The Effect of Study Time Distribution on Learning and Retention: A Goldilocks Principle for Presentation Rate

Mario de Jonge, Huib K. Tabbers, Diane Pecher, and René Zeelenberg
Erasmus University Rotterdam

In 2 experiments, we investigated the effect of presentation rate on both immediate (5 min) and delayed (2 days) cued recall of paired associates. Word pairs were presented for a total of 16 s per pair, with presentation duration of individual presentations varying from 1 to 16 s. In Experiment 1, participants studied word pairs with presentation rates of 16 × 1 s, 8 × 2 s, 4 × 4 s, 2 × 8 s, or 1 × 16 s. A nonmonotonic relationship was found between presentation rate and cued recall performance. Both short (e.g., 1 s) and long (e.g., 16 s) presentation durations resulted in poor immediate and delayed recall, compared with intermediate presentation durations. In Experiment 2, we replicated these general findings. Moreover, we showed that the 4 s condition resulted in less proportional forgetting than the 1 s and the 16 s conditions.

Keywords: presentation rate, study time distribution, paired-associate learning, cued recall, forgetting

One major factor that affects memory performance is the amount of time available for study. It is generally agreed upon that if one studies for a longer period of time then more is learned (Ebbinghaus, 1885/1964). Unfortunately, in reality, students do not have an unlimited amount of study time at their disposal, and even if they did, they would probably never spend it all studying. Since time is such a precious thing and often in short supply, it only makes sense that researchers everywhere spend heaps of it investigating the conditions under which learning is optimal. In the present study, we asked ourselves the following question: If one only has a limited amount of time to study, how should the available time be divided to be optimally effective? Or, more specifically, when learning new information with a fixed amount of time available per item, what would be the most efficient rate of presentation?

Past research on the issue of optimal presentation rates in paired associate learning has led to different opinions on the matter. The most extreme position is probably held by researchers advocating that the amount learned is solely affected by the total study time available (e.g., Bugelski, 1962; Murdock, 1960). This idea, often referred to as the *total time hypothesis*, states that the amount of time necessary to learn a specific amount of information is fixed and does not vary as a function of the individual presentation durations into which the available time is divided (see Cooper & Pantle, 1967, for a review of the literature). There is no doubt that total study time plays an important role in determining the amount that can be learned. However, more recent studies have shown that total time is not the sole determinant of learning. For instance, there is a vast amount of research on the spacing effect, showing that spaced presentations of materials will generally result in superior recall, compared with massed presentations (for a review see Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006). Clearly, these findings pose a serious problem for the total time hypothesis (Dempster, 1988; Melton, 1970). Even though the total time hypothesis might have fallen from grace as a theory for understanding human learning, some studies directly testing this hypothesis have led to interesting findings concerning the effect of presentation rate on the learning of verbal material.

In one such study (Johnson, 1964), participants learned a list of paired associates consisting of consonant–vowel–consonant (CVC) nonsense syllables paired with digits (e.g., *FAW–7*). Both total presentation time as well as presentation rate were manipulated between subjects. Items were studied for a total study time of 10, 20, 40, or 80 s and presented 1, 5, 10, or 20 times. Upon completion of the study phase, participants received an immediate recall test. Not surprisingly, the results showed that the total study time had a significant effect on recall. More important, however, a nonmonotonic relationship was found between presentation rate and recall performance when total presentation time was held constant. Although the relevant statistics were not always provided, the general pattern of results seems to indicate that both short and long presentation durations resulted in suboptimal learning, with optimal learning occurring at an intermediate presentation rate somewhere between 2 and 4 s per item.

The results from the Johnson (1964) study suggest that both the total study time and the duration of individual presentations exert an influence on memory performance. However, the results are not that unequivocal. Johnson (1964) used a fixed 4 s intertrial interval in his study. Because conditions consisting of more exposures

automatically received more 4 s intertrial intervals, the differences in presentation rate between conditions also resulted in substantial differences in total time available for study (Cooper & Pantle, 1967). In a follow-up study by Stubin, Heimer, and Tatz (1970), an attempt was made to eliminate this confound of presentation rate and total study time. Paired associates (pairs of CVC nonsense syllables) were presented via a slide projector, and it took 0.8 s for the projector to change slides. So, even though measures were taken to eliminate the problem with intertrial interval, there still was an effective 0.8 s lag between trials resulting in differences in total study time between conditions. Still, the results from the Stubin et al. (1970) study were largely in agreement with those obtained by Johnson (1964), even though they used an intertrial interval that was considerably shorter. A nonmonotonic relationship was found between presentation rate and subsequent recall: both slow ($\geq$ 10 s) and fast (2 s) presentation rates resulted in inferior recall performance, compared with an intermediate 5 s presentation rate.

The results from Johnson (1964) and Stubin et al. (1970) suggest that with total time held constant, the presentation duration of individual exposures to study materials influences the extent to which new information is learned. We believe these studies have important implications for both theoretical as well as educational purposes. Quite undeservedly, however, these studies have been largely neglected in the literature, and there has been virtually no follow-up since the total time era ended.

In the two experiments reported here, we further investigated the effect of presentation rate on paired associate learning. Our first objective was to replicate the Johnson (1964) and Stubin et al. (1970) studies, controlling for the methodological confound discussed earlier. We incorporated the intertrial interval within the presentation duration to make sure no differences in total time available for study would arise between study conditions. So, for instance, a 2 s presentation consisted of a 1.75 s presentation and a 0.25 s intertrial interval.

