Made available courtesy of teh American Psychological Association: http://www.apa.org/pubs/journals/xlm/index.aspx

Rehearsal Strategies Can Enlarge or Diminish the Spacing Effect : Pure Versus Mixed Lists and Encoding Strategy*

By: Peter F. Delaney and Peter P. J. L. Verkoeijen

Delaney, P. F., & Verkoeijen, P. P. J. L. (2009). Rehearsal strategies can enlarge or diminish the spacing effect: Pure versus mixed lists and encoding strategy. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *35*, 1148-1161. DOI: 10.1037/a0016380

This article may not exactly replicate the final version published in the APA journal. It is not the copy of record.

*****Note: Figures may be missing from this format of the document**

Abstract:
Using 5 experiments, the authors explored the dependency of spacing effects on rehearsal patterns. Encouraging rehearsal borrowing produced opposing effects on mixed lists (containing both spaced and massed repetitions) and pure lists (containing only one or the other), magnifying spacing effects on mixed lists but diminishing spacing effects on pure lists. Rehearsing with borrowing produced large spacing effects on mixed lists but not on pure lists for both free recall (Experiment 1) and recognition (Experiment 2). In contrast, rehearsing only the currently visible item produced spacing effects on both mixed lists and pure lists in free recall (Experiment 3) and recognition (Experiment 4). Experiment 5 demonstrated these effects using a fully within-subjects design. Rehearse-aloud protocols showed that rehearsal borrowing redistributed study from massed to spaced items on mixed lists, especially during massed presentations.
Keywords: rehearsal, spacing effect, long-term memory, list strength effect, output interference

Article:
An important question in memory research is how people process repetitions of the same item on multiple occasions. Repeated items may occur either immediately in succession (called *massed* repetitions) or with other items intervening between them (called *spaced* repetitions). The memory advantage of spaced repetitions over massed repetitions is well known and is termed the *spacing effect*. The spacing effect has been widely replicated by researchers, using many different memory tests, under both incidental and intentional encoding conditions (for reviews, see Dempster, 1996; Greene, 1989; Raaijmakers, 2003; for meta-analyses, see Donovan & Radosevich, 1999; Janiszewski, Noel, & Sawyer, 2003). However, a complete explanation of the cause of the spacing effect has remained elusive.

The current article is focused on the importance of rehearsal borrowing in spacing effect studies. Our studies were not designed to argue that rehearsal is the cause of spacing effects but rather to examine how rehearsal instructions affect the magnitude of the spacing effect in different types of designs. Specifically, we suggest that the most commonly used rehearsal strategy has different effects on pure-list and mixed-list spacing designs.

The Size of the Spacing Effect Depends on Encoding Strategy

* Acknowledgement: This study was supported in part by an Erasmus University TopTalent grant to Peter P. J. L. Verkoeijen. We are grateful to Tiffiny Alexander, Jennifer Brown, Liliana Camara, Heather Coby, Carolina Collado, Jared Cook, Krycya Flores-Rojas, Evan Holst, Brittany Hopkins, Jennifer Hynes, Satish Kunisi, Elizabeth Lancaster, Torcia Lee, Zachary Morrison, Frances “Frankie” Powell, Justin Rundle, Catherine Sawyer, Teisha Seabrook, Andrew Simpson, Brett Sites, Patricia Villafaña, Virginia “Ginny” Wilmoth, and Vincent Woolfolk for their help in collecting the U.S. data; to Tom Zandwijk for his assistance in collecting the data from the Netherlands; and to Tiffiny Alexander, Dalia Avidor, Brian Cline, Sky Fillipini, Brittany Fulbright, Jenna Haddock, Travis Johnson, Goldie Kaufenberg, Melissa Lehman, Andrew LeRoux, Zachary Morrison, Frances “Frankie” Powell, and Rosa “Roxy” Toledo for help transcribing and coding rehearse-aloud protocols. We thank Thomas C. Toppino for providing his word lists, and Lili Sahakyan for helpful comments on a draft of the article.

In a recent article, Delaney and Knowles (2005) argued that one factor contributing to the difficulty of isolating the causes of the spacing effect is that most studies have used intentional learning procedures that do not control encoding strategies. They found that 70% of their participants initially approached list-learning experiments by using a *rehearse-together strategy*—which in other articles we have called a rote rehearsal strategy (see also Sahakyan & Delaney, 2003). The rehearse-together strategy involves reading new items as they appear and using any extra time to rehearse previously studied items.

Delaney and Knowles (2005) had people study several lists, each of which was followed by a free recall test. The lists consisted either of all spaced repetitions or of all massed repetitions. In their first experiment, participants were free to study the lists using any method. Most frequently, participants first adopted the rehearse-together strategy first. However, after encountering several lists with a test after each, rehearse-together strategies were often abandoned for more effective relational encoding strategies, such as making up a story with to-be-learned items as they practiced studying and being tested on lists (see also Sahakyan & Delaney, 2003; Sahakyan, Delaney, & Kelley, 2004). Participants who used the rehearse-together strategy throughout the entire experiment showed a negligible spacing benefit, whereas participants who switched to a relational encoding strategy showed reliable spacing effects. In a second experiment, Delaney and Knowles controlled the encoding strategy participants used and again found that the rehearse-together strategy resulted in no overall spacing benefits, whereas using the story mnemonic produced a spacing effect. Detailed analyses of how different study strategies impact the magnitude of the spacing effect therefore seem warranted. Furthermore, spacing studies can yield inconsistent results when participants choose their own study strategies, suggesting that researchers should ideally track or control encoding strategies.

Rehearsal, Rehearsal Borrowing, and Spacing Effects: The Controversy

When people expect a test, additional rehearsal time enhances learning. Rundus (1977), for example, asked people to study a number; then rehearse a word for 4, 8, or 12 s; and then recall the number. Additional rehearsal time had no effect on a later free-recall test when participants did not expect the test, but additional rehearsal time improved recall when participants expected a test. In a list-learning study, when participants rehearse word lists aloud to learn them, there is a close correspondence between the number of rehearsals an item receives and its probability of recall on a subsequent free-recall test (Rundus, 1971). Modigliani and Hedges (1987) later showed that the largest benefits of rehearsal came from *distributed rehearsals*, which were defined as rehearsals of the same item occurring more than 5 s apart. Even a single distributed rehearsal boosted the probability of recalling an item by 46%. (It is possible that only very memorable items would receive such a delayed rehearsal, which might inflate this relationship. However, for evidence against this interpretation, see Murdock & Metcalfe, 1978).

Memory researchers have long suspected a mediating role for rehearsal in the spacing effect. On one hand, the presence of large spacing effects following incidental learning or using orienting tasks that discourage rehearsal (e.g., Glenberg & Smith, 1981; Greene, 1989, Experiment 1; Jensen & Freund, 1981; Rose & Rowe, 1976; Toppino & Bloom, 2002), and following intentional learning in populations that typically do not rehearse, such as younger children (Rea & Modigliani, 1987; Toppino, 1991), implies that rehearsal cannot fully account for the spacing effect. On the other hand, there is also evidence that rehearsal may be an important cause of spacing effects in other circumstances. Rundus (1971) first proposed that the spacing effect could be explained via rehearsal patterns, with spaced repetitions receiving more rehearsal than massed repetitions. Because people rehearse spaced items more often, they should be stronger and therefore better recalled than massed items. Consistent with his argument, Rundus (1971, Experiment 3) conducted a free recall study in which he presented participants with a mixed list that contained some massed and some spaced items on the same list. People were asked to rehearse out loud, and the number of rehearsals of spaced and massed items during their second presentation was coded. Consistent with his rehearsal hypothesis, Rundus found that participants rehearsed the spaced items more frequently than the massed items, leading to a robust spacing effect. Other related research has shown that when rehearsal is controlled by instructing participants to repeat each item in isolation out loud, the spacing effect in free recall is attenuated or eliminated (Glenberg, 1977, Experiment 2; Wright & Brelsford, 1978).

A controversial article by Hall (1992) proposed that rehearsal patterns are the sole determinant of the spacing effect in intentional learning. In his Experiment 3, Hall asked participants to study three types of lists—one of which was a mixed list (containing both spaced items and massed items), and the other two types were pure lists (one consisting only of massed items and one consisting only of spaced items). Hall suspected that the spacing effects on mixed lists emerged because people redistributed their rehearsal time away from massed items in favor of spaced items—his *rehearsal borrowing hypothesis*. If so, he reasoned, then using pure lists should eliminate the ability to redistribute rehearsal time from one type of item to another (because the items were no longer on the same list). Consistent with his predictions, when he gave participants pure lists of spaced or massed items, he indeed observed no spacing effects. This finding was consistent across three different presentation rates (1 s per item, 2.5 s per item, and 5 s per item). Similar null results with pure lists were obtained using recognition testing by Greene (1990, Experiment 4) and cued-recall by Greeno (1970). However, Hall's main results were not replicated by Toppino and Schneider (1999) or Kahana and Howard (2005), with both studies obtaining a significant spacing effect in free recall using pure lists.

