Journal of Memory and Language 101 (2018) 153–161

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Divided attention during encoding causes separate memory traces to be encoded for repeated events

Lili Sahakyan[a,b,*], Kenneth J. Malmberg[c]

[a] Department of Psychology, University of Illinois at Urbana-Champaign, United States
[b] Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, United States
[c] Department of Psychology, University of South Florida, United States

ARTICLE INFO

Keywords:
List-strength effect
Differentiation
Forgetting
Divided attention
Free recall
Recognition

ABSTRACT

Strengthening some items on a list adversely affects memory for the remaining items on the list – a phenomenon known as the list-strength effect (LSE; e.g., Tulving & Hastie, 1972). Whether the LSE is observed depends on how memory is tested and how items are strengthened (Malmberg & Shiffrin, 2005; Ratcliff, Clark, & Shiffrin, 1990), with free recall producing robust LSE, whereas recognition test typically producing null LSE. In this report, we examined the LSE in free recall and recognition when items were learned with full attention or under divided attention at encoding. In free recall, the results showed a robust LSE under full attention, but a null LSE in divided attention. In contrast, in recognition a null LSE was observed under full attention, but a positive LSE emerged under divided attention. Within REM theoretical framework, the combination of these findings suggests that DA reduces the tendency to accumulate information across repetitions in a single trace, thereby reducing the influence of differentiation.

Introduction

During the study of a list of items, we encode the meaning and physical properties of the items, form inter-item associations between them, and create associations between the items and the episodic context (Anderson & Bower, 1972; Humphreys, Bain, & Pike, 1989; Lehman & Malmberg, 2013; Murdock, 1983; Raaijmakers & Shiffrin, 1980). What features of the event are encoded, and the extent to which they are encoded is determined by task demands and the goals of the participant (Atkinson & Shiffrin, 1968). For instance, an instruction to attend to semantic content of the words enhances the encoding of a word's meaning (e.g., Craik & Lockhart, 1972). Likewise, instructions to create sentences out of temporally adjacent words increases the tendency to encode inter-item associations and to recall the words in the order in which they were studied (e.g., Lehman & Malmberg, 2013). Hence, different tasks may orient attention to the encoding of item and inter-item content of new memory traces.

Less is known about the attentional demands of contextual encoding. Explicit instructions to integrate items and context during study enhance context-dependent memory (Murnane, Phelps, & Malmberg, 1999). Without explicit instructions to integrate items with their context, item-to-context associations are created and enhanced by distributed or spaced repetitions of the items (Murnane & Phelps, 1995). Taken together, these results suggest that both item and context information may be encoded automatically, but additional attentional resources are beneficial to both when applied.

Although encoding under divided attention (DA) conditions impairs memory (e.g., Mulligan, 2008), the mechanisms underlying the disruption are not well understood. DA could disrupt the encoding of some or all types of information comprising an event. DA could impair the encoding of items, without perhaps affecting the amount of context stored in episodic trace, and/or it could impair the degree of contextual encoding. It is also possible that DA disrupts the ability or tendency to access prior episodic traces, which impedes their ability to be updated and increases the tendency for repetitions to be represented by multiple independent traces as opposed to accumulating in the original trace (Flexser & Bower, 1974; Hintzman, 2010 for a review).

List-strength manipulations

In our approach to these questions, we used a mixed-pure lists paradigm, in which memory is tested for three lists of items, varying in composition. The term *mixed list* refers to a list that contains a mixture of strong and weak items, whereas the term *pure strong* or *pure weak* refer to separate lists, where all items are either strongly or weakly encoded. When the strengthening operation involves distributed or

* Corresponding author at: Department of Psychology, University of Illinois at Urbana-Champaign, 603 E. Daniel St., Champaign, IL 61820, United States.
E-mail address: LSAHAKY@illinois.edu (L. Sahakyan).

https://doi.org/10.1016/j.jml.2018.04.004
Received 9 December 2017; Received in revised form 24 March 2018

spaced repetitions of the items and memory is tested via free recall, strong items are better recalled for mixed lists compared to pure-strong lists, whereas weak items are better recalled for a pure-weak list compared to the mixed list (Malmberg & Shiffrin, 2005; Tulving & Hastie, 1972). In other words, although strong items are better recalled than weak items, the effect is magnified on the mixed list compared to the pure lists. This is known as *a positive list-strength effect or LSE*, which is commonly observed when the strengthening operation involves spaced repetitions. In contrast, when memory is tested with recognition or cued-recall, typically a null or slightly negative LSE is observed, where the strength of other items on the list has little or perhaps a beneficial effect on memory of remaining items – strong items show advantage over weak items, but the magnitude of difference is typically similar for pure and mixed lists, and sometimes it is slightly smaller on the mixed list, known as the negative LSE (Ratcliff et al., 1990; Wilson & Criss, 2017).

A comprehensive account of list-strength effects

Within the framework of the search of associative memory theory (SAM, Shiffrin, Ratcliff, & Clark, 1990) and the retrieving effectively from memory theory (REM, Shiffrin & Steyvers, 1997), there are two primary sources of information – item information, and context information, and each type of information can be used to probe memory. Free recall and recognition memory are probed with different types of information – context information is used in free recall, whereas item (plus context) information is used to probe recognition. Strengthening operations are assumed to increase the amount of information stored in a memory trace representing the occurrence of the item itself and the context in which it occurred. A well-known strengthening operation is spacing or distributing the repetitions of list items, leading to better memory for items repeated in spaced fashion compared to massed fashion (Crowder, 1976; and Hintzman, 1974 for reviews). Spaced repetitions produce traces containing additional information about the item and additional information about the encoding context compared to traces of items that were not repeated or repeated in a massed fashion (Malmberg & Shiffrin, 2005).

Free recall

In the present REM model, the positive LSE for free recall is explained in terms of retrieval competition arising from the strength of contextual information in the memory trace (Malmberg & Shiffrin, 2005). Free recall is assumed to be initiated with context cues, and the match between the context stored in the traces and context used to probe memory determines which traces are sampled. The traces that contain more context features that match context cue are more likely to be sampled compared to traces that contain fewer context features that match the context cue. According to the one-shot hypothesis for context encoding, all traces have about the same amount of context stored in them after a pure list is studied, and thus all other things being equal have the same chance to be sampled (Malmberg & Shiffrin, 2005). The situation is different after a mixed list is studied, because items strengthened via spaced repetitions have more context features stored than weakly encoded items. Hence, the traces of items studied more than once in a spaced fashion dominate the traces of item studied only once (or presented in a massed fashion) in the retrieval competition. This produces the positive LSE in free recall. Thus, the relative magnitude of LSE can be indicative of how a particular manipulation affects the strength of contextual encoding.

Recognition

The same encoding assumptions that predict a positive LSE for free recall, predict a slightly negative or null LSE for item recognition (Shiffrin & Steyvers, 1997, 1998; Criss, 2006). One of the assumptions in REM is that the items are stored in memory probabilistically, which means that the memory traces of items are incomplete. During encoding, each feature is stored with some probability, and if a feature is not stored, that feature value in the trace is zero, and indicates a lack of information. As items are strengthened (via increased study time, depth of encoding, or repetitions), the memory traces of those items become updated (by replacing the zero features), which makes them more complete and informative. As additional item features accumulate in the memory trace, it decreases the similarity between the contents of different traces, known as *differentiation* (Criss, 2006, 2009; Kilic, Criss, Malmberg, & Shiffrin, 2017).

