# Encoding strategy changes and spacing effects in the free recall of unmixed lists

By: Peter F. Delaney and Martin E. Knowles

Delaney, P. F., & Knowles, M. E. (2005). Encoding strategy changes and spacing in free recall. *Journal of Memory and Language, 52*, 120-130. doi:10.1016/j.jml.2004.09.002

Made available courtesy of Elsevier: http://www.sciencedirect.com/science/journal/0749596X

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

**Abstract:**
Memory for repeated items often improves when repetitions are separated by other items—a phenomenon called the spacing effect. In two experiments, we explored the complex interaction between study strategies, serial position, and spacing effects. When people studied several unmixed lists, they initially used mainly rote rehearsal, but some people eventually adopted relational encoding strategies like creating a story from the items (Experiment 1). We observed overall spacing effects when participants used the story mnemonic, but not when they employed rote rehearsal strategies (Experiments 1 and 2). This occurred in part because the story mnemonic reduced or eliminated the usual recall advantage of immediately repeated items at the beginning of lists (Experiment 2).
Keywords: Spacing; Massed practice; Spaced practice; Strategy change; Long-term memory; Serial position effects

**Article:**
A major contribution of cognitive psychology has been to build a systematic body of knowledge about how people learn, and to identify techniques that can enhance learning. One of the most straightforward approaches to enhancing learning is to capitalize on the spacing effect. The *spacing effect* is the finding that memory performance is better when study repetitions are separated by other events or items (*spaced* items) than when repetitions immediately follow one another (*massed* items). Although massed items and spaced items receive the same total presentation time, the spaced items are nonetheless recalled better than massed items. The spacing effect is extremely robust and occurs with a variety of testing procedures, materials, and participant populations (for recent reviews, see Dempster, 1996, Greene, 1989 and Raaijmakers, 2003). In fact, the effect is so robust that some researchers have even argued that failures to find spacing effects may be theoretically informative because of the phenomenon's resistance to experimental manipulation (e.g., Challis, 1993, Greene, 1989, Hintzman, 1974, Russo et al., 2002 and Toppino et al., 2002).

Our paper focused on free recall. Free recall behaves differently from cued tests, indicating that different underlying mechanisms may be involved under different testing methods (e.g., Greene, 1989 and Kahana and Greene, 1993). Early accounts of spacing assumed that all testing methods produced spacing effects from a single mechanism, but they were all ultimately unsuccessful in explaining the full range of spacing data. Nonetheless, many of the early accounts still form components of more complex modern accounts of spacing. Hintzman (1974) clustered the most likely candidate mechanisms into two groups: deficient processing accounts and contextual variability accounts. *Deficient processing accounts* propose that spacing effects occur because massed items receive less processing on their second occurrences than spaced items do (e.g., Challis, 1993, Cuddy and Jacoby, 1982, Greeno, 1970 and Hintzman, 1976). It is well established that deficient processing occurs in specific circumstances. For example, if the same anagram is presented twice at different points in time then people usually solve it twice. Each presentation enhances later memory for the solution word. However, if the two presentations are massed, the answer may be retrieved from memory during the second presentation instead of solving the anagram again, which does not enhance memory for the anagram later (Cuddy and Jacoby, 1982 and Jacoby, 1978). Thus, in this situation, spacing effects emerge because of deficient processing in the massed condition. *Contextual variability accounts* instead suggest that spaced items are more likely than massed items to be encoded in several different ways, resulting in more retrieval routes, and consequently producing higher spaced item recall than massed item recall (e.g., Glenberg, 1979, Landauer, 1969, Madigan, 1969, Melton, 1967, Melton, 1970 and Verkoeijen et al., 2004). However, most authors now agree that contextual variability accounts alone are not sufficient to explain the full range of spacing phenomena either (Braun and Rubin, 1998, Challis, 1993, Greene, 1989, Russo et al., 1998 and Russo et al., 2002).

Several of the most recent accounts of spacing effects share similar explanatory mechanisms concerning mechanisms underlying spacing effects for intentional learning of unrelated items (e.g., Braun and Rubin, 1998 and Greene, 1989). One relevant mechanism is a form of deficient processing in which people pay less attention to the second presentation of a massed item. Another mechanism involves study-phase retrieval whereby the second presentation of a spaced item causes retrieval of the first presentation from long-term memory (see also Raaijmakers, 2003, Russo et al., 2002 and Toppino and Bloom, 2002). Because distributed retrieval practice is known to enhance recall, study-phase retrieval enhances memory for spaced items relative to massed items (Carrier & Pashler, 1992). Finally, the Braun and Rubin (1998) account also includes the assumption that the amount of time an item spends in working memory enhances recall. They argued that during study-phase retrieval, spaced items spend more time in working memory because the second presentation of an item causes people to try to retrieve the previous presentation. Retrieval of the first presentation occurs immediately for massed items, while retrieval is longer for spaced items, resulting in a longer time in working memory for spaced items.

Curiously, there has been little interest in how study strategies change over time and whether this influences the spacing effect. Some of the mechanisms proposed to account for the spacing effect would probably function independently of strategy changes, but others would be rather strategy-specific. Therefore, an understanding of how encoding strategies change with practice may be relevant to predicting the magnitude and cause of spacing effects in a variety of circumstances.

