trials to repeated study trials have occasionally been criticised. Carrier and Pashler (1992) have pointed out the difficulty inherent in comparing the effects of studying and testing, as it is hard to experimentally control the cognitive strategies of participants in the various encoding conditions. Some of theire criticisms are relevant to the present experiments. One potential problem stems from the fact that, on intervening study trials, all target items will be familiar to participants. They may therefore believe that they have already successfully learned the items, and decide to ignore the repeated presentations. In contrast, repeated test instructions require participants to attempt recall for each item on every intervening test trial. Therefore, performance following test trials may exceed that following study trials simply because, in the former, participants were more likely to be thinking about target items. The criticism is reasonable enough but, at least in our experiments, it is not consistent with our observations. If anything, the participants taking repeated study trials appeared to be much engaged in their assigned tasks than those in the intervening test conditions. Perhaps more persuasively, it is not clear how such a possibility (greater overall effort in repeated study than repeated test conditions) could explain the consistent *interaction* between encoding condition and retention interval.

An additional point raised by Carrier and Pashler (1992) was that participants in intervening study trials have the opportunity to consider all of the target items, whereas those in intervening test trials can only consider a subset of the items (e.g., the items that could be successfully recalled). Because there are very few words recalled in the last several seconds of each intervening recall trial, this allows participants to spend additional time looking at, rehearsing, elaborating upon, or otherwise considering the items that they have already recalled. In contrast, repeated study conditions are experimenter-paced, and participants typically are allowed only a few seconds per item.

Thus, by this interpretation, the beneficial effects of repeated testing compared to repeated studying may not come from the actual act of retrieval; rather, the learning may take place after the items are retrieved on the intervening tests.

This potential confound could also be extended to both of the current experiments. It is difficult to prevent participants from considering the items that they have already recalled and, in some cases, the participants may have even tried to use these items as retrieval cues to recall other items. The best evidence in rebuttal of this argument actually comes from a relatively old experiment (Thompson et al., 1978, Experiment 3) that has rarely been cited. Although both repeated study and repeated test conditions were included, this experiment is not vulnerable to the criticism that participants in repeated test conditions may have concentrated on a small number of target items. Using multiple short lists, and a cued-recall task during intervening tests, participants in repeated test conditions were able to recall *virtually all of the target items* on the intervening tests. Also, because the intervening tests were extremely short (20 seconds), participants did not have time to mentally elaborate upon recalled words. At a relatively short retention interval of 5 minutes, the study and test groups recalled an equal number of words. After a delay of 48 hours, the repeated test group recalled significantly more target items than the repeated study group.[1] It would be extremely difficult to explain these findings in terms of differential concentration or rehearsal, as participants in both encoding conditions considered all of the items for virtually the same amount of time. Again, it is reasonable to conclude that the act of retrieving facilitated later recall.

Results of these experiments have implications for both basic and applied pursuits. With respect to basic memory research, it now appears that conclusions drawn about memory and learning at one interval may not always be valid at other intervals. Suppose researchers were interested in comparing the relative effectiveness of multiple study versus multiple test trials on free recall, and also that only a single retention interval was chosen for the experiment. If the delay interval were 5 minutes, the conclusion might be that repeated study opportunities were clearly the more beneficial for eventual recall. With a retention interval of 7 days or longer, the opposite conclusion would likely be drawn.

Although these experiments fit squarely within the domain of cognitive memory research, it is interesting that the data do not have implications for many existing cognitive theories. Indeed, there are very few psychological theories about forgetting. Results are, however, consistent with an idea proposed by Bjork and Bjork (1992), who distinguished between two hypothetical factors, ‘‘storage strength’’ and ‘‘retrieval strength’’, which were said to influence the recallability of items. The act of re-studying an item, or successfully recalling the item, would increase its storage strength, or the extent to which it is well learned. But successful retrieval, especially recall, has a powerful and beneficial effect upon retrieval strength, or the ease with which the item can be accessed. Although they (Bjork & Bjork, 1992) were not directly concerned with manipulations of the retention interval, the theory can easily be applied to the present experiments, with the additional assumption that increases to the retrieval strength of an item are relatively long-lasting.