Our hypothesis in the current article is that the rehearse-together strategy serves two opposite functions in different spacing designs. First, as Modigliani and Hedges (1987) suggested, the rehearse-together strategy creates distributed rehearsals of items on the list. We therefore expect that the rehearse-together strategy should convert massed presentations into functionally spaced presentations, reducing the size of the spacing effect. On pure lists containing only spaced or only massed repetitions, the rehearse-together strategy would therefore tend to reduce the magnitude of the spacing effect. As the number of spaced repetitions increases, there are diminishing returns such that each repetition provides a smaller memory advantage than the repetition before (e.g., Greene, 1989). Therefore, if the rehearse-together strategy converts massed items into functionally spaced items, it would tend to benefit the massed items (because of adding a few spaced repetitions) but would tend to have less of an effect on items that were already spaced (because additional spaced repetitions show diminishing returns). The ultimate result would be to diminish the spacing effect when people use the rehearse-together strategy to learn pure lists relative to conditions that discourage rehearsal borrowing.

Second, as Rundus (1971) suggested, when massed and spaced items are presented together on the same list, the rehearse-together strategy serves to redistribute study time away from massed items and to spaced items. The result of this rehearsal redistribution is to enhance the impact of the spacing effect. This is essentially the prediction of Hall's (1992) rehearsal borrowing hypothesis but applied to the special case in which participants rely on the rehearse-together strategy to learn the lists.

In Experiments 1 and 2, we explored how rehearsal redistributes study time on mixed lists. In Experiments 3 and 4, we asked whether a kind of rehearsal that eliminates borrowing—rehearsing only the currently visible item, which we termed the *rehearse-alone strategy*—would have the same effect as the rehearse-together strategy. To preview our results, we obtained broad support for our argument that spacing effects emerge on both pure and mixed lists when rehearsal borrowing is absent; however, when rehearsal borrowing is present, then the magnitude of the spacing effect depends on whether pure- or mixed-list designs are used.

Experiments 1A and 1B

Many spacing studies mix massed items and spaced items on the same list. Experiment 1 was designed to compare the magnitude of spacing effects on pure lists of spaced or massed items with the magnitude of spacing effects on mixed lists containing some spaced and some massed items. Participants studied one mixed list, one pure-spaced list, and one pure-massed list, receiving a free recall test after each list. We controlled encoding strategy by asking participants to rehearse the words aloud over and over, adding more words to their rehearsal set as they moved through the list.

We expected that rehearsal borrowing would occur, magnifying the spacing effect on mixed lists relative to the pure lists. Specifically, our prediction was based on Rundus's (1971) results, which showed that on mixed lists the rehearse-together strategy resulted in more rehearsals of spaced items compared with massed items. This could create a larger spacing effect because of redistribution of practice away from massed items and toward spaced items—which is essentially the prediction of Hall's (1992) rehearsal borrowing hypothesis. This experiment therefore constitutes a test of Hall's account under the conditions that constrained the encoding strategy to the rehearse-together strategy (people's most commonly chosen strategy the first time they study lists).

Method

Participants

Participants in Experiment 1A were 36 University of Florida students who participated for course credit. They were tested individually. An additional 14 University of North Carolina at Greensboro (UNCG) students participated in Experiment 1B.

Materials and counterbalancing

We created three word sets consisting of 32 nouns each using the materials from Toppino and Schneider (1999) and additional medium-frequency words selected from Francis and Kucera's (1982) norms. Three types of lists were constructed for each word set—a mixed list, a pure-massed list, and a pure-spaced list. On each type of list, all 32 words from a word set appeared twice, yielding a total of 64 presentations. The *mixed* form of the list had four massed and four spaced words in each list quadrant. A second version of each mixed list was created for counterbalancing purposes such that one list began with a massed item, and one version began with spaced items. To construct the *pure-massed* forms of the list, we left the massed items from the mixed list in the same serial positions, and we replaced the spaced items with massed items. To construct the *pure-spaced* forms of the list, we left the spaced items from the mixed list in the same serial positions, and then we replaced the massed items with spaced items. In this fashion, we could compare the same words at the same serial positions on the pure lists and the mixed list.

The order of words on each list was constrained so that obvious associates were not near one another. Spaced repetitions were separated by an average of seven words.

Procedure

The only difference between Experiments 1A and 1B was that in Experiment 1A, we did not record the rehearse-aloud protocols. The 14 participants in Experiment 1B were added to analyze a sample of concurrent rehearse-aloud protocols.

Each participant studied three types of lists—one mixed list, one pure-spaced list, and one pure-massed list. Order of the three list types was fully counterbalanced, and we rotated the order of the three word sets to ensure that every word set appeared equally often in each position. During study, the words appeared one by one in large black print on the computer screen for 2 s, with 1 s between presentations. Participants were instructed to study by rehearsing the words out loud, reading each word as it appeared and then using any remaining time to rehearse the word together with earlier items. They were told not to become alarmed if they forgot some of the words and just to keep rehearsing what they remembered. They were given some practice with this procedure to ensure that they understood the instructions, and an experimenter monitored compliance.

After studying the first list, participants solved three-digit by two-digit multiplication problems for 60 s, and then they had 3 min to write down as many items as they could remember from the list (i.e., a free recall test). The second and third lists were tested in the same fashion.

Results and Discussion

We report the main results using only participants from Experiment 1A (so that an equal number of participants were in each counterbalancing cell for the analyses). The rehearse-aloud protocols for Experiment 1B are then discussed.

The crucial prediction of the rehearsal borrowing hypothesis was that massed items would be recalled less well on mixed lists than pure lists, but spaced items would be recalled better on mixed lists than pure lists (presumably because borrowing would redistribute study time away from massed items to spaced items). To test this prediction, we performed a List Type (mixed vs. pure) × Spacing (spaced vs. massed) within-subjects analysis of variance (ANOVA) on the proportion of words recalled from each list. Figure 1 shows the mean recall rates. We observed no main effect of list type, $F(1, 35) = 1.32$, $MSE = 0.012$, $\eta^2 = .036$, which indicates that neither mixed nor pure lists enjoyed a recall advantage. There was a main effect of spacing, $F(1, 35) = 19.05$, $MSE = 0.015$, $p < .001$, $\eta^2 = .352$, and it was qualified by a significant List Type × Spacing interaction, $F(1, 35) = 18.93$, $MSE = 0.014$, $p < .001$, $\eta^2 = .351$, indicating that the spacing effect was not of comparable magnitude on each type of list.

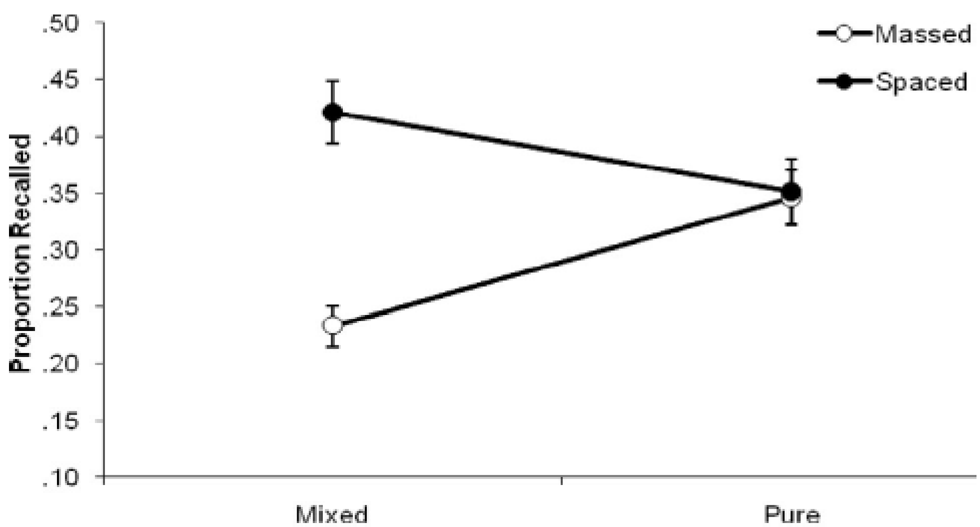

Figure 1. Proportion of massed and spaced words recalled from mixed and pure lists, Experiment 1A. Error bars represent ±*SE*s.

The follow-ups to this interaction test predictions of the rehearsal borrowing hypothesis. The first prediction was that spacing effects should emerge on mixed lists but not on pure lists. Consistent with this prediction, paired *t*-tests revealed a significant spacing advantage on the mixed list, $t(35) = 6.26$, $p < .001$, but not on the pure lists, $t < 1$. A second prediction of the rehearsal borrowing hypothesis is that mixing spaced and massed items on the same list should enhance recall of spaced items and reduce recall of massed items relative to pure lists (as people rehearse spaced items at the expense of massed items). This prediction was also supported, as spaced items were recalled *more* frequently on the mixed lists than on the pure lists, $t(35) = 2.34$, $p < .05$, whereas massed items were recalled *less* frequently on the mixed lists than on the pure lists, $t(35) = 4.10$, $p < .001$. In sum, both predictions of the rehearsal borrowing hypothesis were confirmed.