The key predictions concern the effect of differentiation on item recognition. According to REM models, strengthening operations increase the hit rates and lower false-alarm rates for strong items. Retrieval is assumed to involve a global-matching process, which compares the test item to the contents of all the traces stored in memory during the study phase of the experiment. The more similar a trace is to the test item, the more positive evidence it provides. Hence, strengthening operations increase the hit rates. Importantly, increasing the strength with which a trace is encoded decreases its similarity to representations of other items, which lowers false-alarm rates. When strength is manipulated between lists, the increase in hit rates and decrease in false-alarm rates is referred to as the *strength-based mirror effect* (Wixted & Stretch, 2004).

Spaced repetitions produce an increase in the number of item features stored compared to massed repetitions (Malmberg & Shiffrin, 2005). An important finding is that when strength is manipulated via spaced repetitions, recognition accuracy is unaffected by the composition of the study lists, unlike free recall where a positive LSE is observed. In fact, slightly negative LSEs for item recognition have often been reported (Ratcliff et al., 1990). This is predicted by the REM differentiation model because adding strong item representations to memory reduces the noise associated with the global-matching process, on the assumption that at least half of the repetitions result in the accumulation of features in a trace representing a previous event (Shiffrin & Steyvers, 1997; Malmberg, Holden, & Shiffrin, 2004). If trace accumulation does not occur or rarely occurs, then a positive LSE is predicted (e.g., Murnane & Shiffrin, 1991a, 1991b).

The interaction between list-strength and divided attention

The literature suggests that concurrent tasks performed at encoding not only impair item memory, but they also impair the ability to identify the source or the context in which those items were presented (e.g., Naveh-Benjamin, Guez, & Marom, 2003; Naveh-Benjamin, Guez, & Shulman, 2004; Naveh-Benjamin, Hussain, Guez, & Bar-On, 2003; Smyth & Naveh-Benjamin, 2016). Thus, DA at encoding impairs explicit retrieval of context. If encoding of context is impaired by concurrent tasks, it would lead to difficulty retrieving the context during the test phase. Although deficits in source identification due to DA can be viewed as consistent with the hypothesis that encoding of context demands attention, such findings do not provide direct evidence for it. Source identification involves a complex attributional process that can be influenced by many factors, including guessing biases (Chalfonte & Johnson, 1996). Also, source identification requires retrieval of context given the item, whereas the current investigation examines the opposite problem – retrieval of items given a context cue, allowing us to examine how context *encoding* is affected by DA.

Specifically with respect to the LSE, Malmberg and Shiffrin (2005) speculated that attentional factors may negatively affect updating of existing memory traces when items are repeated in a spaced fashion. For instance, it is likely that attentional demands limit the ability to encode item features representing an event. Moreover, it is possible that under taxing loads, retrieval of a trace stored previously is impaired, making it difficult to update it with item and/or context features representing the second (or subsequent) encoding attempts. If so, overall free recall and recognition performance should be harmed, the positive LSE for free recall should be severely diminished, and the slightly negative or null LSE for recognition should be reversed.

The current study examines the degree of contextual encoding by implementing the LSE paradigm under conditions of DA an FA. If DA impairs mostly the strength of the item without significantly affecting the amount of context encoded, there would be a similar magnitude of the LSE across the DA and FA conditions (in free recall). However, if DA impairs not only the strength of the item, but also the strength of context that is stored with that item, we would expect reduced LSE compared to FA condition (in free recall).

Experiment 1

In this experiment, we assessed the interaction between a list-strength manipulation and divided attention during study using a free recall paradigm. The strengthening manipulation was between items studied two times in a massed fashion and items studied two times in a spaced fashion. In the control condition subjects study pure weak lists (all massed repetitions), pure strong lists (all spaced repetitions), and mixed lists. In the divided, attention condition, subjects performed a concurrent digit-monitoring task during encoding. If taxing attentional resources impairs participants' ability to bind item and context information and/or update previously stored traces for items repeated a spaced fashion, then the positive LSE predicted in the control condition should be disrupted or eliminated in the divided attention condition.

Method

Participants

A total of 72 UNCG undergraduates from Introductory Psychology courses participated in this study for course credit. Demographic information was not collected as part of the study. Sample size was based on previous studies on LSE with free recall from our lab and other published findings on this topic. All participants were randomly assigned to full attention (FA) or divided attention (DA) conditions.

Design

The design was List (pure vs. mixed) × Strength (strong vs. weak) × Group (FA vs. DA) mixed factorial design, with Group as the only between-subjects factor.

Materials

The stimuli included 72 unrelated English words, 4–7 letters long, with a median frequency of 29.00, and an interquartile range of 56.25 (according to Kucera-Frances norms), randomly divided to create three study lists for each participant. A *pure-strong* list contained 24 items presented twice in a spaced fashion. The lag between the spaced presentations was always seven words, and items were repeated in the same order in which they appeared during the first presentation. A *pure-weak* list contained 24 items presented twice back-to-back (i.e., all-massed list). A *mixed-list* contained 12 strong and 12 weak items each presented twice either in a spaced or massed fashion (i.e., half-spaced and half-massed). Each tercile of the mixed list contained an equal number of strong and weak items. The mixed list began with a strong or a weak item equally often. The presentation order of the three lists was fully counterbalanced.

Procedure

Participants were told that they were going to study and recall three different lists of words, and were warned that words would be repeated within each list. Words were presented for 6 s each during which participants were asked to do a rating task. During the first presentation of the word, participants made a pleasantness judgment by clicking on a "+" or "−" sign that appeared on the screen underneath the word, for pleasant and unpleasant, respectively. During the second presentation, they made a graded pleasantness judgment, using a 1–5 scale, where higher numbers indicated stronger liking. Two different versions of the same orienting task (as opposed to two different orienting tasks) were selected to reduce confusion from switching between the orienting tasks and studying the words, especially on the mixed lists.

The presentation of each list was followed by a brief filler task for 30 s, during which series of letters appeared on a screen (one at a time) along with a blank window, and the task was to fill in the blank window with the next sequential letter of the alphabet. After the distractor task, participants were given 90 s to recall as many words as they could remember from the given list by typing the words into computer. The procedure was repeated until all three lists (Pure-Strong, Pure-Weak, and Mixed-List) had been studied and recalled.

Participants in the DA condition engaged in the same procedure, with an exception that they also performed a concurrent digit-monitoring task during encoding. Specifically, a random sequence of single digits (from 1 to 9) was spoken by a female voice at a rate of 2 s per digit, and participants were told to monitor the sequence and indicate whenever they detected three odd digits in a row.

Results

The results are summarized in Fig. 1. A mixed ANOVA the proportion of items correctly recalled was conducted, with List (pure vs. mixed) and Strength (strong vs. weak) as within-subjects factors, and Group (FA vs. DA) as a between-subjects factor. There was a significant main effect of Group, indicating that the DA group performed worse than the FA group, $F(1,70) = 5.10$, $MSE = 0.044$, $p = .027$, $\eta^2 = 0.06$. A significant main effect of Strength confirmed that strong items were better recalled than weak items, $F(1,70) = 52.72$, $MSE = 0.014$, $p < .001$, $\eta^2 = 0.43$. A positive LSE is indicated by a List × Strength interaction, $F(1,70) = 18.88$, $MSE = 0.013$, $p < .001$, $\eta^2 = 0.21$. Importantly, however, there was a 3-way interaction, indicating that the magnitude of the LSE varied across the groups, $F(1,70) = 8.37$, $MSE = 0.013$, $p = .005$, $\eta^2 = 0.11$.