The second objective of our study was to extend the general findings from earlier studies. In the present experiments, presentation rate was manipulated within subjects (as opposed to between-subjects manipulations in previous studies). Also, we used more meaningful materials (words pairs instead of CVC nonsense syllables or digits). More important, we also wanted to look at longer retention intervals than those used in the earlier studies. In both the Johnson (1964) and the Stubin et al. (1970) studies, only a single short retention interval was used. In the Johnson (1964) study, a test was given immediately after learning, and in the Stubin et al. (1970) study, participants received a final test only 20 s after the learning phase was completed. One could argue that to some degree, short-term memory was being compared with long-term memory (Bugelski & McMahon, 1971). That is, the contribution of short-term memory to performance in the final test may have been different, depending on the presentation rate. In conditions with a relatively slow presentation rate, items would on average be recalled a couple of minutes later on a final test, while in conditions with relatively fast presentation rates, this would only be a matter of seconds. In the present study, participants first worked on a 5-min distractor task before taking a final recall test.

Another limitation of the use of a single short retention interval in earlier studies is that these studies do not inform us about the effect of presentation rate on forgetting. In the present study, we therefore also included a retention interval of 2 days. Studies have shown that conditions that result in superior performance on an immediate recall test do not always benefit performance at a longer retention interval (e.g., Rawson & Kintsch, 2005; Roediger & Karpicke, 2006). On a related note, it has been suggested that learning conditions that slow down initial learning can actually benefit long-term retention because these conditions introduce *desirable difficulties* during learning (Bjork, 1994, 1999). It remains to be seen whether presentation rates that are optimal for performance at short retention intervals are also optimal for performance at longer retention intervals.

Experiment 1

Method

Participants.   Forty-two students from the Erasmus University Rotterdam participated in partial fulfillment of course requirements. Data from two participants were excluded from analyses because these participants failed to show up for the 2-day final test.

Materials and design.   Eighty unrelated Dutch word pairs (e.g., *hamer–lift* [*hammer–elevator*], *spin–balkon* [*spider–balcony*]) were used in the experiment. All words were between four and six letters long and consisted of either one or two syllables. The mean word length was 4.87 ($SD = 0.79$). The mean word frequency per million (Keuleers, Brysbaert, & New, 2010) was 16.54 ($SD = 44.62$). Word pairs were divided over five lists of 16 items each. The computer application E-prime (Psychology Software Tools, Pittsburgh, PA) was used to create and run the experiment.

A 2 $\times$ 5 mixed-factorial design was used in the experiment, with study condition as within-subjects factor, retention interval as between-subjects factor, and cued recall score as dependent variable. Participants were randomly assigned to one of two retention interval conditions. One group of participants received a final cued recall test 5 min after the study phase ended, and the other group received the cued recall test 2 days later. Both groups were required to return for the 2 day session regardless of the retention interval condition to which they were assigned.

Participants studied word pairs under five different study conditions; 16 $\times$ 1 s, 8 $\times$ 2 s, 4 $\times$ 4 s, 2 $\times$ 8 s, and 1 $\times$ 16 s. In the 16 $\times$ 1 s condition, a list of word pairs was presented 16 times with a presentation rate of 1 s per pair. In the 8 $\times$ 2 s condition a list of word pairs was presented eight times with a presentation rate of 2 s per pair. The 4 $\times$ 4 s condition consisted of four list presentations of 4 s per pair; the 2 $\times$ 8 s condition consisted of two list presentations, with 8 s per pair; and in the 1 $\times$ 16 s condition, a list of word pairs was presented once, with a presentation rate of 16 s per pair. For each of these conditions, all pairs on the list were presented once in a random order before the pairs were presented again in a different random order (except, of course, for the 1 $\times$ 16 s condition, in which all pairs were presented only once). By manipulating the presentation rate of word pairs in this manner, we kept the total study time for each word pair constant across all conditions. Table 1 shows the average spacing in seconds between repetitions of the same pair as well as the average spacing in seconds between the first and last presentations of the same pair. The manipulation of presentation rate in the present experiment

Table 1

Average Spacing in Seconds for the Five Study Conditions in Experiment 1

Condition	Interrepetition spacing	Total spacing
16×1 s	15	239
8×2 s	30	222
4×4 s	60	188
2×8 s	120	120
1×16 s	—	—

Note. Interrepetition spacing refers to the average number of seconds between repetitions of the same pair. Total spacing refers to the average number of seconds between the first and last presentation of the same pair. A dash indicates that pairs were presented once (i.e., there was no spacing).

also resulted in differential spacing between conditions. However, as we explain in the General Discussion, our results are not readily accounted for by these differences in spacing.

A total of 10 counterbalanced versions were created according to a scheme proposed by Lewis (1989) using a pair of Latin squares. Both the assignment of word pairs to conditions and the order in which conditions were administered during the study phase were counterbalanced. Across participants, all word pairs appeared equally often in each study condition, and all word pairs and study conditions were presented equally often in each of five blocks in the presentation order of conditions. Furthermore, immediate sequential effects were counterbalanced so that each condition was preceded as well as followed by each other condition equally often across participants. In the test phase, cue words were presented in a random order; items from the different study conditions were randomly intermixed.