Enhanced primacy analyses

Previous studies have generally found that when people learn lists using primarily the rehearse-together strategy methods, there is a *reverse spacing effect* in the beginning of the lists (e.g., Delaney & Knowles, 2005; Hall, 1992; Toppino & Schneider, 1999). The reverse spacing effect refers to an advantage for massed items over spaced items at the start of the lists. The reverse spacing effect in the first portion of the list is usually balanced by a smaller spacing advantage throughout the rest of the list, and for this reason is also called the *enhanced primacy effect*.

Because we fixed the order of words on the lists, serial position analyses are not fully appropriate given that items are confounded with serial position. However, in the interests of ensuring that we replicated earlier studies, we performed a 4 Quadrant × 2 Spacing (spaced vs. massed) within-subjects ANOVA for the pure lists on proportion recall. [1] Quadrant-by-quadrant means are given in Table 1. There was a reliable effect of quadrant indicating that some parts of the list were recalled better than others, $F(3, 105) = 16.05$, $MSE = 0.060$, $p < .001$, $\eta^2 = .314$. There was no overall spacing effect, $F < 1$, $\eta^2 = .007$, consistent with our earlier analyses. Most importantly, there was a significant Quadrant × Spacing interaction, $F(3, 105) = 2.92$, $MSE = 0.033$, $p < .05$, $\eta^2 = .077$, indicating that the spacing effect was of different magnitudes in different parts of the lists. We followed up the interaction with four paired *t*-tests comparing the spaced with massed items in each of the four quadrants. There was no spacing effect in Quadrants 2 or 3, both *t*s < 1. There was a reliable reverse spacing effect in Quadrant 1, $t(35) = 3.07$, $p < .05$, replicating earlier studies. Although it was not significant, there was also a marginal spacing effect in Quadrant 4, $t(35) = 1.87$, $p = .07$.

Table 1
Proportion of Words Recalled by Quadrant, Experiment 1A

	List quadrant			
List type	First	Second	Third	Fourth
Pure-massed	0.50 (0.04)	0.19 (0.03)	0.29 (0.04)	0.35 (0.04)
Pure-spaced	0.40 (0.04)	0.19 (0.03)	0.29 (0.05)	0.42 (0.04)
Mixed-massed	0.32 (0.03)	0.12 (0.03)	0.15 (0.04)	0.33 (0.05)
Mixed-spaced	0.56 (0.06)	0.28 (0.04)	0.28 (0.04)	0.48 (0.06)

Note. Values in parentheses represent ±*SE*s.

Taken together, the results are consistent with earlier research showing an enhanced primacy effect on massed lists compared with spaced lists, and a small spacing effect in the rest of the list.

For completeness, a Quadrant × Spacing analysis for the mixed lists yielded a significant spacing effect, $F(1, 35) = 39.84$, $MSE = 0.051$, $p < .001$, $\eta^2 = .532$, and a main effect of quadrant, $F(3, 105) = 15.29$, $MSE = 0.073$, $p < .001$, $\eta^2 = .304$. There was no interaction, $F < 1$.

Output order by spacing

The enhanced primacy effect could be the result of different output order biases for massed lists than for spaced lists. For example, if massed lists led people to output Quadrant 1 items first and spaced lists led people to output Quadrant 4 items first, we would observe enhanced primacy on massed lists and enhanced recency on spaced lists. To address whether enhanced primacy could be caused by output order biases, we computed output percentiles on the basis of a method used by Bjork and Whitten (1974). For each person, we calculated eight output percentiles—one for each list quadrant on the pure-spaced list, and one for each quadrant on the pure-massed list. To calculate an output order, one should first assign each recalled word a nominal position starting from 1. Thus, the first word is Word 1, the second recalled word is Word 2, and so on. Then, for a given list quadrant, the nominal positions of all of the words from that quadrant that were recalled are summed. The output percentile is the sum of the nominal positions of words from that quadrant divided by the number of words on the list. In this fashion, higher output percentiles reflect later average output of words from that list quadrant. Table 2 gives descriptive statistics on output percentiles by list and quadrant (including the mixed lists).

Table 2
Output Percentiles by Quadrant and List Type, Experiment 1A

List quadrant	Massed lists		Spaced lists		Mixed lists	
	Percentile	*N*	Percentile	*N*	Percentile	*N*
First	45.0 (16.8)	34	39.0 (18.8)	31	41.6 (20.6)	34
Second	64.9 (20.6)	28	64.2 (20.5)	24	65.5 (22.3)	29
Third	67.6 (17.9)	30	66.8 (15.4)	27	67.1 (12.8)	25
Fourth	62.3 (22.1)	33	64.9 (20.5)	35	64.0 (16.8)	34

Note. Values in parentheses reflect *SD*s.

To determine whether output biases could explain the enhanced primacy effect for massed items, we compared Quadrant 1 output percentiles for pure-spaced and pure-massed lists using a paired *t*-test. (Data from 4 people with missing values were excluded.) We observed no differences, $t(31) = 1.21$, $p = .23$. It seems unlikely that the enhanced primacy effect was due to output bias, because if it were, we would have expected to see Quadrant 1 items being output earlier on the pure-massed list than on the pure-spaced list; in fact, the trend was toward the reverse pattern. On both list types, people tended to output Quadrant 1 items early.

There were also no differences between pure-massed and pure-spaced lists in any other quadrant, all *t*s < 1. Output bias did not seem to account for the enhanced primacy effect on pure-massed lists.

A final question was whether on mixed lists people tended to output the spaced items before the massed items. The strength advantage of spaced items tends to produce an output bias such that they are output first (although other ways of strengthening items, such as deeper processing instructions, do not affect output order; see Sahakyan, Delaney, & Waldum, 2008). We therefore calculated output percentiles for spaced versus massed items on the mixed lists. Two participants who recalled only one type of item were excluded from analysis. The output percentiles showed that spaced items were output earlier ($M = 0.52$, $SD = 0.07$) than massed items ($M = 0.59$, $SD = 0.13$), $t(33) = 2.12$, $p < .05$. It is therefore possible that output interference could partially explain the larger advantage of spaced over massed items on mixed lists.

Experiment 1B: Rehearse-aloud protocols

To ensure that our results replicated Experiment 1A, we conducted a List Type (mixed vs. pure) × Spacing (spaced vs. massed) within-subjects ANOVA on the proportion of words recalled from each list. There was a main effect of spacing, $F(1, 13) = 26.53$, $MSE = 0.01$, $p < .001$, $\eta^2 = .671$, and no effect of list type, $F < 1$. As in Experiment 1A, the interaction was significant, $F(1, 13) = 11.36$, $MSE = 0.01$, $p < .01$, $\eta^2 = .466$. On mixed lists, spaced words ($M = 0.41$, $SD = 0.15$) were recalled better than massed words ($M = 0.21$, $SD = 0.08$), $t(13) = 4.37$, $p < .001$. However, on pure lists, the spaced words ($M = 0.35$, $SD = 0.09$) were recalled at about the same rate as the massed words ($M = 0.33$, $SD = 0.11$), $t < 1$.

The rehearsal-borrowing hypothesis suggests that people rehearse previously studied spaced items more often than previously studied massed items. Rundus (1971) found that spaced items are generally rehearsed more often than massed items. To replicate his result, we transcribed the rehearse-aloud protocols from mixed lists for 10 participants selected at random. Beeps were inserted at constant intervals to allow coders to know when a new word was presented. Research assistants then scored the protocols on the basis of what words were rehearsed during the presentation of each individual item. Table 3 shows what type of item people were rehearsing as a function of what was on screen (i.e., how many times they rehearsed the onscreen item, an earlier massed item, or an earlier spaced item while that kind of item was onscreen).

Table 3
Total Number of Rehearsals of the Onscreen Item, Earlier Spaced Items, and Earlier Massed Items by Presentation Number and Item Type on Mixed Lists, Experiment 1B

Rehearsal type	First presentation	Second presentation	Total
Self-rehearsals			
Viewing spaced item	18.3 (2.5)	17.4 (2.8)	35.7
Viewing massed item	15.9 (2.3)	15.1 (1.2)	31.0
Total	34.2	32.5	66.7
Rehearsals of spaced items			
Viewing spaced item	21.0 (3.6)	20.2 (2.5)	41.2
Viewing massed item	21.3 (2.9)	22.8 (2.4)	44.1
Total	42.3	43.0	85.3
Rehearsals of massed items			
Viewing spaced item	11.7 (1.1)	11.3 (1.9)	23.0
Viewing massed item	11.9 (1.7)	12.4 (1.5)	24.3
Total	23.6	23.7	47.3

Note. Values in parentheses represent ±*SE*s.