In the FA group, there was also a main effect of Strength, $F(1,35) = 31.51$, $MSE = 0.014$, $p < .001$, $\eta^2 = 0.47$, and a significant List × Strength interaction confirming the presence of LSE, $F(1,35) = 29.26$, $MSE = 0.012$, $p < .001$, $\eta^2 = 0.46$. Strong items were better recalled on a mixed list than on a pure list, $t(35) = 3.97$, $p < .001$, Cohen's $d = 0.68$, whereas weak items were better recalled on a pure list than on a mixed list, $t(35) = 3.36$, $p = .002$, Cohen's $d = 0.56$. These results indicate that both the strengthening component and the weakening component of the LSE were obtained in the FA group.

In the DA group, there was a significant main effect of Strength, $F(1,35) = 21.43$, $MSE = 0.013$, $p < .001$, $\eta^2 = 0.38$, but no List × Strength interaction, $F(1,35) = 0.95$, $MSE = 0.015$, $p = .36$, $\eta^2 = 0.03$. Strong items were *not* better recalled on a mixed list compared to a pure list, $t(35) = 0.19$, $p = .85$, Cohen's $d = 0.04$, and weak

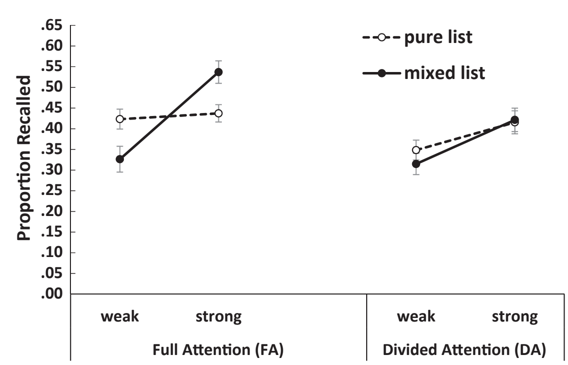

Fig. 1. Proportion recalled as function of list type, strength, and group in Experiment 1. Error bars represent SE of the mean.

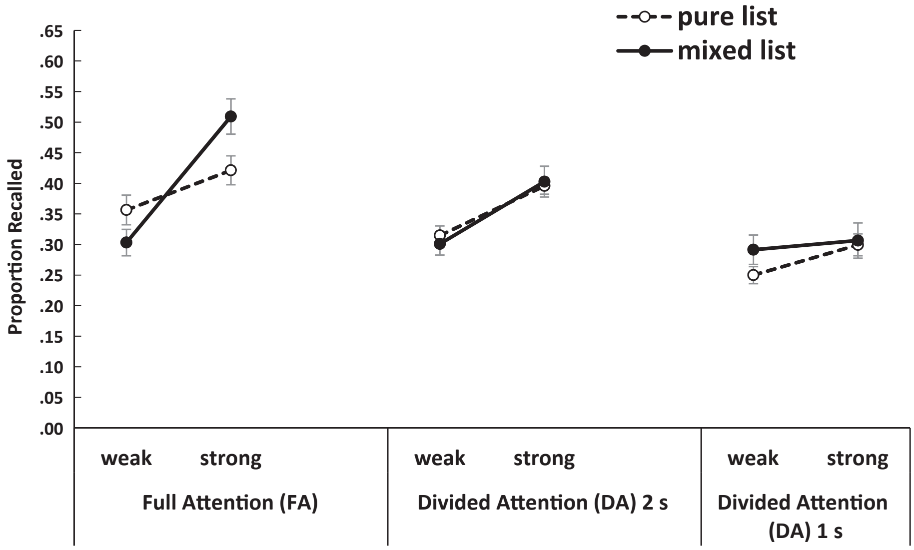

Fig. 2. Proportion recalled as function of list type, strength, and group in Experiment 2. Error bars represent SE of the mean.

items were *not* better recalled on a pure list compared to a mixed list, $t(35) = 1.38$, $p = .18$, Cohen's $d = 0.24$. Thus, neither the strengthening nor the weakening component of the LSE were observed in the DA group, despite obtaining the main effect of strength.

We also computed Bayes factors to assess the evidence for or against the null hypothesis of no List × Strength interaction in FA and DA conditions respectively (using default priors in JASP, JASP Team, 2016). We compared the evidence for a model that includes only the main effects of Strength and List against an alternative model that additionally includes the List × Strength interaction. The Bayes factor BF_{10} may be interpreted as the ratio of evidence in favor of a model that includes the interaction in contrast to a "null" model that includes only main effects of List and Strength.

In the FA group, $BF_{10} = 287.12$, which means that the data are 287 times more likely to be the outcome of a model that includes an interaction than one without the interaction included. In the DA group, $BF_{10} = 0.29$, which suggests that the data were 3.41 times more likely to be the outcome of a model that does not include an interaction.

Experiment 2

Experiment 1 was the first study examining the LSE with divided attention. Dividing attention decreased free recall performance. In addition, both groups benefited from distributed presentations as evidenced by the main effect of strength in both conditions. However, the most important finding concerned the LSE. Whereas LSE was robust in FA group, it was significantly reduced in the DA group and perhaps eliminated. As discussed above, this suggests that DA during encoding reduces the tendency to accumulate information across repetitions in a single trace and increases the tendency to store multiple independent traces representing the repetitions. We will put aside a more detailed discussion of these mechanisms for the moment because given the theoretical and practical importance of these findings, the purpose of Experiment 2 was to replicate them. In addition to including the FA and DA conditions employed in the previous experiment, we also included another group of DA participants, who performed the same digit monitoring with faster presentation rate of the digits. The more challenging divided attention condition was used to determine if a more robust manipulation would produce a more robust interaction between list-composition and attentional demands.

Method

Participants

Participants were 108 UIUC undergraduates from Introductory Psychology courses who participated for course credit. Demographic information was not collected as part of the study. There were 36 participants in the FA condition, and 72 participants in two versions of the DA condition, with 36 participants in slow and fast versions of DA condition.

Materials and procedure

The stimuli, design, and the procedures were similar to Experiment 1. The same control condition was used, and the same DA condition was used, in which the presentation rate of the digit monitoring task was 2 s per digit. However, an additional DA condition, which was aimed to be more challenging, was also included. It had a digit presentation rate of 1 s per digit.

Design

The design was List (pure vs. mixed) × Strength (strong vs. weak) × Group (FA vs. DA-2 s vs. DA-1 s) mixed factorial design, with Group as the only between-subjects factor.

Results

The results are summarized in Fig. 2. Proportion correct recall was analyzed with mixed ANOVA, using List (pure vs. mixed) and Strength (strong vs. weak) as within-subjects factors, and Group (FA vs. DA-2 s vs. DA-1 s) as the between-subjects factor.