Most participants studying word lists appear to use rote rehearsal strategies, but not everyone does, and there is evidence that practice may lead participants to abandon shallow encoding strategies like rote rehearsal in favor of more elaborative encoding strategies like making up a story (Sahakyan and Delaney, 2004 and Sahakyan et al., 2004). Sahakyan et al. (2004) found that participants who studied a list of words and were subsequently instructed to immediately recall the list improved their performance on a subsequent study list due to switching to deeper encoding strategies. The strategy-change effect occurs with recall (Sahakyan et al., 2004) and may occur with recognition as well (Sahakyan & Delaney, 2004). An implication of this research is that within-subjects designs in which multiple lists are studied and immediately tested may be affected by changes in participants' encoding strategies across trials. To the extent that spacing effects interact with strategy, strategy changes would complicate the interpretation of experiments on spacing.

Most of the mechanisms proposed to account for spacing ought to produce superior performance on spaced items relative to massed items regardless of what proportion of the to-be-learned items are spaced. Therefore, from a practical perspective, applying the spacing effect to learning situations would probably involve spacing all to-be-learned items rather than only some of the items. However, most studies on spacing have employed mixed lists containing some spaced items and some massed items. Lists containing all spaced or all massed items are called *unmixed* lists and overall spacing effects on unmixed lists are sometimes small or even absent (Hall, 1992 and Toppino and Schneider, 1999). One reason why unmixed lists often show small spacing effects is that massed lists apparently produce larger primacy effects than spaced lists do (Toppino & Schneider, 1999). The result is that massed items are better learned at the beginning of a list, producing a kind of inverse spacing effect in the beginning of a list, while spaced items are better learned in later portions, producing the standard spacing effect throughout the rest of the list. The total benefit to the learner of spaced practice, which we call the *overall spacing effect*, is the sum of the performance gain in the latter portion of the serial position curve from spacing items and the performance loss in the early portion of the serial position curve from spacing items. In some cases, no overall spacing effect emerges because the magnitude of the spacing benefit in the rest of the learning curve is balanced out by the corresponding inverse spacing effect in the primacy region (Hall, 1992). In other cases, overall spacing effects still emerge, but they are often rather small (Kahana and Howard, in press and Toppino and Schneider, 1999).

That overall spacing effects are so small on unmixed lists would seem to limit the applicability of the spacing effect for learning. The goal is not to enhance the learning of some items at the expense of others, but to increase total learning. If the magnitude of the primacy effect could be reduced somehow, then perhaps larger overall spacing effects would emerge. One possible way to alter serial position functions is to alter participants' encoding strategies (e.g., Herrmann, Geisler, & Atkinson, 1973)—something we suggested may happen without intervention anyway as people become more experienced at studying. If strategy change affects the primacy effect, perhaps it can also reduce or eliminate the inverse spacing effect in the primacy region, and enhance overall spacing effects.

In summary, our main interest was in understanding how strategies change with practice and the impact of such strategy changes on spacing effects. Different encoding strategies result in different patterns of primacy and recency, which could lead to changes in the overall effectiveness of spacing manipulations. Specifically, because some strategies show reduced primacy effects, one might expect that this could interact with spacing and produce greater overall benefits of spacing.

## Experiment 1

Experiment 1 was designed to test our hypothesis that people switch from shallow to deep encoding strategies during multi-list spacing experiments, and that this strategy switching alters which items benefit most from the spacing effect. We asked participants to study four lists, two of which were all-spaced and two of which were all-massed. We ensured that participants received one all-spaced and one all-massed list in each half of the experiment so that we could assess whether practice would lead to greater spacing over time. Retrospective verbal reports were taken to assess strategy changes by asking participants about their study techniques on Lists 1 and 4 using the standard instructions from Ericsson and Simon (1993).

We expected to find that participants would tend to switch from shallower encoding strategies (like rote rehearsal) to deeper encoding strategies (like using the story mnemonic) over successive trials. This should tend to reduce the magnitude of the primacy effect across trials, and hence tend to enhance the overall spacing effect over trials.

## Method

### Participants

Participants were 96 University of Florida undergraduates enrolled in an introductory psychology class who participated for course credit. They were tested individually. None of them reported in a post-session questionnaire having participated in another memory experiment, and most had completed less than 1 month of the course. None of their instructors reported having covered mnemonic strategies.

### Materials

Four sets of words were created using the materials from Toppino and Schneider (1999) and additional medium-frequency words selected from the Francis and Kucera (1982) norms. For each set, one all-spaced and one all-massed list were constructed. Each list was broken into four sets of eight items (quadrants) whose order was counterbalanced using a latin square design. Each word appeared equally often in each quadrant of the list and for each massed order a corresponding spaced order was constructed that had all of the same words in the same quadrant. For spaced lists, the average number of words between the two spaced presentations was 4.75, with a range of 4–8. The order of words on the lists was constrained so that obvious associates were not near one another.

### Procedure

Participants studied four lists. Lists alternated between all-massed and all-spaced such that in each half of the experiment participants studied one spaced and one massed list. During study, each word appeared in large black print on the computer screen for 2 s with a 1 s interstimulus interval. After each list was studied, participants solved three-digit by two-digit multiplication problems for 60 s, followed by a 3-min free recall test.

Immediately following recall of the fourth list, we asked participants to give a retrospective verbal report on how they studied that list. Participants were asked to remember everything they said or did as they studied the list. If they indicated that they used a story creation strategy or other deep encoding, we asked them to recall the story or associations they drew. If they indicated that they repeated items over and over, we asked them to demonstrate how they did it. We then reminded them of the first three items from List 1, and asked them to give a retrospective report concerning their study of that list in the same fashion. Finally, we prompted them to tell us if they did anything different on the two lists. We did not probe strategies after each list because of concerns about reactivity.