With regard to rehearsing the item on screen (self-rehearsals), a 2 Item Type (spaced vs. massed) × 2 Presentation (first vs. second) within-subjects ANOVA on the number of self-rehearsals was conducted. We found no differences across item types, $F(1, 9) = 1.17$, $MSE = 13.35$, $p = .30$, $\eta^2 = .115$; or across presentation number, $F(1, 9) = 1.68$, $MSE = 3.35$, $p = .23$, $\eta^2 = .157$; and no interaction, $F(1, 9) = 1.37$, $MSE = 5.28$, $p = .27$, $\eta^2 = .132$. As in Experiment 1B, we obtained no evidence for a massed item attention deficit.

Turning to rehearsal of previously studied items, which addresses rehearsal borrowing, we conducted a 2 Rehearsal Type (spaced item rehearsal vs. massed item rehearsal) × 2 Item Type (looking at massed item vs. looking at spaced item) × 2 Presentation (first vs. second) within-subjects ANOVA on how many times people rehearsed previously studied items. The interesting results involved how people allocated their rehearsals to previously studied items. As in Experiment 1B, there was a main effect of rehearsal type, $F(1, 9) = 27.18$, $MSE = 138.11$, $p < .001$, $\eta^2 = .751$, indicating that people rehearsed spaced items about twice as often as massed items. The greater overall rehearsal of spaced items did not depend on whether people were looking at the first or second presentation, $F(1, 9) = 1.31$, $MSE = 27.87$, $p = .28$, $\eta^2 = .127$. However, there was an Item Type × Rehearsal Type interaction, $F(1, 9) = 35.13$, $MSE = 15.99$, $p < .001$, $\eta^2 = .796$. Although people always rehearsed spaced items more than massed items, the borrowing was largest when viewing a massed item. Compared with rehearsal during a spaced item, rehearsal during a massed item led to greater rehearsal of earlier spaced items, $t(9) = 5.65$, $p < .001$, and lesser rehearsal of massed items, $t(9) = 3.35$, $p < .01$. These results are consistent with a strategy of rehearsing spaced items at the expense of massed items, particularly during the presentation of massed items. That could produce lower recall and recognition of massed items than of spaced items. None of the other main effects or interactions were significant, all Fs < 1.

Next we performed a 2 List Type (massed vs. spaced) × 2 Presentation (first vs. second) within-subjects ANOVA on the number of displaced rehearsals on pure lists. People rehearsed earlier-presented items on massed lists ($M = 88.3$) about as often as on spaced lists ($M = 90.3$), $F < 1$. There was a tendency to rehearse more previously studied items on the first presentation ($M = 93.0$) than on the second presentation ($M = 85.6$),

$F(1, 9) = 6.71$, $MSE = 82.75$, $p < .05$, $\eta^2 = .427$. The interaction was not significant, $F(1, 9) = 1.09$, $MSE = 109.53$, $\eta^2 = .108$.

Conclusions

The results are similar to Experiment 1: Using the rehearse-together instructions, there was a small, nonsignificant spacing effect on pure lists and a large spacing effect on mixed lists. Also, compared with the recognition of mixed-list items, massed items were more frequently recognized when presented on a pure list, whereas spaced items were less frequently recognized when presented on a pure list. Furthermore, the rehearse-aloud protocols suggested that the large spacing effect on mixed lists was associated with substantial rehearsal borrowing such that massed items received less practice than spaced items. In contrast, massed and spaced items were rehearsed approximately equally often on pure lists.

Experiment 3

Experiment 1 and 2 both obtained much larger spacing effects on mixed lists than on pure lists when people studied using rehearse-together instructions. We argued that rehearsal works to diminish spacing effects on pure lists and to magnify spacing effects on mixed lists. We attributed these results to borrowing, but they could be due merely to the use of rehearsal strategies per se. We therefore conducted an experiment in which participants were given rehearse-alone instructions—that is, they were instructed to rehearse only the item that was currently visible on the screen. If borrowing was the cause of the pattern of results in Experiments 1 and 2, we would expect to find that when borrowing was not allowed we should obtain a significant spacing effect on pure lists but a smaller spacing effect than before on mixed lists.

Method

Participants

Participants were 36 UNCG students who received course credit.

Materials

The same lists (pure/spaced, pure/massed, and mixed) were used as in Experiment 3.

Procedure

The procedure was similar to that of Experiment 1, with three lists and free recall testing. However, participants in Experiment 3 were instructed to rehearse only the item that was currently visible on the screen and not to practice earlier viewed items. The rehearse-aloud protocols were recorded to verify compliance with the instructions and to count rehearsals.

Results and Discussion

As in Experiments 1A and 1B, we conducted a List Type (pure vs. mixed) × Item Type (spaced vs. massed) interaction. A within-subjects ANOVA obtained a main effect of item type, $F(1, 35) = 11.70$, $MSE = 0.010$, $p < .01$, $\eta^2 = .25$, with better recall of spaced items (0.34) than massed items (0.29). There was also a main effect of list type, $F(1, 35) = 8.83$, $MSE = 0.008$, $p < .01$, $\eta^2 = .20$, reflecting unexpectedly better recall on pure lists (0.34) than on mixed lists (0.29). The interaction—shown as Figure 3—was not significant, $F(1, 35) = 2.59$, $MSE = 0.007$, $p = .12$, $\eta^2 = .07$.

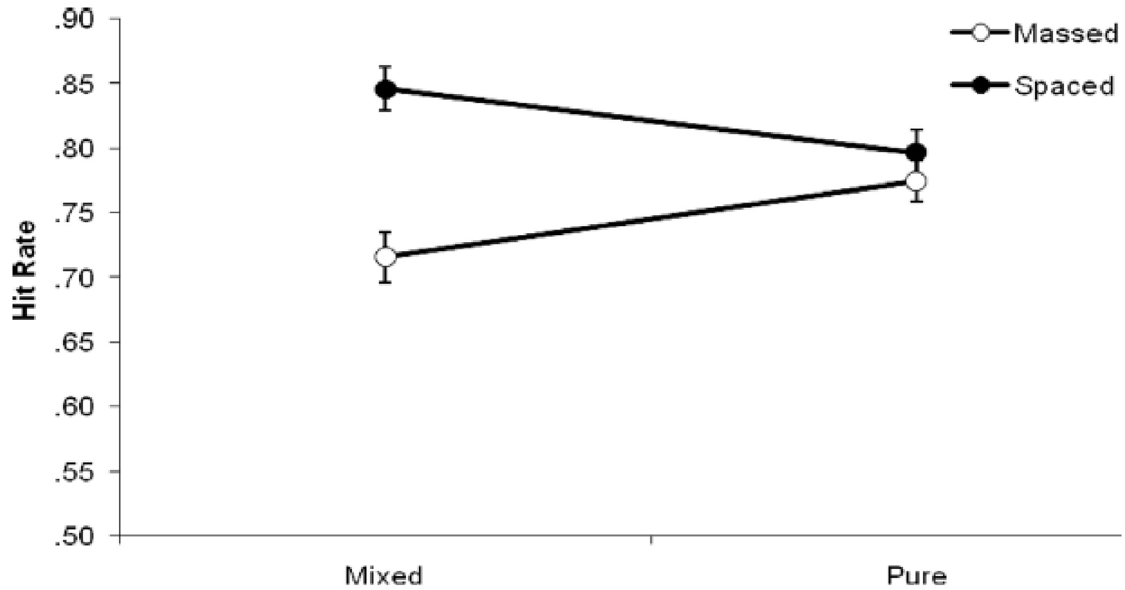

Figure 2. Hit rates for massed and spaced items from mixed and pure lists, Experiment 2. Error bars represent ±*SE*s.

Paired *t*-tests indicated a spacing effect for both pure lists, $t(35) = 2.21$, $p < .05$, and mixed lists, $t(35) = 3.02$, $p < .01$. Spacing effects emerged on both pure lists and mixed lists, in contrast with the rehearse-together strategy results from Experiments 1 and 2. The nonsignificant interaction probably reflected a small list-strength effect combined with a massed item deficit on mixed lists. Compared with pure lists, the massed items suffered on mixed lists, $t(35) = 3.71$, $p < .01$, although there was no difference between spaced items on pure and mixed lists, $t < 1$.

Enhanced primacy analyses

We have argued that the enhanced primacy effect on pure lists is largely the result of rehearsal patterns. If so, we should expect no enhanced primacy on pure/massed lists when we control rehearsal borrowing. A 2 Item Type (spaced vs. massed) × 4 Quadrant repeated-measures ANOVA was conducted (for means, see Table 6). We obtained the main effect of item type, $F(1, 35) = 5.38$, $MSE = 0.018$, $p < .05$, $\eta^2 = .13$, reflecting higher recall of spaced ($M = 0.35$) than massed ($M = 0.31$) items. There was also a main effect of quadrant, $F(3, 105) = 6.21$, $MSE = 0.034$, $p < .001$, $\eta^2 = .15$. Post hoc *t*-tests showed that Quadrants 3 and 4 did not reliably differ, but they both were better recalled than Quadrants 1 and 2, which did not reliably differ. Thus, there was a recency advantage. The interaction was not significant, $F(3, 105) = 1.99$, $MSE = 0.041$, $\eta^2 = .05$.