There was a significant main effect of Strength, indicating that strong items were better recalled than weak items, $F(1, 105) = 59.90$, $MSE = 0.014$, $p < .001$, $\eta^2 = 0.36$. There was also a significant main effect of Group, $F(2, 105) = 12.18$, $MSE = 0.034$, $p < .001$, $\eta^2 = 0.19$. Namely, the FA group recalled the most (0.40), followed by the DA-2s group (0.35), and the DA-1s group recalling the least (0.29). There was also a significant Strength × Group interaction, $F(2, 105) = 6.98$, $MSE = 0.014$, $p = .001$, $\eta^2 = 0.12$. In the FA group, strong versus weak item recall difference was 13% (0.46 vs. 0.33) and was significant, $t(35) = 6.51$, $p < .001$, Cohen's $d = 1.07$. In the DA-2s group, strong versus weak difference was 9% (0.40 vs. 31), and was significant, $t(35) = 5.33$, $p < .001$, Cohen's $d = 1.02$. In the DA-1s, strong versus weak difference was attenuated to 4% (0.31 vs. 0.27), and the difference was marginally significant, $t(35) = 2.01$, $p = .05$, Cohen's

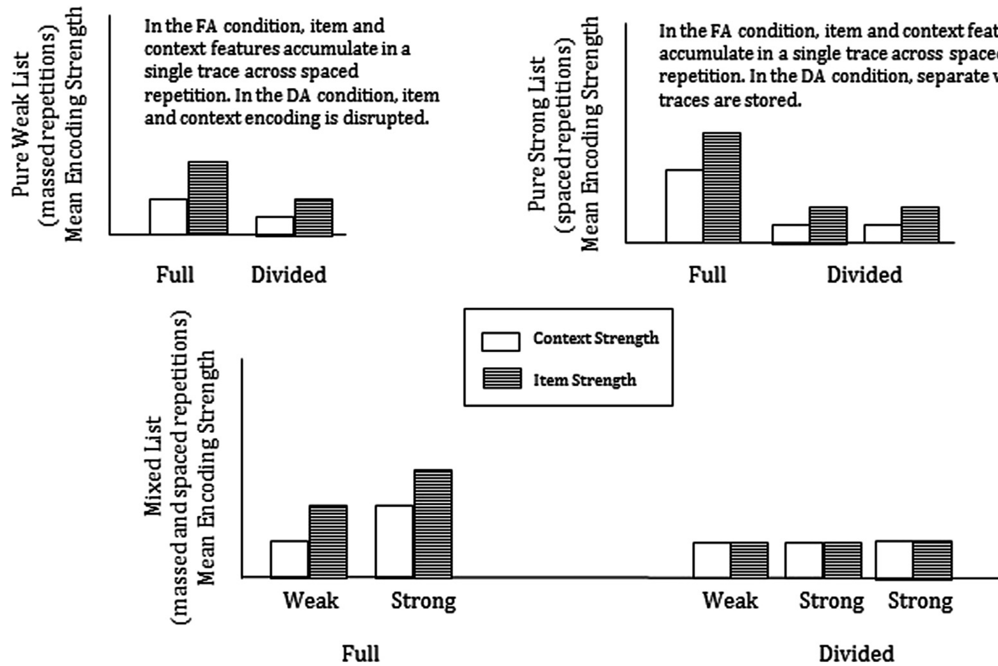

Fig. 3. Conceptual representation of the effect of divided attention on the encoding of item and context information.

$d = 0.34$. Thus, as the DA task became more demanding, both the overall level of recall was reduced and the overall strength effect was reduced, presumably because participants may have missed noticing some of the repetitions. Importantly, consistent with the previous experiment, we observed a 3-way interaction, $F(2, 105) = 4.36$, $MSE = 0.014$, $p = .015$, $\eta^2 = 0.08$. To follow up the interaction, we performed a List × Strength repeated ANOVA in each group separately.

In the FA group, there was a main effect of Strength, $F(1, 35) = 28.41$, $MSE = 0.024$, $p < .001$, $\eta^2 = .45$, and an interaction with List, $F(1, 35) = 9.93$, $MSE = 0.014$, $p = .003$, $\eta^2 = 0.22$. Strong items were better recalled on the mixed list than pure list, $t(35) = 2.19$, $p = .035$, Cohen's $d = 0.40$, whereas weak items were better recalled on the pure list than on the mixed list, $t(35) = 2.28$, $p = .029$, Cohen's $d = 0.37$. Thus, both the strengthening and the weakening components of the LSE were present in the FA group. The Bayesian analyses in the FA group revealed $BF_{10} = 5.04$, which means that the data are 5.04 times more likely to be the outcome of a model that includes the List × Strength interaction than a model without the interaction.

In the DA-2 s group, we observed a main effect of Strength, $F(1, 35) = 42.39$, $MSE = .007$, $p < .001$, $\eta^2 = .55$, but there was no interaction with List, $F < 1$. The Bayesian analyses revealed $BF_{10} = 0.18$, which mean the results were 5.71 more likely to be an outcome of a model that does not include an interaction, rather than one with an interaction term. Strong items were *not* better recalled on a mixed list compared to a pure list, $t(35) = 0.27$, $p = .79$, Cohen's $d = 0.06$, and weak items were *not* better recalled on a pure list compared to a mixed list, $t(35) = 0.64$, $p = .53$, Cohen's $d = 0.14$.

In the DA-1s group, there was a marginal main effect of strength, $F(1, 35) = 4.03$, $MSE = 0.011$, $p = .053$, $\eta^2 = 0.10$, but no interaction with List, $F < 1$. The Bayesian analyses revealed $BF_{10} = 0.09$, suggesting that the results were 10.2 times more likely to be an outcome of a model that does not include an interaction. Strong items were *not* better recalled on a mixed list compared to a pure list, $t(35) = 0.09$, $p = .93$, Cohen's $d = 0.02$, and weak items were *not* better recalled on a pure list compared to a mixed list, $t(35) = 1.44$, $p = .16$, Cohen's $d = 0.34$.

Finally, if the FA-group is excluded from the analyses, and the two DA groups are compared using List × Strength × Group mixed ANOVA, only the main effect of Strength ($F(1, 70) = 32.22$, $p < .001$), the main effect of Group ($F(1, 70) = 12.09$, $p = .001$), and the Strength × Group interaction ($F(1, 70) = 6.60$, $p = .012$) emerged as discussed previously. There was neither a List × Strength interaction ($F < 1$; $BF_{10} = 0.07$, suggesting that the results are 13 times more likely to be the outcome of a model that does not include an interaction), nor a 3-way interaction ($F(1, 70) = 1.12$, $p = .29$; $BF_{10} = 0.04$, suggesting that the results are 25 times more likely to be an outcome of a model that assumes no 3-way interaction). Overall, these results indicate the null LSE across both DA groups.

Discussion

These findings replicate the results of the previous experiment showing that DA disrupts the positive LSE typically observed in free recall. There are several potential explanations for null LSEs in the DA conditions. If information accumulates in the original trace with each repetition, then in addition to impairing item strength, DA is likely to also impair context strength (e.g., Naveh-Benjamin, Guez, et al., 2003, Naveh-Benjamin, Hussain, et al., 2003; Naveh-Benjamin, Guez, & Shulman, 2004; Smith & Naveh-Benjamin, 2016). If the disruption of context storage is disproportional, affecting strong traces more than weak traces, then the sampling advantage for strong items would be reduced, explaining why the LSE was reduced in the DA conditions. However, the data suggest that the disruption to the LSE is more extensive than this model would predict.