Present results are also consistent with earlier ideas about the role of recall practice in the retrieval of semantic memory. Research by Harry Bahrick (Bahrick & Hall, 1991; Bahrick & Phelps, 1988) has stressed the fluctuations in access for individual targets (episodic or semantic) over time. Although the amount that can be recalled from a particular study list generally decreases with time, accessibility to individual items can fluctuate either up or down. Successful recall has been considered as *preventive access maintenance* because such recall prevents, or at least delays, subsequent downward fluctuation (Bahrick & Hall, 1991). The idea is doubtless true at some level, but is also fiendishly difficult to test rigorously.

It may be safest simply to conclude that the act of successfully recalling information has a qualitatively different effect from the act of re-studying information. While it is well-known that effortful, elaborative rehearsal leads to enhanced learning and memory, it may be that the retention of that information over a matter of day, weeks, and months is influenced largely by prior successful acts of retrieving the information, and not solely from the strength or quality of the original

[1] Although this comparison reached traditional levels of statistical significance ($p < .05$), the authors (Thompson et al., 1978) had adopted a more stringent criterion of .01. Therefore, they did not interpret the delayed recall differences between the repeated test and repeated study conditions as being significantly different, although they did acknowledge the trend.

encoding episode. An act of retrieval, especially recall, does not simply ‘‘strengthen’’ a representation in memory; rather, it enhances some aspect of the retrieval process itself. Such enhancement helps to buffer against subsequent forgetting. Within the verbal learning tradition, a compatible proposition was advanced by Allen et al. (1969), in regard to experiments with paired-associate learning.

Results also have implications for almost any applied pursuit in which a study, or training, phase is conducted with the goal of influencing performance many days, weeks, or months in the future. One obvious area of application is education; students typically prepare for exams by reading, and re-reading, relevant books and lecture notes. Although repeated study surely enhances the amount that can be retained over relatively short intervals, this approach may lead to poor retention over longer time periods, as well as misleading expectations. Students may believe that they have mastered the target material following study, yet without retrieval practice the forgetting function for such material will be steep. Such a result could disrupt eventual exam performance and, even more likely, reduce the amount of studied information that can be remembered after completion of the course. The use of tests as exam preparation, and also more frequent testing in the classroom, would increase the very long-term retention of the material. Similarly, Schmidt and Bjork (1992) have discussed the value of retrieval practice during training in military and industrial settings. The inclusion of testing would help trainees to retain what they have learned even many months after training is completed, and also to minimise the time and money spent on refresher courses.

To return to the main point, two experiments demonstrated a striking difference in forgetting rates between two different learning conditions. Repeated study of a target list led to a relatively high level of recall at a short retention interval of 5 minutes, and a considerable loss of recall over the next several days. By contrast, a single study opportunity followed by repeated recall tests resulted in a lower level of immediate recall, combined with a much slower rate of forgetting. The act of successfully recalling information may make that information more resistant to forgetting, although additional research will be necessary to determine the precise mechanisms underlying the interaction. Whatever the cause of the phenomenon, the results (especially from Experiment 2) validate the utility of further investigation of the effects of retention interval on encoding manipulations. Also, results strongly reinforce the idea that retrieval practice, especially in the form of recall tests, has a powerful, beneficial effect on long-term retention. It is interesting that more than 100 years after Ebbinghaus (1885/1964) published his groundbreaking work on memory and forgetting, psychologists are still discovering relatively basic facts about the mechanics of forgetting. It may ultimately turn out that the factors influencing forgetting are very bit as complex as those affecting learning.

Manuscript received 9 September 2001
Manuscript accepted 10 July 2002

REFERENCES

Allen, G.A., Mahler, W.A., & Estes, W.K. (1969). Effects of recall tests on long-term retention of paired associates. *Journal of Verbal Learning and Verbal Behavior*, *8*, 463–470.