### Strategy report coding

Strategy reports were coded into three categories using a system similar to that used by Goodwin, 2003 and Perfect and Dasgupta, 1997, and Sahakyan and Delaney (2003). *Shallow* encoding strategies included maintenance rehearsal of each individual item, rhyming, and rehearsing only the first letter of each word. The vast majority (90%) of these shallow encoding reports consisted of rehearsing the items in small groups or rehearsing whatever words could be remembered (i.e., rote rehearsal strategies involving displaced rehearsal of earlier-studied items). The rest consisted of trying to recall the first letter of the words (4%) and just staring at the words (6%). Every participant in the experiment eventually abandoned the first letter and staring strategies in favor of something else (often, for rehearsal). Creating individual sentences with the words, using extensive self-referencing, creating a story using all the items on the list, and creating interactive images were considered to be *deep* encoding strategies. Finally, visualizing individual items or using a mixture of shallow and deep strategies on the same list were called *intermediate* encoding strategies. Individual raters were blind to the hypotheses of the study and agreement on such ratings in previous studies exceeded 90%.

## Results and discussion

Our first question was whether we would replicate Toppino and Schneider's (1999) finding of an overall spaced advantage with unmixed lists. We therefore conducted a 2 Spacing (spaced vs. massed) × 2 Block (first two lists vs. last two lists) ANOVA on the arcsine proportion of items recalled, ignoring strategy. There was a main effect of block, $F$ (1, 95) = 4.99, $MSE$ = .008, $p$ < .05, indicating slightly better performance on the second two lists (.44) than on the first two lists (.42), but no main effect of spacing, $F$ (1, 95) = 1.70, $MSE$ = .013, $p$ = .20. The interaction approached significance, $F$ (1, 95) = 3.30, $MSE$ = .011, $p$ = .07. Planned comparisons aimed at testing the hypothesis of an overall effect of spacing in each block suggested that there was no spacing advantage in the first block, $t$ < 1. In the second block, spaced items were better recalled (.46) than massed items (.43), $t(95) = 2.20$, $p < .05$. In summary, although we did not find an effect of spacing in the first block, the second block replicated Toppino and Schneider's (1999) finding of an overall spacing effect.

### Verbal reports analyses

The analysis of participants' retrospective verbal reports indicated that most participants (78%) began with a shallow strategy on List 1. Our hypothesis was that participants might be switching to a deeper encoding strategy as they studied more and more lists. Consistent with our argument, of the participants who initially used shallow encoding, a large proportion (35%) reported switching to a deep encoding strategy before reaching List 4. The majority (56%) continued to use shallow encoding, and a few participants (9%) switched to some intermediate strategy (such as making up some sentences from words while rehearsing the rest). These results are supportive of earlier work in our lab suggesting that repeatedly testing participants results in increasing use of deep, relational encoding strategies (Sahakyan and Delaney, 2003 and Sahakyan et al., 2004).

Of the participants who did not use shallow encoding on the first list, 16% used a deep strategy (of which 1 participant switched to shallow encoding by the final list) while 6% used intermediate on the first list (of which 1 participant switched to deep encoding by the final list).

### Strategy-specific analyses

Following Sahakyan and Delaney (2003), we restricted our further analyses to those participants who used either shallow encoding throughout (*shallow/shallow* participants) and those who switched from shallow to deep encoding (*shallow/deep* participants), as they composed the two largest subgroups. Other strategy combinations were too rare for meaningful statistics. Our main purpose was to show that the spacing effects on unmixed lists emerged only after participants had tried the task and had the opportunity to change strategies, and then only in the case when participants actually switched to relational encoding. We therefore predicted a three-way interaction between spacing, List 2 strategy (i.e., whether people switched or not), and block. It should be noted that item-related effects may have contaminated these analyses, because the counterbalancing scheme was violated (an issue we addressed in Experiment 2). However, the participant-specific nature of strategy choice required analyzing the experiment in this fashion.

We conducted a 2 Spacing (spaced vs. massed) × 2 Block (first two lists vs. last two lists) × 2 Strategy (shallow/shallow vs. shallow/deep) mixed ANOVA on the arcsine proportion of words recalled. Only strategy was a between-subjects factor. We did not include quadrant-by-quadrant analyses because quadrant did not interact with spacing or strategy, and because of concerns about item effects with smaller samples. Fig. 1 shows the mean proportion recalled by spacing, list, and strategy. There was a main effect of block, $F(1, 67) = 10.16$, $MSE = .010$, $p < .005$, and a main effect of strategy, $F(1, 67) = 36.50$, $MSE = .036$, $p < .001$. The main effect of spacing approached significance, $F(1, 67) = 3.41$, $MSE = .011$, $p = .07$. There was also a significant Block × Spacing interaction, $F(1, 67) = 5.89$, $MSE = .012$, $p < .05$, and a significant Block × Strategy interaction, $F(1, 67) = 16.43$, $MSE = .010$, $p < .001$. The Spacing × Strategy interaction was not significant,

$F < 1$. However, as predicted all these effects were moderated by a significant a three-way interaction, $F$ (1, 67) = 5.84, $MSE$ = .012, $p < .05$.