Procedure. Participants were tested either alone or in small groups during two sessions. In the first session, participants received verbal as well as onscreen instructions about the experiment. They were told that they would study word pairs at different presentation rates during five consecutive study blocks and that they would receive a memory test afterwards to assess their performance. Participants were also told that they were not allowed to cover part of the computer screen with their hand in order to test themselves during study. This was explicitly stated because during a pilot study we observed a number of participants using this strategy during study. To control for any undesirable effects that might occur as a result of self-testing, we stressed that this was not allowed.

Before each study block, participants received onscreen instructions telling them in which way the materials would be presented (how many times and at what rate). During each study block, word pairs were presented on a computer screen in a different random order for each cycle. The two words of a pair were presented simultaneously, one above the other on the center of the screen. Upon completion of the study phase, participants worked on Sudoku puzzles for a period of 5 min as a distractor task. After the 5-min distractor task, half the participants received an immediate self-paced cued recall test. The remaining participants were dismissed after the distractor task and received a self-paced cued recall test 2 days later.

Results and Discussion

Figure 1 shows the mean proportions of correctly recalled words for both the 5-min group and the 2-day group as a function of study condition. At both delays, presentation rate and performance showed an inverted u-shape. Performance was optimal in the 4×4 s condition and dropped off with higher and lower presentation rates. On the 5-min test, the mean percentages of correctly recalled items were 22%, 41%, 52%, 49%, and 37% for the 16×1 s, 8×2 s, 4×4 s, 2×8 s, and the 1×16 s condition, respectively. Two days later, recall was considerably lower: 4%, 18%, 20%, 17%, and 9%, respectively (for the same five study conditions). The data were analyzed using a 2×5 repeated-measures analysis of variance (ANOVA) with retention interval as a between-subjects factor, study condition as within-subjects factor and recall score as dependent variable. The analysis showed a significant effect of retention interval on final test score, $F(1, 38) = 16.47$, $p < .001$, $\eta_p^2 = .64$. Recall scores were considerably lower on the 2-day test (14% correct), compared with recall on the 5-min test (40% correct). More important, study condition also had a significant effect on cued recall performance, $F(4, 152) = 14.22$, $p < .001$, $\eta_p^2 = .27$, indicating that the different rates of presentation during study resulted in differences in final test score. The interaction between retention interval and study condition was not significant, $F(4, 152) = 1.38$, $p > .20$.

We performed a subsequent repeated contrast analysis to determine whether performance for each presentation rate was significantly different from the next slower presentation rate. This analysis showed that studying word pairs eight times with a presentation duration of 2 s per pair (the 8×2 s condition) resulted in superior recall, compared with studying word pairs 16 times with 1 s per pair (the 16×1 s condition), $F(1, 57) = 31.26$, $p < .001$, $\eta_p^2 = .38$. Studying word pairs four times with 4 s per pair (the 4×4 s condition) resulted in superior recall, compared with the 8×2 s condition, $F(1, 57) = 6.23$, $p < .05$, $\eta_p^2 = .10$. The difference between the 4×4 s and the 2×8 s condition was not significant ($F < 1$). Finally, studying word pairs once for 16 s (the 1×16 s condition) resulted in inferior recall, compared with the 2×8 s condition, $F(1, 57) = 30.76$, $p < .001$, $\eta_p^2 = .29$. The general pattern of results bears a striking resemblance to the findings of Johnson (1964) and Stubin et al. (1970). Indeed, there appears to be a nonmonotonic relationship between presentation rate and recall of paired associates. Participants recalled few words for presentation durations of 1 s, but performance increased as presenta-
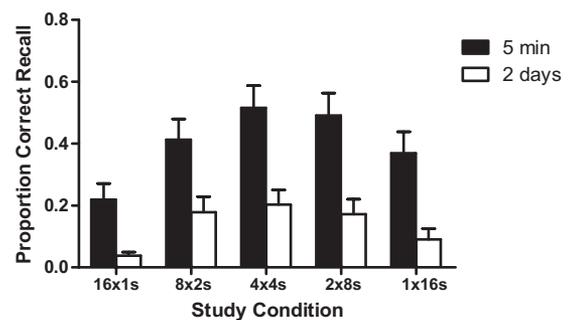
Figure 1. Proportion of words recalled on the 5-min and 2-day cued recall test as a function of study condition in Experiment 1. Error bars represent standard errors of the mean.

tion duration increased. However, this trend stalled for presentations of 4 to 8 s and reversed for presentations of 16 s.

Another point of interest in the present experiment was whether or not the general findings would extend over a longer retention interval. Or in other words, does the general pattern of results change over time? As can be seen in Figure 1, the general pattern of results persisted over the 2-day interval. The lack of an interaction between retention interval and study condition reported above supports this observation. If we would interpret the absolute difference in performance between the immediate and 2-day recall test for the different study conditions as an indication of forgetting, then we would have to conclude that the different presentation rates did not result in different rates of forgetting. However, research on forgetting suggests that the course of forgetting is best described by a power function (Wixted & Carpenter, 2007; Wixted & Ebbesen, 1991).[1] A power function of forgetting measures forgetting as a proportional decline of the amount of information that was originally stored in memory (Carpenter, Pashler, Wixted, & Vul, 2008). In this respect a proportional measure of forgetting would be a more appropriate way of looking at the rate of forgetting. Proportional forgetting can sometimes lead to different conclusions about forgetting, compared with an absolute measure (e.g., Loftus, 1985). Figure 2 shows the proportional forgetting measures for all five study conditions. As is clear from Figure 2, study conditions that resulted in poor initial recall also resulted in high proportional forgetting, 83% in the 16 × 1 s condition and 75% in the 1 × 16 s condition. However, study conditions that resulted in superior recall on the 5-min test resulted in less proportional forgetting; 57% in the 8 × 2 s condition, 60% in the 4 × 4 s condition, and 65% in the 2 × 8 s condition.