In addition, DA might have decreased the tendency to encode inter-item associations between adjacent stimuli, and there is evidence that under these conditions a smaller LSE is observed in free recall (e.g., Malmberg & Shiffrin, 2005). In the current Experiments 1 and 2, spaced items were always repeated in the same order, which is a condition that favors the encoding of relatively strong inter-item associations, if they can be updated. In the previous research examining LSE in free recall, Malmberg and Shiffrin (2005) included two spaced-repetition conditions – in one condition, spaced items were always repeated in the same order, whereas in the other condition, items were always repeated in a different order. In both conditions, a positive LSE was observed, but it was far more robust when items were repeated in the same order. This suggested that inter-item associations are used when available as retrieval cues during the course of free recall, and if DA in the current experiments interfered with inter-item associations, this could also reduce the magnitude of LSE.

A third effect of DA is conceptualized in Fig. 3. DA might reduce the tendency to accumulate information in the original memory trace and increase the tendency for a new trace to be stored for each repetition. Note that in SAM and REM models, successful retrieval requires the sampling of a trace from memory *and* the recovery of its contents. Sampling is determined by the *relative* strength of the item in comparison with the total strength of all other items in memory, whereas recovery is determined by the *absolute* strength of the item, regardless of the other items in memory. Weak traces are less likely to be recalled both because they are less likely to be sampled *and* because their contents are less likely to be recovered. Hence, items repeated in a spaced fashion tend to be represented by multiple traces under DA, accounting for a main effect of strength, but the amount of context and item information stored in weak and strong traces will be approximately the same all else being equal, reducing the ability recover the trace contents under DA. In the mixed list condition, although there will be sampling advantage for strong items because there are multiple traces stored for spaced items, this advantage is attenuated because the traces representing repetitions are relatively weak and therefore their contents are less likely to recovered (Gillund & Shiffrin, 1981; Raaijmakers & Shiffrin, 1980).

Experiment 3

The aforementioned hypotheses need not be mutually exclusive. However, the hypothesis that states DA disrupts the accumulation of storage across spaced item repetitions leads to one testable hypothesis. If information does not accumulate in the original memory trace, then memory traces will not be differentiated. If so, there should be two important outcomes. First, a positive LSE for single item recognition should be observed under DA conditions. Recall that more frequently a null LSE or even slightly negative LSE is found in recognition (e.g., Hirshman, 1995; Murnane & Shiffrin, 1991a; Ratcliff et al., 1990; Wilson & Criss, 2017; Yonelinas, Hockley, & Murdock, 1992). Murnane and Shiffrin (1991b) were the first to obtain positive LSE in recognition, but only when each repetition of a word was presented in a unique sentence. They suggested that presentations in a unique sentence context are stored as separate memory traces. Thus, if separate traces are stored with repetitions under DA conditions, then we would expect to obtain LSE in recognition. In FA condition, we should replicate the null or slightly negative LSE. If null LSE is obtained both in FA and DA conditions, it would indicate that DA does not affect differentiation; instead, information accumulates in the original trace, but DA hurts the accumulation of both item information and context information.

The second key prediction is that the ubiquitous strength based mirror effect should be disrupted in a specific way. Namely, the false alarm rate advantage for strong items should be reversed such that false alarms on pure strong lists should be greater than the false alarms for pure weak lists. To the best of our knowledge, this would be another novel outcome.

Method

Participants

A total of 90 UIUC undergraduates recruited from Introductory Psychology courses participated in this study for course credit. The demographic information was not collected as part of the study, and sessions were scheduled to meet a sample size of a minimum of 45 in each condition, which is greater than the previous studies on LSE in recognition. All participants were randomly assigned to FA or DA conditions.

Materials

The stimuli included 72 words from Experiment 1, and an additional set of 72 words was created that matched the original set on word length and frequency. Both sets served equally often as the target and the distractor list.

Procedures

The presentation order of the items within each study list, as well as the presentation order of the study lists followed the procedures used in the previous experiments, with a few changes. During encoding, items were presented at a rate of 2 s per word, and there was no orienting task during encoding. These changes were implemented to maintain the list length and repetition conditions similar to the previous experiments while offsetting ceiling performance. Participants were told to study the items for the upcoming memory test, but they were not informed how memory would be tested. At the end of each of the three study lists, participants performed the same distractor task as in the previous experiments for 30 s, and then took a Yes/No single-item recognition test on each study lists. Testing was self-paced, and the test list randomly intermixed targets and distractors for each participant. Concurrent digit monitoring task was employed during encoding, with digits presented at a rate of 1 s per digit.

Design

The design was List (pure vs. mixed) × Strength (strong vs. weak) × Group (FA vs. DA) mixed factorial design, with Group as the only between-subjects factor.

Results and discussion

Recognition accuracy (d') was computed after hits and false alarms were transformed using a loglinear correction (Hautus, 1995; Stanislaw & Todorov, 1999). The raw hits and false alarms are reported in Table 1. Discrimination accuracy was analyzed using a List × Strength × Group mixed ANOVA, and the results are summarized in Fig. 4.

There was a significant main effect of Group, indicating that overall accuracy was worse under DA than FA condition, $F(1, 88) = 114.05$, $MSE = 1.16$, $p < .001$, $\eta^2 = 0.56$. There was also a significant main effect of Strength, confirming that strong items were recognized better than weak items, $F(1, 55) = 10.86$, $MSE = 0.215$, $p = .002$, $\eta^2 = 0.17$. There was also a significant 3-way interaction, $F(1, 55) = 4.30$, $MSE = 0.262$, $p = .043$, $\eta^2 = 0.07$. To follow-up the interaction, we used List × Strength repeated measures ANOVA in FA and DA groups separately.

In the FA group, there was only a main effect of Strength, $F(1, 44) = 16.36$, $MSE = 0.204$, $p < .001$, $\eta^2 = 0.20$. The interaction was not significant, $F(1, 44) = 0.93$, $p = .34$. Bayesian analyses revealed $BF_{10} = 0.20$, which suggests that the results are five times more likely to be the outcome of a model that includes no interaction term

Table 1
Untransformed hits and false alarms rates across List, Strength, and Group in Experiment 3. Numbers in brackets represent SE of the mean.