Table 4
Hit Rate by Quadrant for Pure-Massed and Pure-Spaced Lists, Experiment 2

	List quadrant			
List type	First	Second	Third	Fourth
Massed	0.80 (0.03)	0.73 (0.03)	0.78 (0.03)	0.80 (0.02)
Spaced	0.77 (0.03)	0.80 (0.03)	0.79 (0.03)	0.83 (0.03)

Note. Values in parentheses represent ±*SE*s.

Rehearse-aloud protocols

A sample of 10 participants' rehearse-aloud protocols were selected for transcription. Most participants adopted a strategy of repeating each item the same number of times on each presentation. Starting with the mixed lists, we analyzed the number of self-rehearsals using a 2 Item Type (spaced vs. massed) × 2 Presentation (first repetition vs. second repetition) within-subjects ANOVA. On mixed lists, we obtained no effect of presentation, $F < 1$; no main effect of item type, $F(1, 9) = 2.55$, $MSE = 10.04$, $p = .15$, $\eta^2 = .22$; and no interaction, $F < 1$. The nonsignificant effect of item type favored spaced items (105.8 total rehearsals, summing across both presentations) over massed items (102.5 total rehearsals). It appears that we were reasonably successful in equating the number of rehearsals across item types, as the difference between the two item types amounted to saying three words more during spaced repetitions throughout the entire list.

On pure lists, there was again no main effect of presentation, $F < 1$, or interaction, $F(1, 9) = 1.15$, $MSE = 44.01$, $\eta^2 = .11$. However, there was a marginal main effect of item type, $F(1, 9) = 3.46$, $MSE = 328.06$, $p = .10$, $\eta^2 = .28$. The effect reflected more total rehearsals on spaced lists (221.7 total rehearsals) compared with massed lists (200.4). There was a general tendency for participants to rehearse more often on the spaced lists than on the massed lists, but the real difference is probably smaller than our estimate, because one participant switched from rehearsing each presentation three times (on the pure-massed list) to rehearsing each presentation four times (on both other lists). This participant accounted for roughly half of the spaced list rehearsal rate advantage.

Output order analyses

We conducted the same output percentile analyses as in Experiment 1A. Table 7 shows the output percentiles by list type and quadrant. Output order was relatively similar across all three lists. For the pure lists, there was no difference between the pure-massed and pure-spaced list in any of the four quadrants, all *t*s < 1. The recall rate differences between these lists seemingly did not reflect dramatically different output orders.

Table 7
Output Percentiles by Quadrant and List Type, Experiment 3

	Massed lists		Spaced lists		Mixed lists	
List quadrant	Percentile	*N*	Percentile	*N*	Percentile	*N*
First	54.3 (25.5)	34	54.9 (26.0)	33	52.6 (20.1)	31
Second	62.0 (19.6)	26	60.1 (22.6)	32	64.5 (19.0)	30
Third	58.5 (16.2)	34	57.4 (20.2)	33	56.5 (19.3)	31
Fourth	51.9 (16.1)	33	49.9 (20.4)	35	50.3 (18.1)	34

Note. Values in parentheses reflect *SD*s.

For the mixed lists, there was a significant output bias, $t(34) = 2.33$, $p < .05$. The means were similar to those observed in Experiment 1A, with spaced items ($M = 0.53$, $SD = 0.06$) output slightly earlier than massed items ($M = 0.59$, $SD = 0.09$). Others have obtained an output bias of a similar magnitude favoring spaced items on incidentally learned lists (Sahakyan et al., 2008). On all three lists, the output order was different than in Experiment 1A. People tended to favor starting recall from the end of the list rather than from the beginning when rehearsal was constrained to currently visible item.

Experiment 4

In Experiment 2, we obtained an interaction whereby spacing effects were large for mixed lists but absent for pure lists. We conducted Experiment 4 to confirm the result of Experiment 3 that rehearsing only the current item would produce spacing effects for both pure and mixed lists, but this time we used recognition testing.

Method

Participants

Participants were 36 UNCG students who participated for course credit.

Materials

The same lists (pure/spaced, pure/massed, and mixed) were used as in Experiment 1.

Procedure

The procedure was identical to Experiment 3, except that instead of a free recall test, participants received an old/new recognition test (as in Experiment 2). There were an equal number of distractors and targets on each list.

Results and Discussion

As in Experiment 2, false alarm rates were low. A one-way within-subjects ANOVA obtained no reliable differences between the false alarm rates across the three lists, $F(2, 79) = 2.33$, $MSE = 0.002$, $\eta^2 = .06$, $p > .10$. The false alarm rates were, respectively, 0.04 for spaced lists ($SD = 0.04$), 0.06 for massed lists ($SD = 0.06$), and 0.05 for mixed lists ($SD = 0.06$).

A 2 Spacing (spaced vs. massed) × 2 List Type (mixed vs. pure) within-subjects ANOVA on hit rate obtained only a main effect of spacing, $F(1, 35) = 32.77$, $MSE = 0.007$, $p < .001$, $\eta^2 = .48$. Thus, spaced items ($M = 0.87$) were better recognized than massed items ($M = 0.79$). The main effect of list type was not significant, $F < 1$, and neither was the interaction, $F < 1$. Figure 4 shows hit rates by condition (for comparison with Experiment 2).

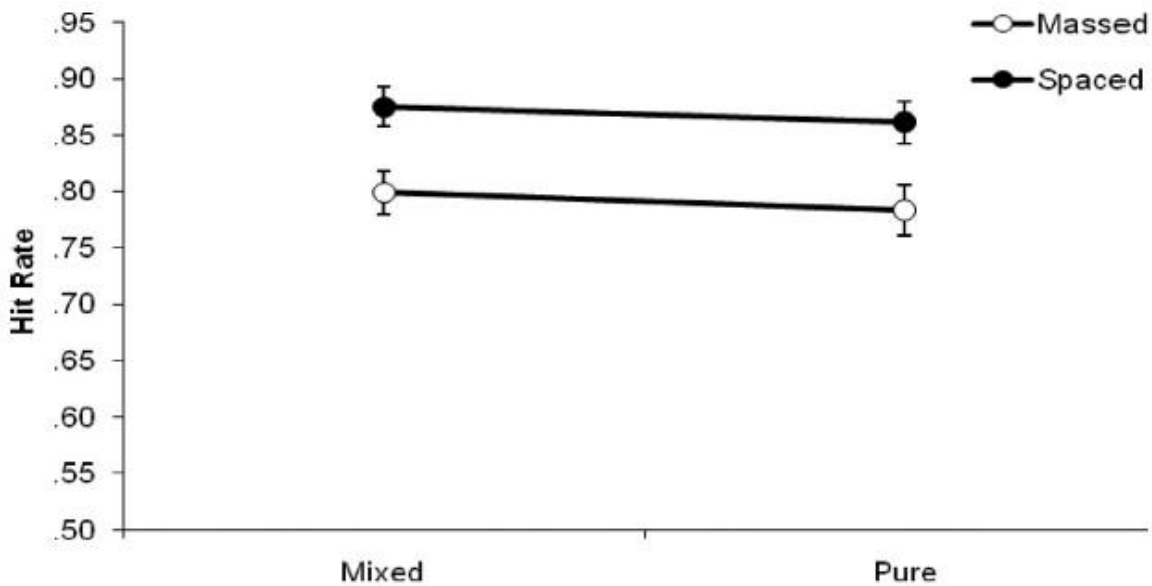

Figure 4. Hit rates for massed and spaced items from mixed and pure lists, Experiment 4. Error bars represent ±*SE*s.

Consistent with Experiment 3 and our predictions, when people rehearsed only the item on screen, they showed a comparable spacing effect for both pure and mixed lists.

Experiment 5

Experiment 5 combined the recognition procedures used in Experiments 2 and 4 within a single experimental design. Taken together, our previous experiments implied a Spacing × List Type × Study Instruction interaction should emerge in recognition. Specifically, we expected to obtain different results with rehearse-alone and rehearse-together strategies. Rehearse-alone should produce equivalent spacing effects on both pure and mixed lists, as in Experiments 3 and 4. Rehearse-together, on the other hand, should produce no spacing effect on pure lists and a large spacing effect on mixed lists, as in Experiments 1 and 2.

Method

Participants

Participants were 48 UNCG students who participated for course credit.

Materials

We supplemented the three lists from Experiment 1 with three new 32-word lists (and their corresponding distracter items) constructed using words drawn from published word norms. The mixed, pure-spaced, and pure-massed versions of these lists were constructed according to the same scheme as in Experiment 1.