Experiment 2

Experiment 2 was designed to extend the findings from Experiment 1 and to allow us to further investigate proportional forgetting. In Experiment 1, we assessed forgetting by comparing average performance on the immediate test with average performance on a delayed test across different groups of subjects. This made it impossible to perform standard statistical analysis on proportional forgetting in Experiment 1. In our second experiment, both study condition and retention interval were manipulated within subjects. This enabled us to perform statistical analyses comparing propor-

tional forgetting in the different study conditions. Because including all five study conditions present in Experiment 1 would result in a somewhat tedious experiment from the participants' perspective, we only compared the most extreme study conditions from Experiment 1 (the 16 × 1 s, the 4 × 4 s, and the 1 × 16 s conditions).
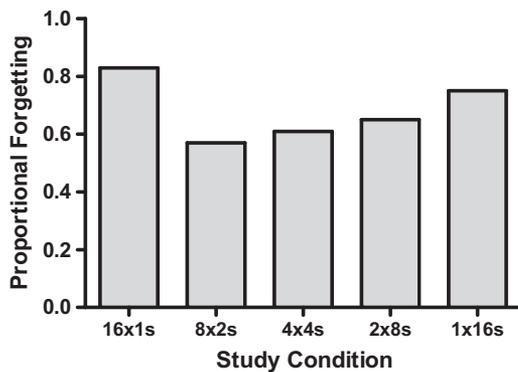
Method

Participants. Thirty students from the Erasmus University Rotterdam participated in partial fulfillment of course requirements. Six participants were excluded from the analysis because of insufficient performance on the 5-min memory test (recall scores of zero on the 5-min test made assessment of subsequent proportional forgetting impossible). None of the participants had participated in our first experiment.

Materials and design. A 2 × 3 factorial design was used with both retention interval and study condition as within-subject factors and recall score as dependent variable. Participants studied word pairs under three different study conditions (16 × 1 s, 4 × 4 s, and 1 × 16 s). Participants were tested on half of the word pairs on the 5-min test, and the other half was tested after a 2-day interval.

Ninety-six word pairs were used in the experiment. Eighty word pairs were identical to the word pairs used in Experiment 1, and 16 new word pairs were compiled to supplement the original 80 word pair list. The mean word length was 4.86 ($SD = 0.78$). The mean word frequency per million (Keuleers et al., 2010) was 16.77 ($SD = 43.14$). Word pairs were divided over six lists of 16 items each. Word pairs were assigned to 16-item word pair lists in such a fashion that every list would include an approximately equal number of new items. Six counterbalanced versions were created in the same general manner as in Experiment 1.

Procedure. The procedure was similar to the procedure in Experiment 1. Participants studied word pairs under three different study conditions during three consecutive study blocks. Upon completion of the 25-min study phase, participants received a distractor task (Sudoku puzzles), followed by the 5-min cued recall test on half of the word pairs from each condition. All participants returned for the cued recall test on the remaining word pairs 2 days later.

Results and Discussion

Figure 3 shows the mean proportions of correctly recalled words on the 5-min and the 2-day recall tests as a function of study condition. On the 5-min test, the mean percentages of correctly recalled items were 43%, 73%, and 60% for the 16 × 1 s condition, 4 × 4 s condition, and the 1 × 16 s condition, respectively.

Figure 2. Proportional forgetting during the 2-day interval as a function of study condition in Experiment 1.

---

[1] The question of which function provides the best description of the forgetting curve has been debated in the literature. Both power and exponential functions (as well as other functions) often provide excellent fits of forgetting data, and it has proved hard to draw firm conclusions about the mathematical form of empirical forgetting functions. Nevertheless, based on different sets of data and different approaches, in recent studies researchers have argued that power functions provide the best description of forgetting. For elaborate discussions of this issue, we refer to Averell and Heathcote (2011) and Wixted (2004).
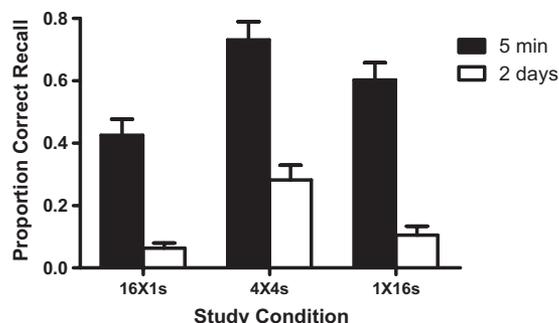
Figure 3. Proportion of words recalled on the 5-min and 2-day cued recall test as a function of study condition in Experiment 2. Error bars represent standard errors of the mean.

Two days later, recall for the three conditions dropped to 6%, 28%, and 11%, respectively. Thus, after a 2-day delay, cued recall test performance in the 4×4 s condition was 348% and 169% higher, compared with that in the 16×1 s and 1×16 s conditions, respectively.