	Hits			
	Full attention (FA)		Divided attention (DA)	
	Weak	Strong	Weak	Strong
Pure list	0.83 (0.02)	0.89 (0.01)	0.65 (0.02)	0.67 (0.02)
Mixed list	0.85 (0.02)	0.89 (0.02)	0.61 (0.03)	0.72 (0.02)
	False alarms			
	Full attention (FA)		Divided attention (DA)	
Pure strong	0.08 (0.01)		0.26 (0.02)	
Pure weak	0.12 (0.02)		0.23 (0.02)	
Mixed list	0.12 (0.02)		0.22 (0.02)	

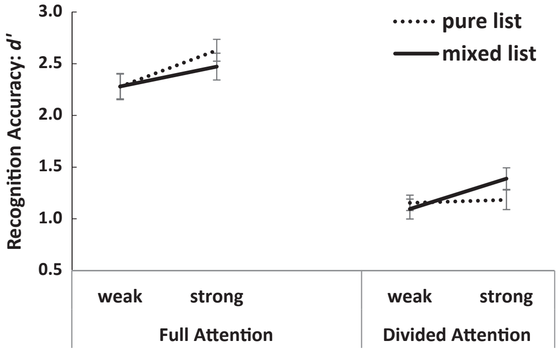

Fig. 4. Recognition accuracy as function of list type, strength, and group in Experiment 3. Error bars represent SE of the mean.

compared to a model that assumes there is a List × Strength interaction. Strong items were *not* recognized better on mixed list than on pure lists, $t(44) = 1.16$, $p = .25$, and if anything, there was a small effect in the opposite direction, with numerical advantage for strong items on pure list than on mixed list (Cohen's $d = 0.20$). Also, weak items were *not* recognized better on pure list than on mixed list, $t(44) = 0.01$, $p = .99$, Cohen's $d = 0$. Overall, these results indicate a null LSE in recognition, replicating the established findings in the literature.

In contrast, in DA group, there was a main effect of Strength, $F(1, 44) = 6.31$, $MSE = 0.186$, $p = .016$, $\eta^2 = 0.13$, and a significant interaction with List, $F(1, 44) = 6.13$, $MSE = 0.129$, $p = .017$, $\eta^2 = 0.12$, $BF_{10} = 2.19$. Strong items benefited from being on the mixed list compared to the pure list, although this small-to-medium size effect (Cohen's $d = 0.31$) fell shy of statistical significance, $t(44) = 1.95$, $p = .057$. Importantly, this "strengthening component of the LSE" is the opposite of what was observed in the FA group, where a numerical advantage for strong items was present on the pure list compared to the mixed list. In the DA group, the difference between the weak items on the pure list versus mixed list was not statistically significant ($t < 1$), but it was in the direction of the positive LSE, with impairment of weak items on the mixed list compared to the pure weak list (Cohen's $d = 0.11$). Typically, the strengthening component of the LSE is more robust than the weakening component, and the current results are consistent with it. Importantly, this weakening component emerged in the DA group, whereas in the FA group, it was completely absent with identical performance across pure weak and mixed list on the weak items. Overall, these results indicate that whereas LSE is not obtained in recognition with full attention, it emerged (albeit weakly) under the divided attention condition.

The hits and false alarms are summarized in Table 1. To examine the impact of divided attention on the strength-based mirror effect, hits and false alarms were analyzed separately using Group (FA vs. DA) × List-type (pure strong vs. pure weak) mixed ANOVA. The results are summarized in Fig. 5.

For hits, there was a main effect of group, indicating higher hit rates in the FA group than the DA group, $F(1, 88) = 79.49$, $MSE = .022$, $p < .001$, $\eta^2 = 0.48$. There was also a main effect of strength, indicating higher hit rates for strong items than weak items, $F(1, 88) = 8.27$, $MSE = 0.010$, $p = .005$, $\eta^2 = 0.09$. There was no interaction, $F(1, 88) = 2.14$, $p = .15$. For false alarms, there was a significant main effect of group, indicating that there were more false alarms in the DA group than the FA group, $F(1, 88) = 42.31$, $MSE = 0.022$, $p < .001$, $\eta^2 = 0.33$. There was also a significant interaction, $F(1, 88) = 4.52$, $MSE = 0.008$, $p = .036$, $\eta^2 = 0.05$. In the FA group, false alarm rates were lower on the pure-strong list compared to the pure-weak list $t(44) = 2.39$, $p = .02$. Taken with greater hit rates, this supports a strength-based mirror effect in the FA condition. The opposite pattern was present in the DA group, where false alarm rates were numerically higher, but not significantly so, on the pure strong list

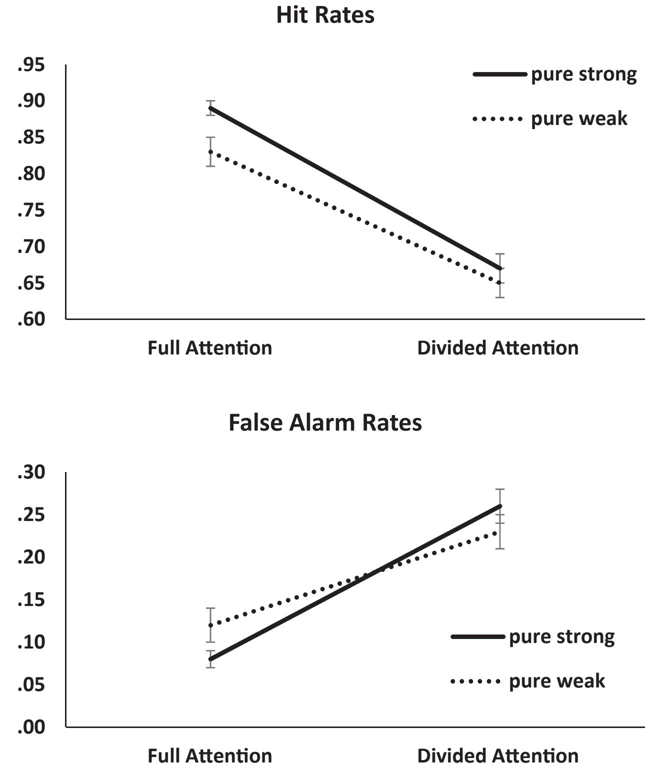

Fig. 5. Hits and false alarm rates on the pure strong and pure weak lists as a function of full and divided attention in Experiment 3. Error bars represent SE of the mean.

compared to the pure weak list, $t(44) = 1.13$, $p = .26$, indicating that DA disrupted the strength-based mirror effect.

General discussion

DA during encoding produced a null LSE for free recall and a positive LSE for recognition. The results in DA seem striking given that the opposite results are commonly reported (i.e., positive LSE for free recall and a null LSE for recognition under full attention), and those findings were replicated here under FA. The pattern of LSE findings in the DA conditions was obtained along with observing the main effect of strength, suggesting that spaced repetitions were stored in memory and they enhanced both free recall and recognition compared to massed repetitions, consistent with earlier reports in the literature (e.g., Benjamin & Craik, 2001; Glenberg, 1979; Hintzman, Summers, & Block, 1975; Sahakyan, Abushanab, Smith, & Gray, 2014; Toppino, Kasserman, & Mracek, 1991). The specific pattern of the LSE observed in the DA group across free recall and recognition cannot be explained by low performance levels because the strengthening manipulations were effective in the DA condition, suggesting that performance could not possibly be at floor. Furthermore, the opposite findings were observed across free recall and recognition, with null LSE in recall, but a positive LSE in recognition in the DA condition. Thus, low performance cannot explain the full pattern of data. Although it is unclear whether DA eliminated or sharply attenuated the positive LSE for free recall, these results suggest that DA may have reduced the tendency to store repeated items in a single memory trace, and this hypothesis is supported by the positive LSE observed for item recognition. Conceptually, the effects DA also similar to increasing the list length of the study list because repetitions of items become stored as separate traces versus accumulating in the original trace. DA likely reduced the amount of item, context, and inter-item associative information in both the weak-item and strong-item conditions. We discuss each of these four possibilities and their implications for memory theory in turn.