Anderson, J.R., & Schooler, L.J. (1991). Reflections on the environment in memory. *Psychological Science*, *2*, 396–408.

Bahrick, H.P., & Hall, L.K. (1991). Preventive and corrective maintenance of access to knowledge. *Applied Cognitive Psychology*, *5*, 1–18.

Bahrick, H.P., & Phelps, E. (1988). The maintenance of marginal knowledge. In U. Neisser & E. Winograd (Eds.), *Remembering reconsidered: Ecological and traditional approaches to the study of memory* (pp. 178–192). New York: Cambridge University Press.

Bjork, R.A., & Bjork, E.L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In A. Healy, S.M. Kosslyn, & R.M. Shiffrin (Eds.), *From learning processes to cognitive processes: Essays in honor of William K. Estes*, (Vol. 2, pp. 35–67). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Box, G.E.P. (1954). Some theorems on quadratic forms applied in the study of analysis of variance problems, II: Effects of inequality of variance and of correlation between errors in the two-way classification. *Annals of Mathematical Statistics*, *25*, 484–498.

Carrier, M., & Pashler, H. (1992). The influence of retrieval on retention. *Memory & Cognition*, *20*, 633–642.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates Ltd.

Ebbinghaus, H. (1885/1964). *Memory: A contribution to experimental psychology* (H.A. Ruger & C.E. Bussenius, Trans.). New York: Dover.

Gates, A.I. (1917). Recitation as a factor in memorizing. *Archives of Psychology*, *6*, *No. 40*.

Glover, J.A. (1989). The ‘‘testing’’ phenomenon: Not gone but nearly forgotten. *Journal of Educational Psychology*, *81*, 392–399.

Hockley, W.E. (1992). Item versus associative information: Further comparisons of forgetting rates. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 1321–1330.

Hogan, R.M., & Kintsch, W. (1971). Differential effects of study and test trials on long-term recognition and recall. *Journal of Verbal Learning and Verbal Behavior*, *10*, 562–567.

Howe, M.L., & Hunter, M.A. (1986). Long-term memory in adulthood: An examination of the development of storage and retrieval processes at acquisition and retention. *Developmental Review*, *6*, 334–364.

Landauer, T.K., & Bjork, R.A. (1978). Optimum rehearsal patterns and name learning. In M.M. Gruneberg, P.E. Morris, & R.N. Sykes (Eds.), *Practical aspects of memory* (pp. 625–632). London: Academic Press.

Loftus, G.R. (1985). Evaluating forgetting curves. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 397–409.

Nelson, T.O., & Vining, S.K. (1978). Effects of semantic versus structural processing on long-term retention. *Journal of Experimental Psychology: Human Learning & Memory*, *4*, 198–209.

Rea, C.P., & Modigliani, V. (1985). The effects of expanded versus massed practice on the retention of multiplication facts and spelling lists. *Human Learning*, *4*, 11–18.

Rubin, D.C., & Wenzel, A.E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review*, *103*, 734–760.

Runquist, W. (1983). Some effects of remembering on forgetting. *Memory & Cognition*, *11*, 641–650.

Schmidt, R.A., & Bjork, R.A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*, *3*, 207–217.

Slamecka, N.J., & McElree, B. (1983). Normal forgetting of verbal lists as a function of their degree of learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 812–816.

Spitzer, H.F. (1939). Studies in retention. *Journal of Educational Psychology*, *30*, 641–656.

Thompson, C.P., Wenger, S.K., & Bartling, C.A. (1978). How recall facilitates subsequent recall: A reappraisal. *Journal of Experimental Psychology: Human Learning and Memory*, *4*, 210–221.

Wheeler, M.A. (2000). A comparison of forgetting rates in older and younger adults. *Aging, Neuropsychology, and Cognition*, *7*, 179–193.

Wheeler, M.A., & Roediger, H.L. (1992). Disparate effects of repeated testing: Reconciling Ballard's (1913) and Bartlett's (1932) results. *Psychological Science*, *3*, 240–245.

Wixted, J.T. (1990). Analyzing the empirical course of forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 927–935.