A 2×3 repeated-measures ANOVA revealed that the effect of retention interval on recall score was significant, $F(1, 23) = 168.37$, $p < .001$, $\eta_p^2 = .88$, indicating that forgetting occurred during the 2-day interval (59% correct on the 5-min test vs. 15% correct on the 2-day test). Also, as in Experiment 1, there was a significant effect of study condition on recall score, $F(2, 46) = 26.08$, $p < .001$, $\eta_p^2 = .53$, indicating that the different presentation rates resulted in different recall scores. The interaction between retention interval and study condition did not reach the conventional level of significance, $F(2, 46) = 2.74$, $p > .05$.

Follow-up analyses showed that the 4×4 s condition resulted in superior recall, compared with both the 16×1 s condition, $F(1, 23) = 41.62$, $p < .001$, $\eta_p^2 = .64$, and the 1×16 s condition, $F(1, 23) = 29.23$, $p < .001$, $\eta_p^2 = .56$. To summarize, the general findings from Experiment 1 were replicated in Experiment 2, showing that presentation rate exerted a large influence on cued recall performance.

A more important question addressed by the present experiment was whether or not the different study conditions would result in different proportional forgetting. Figure 4 shows proportional forgetting as a function of study condition. As can be seen, the 16×1 s and the 1×16 s condition resulted in similar proportional forgetting (88% and 86%, respectively), while studying word pairs in the 4×4 s conditions resulted in less proportional forgetting (64%). A repeated-measures ANOVA revealed that the effect of presentation rate on proportional forgetting was significant, $F(2, 46) = 12.46$, $p < .001$, $\eta_p^2 = .35$. Follow-up analyses showed that the 4×4 s condition resulted in less proportional forgetting, compared with both the 16×1 s condition and the 1×16 s condition, $F(1, 23) = 13.79$, $p < .005$, $\eta_p^2 = .38$ and $F(1, 23) = 19.32$, $p < .001$, $\eta_p^2 = .46$, respectively. So, the optimal presentation rate in the present experiment (the 4×4 s condition) did result not only in superior recall but also in less proportional forgetting.

General Discussion

In two experiments, we investigated the effect of presentation rate on the learning and retention of paired associates. With total study time kept constant, we found a nonmonotonic relationship between the presentation rate of word pairs and subsequent recall. Performance was poor for short (e.g., 1 s) and long (e.g., 16 s) presentation durations and much better for intermediate (e.g., 4 s) presentation durations. These results indicate that the presentation duration of individual exposures has a large effect on memory performance even when the total study time is kept constant. Our findings extend earlier studies by Johnson (1964) and Stubin et al. (1970) by eliminating the methodological problems present in their studies and by using meaningful stimuli rather than CVC nonsense syllables. Furthermore, we showed that the effect of presentation rate is not only apparent on an immediate test but also extends to a longer retention interval of 2 days. In Experiment 2, we replicated the general pattern of results and extended the findings by looking at proportional forgetting. We showed that presentation rates that resulted in poor immediate recall also resulted in more proportional forgetting.

In the present study, using unrelated word pairs, we found that a presentation rate of around 4 s resulted in optimal performance. Johnson (1964) and Stubin et al. (1970) obtained similar optimal presentation rates with different types of stimuli (CVC–digit pairs and CVC–CVC pairs, respectively). Nevertheless one should exercise caution in generalizing these optimal presentation rates to other materials. Although we would expect that the same general pattern would emerge across different kinds of materials, the optimal presentation rate might shift depending on the kind of materials used. For example, with more difficult materials, a longer presentation rate might turn out to be optimal. Individual differences among learners may also affect the optimal rate of presentation. Also, as noted earlier, total study time is an important factor determining learning outcomes: When more time is available for learning, more can be learned. Both Johnson (1964) and Stubin et al. (1970) found that doubling the amount of time available for study resulted in substantial increases in cued recall performance. Thus we do not deny that total study time is an important determinant of learning. However, how the available time is divided up into study episodes is at least as important a factor.

Our results provide an intriguing puzzle for theoretical accounts of optimal study routines. There is a large body of literature on theoretical frameworks explaining a variety of distribution of practice phenomena like the spacing effect (see Cepeda et al.,
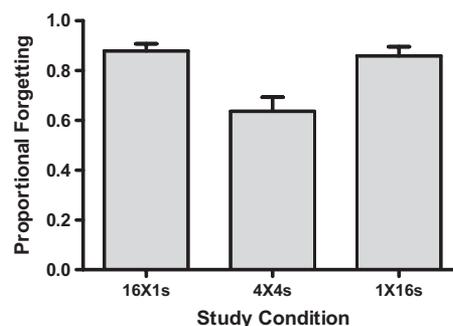
Figure 4. Proportional forgetting during the 2-day interval as a function of study condition in Experiment 2. Error bars represent standard errors of the mean.