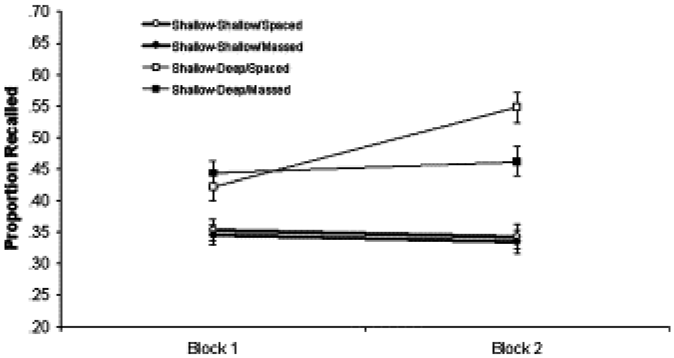


Fig. 1. Proportion of items recalled by spacing, two-list block, and self-reported study strategies, Experiment 1. Error bars represent ±*SE*.

To follow up the three-way interaction, we evaluated two separate within-subjects Block × Spacing interactions, one for shallow/shallow participants and one for shallow/deep switching participants. For shallow/shallow participants, we found no main effects of block, spacing, or Block × Spacing interaction, all $F$s < 1. It appears that participants who reported using shallow encoding did not show an unmixed list spacing effect—or at least we failed to detect one here using lists very similar to those originally used to demonstrate unmixed list spacing effects (cf. Toppino & Schneider, 1999). Shallow/shallow participants also did not improve their recall performance with practice.

In contrast, for the participants who reported switching encoding strategies from one of the shallow strategies to one of the deep strategies, we observed a main effect of block, $F$ (1, 21) = 12.57, $MSE$ = .016, $p < .005$, reflecting improved performance on later lists. Although the main effect of spacing only approached significance, $F$ (1, 21) = 2.86, $MSE$ = .015, $p$ = .10, there was a significant Spacing × Block interaction, $F$ (1, 21) = 6.26, $MSE$ = .018, $p < .05$. Post hoc paired $t$ tests revealed no effect of spacing in the first block, $t < 1$, but a significant spacing advantage in the second block, $t(21) = 3.02$, $p < .01$. Consistent with our argument, there was little evidence for a spacing advantage in the first block, but by the second block, when participants reported using deep encoding on at least the fourth list, there was a spacing advantage. The spacing advantage that emerged in our second block resembled the one obtained by Toppino and Schneider (1999).

One might wonder why the shallow/deep group recalled more items than the shallow/shallow group on Block 1. Many people switched strategies following List 1 (cf. Sahakyan & Delaney, 2003), and so our Block 1 points reflect a mixture of shallow and deep encoding. For comparison, when we consider only the first list studied by the shallow/deep group, on which everyone used shallow encoding, then the corresponding List 1 recall rates were. 42 (*SD* = .08) for massed and. 39 (*SD* = .07) for spaced. These values are rather close to the List 1 recall rates for shallow/shallow participants (.39 and. 36, respectively, both *SD* = .09).

## Conclusions

Participants used a variety of strategies to study the words, although most relied on either rote rehearsal or story creation. As people studied and were tested on several lists in sequence, they sometimes switched to more effective encoding strategies. Consistent with our view that overall spacing effects are small or absent in rote rehearsal, we observed no spacing effects on unmixed lists whenever participants employed shallow encoding strategies. Only when participants reported using deep encoding strategies (like making up a story using all of the words on the list) did we observe a significant spacing effect.

The increasingly frequent use of deeper encoding strategies with practice could have produced significant unmixed list spacing effects like those seen in Toppino and Schneider's (1999) experiment and Kahana and Howard's (in press) studies. The results suggest that in intentional learning situations where all to-be-learned items are spaced or massed, spacing effects will be much smaller or even absent when participants employ exclusively rote rehearsal strategies. Furthermore, capitalizing on the spacing effect may be more effective if people are first trained to use better study strategies.

## Experiment 2

Experiment 2 was intended to validate our conclusions from Experiment 1 by experimentally controlling participants' encoding strategies. This would help to reduce some of the variance normally observed in free recall tests, allowing us to gain more control over encoding quality. In addition, experimental control of encoding strategy would enable us to validate the results of Experiment 1 that were based on retrospective verbal reports. In particular, we expected to replicate the presence of spacing effects on the lists that were elaboratively encoded and the absence of spacing effects on the lists that were non-elaboratively encoded. We therefore instructed participants to use one of the two most commonly reported strategies from Experiment 1—either verbal rehearsal or the story mnemonic to study two unmixed word lists. Each participant studied one list of all-massed items and one list of all-spaced items. Therefore, item type was manipulated within-subjects, but between lists, maximizing our power to detect spacing effects.

We also explored the pattern of spacing effects across different serial positions. If Toppino and Schneider (1999) were correct that enhanced primacy in massed lists is responsible for the frequently small or absent spacing effects on unmixed lists, then manipulations that reduced the primacy advantage (such as the story mnemonic) would tend to increase the overall magnitude of spacing effects. This would be manifested as a reduction of the massed item advantage in the primacy region when participants used the story mnemonic compared to when they used rote rehearsal strategies to study the list. A larger overall spacing effect should emerge when the primacy effect is reduced in magnitude.

## Method

### Participants

Participants were 128 University of Florida undergraduates enrolled in an introductory psychology course who participated for course credit. They were tested individually.

### Materials

We created two sets of 32 words from the word lists used by Toppino and Schneider (1999). As in Experiment 1, the order was constrained so that obvious associates were not near one another. These words were then used to create eight versions of each list in which each word appeared twice for a total of 64 presentations. Four of the lists contained only massed presentations and the other four contained only spaced presentations. Each word appeared equally often in each quadrant of the list and for each massed order a corresponding spaced order was constructed that had all of the same words in the same quadrant. For spaced lists, the average number of words between the two spaced presentations was 4.75, with a range of 4 to 8, just as in Experiment 1. The placement of spaced items in each quadrant was the same as for Experiment 1 lists.