Procedure

The procedure was similar to Experiment 2, except that participants received six lists rather than three. Half of the participants received the Experiment 2 rehearse-together instructions for their first three lists and the rehearse-alone instructions from Experiment 4 for their last three lists. The other participants saw the rehearse-alone instructions first followed by the rehearse-together instructions.

Results and Discussion

False alarms

False alarms were generally low, as in our earlier experiments. A 3 List (pure/spaced, pure/massed, mixed) × 2 Instruction (rehearse-alone vs. rehearse-together) ANOVA on the number of false alarms resulted in marginally significant main effects of list, $F(2, 94) = 2.61$, $MSE = 0.002$, $p = .08$, $\eta^2 = .05$, and instruction, $F(1, 47) = 3.08$, $MSE = 0.007$, $p = .09$, $\eta^2 = .06$. However, these marginal effects were qualified by a significant interaction, $F(2, 94) = 3.93$, $MSE = 0.002$, $p < .05$, $\eta^2 = .08$. Follow up paired *t*-tests showed that the effect reflected more errors on pure/spaced lists when people rehearsed together than on some of the other lists (see Table 8). No other comparisons were significant.

Table 8
False Alarm Rates by Condition, Experiment 5

Condition	Pure/spaced	Pure/massed	Mixed
Rehearse-together	0.07 (0.09)	0.06 (0.06)	0.04[a] (0.05)
Rehearse-alone	0.04[a] (0.05)	0.05 (0.05)	0.04[a] (0.06)

Note. Values in parentheses represent *SD*s.
[a] Significantly different from the pure/spaced, rehearse-together condition ($p < .05$).

Hit rates

Our next question was whether the predicted three-way interaction would emerge. Including whether rehearse-alone or rehearse-together was first produced no main effect and no interactions, so we collapsed over order. Means are given in Figure 5.

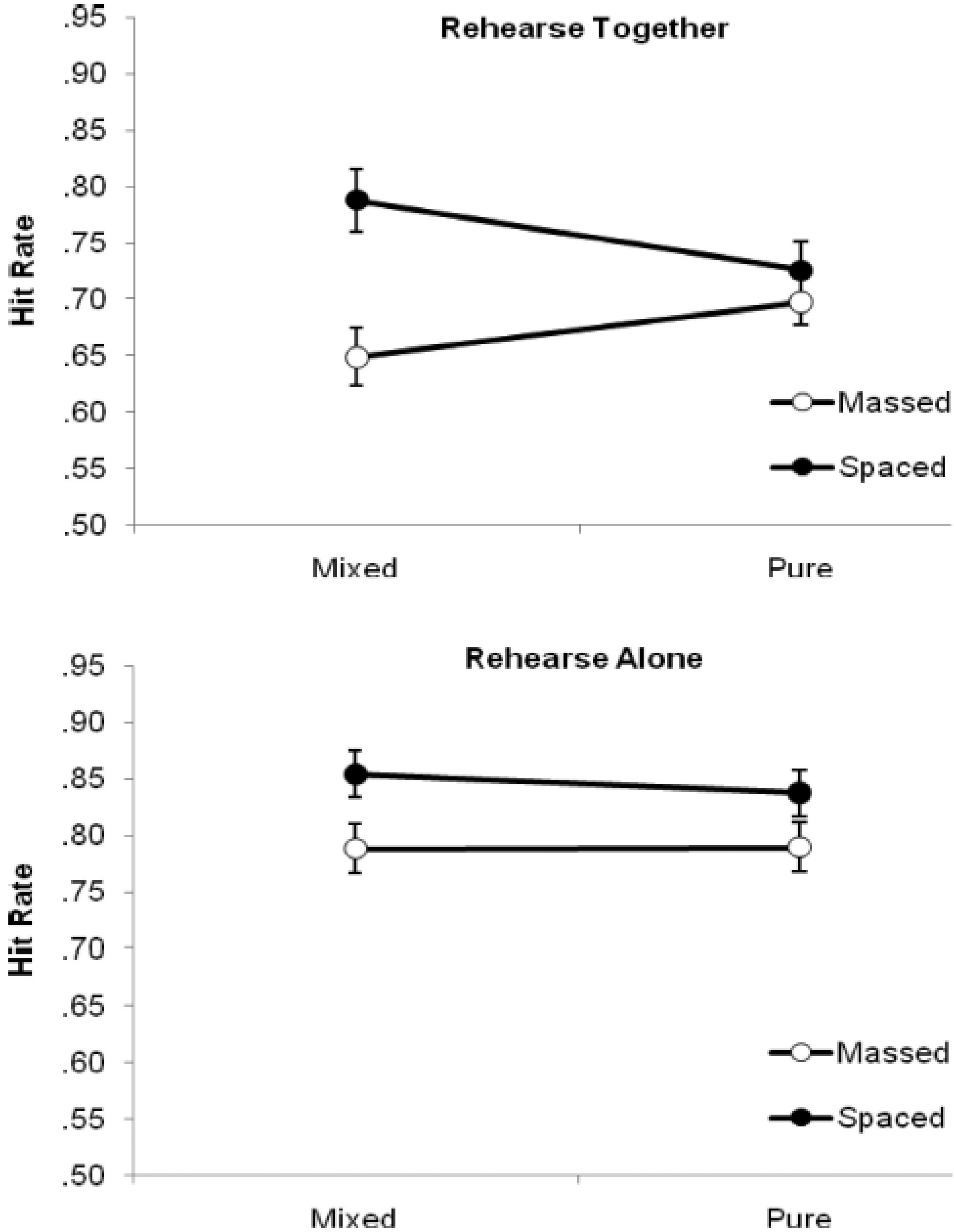

Figure 5. Hit rates for massed and spaced items from mixed and pure lists, split by encoding instructions, Experiment 5. Error bars represent ±*SE*s.

A 2 Spacing (spaced vs. massed) × 2 List Type (mixed vs. pure) × 2 Instruction (rehearse-alone vs. rehearse-together) within-subjects ANOVA on hit rates produced the expected three-way interaction, $F(1, 47) = 8.98$, $MSE = 0.006$, $p < .01$, $\eta^2 = .16$. Those who wish to avoid a flurry of required but uninteresting statistics may skip to the next paragraph. For the rest, there was a main effect of spacing, $F(1, 47) = 41.94$, $MSE = 0.011$, $p < .001$, $\eta^2 = .47$, reflecting an advantage of spaced repetitions ($M = 0.80$) over massed repetitions ($M = 0.73$). There was no main effect of list type, $F < 1$, but there was a large main effect of instruction, $F(1, 47) = 29.87$, $MSE = 0.034$, $p < .001$, $\eta^2 = .39$, reflecting a memory advantage for rehearse-alone ($M = 0.82$) compared with rehearse-together ($M = 0.72$) instructions. As for the two-way interactions, instruction interacted with neither list type, $F < 1$, nor spacing, $F(1, 47) = 1.43$, $MSE = 0.012$, $\eta^2 = .03$. The only two-way interaction was between spacing and list type, $F(1, 47) = 17.30$, $MSE = 0.006$, $p < .001$, $\eta^2 = .27$.

Splitting the three-way interaction by instruction addressed our predictions. For rehearse-together, there was a main effect of spacing, $F(1, 47) = 29.07$, $MSE = 0.012$, $p < .001$, $\eta^2 = .38$, reflecting higher hit rates for spaced items (0.76) than massed items (0.67). The main effect of list type was not significant, $F < 1$. Finally, replicating our earlier experiments, there was a significant Spacing × List Type interaction, $F(1, 47) = 19.14$, $MSE = 0.008$, $p < .001$, $\eta^2 = .29$. The interaction reflected a large spacing effect on mixed lists, $t(47) = 6.97$, $p < .001$, but a small nonsignificant spacing effect on the pure lists, $t(47) = 1.39$, $p = .17$.

For the rehearse-alone instructions, we also replicated our earlier experiments. There was a main effect of spacing, $F(1, 47) = 13.75$, $MSE = 0.011$, $p < .01$, $\eta^2 = .23$, reflecting higher hit rates for spaced items ($M = 0.85$) than massed items ($M = 0.79$). The main effect of list type was not significant, $F < 1$, and likewise there was no Spacing × List Type interaction, $F(1, 47) = 1.04$, $MSE = 0.004$, $\eta^2 = .02$.

To summarize these results, when people were allowed to borrow rehearsal time from one item to help another, they showed poorer overall recognition than when they were forced to attend to each item individually. Furthermore, the rehearse-together instructions resulted in a significant spacing effect on mixed lists but no spacing effect on pure lists. When participants had to rehearse only the currently shown item, they showed a spacing effect of comparable magnitude on both pure and mixed lists.