2006; Delaney, Verkoeijen, & Spirgel, 2010, for recent reviews of the literature). Unfortunately, the relationship between presentation rate and the amount of spacing is not as straightforward as one might presume. Table 1 shows the average spacing in seconds between repetitions of the same pair as well as the average spacing in seconds between the first and last presentations of the same pair. As can be seen, the average interval between two presentations of a word pair increases as the presentation rate decreases. Following this measure of spacing, one would have to conclude that slower presentation rates resulted in more spacing between presentations. On the other hand, one could also consider the total time between the first and the last presentations of a pair as an indication of spacing. Following this measure of spacing, one would conclude that faster presentation rates resulted in more spacing of word pairs. Although both measures of spacing seem reasonable, the problem, of course, is that they lead to different conclusions about which study conditions were more spaced. Of course, this line of reasoning assumes that the evolvement of time is the critical dimension underlying spacing. If one assumes that the number of intervening presentations between repetitions as the critical dimension, the picture is somewhat clearer. In this case, the average spacing between repetitions is identical for all conditions (except for the 1×16 s condition), but the average total spacing (from the first to the last presentation) increases linearly with presentation rate.

Although the manipulation of presentation rate in the present experiment inevitably resulted in differential spacing between conditions, we believe the present results are not that easily explained from a spacing point of view. That is, other factors, besides spacing *per se*, seem to play a role. In both our experiments, we found similar patterns of results after a short and a long retention interval. So, presentation rates resulting in relatively good performance did so regardless of the delay between study and test. This is unlike research on the spacing effect which actually shows that different distributions of practice are optimal for different retention intervals (Cepeda, Vul, Rohrer, Wixted, & Pashler, 2008; Glenberg, 1976). Furthermore, at the 2-day retention interval, the differences in spacing between the different presentation rate conditions of the present study were rather small relative to the length of the retention interval (approximately 170,000 s) and are therefore not expected to have a substantial impact on performance. Yet, large differences in performance were still found. The observed findings are also not simply an effect of massed versus spaced presentations. Numerous studies have shown that spaced presentations result in better performance than massed presentations (see Cepeda et al., 2006, for a review).[2] However, in both Experiment 1 and Experiment 2 of the present study, performance was higher in the massed condition (1×16 s) than in the 16×1 s condition.[3] So, even though spacing of items and pacing of items can both be considered accounts of distributed practice, we believe that there are some fundamental differences between the two.

We are inclined to propose an alternative explanation for the effect of presentation rate on subsequent recall, namely, the *effective study time hypothesis*. It has been argued that some minimal amount of time might be necessary in order to optimally form an association (Stubin et al., 1970). This could explain why a fast presentation rate results in poor recall on a subsequent test. On the other hand, it has also been argued that presentation durations beyond some optimal value might cause inattention, decreased concentration, and boredom (Bugelski & McMahon, 1971). In this way, a 16 s presentation might be inefficient because less time is needed to form an associative link between two unrelated words. As a result, the remaining time beyond some optimum will be utilized in a less efficient way; that is, less additional information will be stored in memory per unit time. This idea is reminiscent of the famous story of Goldilocks and the Three Bears. In the story, Goldilocks successively tries three different bowls of porridge. She finds that one bowl is too cold, the other one is too hot, but the one in the middle is just right. The same principle appears to be true for presentation rates during the learning of paired associates. Presentation durations should be not too long and not too short, but just right.

In the present study, we did not look at the kind of processing that took place during the different presentation rate conditions, so we can only speculate about the strategies participants used during learning. However, it has been argued that elaborative study strategies take a certain amount of time to be effective (Bugelski, 1970). For instance, the results of a study by Bugelski, Kidd, and Segmen (1968) suggested that participants who studied paired associates under imagery instructions needed 4–8 s to form a useful image. At presentation rates of 4 and 8 s, participants in the imagery group outperformed those in the control group. However, at a presentation rate of 2 s, participants in the imagery group failed to outperform those in the control group, suggesting that they were unable to form an effective mental image. Perhaps a relatively fast presentation rate provides too little time for elaborative processing and learners will be forced to rely on less effective learning strategies (e.g., rote rehearsal).

Other related factors may be at play as well and provide a possible account of why intermediate presentation rates enhance initial learning and reduce forgetting. Note that although our effects are not simply the result of spacing, some mechanisms proposed in the spacing literature may provide a (partial) account of our results. One such mechanism is encoding variability. Encoding variability assumes that context fluctuates over time (Glenberg, 1976; Melton, 1967). Furthermore, encoding materials in different contexts enhances memory performance. More diverse contextual elements would be encoded for items presented four times (as in the 4×4 s condition) than for items presented only once (as in the 1×16 s condition). Without additional assumptions, this account would predict optimal performance for the condition with the largest number of presentations, the 16×1 s

---

[2] Some studies have reported better performance for massed practice than spaced practice on recall tests given almost immediately after learning (e.g., Balota, Ducheck, & Paullin, 1989; Peterson, Wampler, Kirkpatrick, & Saltzman, 1963). Allegedly, Endel Tulving has dubbed this rather paradoxical finding the "Peterson paradox" (Roediger, Balota, & Watson, 2001). In their meta-analysis, Cepeda et al. (2006) found that spaced presentations improved final-test performance by 9% when averaging over 96 studies that used a retention interval of less than 1 min. Improvements due to spacing were also found for longer retention intervals. It appears that the beneficial effect of massed practice is limited to studies that used retention intervals of 4–8 s; retention intervals that are much shorter than those in the present study.

[3] In both Experiment 1 and Experiment 2, more words were recalled in the 1×16 s condition than in the 16×1 s condition, $F(1, 38) = 9.85$, $p < .01$, $\eta_p^2 = .21$, and $F(1, 23) = 7.79$, $p < .01$, $\eta_p^2 = .25$, respectively.

condition. This prediction is clearly violated by our results. One could make the additional assumption that context storage takes time, and little context information is stored during brief presentations of word pairs. Such a hypothesis, however, seems to conflict with findings that suggest context information is stored early on in processing (Malmberg & Shiffrin, 2005).