Divided attention disrupts trace accumulation

Of the four likely consequences of DA during encoding, the disruption of the tendency to represent multiple repetitions of an item in a single trace is the novel one, and it has the greatest theoretical importance and pragmatic implications for learning. Trace accumulation is the mechanism by which representations of event differentiate, and differentiation is the key feature of a class of models, known as item-noise models, that allow them to account for the null LSE for item recognition (McClelland & Chappell, 1998; Shiffrin & Steyvers, 1997; Criss, 2009 for a review).[1] The prior generation of memory models all predicted a positive LSE (Shiffrin, Ratcliff, & Clark, 1990) and were disconfirmed by the null or slightly negative LSE (Ratcliff et al., 1990). The prediction of a positive LSE was grounded in the global-matching mechanisms of these models that assumed that as traces became strengthened, the variability in their familiarity increases (Ratcliff & McKoon, 1991; Ratcliff, Sheu, & Gronlund, 1992). To account for the null LSE, memory researchers developed a differentiation account that assumes that as traces are strengthened, they tend to become less similar to each other, which counteracts the increase in variability associated with strengthening. Hence, strong items do not cause more interference than weak items, and may interfere less.

The null or slightly negative LSE is only predicted for item recognition because the study and test lists are typically comprised of items randomly drawn from a corpus, and therefore have randomly similar representations. When the traces cannot be assumed to be randomly similar, differentiation does not occur even though trace accumulation does (Malmberg, Holden, Shiffrin, 2004). Malmberg and Shiffrin (2005) proposed that item and context information is stored with some degree of independence, and that the time course of item and context encoding under normal encoding conditions are different. Whereas item and inter-item information may be strengthened for as long as they are attended to, the association between item and the context is fully encoded within 1–2 s of study time and does not get incremented with prolonged study time or with elaborate encoding performed on the item. In their experiments, elaborate processing or increasing study time led to overall strengthening (i.e., deeper encoding led to better memory than shallow encoding, and longer presentations led to better memory than shorter), but the advantage of strong items over weak items was similar across the pure lists and the mixed list, producing a null LSE. However, when items were repeated via spaced repetitions, Malmberg and Shiffrin (2005) observed a positive LSE in free recall, which they explained in terms of an additional opportunity to store context information with spaced repetitions.

Interestingly, the results in the DA condition in free recall resemble the depth of processing or the extra study time manipulations from Malmberg and Shiffrin (2005) study, which produced robust differences across strong and weak conditions of the experiment, but no interaction with the type of list. Malmberg and Shiffrin (2005) proposed that contextual information is automatically stored at the start of the study, and does not get stronger when the item is processed more elaborately or when the item is presented for longer durations. Those types of manipulations presumably only augment item strength, without affecting the context strength. In contrast, spaced repetitions produce an LSE because repetitions contribute to the incrementing of context strength along with each additional presentation. Importantly, both the item information and the context information accumulate in the original trace (as opposed to being stored in a new trace), and this accumulation produces more differentiated memory traces – items become more dissimilar to other items as a result of differentiation.

Divided attention disrupts item, associative, and contextual encoding

The decrease in item recognition under DA conditions is a clear sign that item encoding is disrupted. The disruption of item encoding also negatively affects the ability to recover the contents of memory traces according to SAM and REM models. This causes a decrease in free recall performance, and attenuates any sampling advantage for spaced items.

There are reasons to suspect that encoding of context may be attentionally demanding, causing LSE to be reduced in DA. Research examining the LSE across working memory capacity demonstrates that low-span participants do not show the LSE, whereas high-spans do (Sahakyan et al., 2014). Given the well documented deficits of low-spans in attentional control (e.g., Engle & Kane, 2004), reduced LSE among low-spans suggests that that contextual encoding may demand attention in order to bind it to item information.

Another set of findings related to attentional demands related to information limitations come from one of Malmberg and Shiffrin's (2005) experiments, where they obtained a partial LSE using a study time manipulation, where weak items were presented at a very fast rate (1s per item), whereas increasing the presentation rate to 2 s led to a null LSE. The authors reasoned that 1 s may not have been enough time to store a full shot of context in the weak traces, which could explain why a partial LSE was observed.

Finally, according to Lehman and Malmberg (2013) model, the ability to rehearse the items in the short-term buffer affects the ability to associate items with the study context. DA could place constraints on the number of items that can be rehearsed together in the buffer, and this would affect their associations with the study context. Based on these assumptions, one would predict reduced LSE in DA, along with lower performance in DA group (in free recall).

As noted above, DA might have decreased the tendency to encode inter-item associations between adjacent stimuli. Malmberg and Shiffrin (2005) observed a more robust positive LSE for free recall when items were repeated in the same order. This suggested that inter-item associations are used when available as retrieval cues during the course of free recall, and if divided attention in the current experiment interfered with inter-item associations, this could also reduce the magnitude of LSE. Whereas some studies found that memory for the components of the episode measured via single item recognition test is impaired to the same extent as memory for the inter-item associations (measured via associative recognition test) under DA condition (e.g., Kilb & Naveh-Benjamin, 2007; Naveh-Benjamin, Guez, et al., 2003, Naveh-Benjamin, Hussain, et al., 2003), other studies indicate that DA impairs item-item binding to a greater extent than impairing memory for the individual items (e.g., Peterson & Naveh-Benjamin, 1997; Reinitz, Morrisey, & Demb, 1994). The disruption of item-item associative information was also observed when pairs versus single-items were encoded (Lehman & Malmberg, 2013). Under the pair encoding condition, strong inter-item association were encoded among members in a given pair but not between members of temporally adjacent pairs. In contrast, when single items were studied, participants formed inter-item associations among items from more temporally distant serial positions. The present results further suggest that an item-item associative encoding strategy is less likely to be used and/or be effective under DA conditions.

Implications for memory models

Differentiation is not the only account of LSEs. Another class of models, known as context-noise models, assume that recognition decision is based on the similarity between the test context and the previous contexts in which the test item was experienced (Dennis & Humphreys, 2001). In these models, interference only arises from contextual confusions, and no comparisons are made between the test item and any other studied items. Briefly, these models assume that each time a word

[1] Trace Accumulation and the differentiation that results is also the basis for the creation of contextually independent general knowledge or lexical/semantic traces according SAM/REM models (Nelson & Shiffrin, 2013).

is encoded, it gets bound to its current context. At the time of test, the test item retrieves its pre-experimental contexts, which get compared to the current experimental context. The recognition decision is influenced by how well the retrieved contexts of the test item match the context at test. Because context is the only factor that affects memory decision in these models, and the strength of other items on the list is never compared to the test item, as a result context-noise models make the general prediction that list-composition manipulations, like those used in mixed/pure list paradigm, should produce null effects. On this prediction, they easily account for previously observed null LSE for item recognition under FA conditions. In Experiment 3, however, we observed a positive LSE, and a disruption of the strength-based mirror effect under DA condition. One might be tempted to amend the context-noise framework to include the assumption that under DA conditions context is more weakly encoded than under FA condition, but this would only produce a main effect of strength and not the list composition by attention interaction observed here, or other finding of positive list-composition effects (e.g., Malmberg & Murnane, 2002).