### Procedure

Participants in the *rehearsal* condition were instructed to rehearse the words out loud, adding each new word to the set already being rehearsed. They were told not to become alarmed if they forgot some of the words and just to keep rehearsing what they remembered. In the *story* condition, participants were told to make up a story using each word on the list. Participants were instructed to ensure that they used each word on the list. Because everyone had to talk aloud as they worked, the experimenters could monitor compliance with the instructions, and the six participants who failed to follow the instructions were replaced. In both conditions, participants were instructed to learn the words.

Each participant studied one all-massed list and one all-spaced list, with the order of the lists counterbalanced. Each word appeared in large black print on the computer screen for 2 s, with a 1 s interstimulus interval. After studying the first list, participants solved three-digit by two-digit multiplication problems for 60 s, and then had 3 min to write down as many items as they could remember from the list. The second list was tested in a similar fashion.

## Results and discussion

In Experiment 2, there were no issues of selection bias to contaminate the counterbalancing, and so we were able to examine participants' recall by list quadrant without worrying about item effects. A 2 List Type (massed vs. spaced) × 4 Quadrant × 2 Strategy (rehearsal vs. story) mixed ANOVA was conducted on arcsine transformed proportion recall. Strategy was a between-subjects factor while the list type and quadrant were within-subjects. Our main interest was in whether the overall magnitude of spacing would vary as a function of strategy. Consistent with our hypothesis, there was a significant List Type × Strategy interaction, $F\ (1, 126) = 3.92$, $MSE = .049$, $p < .05$, indicating that the magnitude of the overall spacing effects depended on the encoding strategies used. The List Type × Strategy interaction is shown as Fig. 2. Follow-up Bonferroni corrected *t* tests uncovered no overall spacing effect for rehearsal, $t < 1$, but a significant spacing advantage for story creation, $t(63) = 2.78$, $p < .01$. Both spaced items and massed items were better recalled in the story condition than in the rehearsal condition, $t(126) = 10.23$ and $t(126) = 8.56$, respectively, both *p*s < .001.

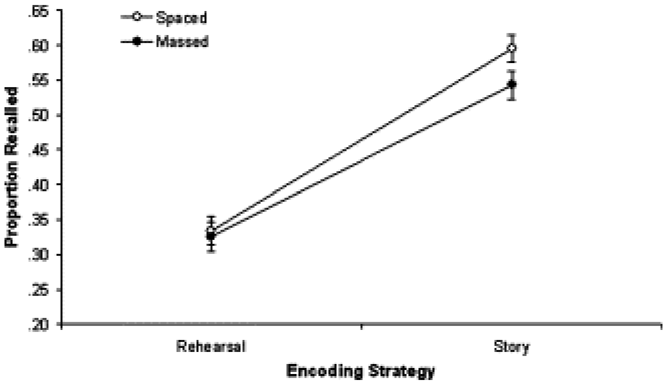


Fig. 2. Proportion of items recalled by spacing and study strategy control instruction, Experiment 2. Error bars represent ±*SE*.

There was a main effect of spacing, $F(1, 126) = 7.26$, $MSE = .049$, $p < .01$, reflecting slightly higher recall for spaced items ($M = .47$) than massed items ($M = .43$). The main effect of strategy, $F(1, 126) = 115.21$, $MSE = .161$, $p < .001$, reflected better recall in the story condition ($M = .57$) than in the rehearsal condition ($M = .33$). The main effect of quadrant, $F(3, 368) = 23.03$, $MSE = .068$, $p < .001$, reflected different recall rates across quadrants ($M$ = .53, .38, .42, and .48, respectively, for Quadrants 1–4). No three-way interaction emerged, $F(3, 378) = 1.14$, $MSE = .064$, $p = .33$. However, several significant two-way interactions emerged. The Spacing × Quadrant interaction, $F(3, 368) = 3.00$, $MSE = .064$, $p < .05$, indicated that the magnitude of the spacing effect varied by quadrant. Post hoc paired *t* tests indicated that there was a non-significant advantage for massed items in the first quadrant ($M$ = .52 vs. .55), $p = .26$; a significant spacing advantage in Quadrant 2 ($M$ = .41 vs. .34), $p < .01$, and Quadrant 3 ($M$ = .45 vs. .38), $p < .05$; and a non-significant spacing advantage in Quadrant 4 ($M$ = .50 vs. .47), $p = .31$. The Strategy × Quadrant interaction, $F(3, 368) = 42.42$, $MSE = .068$, $p < .001$, indicated that the effect of strategy varied by quadrant. The story strategy and the rehearsal strategy did not differ in Quadrant 1, but the story strategy produced higher recall in all other quadrants, all *p* s < .01. These results indicated that the primacy effect was indeed reduced by the story strategy relative to the rehearsal strategy.

We investigated the pattern of spacing in the strategy groups by conducting two separate List Type × Quadrant ANOVAs, one for each strategy (rehearsal and story). The complete data are shown as Table 1.

Table 1.