Another possible explanation is provided by the study phase retrieval account (see Raaijmakers, 2003, for a theory that combines context fluctuation and study phase retrieval to account for spacing effects). This account assumes that when an item is repeated it is retrieved from long-term memory and additional information is stored in the original trace (provided that retrieval of the item is successful). Spacing is beneficial because it results in more contextual information being stored in the memory trace. Like the encoding variability account, the study phase retrieval account could explain why performance in the 4×4 s condition is better than in the 1×16 s condition. More repetitions result in more successful study phase retrievals and, consequently, more contextual information will be stored in the trace. However, without additional assumptions, this account too would predict optimal performance for the condition with the largest number of presentations, the 16×1 s condition. It is plausible though, that successful study phase retrieval depends on the amount of time an item is presented; for brief presentation times of 1 or 2 s, study phase retrieval may not successfully explain poor performance in these conditions. However, to arrive at testable predictions, such an account would have to make specific assumptions about the time course of study phase retrieval. To summarize, spacing theories do not readily account for all aspects of our results. Factors such as encoding variability and study phase retrieval may play a role in our findings, but in order to account for the entire pattern of results spacing theories would need to make additional assumptions.

Recent years have seen a renewed interest in the factors that enhance learning and retention. Of particular importance, these studies have looked at the effects of study manipulations on performance not only on immediate recall but also after retention intervals ranging from several days to several months and even up to a year. Recent studies have shown that testing enhances long-term retention for a variety of materials. For example, Roediger and Karpicke (2006) found that recall of prose passages after a 1-week retention interval was substantially better for subjects who had been tested on those passages after initial study, compared with subjects who received additional study opportunities. Similar benefits of testing over study have been found for the recall of Swahili–English word pairs (e.g., Karpicke & Roediger, 2008; Pyc & Rawson, 2010). In some cases, the advantages of testing over study amounted to improvements in performance of more than 150% (e.g., Karpicke & Roediger, 2008). Spacing also has substantial effects on memory performance. For example, in a very ambitious study Cepeda et al. (2008) investigated the effect of spacing (gap varied from 0 to 105 days) and retention interval (from 7 to 350 days) on cued recall and recognition of trivia facts. They found improvements in cued recall performance of up to 111% for the optimal gap between study trials, as compared with a zero-day gap. Together, these studies and the present one indicate that testing, spacing, and appropriate presentation rates can have a large impact on memory. Not only immediate memory but

also delayed memory can benefit enormously from the right set of study conditions.

In sum, the present study indicates that there is a Goldilocks principle at work with regard to the presentation rate during the learning of paired associates within a fixed amount of time. We showed that presentation rates that are optimal for a short 5-min retention interval also benefit retention after a longer 2-day delay. We believe not just that these results are interesting from a theoretical point of view but that they might also be of particular relevance for educational purposes. For instance, they could be used for optimizing foreign vocabulary learning. Most computer programs for learning foreign vocabulary provide their users with the opportunity for self-paced learning. In the present study, we compared learning conditions with different presentation rates that remained constant during learning. However, when learning foreign vocabulary under self-paced instructions, learners tend to speed up the presentation rate as learning progresses. Even though they employ a reasonable presentation rate the first time through a list, they ultimately devolve to a presentation rate of less then 1 s per item (Kornell & Bjork, 2007). In the present study, we showed that fast presentation rates of 1 s per pair resulted in suboptimal learning. Thus, it is doubtful whether the opportunity for self-paced study will result in an efficient use of study time. Research on metacognition and learning generally shows that students are not very proficient when it comes to allocating self-paced study time (e.g., Nelson & Leonesio, 1988). Since self-pacing often results in suboptimal study time allocation, it could be interesting to look at the usefulness of externally paced study schedules for improving learning and long-term retention.

References

Averell, L., & Heathcote, A. (2011). The form of the forgetting curve and the fate of memories. *Journal of Mathematical Psychology, 55,* 25–35. doi:10.1016/j.jmp.2010.08.009

Balota, D. A., Duchek, J. M., & Paullin, R. (1989). Age-related differences in the impact of spacing, lag, and retention interval. *Psychology and Aging, 4,* 3–9. doi:10.1037/0882-7974.4.1.3

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). Cambridge, MA: MIT Press.

Bjork, R. A. (1999). Assessing our own competence: Heuristics and illusions. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII. Cognitive regulation of performance: Interaction of theory and application* (pp. 435–459). Cambridge, MA: MIT Press.