Conclusions

In conclusion, DA impairs the encoding of item information, inter-item, and context information, causing a decrease in free recall performance and disrupting the LSE in free recall. In addition, DA disrupts trace accumulation, thereby reducing the influence of differentiation in recognition memory. This leads to the positive LSE in recognition, and disrupts the strength-based mirror effect.

References

Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review, 79*, 97–123.
Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed systems and its control processes. In W. E. Spence, & J. T. Spence (Vol. Eds.), *The psychology of learning and motivation: Vol. 2*. New York: Academic.
Benjamin, A. S., & Craik, F. I. M. (2001). Parallel effects of aging and time pressure on memory for source: Evidence from the spacing effect. *Memory & Cognition, 29*, 691–697.
Chalfonte, B. L., & Johnson, M. K. (1996). Feature memory and binding in young and older adults. *Memory & Cognition, 24*, 403–416.
Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671–684.
Criss, A. H. (2006). The consequences of differentiation in episodic memory: Similarity and the strength based mirror effect. *Journal of Memory and Language, 55*, 461–478.
Criss, A. H. (2009). The distribution of subjective memory strength: List strength and response bias. *Cognitive Psychology,* 297–319.
Dennis, S., & Humphreys, M. S. (2001). A context noise model of episodic word recognition. *Psychological Review, 108*, 452–478.
Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Vol. Ed.), *The psychology of learning and motivation: Vol. 44*, (pp. 145–199). New York: Elsevier.
Flexser, A., & Bower, G. H. (1974). How frequency affects recency judgments. *Journal of Experimental Psychology, 103*, 706–716.
Gillund, G., & Shiffrin, R. M. (1981). Free recall of complex pictures and abstract words. *Journal of Verbal Learning and Verbal Behavior, 20*, 575–592.
Glenberg, A. M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition, 7*, 95–112.
Hautus, M. J. (1995). Corrections for extreme proportion and their biasing effects on estimated values of *d'. Behavior Research Methods, Instruments, & Computers, 27*, 46–51.
Hintzman, D. L. (1974). Theoretical implications of the spacing effect. In R. L. Solso (Ed.). *Theories in cognitive psychology: The Loyola Symposium*. Oxford, England: Lawrence Erlbaum.
Hintzman, D. L. (2010). How does repetition affect memory? Evidence from judgments of recency. *Memory & Cognition, 38*, 102–115.
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 paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 302–313.
Humphreys, M. S., Bain, J. D., & Pike, R. (1989). Different way to cue a coherent memory system: A theory of episodic, semantic, and procedural tasks. *Psychological Review, 96*, 208–233.
Kilb, A., & Naveh-Benjamin, M. (2007). Paying attention to binding: Further studies assessing the role of reduced attention resources in the associative deficit of older adults. *Memory & Cognition, 35*, 1162–1174.
Kilic, A., Criss, A. H., Malmberg, K. J., & Shiffrin, R. M. (2017). Models that allow us to perceive the world more accurately also allow us to remember past events more accurately via differentiation. *Cognitive Psychology, 92*, 65–86.
Lehman, M., & Malmberg, K. J. (2013). A Buffer model of encoding and temporal correlations in retrieval. *Psychological Review, 120*, 155–189.
Malmberg, K. J., Holden, J. E., & Shiffrin, R. M. (2004). Modeling the effects of repetitions, similarity, and normative word frequency on judgments of frequency and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 319–331.
Malmberg, K. J., & Murnane, K. (2002). List composition and the word-frequency effect for recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 616–630.
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.
McClelland, J. L., & Chappell, M. (1998). Familiarity breeds differentiation: A subjective-likelihood approach to the effects of experience in recognition memory. *Psychological Review, 105*, 724–760.
Mulligan, N. W. (2008). Attention and memory. In H. L. Roediger (Ed.). *Learning and memory: A comprehensive reference* (pp. 7–22). Oxford: Elsevier.
Murdock, B. B. (1983). A distributed memory model for serial-order information. *Psychological Review, 90*, 316–338.
Murnane, K., & Phelps, M. P. (1995). Effects of changes in relative cue strength on context-dependent recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 158–172.
Murnane, K., Phelps, M. P., & Malmberg, K. (1999). ICE: A theory of context dependent discrimination. *Journal of Experimental Psychology: General, 128*, 403–415.
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.
Naveh-Benjamin, M., Guez, J., & Marom, M. (2003a). The effects of divided attention at encoding on item and associative memory. *Memory & Cognition, 31*, 1021–1035.
Naveh-Benjamin, M., Guez, J., & Shulman, S. (2004). Older adults' associative deficit in episodic memory: Assessing the role of decline in attentional resources. *Psychonomic Bulletin & Review, 11*, 1067–1073.
Naveh-Benjamin, M., Hossain, Z., Guez, J., & Bar-On, M. (2003b). Adult-age differences in memory performance: Further support for an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory and Cognition, 29*, 826–837.
Nelson, A. B., & Shiffrin, R. M. (2013). The co-evolution of knowledge and event memory. *Psychological Review, 120*, 356–394.
Peterson, D. J., & Naveh-Benjamin, M. (1997). The role of attention in item-item binding in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*, 1403–1414.
Raaijmakers, J. G. W., & Shiffrin, R. M. (1980). SAM: A theory of probabilistic search of associative memory. In G. H. Bower (Vol. Ed.), *The psychology of learning and motivation: Advances in research and theory: Vol. 14*, (pp. 207–262). New York: Academic Press.
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., & McKoon, G. (1991). Using ROC data and priming results to test global memory models. In W. E. Hockley, & S. Lewandowsky (Eds.). *Relating theory and data; essays in honor of Bennet B. Murdock* (pp. 279–296). Hillsdale, NJ: Erlbaum.
Ratcliff, R., Sheu, C.-F., & Gronlund, S. D. (1992). Testing global memory models using ROC curves. *Psychological Review, 99*, 518–535.
Reinitz, M. T., Morrisey, J., & Demb, J. (1994). Role of attention in face encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 161–168.
Sahakyan, L., Abushanab, B., Smith, J. R., & Gray, K. J. (2014). Individual differences in contextual storage: Evidence from the list-strength effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 873–881.
Shiffrin, R. M., Ratcliff, R., & Clark, S. E. (1990). List-strength effect: II. Theoretical mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 179–195.
Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM—Retrieving effectively from memory. *Psychonomic Bulletin & Review, 4*, 145–166.
Smyth, A. C., & Naveh-Benjamin, M. (2016). Can DRYAD explain age-related associative memory deficits? *Psychology and Aging, 31*, 1–13.
Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, and Computers, 31*, 137–149.
Toppino, T. C., Kasserman, J. E., & Mracek, W. A. (1991). The effect of spacing repetitions on the recognition memory of young children and adults. *Journal of Experimental Child Psychology, 51*, 123–138.
Tulving, E., & Hastie, R. (1972). Inhibition effects in intralist repetitions in free recall. *Journal of Experimental Psychology, 92*, 297–304.
Wilson, J. H., & Criss, A. H. (2017). The list strength effect in cued recall. *Journal of Memory and Language, 95*, 78–88.
Wixted, J. T., & Stretch, V. (2004). In defense of the signal-detection interpretation of remember/know judgments. *Psychonomic Bulletin & Review, 11*, 616–641.
Yonelinas, A. P., Hockley, W. E., & Murdock, B. B. (1992). Tests of the list-strength effect in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 345–355.