General Discussion

We have not attempted to provide an answer in this article as to the mechanism(s) underlying the spacing effect in conditions that discourage rehearsal. They have been explained via a number of mechanisms, including selective survival of effective encoding strategies (Bahrick & Hall, 2005), less attention to closely spaced repetitions (e.g., Shaughnessy et al., 1974; Zimmerman, 1975), that repetitions of an item are forgotten more quickly when the item is highly active in memory (e.g., Pavlik & Anderson, 2005; Schmidt & Bjork, 1992), that people notice repetitions of items and retrieve the earlier presentation of an item when they encounter it again (e.g., Greene, 1989; Hintzman, Summer, & Block, 1975; Johnston & Uhl, 1976; Thios & D'Agostino, 1976), and that contextual information is stored with an item so that items spaced widely apart should therefore have multiple (or stronger) retrieval routes, leading to better memory (e.g., Glenberg, 1979; Malmberg & Shiffrin, 2005; Raaijmakers, 2003; Sahakyan et al., 2008; Verkoeijen, Rikers, & Schmidt, 2004).

The focus on the current article was on the encoding strategies that people use in the most common type of study that has explored the spacing effect—specifically, designs with mixed lists, short study-test delays, and intentional learning. Most people approach list-learning tasks—at least initially—by using a rehearse-together strategy in which they rehearse earlier presented items as often as possible (Delaney & Knowles, 2005; Sahakyan & Delaney, 2003), so we explored how rehearsal strategies and spacing interact to affect memory in a list-learning task.

The clearest summary of our results is that there is a three-way interaction between the type of list design (pure or mixed), the encoding strategy (rehearse-together or rehearse-alone), and the benefits of spacing, as illustrated in Experiment 5. We replicated subcomponents of this interaction using both free recall (Experiments 1 and 3) and recognition (Experiments 2 and 4). The interaction can be characterized rather simply. First, the spacing effect is of comparable size on pure and mixed lists with the rehearse-alone strategy. Second, for pure-list designs, the rehearse-together strategy results in a much smaller spacing effect than the rehearse-alone strategy. Third, for mixed-list designs, the reverse is true; the rehearse-together strategy results in a much larger spacing effect than the rehearse-alone strategy.

Our explanation for the rehearse-together strategy's magnification of spacing effects on mixed lists was that borrowing redistributes study time away from the weak massed items and favors the strong spaced items. A similar "strong get stronger" argument was proposed by Yonelinas, Hockley, and Murdock (1992) to explain why list-strength effects sometimes seem to emerge in recognition testing. Experiments 1 and 2 focused on the role of rehearsal on mixed lists consisting of some spaced items and some massed items. The main difference between these studies was that Experiment 1 used free recall testing and Experiment 2 used recognition testing.

The rehearsal strategy that we encouraged people to use was based on strategies reported to us by participants in "regular" memory studies and was intended to be a similar as possible to what people would naturally do. Rehearse-aloud protocols collected in Experiments 1 and 2 confirmed that spaced presentations received more rehearsals than massed presentations on mixed lists, and for recognition testing redistribution was especially pronounced during the time that massed items were visible. In other words, people tended to use the time that massed items were presented to go back and study the spaced items. We therefore propose a version of Hall's (1992) displaced rehearsal hypothesis—at least when people are actually using a rehearse-together strategy. Rehearsal on mixed lists created large spacing benefits by taking rehearsal away from weak items and by helping already strong items.

In contrast, on the pure lists studied with rehearse-together instructions, spaced and massed items received comparable numbers of rehearsals, and rehearsal borrowing on pure lists served to reduce the differences between massed and spaced study. A consistent finding in our rehearse-together experiments was that massed items were remembered better on pure lists than on mixed lists, whereas the opposite was true for spaced items (which are remembered better on mixed lists than on pure lists). However, massed items showed a bigger *deficit* on mixed lists than spaced items showed *advantage* on mixed lists. Our explanation for the smaller spacing effect on pure lists when rehearsal borrowing occurs is by necessity more speculative. We argued that perhaps participants spontaneously produce distributed rehearsals of the massed items, which reduces the relative advantage of spaced items.

An alternative explanation for the larger spacing effect on mixed than pure lists that emerged in Experiment 1 is that it reflected a list-strength effect. Strong items are better recalled when their competitors are weak items, whereas weak items are recalled less frequently when their competitors are strong items (e.g., Tulving & Hastie, 1972). If spaced items are strong and massed items weak, then one would expect an interaction similar to the one in Figure 1. However, this explanation is undermined by the presence of a similar interaction in Experiment 2, which used recognition testing. List-strength effects have generally not been found for words with recognition testing (Hirshman, 1995; Murnane & Shiffrin, 1991a, 1991b; Ratcliff, Clark, & Shiffrin, 1990; Ratcliff, Sheu, & Gronlund, 1992; Yonelinas et al., 1992), unless exceptionally strong lures or remember/know judgments are used (Diana & Reder, 2005; Norman, 2002) or people engage in rehearsal borrowing—the latter producing an artifact like what we observed here (e.g., Yonelinas et al., 1992). Furthermore, if list-strength effects were responsible for our results, it is unclear why they emerge following rehearse-together and not rehearse-alone strategies (as in Experiment 5).

Lastly, despite participants' preference for the rehearse-together strategy, Experiment 5 suggested that they are generally better off with a rehearse-alone strategy (i.e., focusing on each item individually), because the rehearse-together strategy produces lower overall recall rates than the rehearse-alone strategy. One can always be concerned that the strategy induced by instructions is different from the spontaneous version of the same strategy, but we note that recall rates when participants spontaneously report using rehearse-together are quite similar to the recall rates when they are induced to use a rehearse-together strategy (see Delaney & Knowles, 2005).

The Enhanced Primacy Effect and Reverse Spacing Effect

A more technical point explored in our article is the interaction between serial position at study and spacing. The issue emerges from studies using intentional learning on pure lists—a typical observation is that in the first quadrant or so of the lists, there is a reverse spacing effect such that items from pure-massed lists are recalled better than items from pure-spaced lists (e.g., Delaney & Knowles, 2005; Hall, 1992; Toppino & Schneider, 1999). We previously attributed the crossover interaction between serial position and spacing to an enhanced primacy effect on pure-massed lists added to a real spacing effect throughout the list. The general idea was that massed items, for whatever reason, receive a much greater primacy benefit than spaced items do. However, the primacy effect is often attributed to a combination of frequency and recency of rehearsal (cf. Tan & Ward, 2000). In the time that the first four items are presented on a pure-massed list, perhaps eight items are presented on the pure-spaced list. The result is that items in the primacy region on pure-massed lists receive disproportionate amounts of study time at the beginning, strengthening them and making them more likely to be retrieved later. Strengthening the early items may also cause proactive interference on later list items, leading to even greater relative strength for primacy items.

The total number of rehearsals did not differ between pure-spaced and pure-massed lists. One can therefore think of rehearsal within a list as being a zero-sum game—more rehearsals of one part of the list implies fewer rehearsals of other parts of the list. Finding more rehearsal of Quadrant 1 on pure-massed lists than on pure-spaced lists implies that the rest of the list should receive correspondingly fewer rehearsals on pure-massed lists than on pure-spaced lists. Thus, the apparent spacing advantage in Quadrants 2–4 may reflect redistribution of practice within the massed list away from Quadrants 2–4 and to Quadrant 1. Consistent with a rehearsal redistribution account of the serial position by spacing interaction, when enhanced primacy effects were small or absent, as they were in Experiment 2, so were the spacing effects throughout the rest of the list. Further evidence for the zero-sum game was presented by Verkoeijen and Delaney (2008), who demonstrated that the Quadrant 1 massing advantage strongly predicts the magnitude of the spacing advantage elsewhere in the list.

One might be able to observe an overall spacing effect on pure lists studied with the rehearse-together strategy if the enhanced primacy effect produced a ceiling effect. For example, using three presentations of each item or extremely long study times might produce ceiling effects in the primacy region (three presentations were used by Kahana & Howard, 2005). If so, then the massed item penalty in Quadrants 2–4 (which we suggest normally produces a small spacing effect) would not be offset by a corresponding massed item benefit in Quadrant 1 (thanks to a ceiling effect).

Conclusions

Our studies—like many other theoretical tests of the mechanisms underlying the spacing effect—have limitations. Using longer delays will be critical to future tests of the effect of encoding strategy on spacing effects, and we have started some investigations along these lines. Furthermore, the use of word lists and highly artificial study situations limits peoples' ability to make metacognitively based adjustments to study (e.g., Benjamin & Bird, 2006; Dunlosky & Hertzog, 1998; Metcalfe, 2002; Son, 2004) and to adjust their strategies on an item-by-item basis in response to experience (e.g., Bahrick & Hall, 2005).

Despite these limitations, a fruitful strategy in memory research has been to provide people with a word list to study and, without any particular direction on how to study, to ask them to learn it for a later memory test. Hundreds of articles have used similar methods to test theories about spacing and testing. Our work suggests that spacing effects in intentional learning emerge (or fail to emerge) as a consequence of the strategy that people use to study (see also Delaney & Knowles, 2005). A detailed analysis of the impact of one study strategy (the rehearse-together strategy) revealed that rehearsal borrowing sometimes enhances and sometimes diminishes the spacing effect. On mixed lists, rehearsal borrowing redistributes study time away from massed items and toward spaced items, producing a larger memory advantage for spaced items. However, rehearsal borrowing also tends to produce small or absent spacing effects on pure lists. The effects attributed to rehearsal borrowing were large enough that they could overwhelm the usual causes of the spacing effect.