Proportion of words recalled by strategy control instruction and quadrant for massed and spaced lists, experiment 2

| List type | List quadrant | | | | Total |
| --- | --- | --- | --- | --- | --- |
| | **First** | **Second** | **Third** | **Fourth** | |
| *Rehearsal strategy* | | | | | |
| Massed | .55 | .20 | .21 | .26 | .33 |
| Spaced | .48 | .26 | .25 | .30 | .33 |
| | | | | | |
| *Story strategy* | | | | | |
| Massed | .48 | .45 | .50 | .62 | .54 |
| Spaced | .52 | .52 | .59 | .63 | .60 |

### Spacing effects in the rehearsal condition

For the rehearsal strategy, there was no main effect of spacing, $F < 1$ ($M = .33$ for both spaced and massed lists). There was a significant main effect of quadrant, $F$ (3, 189) = 63.06, $MSE = .056$, $p < .001$, which was qualified by a significant List Type × Quadrant interaction, $F$ (3, 189) = 3.37, $MSE = .056$, $p < .05$. We followed up the significant interaction with paired $t$ tests. The massed advantage in the first quadrant was significant, $t(63) = 2.00$, $p < .05$. However, the spaced advantage in the second quadrant only approached significance, $t(63) = 1.97$, $p = .05$. The slightly better recall of spaced items in the third quadrant, $t$ (63) = 1.14, $p = .26$, and fourth quadrant, $t$ (63) = 1.27, $p = .21$, were not significant.

### Spacing effects in the story condition

In comparison, a List Type × Quadrant within-subjects ANOVA for the story strategy revealed a main effect of list type, $F$ (1, 63) = 7.73, $MSE = .069$, $p < .01$, indicating an overall spacing advantage. The overall proportion recalled in the massed condition was .54 compared to .60 in the spaced condition. There were also recall differences across quadrants, $F$ (3, 189) = 11.08, $MSE = .079$, $p < .001$. Post hoc paired $t$ tests indicated that the quadrant effect reflected greater recall in the fourth quadrant (.66) than in any other quadrant and greater recall in the third quadrant (.59) than the first quadrant (.52) or second quadrant (.51)—a recency effect. The interaction was not significant, $F$ (3, 189) = 1.06, $MSE = .072$, $p = .37$.

## Conclusions

In summary, when participants used rote rehearsal strategies on all-spaced and all-massed lists, the serial position functions resembled those seen in earlier studies. Specifically, there was a significant recall advantage for massed items in the primacy region of the curve, which was exactly offset by a smaller recall advantage for spaced items throughout the rest of the serial position curve. No overall spacing effect emerged. The same pattern of results was observed by Hall (1992).

However, when participants used the story mnemonic, it eliminated the massed advantage in the primacy region. Thus, in the story condition, a main effect of spacing emerged, with better recall for spaced items compared to massed items. This is consistent with our contention that studies like Toppino and Schneider's (1999) and Kahana and Howard's (in press) obtained overall spacing effects with unmixed lists because of strategy changes that occurred over trials.

## General discussion

Our strongest result was that strategy mediates the emergence of overall spacing effects when people study unmixed lists. Over successive lists, many participants adopted increasingly effective encoding strategies. Robust overall spacing effects occurred when participants employed encoding strategies like making up a story. However, when rote rehearsal was employed, the total magnitude of the spacing effect was reduced or even eliminated (at least on unmixed lists). These changes in the spacing effect interacted with serial position (see Experiment 2). For rote rehearsal based encoding strategies, we replicated earlier results showing an advantage for massed items in the primacy region of the serial position function but a smaller spaced advantage in all other regions (cf. Hall, 1992 and Toppino and Schneider, 1999). However, when participants used the story mnemonic, the massing advantage in the primacy region was not detected, and the spacing advantage throughout the list was enhanced.

In Experiment 1, we further found that the magnitude of spacing effects on unmixed lists grew as more lists were studied because participants changed strategies. The presence of strategy changes implies that within-subjects designs involving multiple lists will tend to produce larger spacing effects across trials. Thus, one would expect that if a study were to use fifteen lists instead of one or two (cf. Kahana & Howard, in press), it would tend to enhance the apparent magnitude of overall spacing effects.

Experiment 2 further showed that strategy changes can be brought on artificially by teaching participants to use the story mnemonic to study the lists, thereby producing an overall spacing effect. Likewise, when participants are constrained to use only rote rehearsal, overall spacing benefits do not emerge on unmixed lists (or at least are much smaller in magnitude than when participants use the story mnemonic). It therefore seems that teaching effective encoding strategies not only increases overall retention but also enhances the benefit of spacing presentations.

We further note that many of the classic studies on spacing effects gave participants several practice trials before recording any results (e.g., Murdock & Okada, 1971). The purpose of this procedure was to eliminate warm-up effects. It may very well be that an unintentional side-effect of this procedure was to induce encoding strategy changes, and thus enhance the magnitude of observed spacing effects.

We used unmixed lists because applying spacing effects to practical learning situations would likely involve spacing all of the to-be-learned items rather than using mixed lists where only some items were spaced. This might seem to make generalizing our results to other spacing studies, which typically involve mixed lists consisting of some spaced items and some massed items, more difficult. However, similar patterns of strategy change across successive study lists ought to be found using other testing procedures and mixed lists as well. Toppino and Schneider (1999) also reported a manipulation that suggested that the advantage for massed items in the primacy region of the serial position curve generalized to mixed lists as well. This would seem to suggest that strategy changes ought to interact with spacing on mixed lists in a similar fashion. Furthermore, on sufficiently long lists, the strategy change phenomenon also produces better memory performance on recognition tests (Sahakyan & Delaney, 2004), and so strategy change effects should also be expected with recognition testing in multi-list spacing studies (e.g., Kahana & Howard, in press).