Bugelski, B. R. (1962). Presentation time, total time, and mediation in paired-associate learning. *Journal of Experimental Psychology, 63,* 409–412. doi:10.1037/h0045665

Bugelski, B. R. (1970). Words and things and images. *American Psychologist, 25,* 1002–1012. doi:10.1037/h0030150

Bugelski, B. R., Kidd, E., & Segman, J. (1968). Image as a mediator in one-trial paired-associate learning. *Journal of Experimental Psychology, 76,* 69–73. doi:10.1037/h0025280

Bugelski, B. R., & McMahon, M. L. (1971). The total time hypothesis: A reply to Stubin, Heurer, and Tatz. *Journal of Experimental Psychology, 90,* 165–166. doi:10.1037/h0031337

Carpenter, S. K., Pashler, H., Wixted, J. T., & Vul, E. (2008). The effects of tests on learning and forgetting. *Memory & Cognition, 36,* 438–448. doi:10.3758/MC.36.2.438

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative

synthesis. *Psychological Bulletin, 132,* 354–380. doi:10.1037/0033-2909.132.3.354

Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (2008). Spacing effects in learning: A temporal ridgeline of optimal retention: Research article. *Psychological Science, 19,* 1095–1102. doi:10.1111/j.1467-9280.2008.02209.x

Cooper, E. H., & Pantle, A. J. (1967). The total-time hypothesis in verbal learning. *Psychological Bulletin, 68,* 221–234. doi:10.1037/h0025052

Delaney, P. F., Verkoeijen, P. P. J. L., & Spirgel, A. (2010). Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. *Psychology of Learning and Motivation, 53,* 63–147. doi:10.1016/S0079-7421(10)53003-2

Dempster, F. N. (1988). The spacing effect: A case study in the failure to apply the results of psychological research. *American Psychologist, 43,* 627–634. doi:10.1037/0003-066X.43.8.627

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

Glenberg, A. M. (1976). Monotonic and nonmonotonic lag effects in paired-associate and recognition memory paradigms. *Journal of Verbal Learning & Verbal Behavior, 15,* 1–16. doi:10.1016/S0022-5371(76)90002-5

Johnson, N. F. (1964). The functional relationship between amount learned and frequency versus rate versus total time of exposure of verbal materials. *Journal of Verbal Learning & Verbal Behavior, 3,* 502–504. doi:10.1016/S0022-5371(64)80022-0

Karpicke, J. D., & Roediger, H. L., III. (2008, February). The critical importance of retrieval for learning. *Science, 319,* 966–968. doi:10.1126/science.1152408

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods, 42,* 643–650. doi:10.3758/BRM.42.3.643

Kornell, N., & Bjork, R. A. (2007). The promise and perils of self-regulated study. *Psychonomic Bulletin & Review, 14,* 219–224. doi:10.3758/BF03194055

Lewis, J. R. (1989). Pairs of Latin squares to counterbalance sequential effects and pairing of conditions and stimuli. In *Proceedings of the Human Factors Society 33rd Annual Meeting* (pp. 1223–1227). Santa Monica, CA: Human Factors Society.

Loftus, G. R. (1985). Evaluating forgetting curves. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 397–406. doi:10.1037/0278-7393.11.2.397

Malmberg, K. J., & Shiffrin, R. M. (2005). The "one-shot" hypothesis for context storage. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 322–336. doi:10.1037/0278-7393.31.2.322

Melton, A. W. (1967, October). Repetition and retrieval from memory. *Science, 158,* 532. doi:10.1126/science.158.3800.532-b

Melton, A. W. (1970). The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning & Verbal Behavior, 9,* 596–606. doi:10.1016/S0022-5371(70)80107-4

Murdock, B. B. (1960). The immediate retention of unrelated words. *Journal of Experimental Psychology, 60,* 222–234. doi:10.1037/h0045145

Nelson, T. O., & Leonesio, R. J. (1988). Allocation of self-paced study time and the "labor-in-vain effect." *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 676–686. doi:10.1037/0278-7393.14.4.676

Peterson, L. R., Wampler, R., Kirkpatrick, M., & Saltzman, D. (1963). Effect of spacing presentations on retention of a paired associate over short intervals. *Journal of Experimental Psychology, 66,* 206–209. doi:10.1037/h0046694

Pyc, M. A., & Rawson, K. A. (2010, October). Why testing improves memory: Mediator effectiveness hypothesis. *Science, 330,* 335. doi:10.1126/science.1191465

Raaijmakers, J. G. W. (2003). Spacing and repetition effects in human memory: Application of the SAM model. *Cognitive Science, 27,* 431–452. doi:10.1207/s15516709cog2703_5

Rawson, K. A., & Kintsch, W. (2005). Rereading effects depend on time of test. *Journal of Educational Psychology, 97,* 70–80. doi:10.1037/0022-0663.97.1.70

Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17,* 249–255. doi:10.1111/j.1467-9280.2006.01693.x

Roediger, H. L., III, Balota, D. A., & Watson, J. M. (2001). Spreading activation and arousal of false memories. In H. L. Roediger, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 95–115). Washington, DC: American Psychological Association.

Stubin, E. J., Heimer, W. I., & Tatz, S. J. (1970). Total time and presentation time in paired-associate learning. *Journal of Experimental Psychology, 84,* 308–310. doi:10.1037/h0029096

Wixted, J. T. (2004). On common ground: Jost's (1897) law of forgetting and Ribot's (1881) law of retrograde amnesia. *Psychological Review, 111,* 864–879. doi:10.1037/0033-295X.111.4.864

Wixted, J. T., & Carpenter, S. K. (2007). The Wickelgren power law and the Ebbinghaus savings function. *Psychological Science, 18,* 133–134. doi:10.1111/j.1467-9280.2007.01862.x

Wixted, J. T., & Ebbesen, E. B. (1991). On the form of forgetting. *Psychological Science, 2,* 409–415. doi:10.1111/j.1467-9280.1991.tb00175.x