Our results suggest that computational models of spacing should take into account not only what is presented but also what participants are rehearsing to themselves. The simple computational model EICL (which incorporates excitation, inhibition, and a closed-loop learning algorithm) assumes that learning from seeing a presentation of an item depends on the difference between a maximum theoretical strength and its current strength, with steady decreases in strength as new items are encountered (Murdock, 2003). It provides an intuitive and straightforward explanation for spacing effects, and it could be viewed as consistent with our results provided that we consider not only what is onscreen at any given moment but what participants are themselves bringing to mind as they study.

We therefore recommend that if intentional learning is used to explore spacing, researchers should consider controlling the strategy that people use and perhaps tracking what is being rehearsed. An alternative approach is to use incidental learning procedures, which—so long as people follow the directions—should provide more accurate information about the magnitude of spacing effects following various manipulations.

Footnotes

[1] Using arcsine proportion recall as recommended by Toppino and Schneider (1999) did not change any of the results, so we used the untransformed proportions.

References

Bahrick, H. P., & Hall, L. K. (2005). The importance of retrieval failures to long-term retention: A metacognitive explanation of the spacing effect. *Journal of Memory and Language*, *52*, 566–577.

Benjamin, A. S., & Bird, R. D. (2006). Metacognitive control of the spacing of study repetitions. *Journal of Memory and Language*, *55*, 126–137.

Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology*, *6*, 173–184.

Delaney, P. F., & Knowles, M. E. (2005). Encoding strategy changes and spacing effects in the free recall of unmixed lists. *Journal of Memory and Language*, *52*, 120–130.

Dempster, F. N. (1996). Distributing and managing the conditions of encoding and practice. In E. L.Bjork & R. A.Bjork (Eds.), *Memory* (pp. 317–344). San Diego, CA: Academic Press.

Diana, R. A., & Reder, L. M. (2005). The list-strength effect: A contextual competition account. *Memory & Cognition*, *33*, 1289–1302.

Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don’t. *Journal of Applied Psychology*, *84*, 795–805.

Dunlosky, J., & Hertzog, C. (1998). Training programs to improve learning in later adulthood: Helping older adults educate themselves. In D. J.Hacker, J.Dunlosky, & A. C.Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 249–275). Mahwah, NJ: Erlbaum.

Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage*. Boston: Houghton Mifflin.

Glenberg, A. M. (1977). Influences of retrieval processes on the spacing effect in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *3*, 282–294.

Glenberg, A. M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition*, *7*, 95–112.

Glenberg, A. M., & Smith, S. M. (1981). Spacing repetitions and solving problems are not the same. *Journal of Verbal Learning and Verbal Behavior*, *20*, 110–119.

Greene, R. G. (1989). Spacing effects in memory: Evidence for a two-process account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 371–377.

Greene, R. G. (1990). Spacing effects on implicit memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 1004–1011.

Greeno, J. G. (1970). How associates are memorized. In D. A.Norman (Ed.), *Models of human memory* (pp. 257–284). New York: Academic Press.

Hall, J. W. (1992). Unmixing the effects of spacing on free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 608–614.

Hintzman, D. L., Summers, J. J., & Block, R. A. (1975). What causes the spacing effect? Some effects of repetition, duration, and spacing on memory for pictures. *Memory & Cognition*, *3*, 287–294.

Hirshman, E. (1995). Decision processes in recognition memory: Criterion shifts and the list strength effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 302–313.

Janiszewski, C., Noel, H., & Sawyer, A. G. (2003). A meta-analysis of the spacing effect in verbal learning: Implications for research on advertising repetition and consumer memory. *Journal of Consumer Research*, *30*, 138–149.

Jensen, T. D., & Freund, J. S. (1981). Persistence of the spacing effect in incidental free recall: The effect of external list comparisons and intertask correlations. *Bulletin of the Psychonomic Society*, *18*, 183–186.

Johnston, W. A., & Uhl, C. N. (1976). The contributions of encoding effort and variability to the spacing effect on free recall. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 153–160.

Kahana, M. J., & Howard, M. W. (2005). Spacing and lag effects in free recall of pure lists. *Psychonomic Bulletin & Review*, *12*, 159–164.

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.

Metcalfe, J. (2002). Is study time allocated selectively to a region of proximal learning?*Journal of Experimental Psychology: General*, *131*, 349–363.

Modigliani, V., & Hedges, D. G. (1987). Distributed rehearsals and the primacy effect in single-trial free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *6*, 426–436.

Murdock, B. (2003). The mirror effect and the spacing effect. *Psychonomic Bulletin & Review*, *10*, 570–588.

Murdock, B., & Metcalfe, J. (1978). Controlled rehearsal in single-trial free recall. *Journal of Verbal Learning and Verbal Behavior*, *77*, 309–324.

Murnane, K., & Shiffrin, R. M. (1991a). Interference and the representation of events in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 855–874.

Murnane, K., & Shiffrin, R. M. (1991b). Word repetitions in sentence recognition. *Memory & Cognition*, *19*, 119–130.

Norman, K. A. (2002). Differential effects of list strength on recollection and familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 1083–1094.

Pavlik, P. I., & Anderson, J. R. (2005). Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect. *Cognitive Science*, *29*, 559–586.

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

Ratcliff, R., Clark, S. E., & Shiffrin, R. M. (1990). List-strength effect: I. Data and discussion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 163–178.

Ratcliff, R., Sheu, C. E., & Gronlund, S. D. (1992). Testing global memory models using ROC curves. *Psychological Review*, *99*, 518–535.

Rea, C. P., & Modigliani, V. (1987). The spacing effect in 4- to 9-year-old children. *Memory & Cognition*, *15*, 436–443.

Rose, R. J., & Rowe, E. J. (1976). Effects of orienting task and spacing of repetitions on frequency judgments. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 142–152.

Rundus, D. (1971). Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology*, *89*, 63–77.

Rundus, D. (1977). Maintenance rehearsal and single-level processing. *Journal of Verbal Learning and Verbal Behavior*, *16*, 665–681.

Sahakyan, L., & Delaney, P. F. (2003). Can encoding differences explain the benefits of directed forgetting in the list-method paradigm?*Journal of Memory and Language*, *48*, 195–201.

Sahakyan, L., Delaney, P. F., & Kelley, C. M. (2004). Self-evaluation as a moderating factor in strategy change in directed forgetting benefits. *Psychonomic Bulletin & Review*, *11*, 131–136.

Sahakyan, L., Delaney, P. F., & Waldum, E. R. (2008). Intentional forgetting is easier after two "shots" than one. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 408–411.

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.

Shaughnessy, J. J., Zimmerman, J., & Underwood, B. J. (1974). The spacing effect in the learning of word pairs and the components of word pairs. *Memory & Cognition*, *2*, 742–748.

Son, L. K. (2004). Spacing one's study: Evidence for a meta-cognitive control strategy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 601–604.

Tan, L., & Ward, G. (2000). A recency-based account of the primacy effect in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1589–1625.

Thios, S. J., & D'Agostino, P. R. (1976). Effects of repetition as a function of study-phase retrieval. *Journal of Verbal Learning and Verbal Behavior*, *15*, 529–536.

Toppino, T. C. (1991). The spacing effect in young children's free recall: Support for automatic-process explanations. *Memory & Cognition*, *19*, 159–167.

Toppino, T. C., & Bloom, L. C. (2002). The spacing effect, free recall, and two-process theory: A closer look. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 437–444.

Toppino, T. C., & Schneider, M. A. (1999). The mix-up regarding mixed and unmixed lists in spacing-effect research. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1071–1076.

Tulving, E., & Hastie, R. (1972). Inhibition effects of intralist repetition in free recall. *Journal of Experimental Psychology*, *92*, 297–304.

Verkoeijen, P. P. J. L., & Delaney, P. F. (2008). The rehearse-together strategy and spacing effects in the free recall of pure and mixed lists. *Journal of Memory and Language*, *58*, 35–47.

Verkoeijen, P. P. J. L., Rikers, R. M. J. P., & Schmidt, H. G. (2004). Detrimental influence of contextual change on spacing effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 796–800.

Wright, J., & Brelsford, J. (1978). Changes in the spacing effect with instructional variables in free recall. *American Journal of Psychology*, *91*, 631–643.

Yonelinas, A. P., Hockley, W. G., & Murdock, B. B. (1992). Tests of the list-strength effect in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 345–355.

Zimmerman, J. (1975). Free recall after self-paced study: A test of the attention explanation of the spacing effect. *American Journal of Psychology*, *88*, 277–291.