**Why are spacing effects enhanced when using the story mnemonic?**

There are a number of possible reasons why the story mnemonic could produce larger spacing effects than rote rehearsal. Toppino and Schneider (1999) suggested that massed items benefit more from primacy effects than spaced items do, and that therefore all-massed lists would show a large massing advantage in the primacy region of the serial position function. In our Experiment 2, we also found that enhanced primacy effects occurred in the all-massed condition relative to the all-spaced condition when people used rote rehearsal. However, in Experiment 2 we found that deep encoding reduced the magnitude of the primacy effect. If one effect of strategy changes is that the primacy advantage tends to be reduced over trials, then one plausible interpretation of our results is that practice leads to reduction in primacy advantages (see also Goodwin, 1976), and that this in turn reduces the customarily observed recall advantage for massed items in the early portion of the serial position curve. On the first or second study lists, the overall benefit of spacing presentations may be small or even absent, but with practice the spacing benefit will tend to grow as participants adopt increasingly deep, relational encoding strategies because the massed item advantage in the primacy region of the serial position curve is reduced or eliminated.

Our present results were inconclusive as to whether the story mnemonic or other related encoding strategies also change the underlying mechanisms of the spacing effect. It seems plausible that mechanisms that typically operate only in paired associates learning may also operate when the story mnemonic is employed. The story mnemonic involves generating associations between sequential items, which is similar to what happens in paired associates learning. According to Greene (1989), during paired associates learning there are rehearsal biases that lead to diminished processing of massed items. Several competing explanations of the spacing effect in the literature rely on generating associations (e.g., Bower, 1972, Dempster, 1996, Glenberg, 1976, Glenberg, 1977, Glenberg, 1979, Madigan, 1969, Melton, 1967 and Melton, 1970), and perhaps these mechanisms may play a role when people create stories from the words on a list (but not for the more commonly-used rote rehearsal strategies). Dempster, for example, argued that when participants encounter a spaced item, they either retrieve their earlier encoding (which may strengthen it) or, if retrieval fails, generate a new encoding (see also Raaijmakers, 2003). Thus, the second spaced presentation virtually guarantees a higher chance of recalling spaced items. In contrast, when encountering a massed item, new associations are not generated on the second occurrence and there is no opportunity to practice retrieving the earlier association. Conditions that maximize the strengthening of individual items, such as problem solving, seem especially likely to show analogous effects because the processing that would normally be required can be bypassed on massed items by accessing the item directly from short-term or working memory (Braun and Rubin, 1998 and Cuddy and Jacoby, 1982).

One final unlikely possibility, raised most recently by Hall (1992) after his repeated failure to find overall spacing effects with unmixed lists, is that when participants employ rote rehearsal strategies they tend to allocate more rehearsal time to spaced items at the expense of massed items. This *displaced rehearsal* could occur if when participants encounter an item, they read it and then use the remaining presentation time to rehearse the item together with earlier items. Recently-presented items would tend to be more available in memory and, because spaced items occur at two or more locations on the list, they would tend to be overrepresented in recent presentations and thus benefit from rote rehearsal to a greater extent than massed items, which occur in only one location (Rundus, 1971). The result would be that spaced items would receive extra rehearsal time at the expense of massed item time on mixed lists. However, for unmixed lists, any time spent rehearsing a spaced item would be taken from another spaced item—and not from a massed item—yielding no overall spacing effects. Toppino and Schneider (1999) ruled out this explanation by showing overall spacing effects even with unmixed lists, but if, as we have shown, participants often changed encoding strategies, then even unmixed lists might show some overall spacing effects unless strategy was controlled. Hall's (1992) displaced rehearsal account would also rather naturally predict the enhanced primacy effect for massed items, which results in the massed advantage in the primacy region of the serial position curve, because people would have had few spaced items presented early in the list and so massed items would tend to receive a lot of rehearsal time at that point compared to spaced items. As the primacy effect is usually thought to reflect either participants' bias to rehearse items from the start of the list or their greater opportunities to do so (e.g., Atkinson and Shiffrin, 1968 and Tan and Ward, 2000), a displaced rehearsal account would provide an explanation for the inverse spacing effect in the primacy region. While the displaced rehearsal account is compelling in many ways, there are a number of reasons why others have not adopted Hall's view (reviewed in Greene, 1989). For example, spacing effects emerge in young children who would be unlikely to use displaced rehearsal strategies (Toppino, 1991). Spacing effects also emerge in incidental learning studies, where people should not show displaced rehearsal (Glenberg and Smith, 1981, Greene, 1989, Jensen and Freund, 1981 and Toppino and Bloom, 2002). The strategy change effect also does not appear to occur for incidentally learned lists (Sahakyan & Delaney, 2004).

## Conclusions

Much of the evidence for study-phase retrieval comes from the fact that it is very hard to eliminate the spacing effect in free recall (for a review, see Greene, 1989). However, it may be that people use a variety of strategies even in the same experiment. If the cause of spacing effects depends on particular encoding strategies, as we observed here, then eliminating one cause of the spacing effect would still leave other possible sources of spacing owing to the diverse strategies employed by participants. If participants adopt more elaborative encoding strategies like making up a story using all the words on a list, then the spacing effect emerges on unmixed lists. With rote rehearsal, however, spacing effects seem to be small or even absent.

That spacing effects interact with serial position and encoding strategies suggests that from a practical perspective, it may be important to instruct participants in appropriate study strategies to gain maximum benefit from spacing instructions. It also implies that theories of spacing ought to take into account the possibility that different strategies rely on different underlying mechanisms, and that it is study strategy rather than the type of material per se that governs the appropriate explanation for spacing effects.

## Acknowledgments

We are grateful to Alice Azzalin, Carolina Collado, Kelly Johnson, Katie Logan, Satish Kunisi, Sejal Mehta, Jennifer Meisel, Justin Rundle, Teisha Seabrook, Patricia Villafaña, and Valentina Weis for their help in collecting the data. We are also grateful to Thomas Toppino for providing his word lists. Thanks also to Ira Fischler and Lili Sahakyan, who provided useful feedback on an earlier version of the manuscript.

## References

Atkinson and Shiffrin, 1968 R.C. Atkinson and R.M. Shiffrin, Human memory: A proposed system and its control processes In: K.W. Spence and J.T. Spence, Editors, *The psychology of learning and motivation*, Academic Press, New York (1968), pp. 89–105.

Bower, 1972 G.H. Bower, Stimulus-sampling theory of encoding variability In: A.W. Melton and E. Martin, Editors, *Coding in human memory*, Winston, New York (1972), pp. 85–124.

Braun and Rubin, 1998 K. Braun and D.C. Rubin, The spacing effect depends on an encoding deficit, retrieval, and time in working memory: Evidence from once-presented words, *Memory* 6 (1998), pp. 37–65.

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

Challis, 1993 B.H. Challis, Spacing effects on cued-memory tests depend on level of processing, *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19 (1993), pp. 389–396.

Cuddy and Jacoby, 1982 L.J. Cuddy and L.L. Jacoby, When forgetting helps memory: An analysis of repetition effects, *Journal of Verbal Learning and Verbal Behavior* 21 (1982), pp. 451–467.

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

Ericsson and Simon, 1993 K.A. Ericsson and H.A. Simon, Protocol analysis: Verbal reports as data, MIT Press, Cambridge, MA (1993).

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

Glenberg, 1976 A.M. Glenberg, Monotonic and nonmonotonic lag effects in paired-associate and recognition memory paradigms, *Journal of Verbal Learning and Verbal Behavior* 15 (1976), pp. 1–16.

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

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

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

Goodwin, 1976 C.J. Goodwin, Changes in primacy and recency with practice in single-trial free recall, *Journal of Verbal Learning and Verbal Behavior* 15 (1976), pp. 119–132.

Goodwin, 2003 Goodwin, K. A. (2003). *Encoding and retrieval processes in false recall.* Poster presented at the 15th Annual Convention of the American Psychological Society, Atlanta, GA, May 29–June 1, 2003.

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

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

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

Herrmann et al., 1973 D.J. Herrmann, F.V. Geisler and R.C. Atkinson, The serial postion function for lists learned by a narrative-story mnemonic, *Bulletin of the Psychonomic Society* 2 (1973), pp. 377–378.

Hintzman, 1974 D.L. Hintzman, Theoretical implications of the spacing effect In: R.L. Solso, Editors, *Theories in cognitive psychology: The Loyola symposium*, Erlbaum, Hillsdale, NJ (1974), pp. 77–99.

Hintzman, 1976 D.L. Hintzman, Repitition and memory In: G.H. Bower, Editors, *The psychology of learning and motivation: Advances in research and theory* Vol. 10, Academic Press, New York (1976), pp. 47–91.

Jacoby, 1978 L.L. Jacoby, On interpreting the effects of repetition: Solving a problem versus remembering a solution, *Journal of Verbal Learning and Verbal Behavior* 17 (1978), pp. 649–667.

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

Kahana and Greene, 1993 M.J. Kahana and R.L. Greene, Effects of spacing on memory for homogenous lists, *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19 (1993), pp. 159–162.

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

Landauer, 1969 T.K. Landauer, Reinforcement as consolidation, *Psychological Review* 76 (1969), pp. 82–96.

Madigan, 1969 S.A. Madigan, Intraserial repetition and coding processes in free recall, *Journal of Verbal Learning and Verbal Behavior* 8 (1969), pp. 828–835.

Melton, 1967 A.W. Melton, Repetition and retrieval from memory, *Science* 158 (1967), p. 532.

Melton, 1970 A.W. Melton, The situation with respect to the spacing of repetitions and memory, *Journal of Verbal Learning and Verbal Behavior* 9 (1970), pp. 596–606.

Murdock and Okada, 1971 B.B. Murdock and R. Okada, Interresponse times in single-trial free recall, *Journal of Experimental Psychology* 86 (1971), pp. 263–267.

Perfect and Dasgupta, 1997 T.J. Perfect and Z.R.R. Dasgupta, What underlies the deficit in reported recollective experience in old age?, *Memory & Cognition* 25 (1997), pp. 849–858.

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

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

Russo et al., 2002 R. Russo, N. Mammarella and S.E. Avons, Toward a unified account of spacing effects in explicit cued-memory tasks, *Journal of Experimental Psychology: Learning, Memory, and Cognition* 28 (2002), pp. 819–829.

Russo et al., 1998 R. Russo, A.J. Parkin, S.R. Taylor and J. Wilks, Revising current two-process accounts of spacing effects in memory, *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24 (1998), pp. 161–172.

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

Sahakyan and Delaney, 2004 Sahakyan, L., & Delaney, P. F. (2004). Directed forgetting in incidental learning and recognition testing: Support for a two-factor account. Manuscript submitted for publication.

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

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

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

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

Toppino et al., 2002 T.C. Toppino, Y. Hara and J. Hackman, The spacing effect in the free recall of homogeneous lists: Present and accounted for, *Memory & Cognition* 30 (2002), pp. 601–606.

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

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