2013

The Effect of Multidimensional Context Reinstatement on False Recognition

Anthony N. Correro II

The Effect of Multidimensional Context Reinstatement on False Recognition

Anthony N. Correro II

Louisiana State University

Abstract

In two experiments, people visually studied lists of associatively-related items, and multidimensional context reinstatement (font style and screen location) was manipulated in such a way to provide participants with the best conditions to employ the use of recall-to-reject processes. Associative lists of either one or six items were presented individually in distinct font styles and locations. In a later recognition task, studied items and critical lures were presented in a matching or switched font style and a matching or switched location. In Experiment 2, participants also provided rationales for their old/new responses. Results were highly consistent between Experiment 1 and 2. False recognition was greatest in the long list condition as the number of reinstated contexts increased. However, in this same condition, the rate of recollection-based rejection claims for items called "new" was greatest. These data suggest that context reinstatement not only increases activation for false memories but also enables participants to monitor and reject false memories. Global matching models cannot wholly explain these results. However, dual process theories of memory can account for the binary role of context reinstatement on the creation and rejection of false memories.

The Effect of Multidimensional Context Reinstatement on False Recognition

False memory is defined as when individuals incorrectly remember pieces of information that were never actually learned (Radvansky, 2011). One of the most effective ways to elicit false memories is through the utilization of the Deese-Roediger-McDermott (DRM) paradigm (Deese, 1959; Roediger & McDermott, 1995). In this paradigm, participants study a list of words with strong semantic associations to a nonpresented critical lure. On a later recognition test, individuals often misremember studying the critical lure and will claim that the word is "old." For example, for the list with the critical lure "sleep," individuals will encode the words "bed, rest, awake, tired," etc. but will not be shown the word "sleep." During the recognition test, participants have the tendency to incorrectly identify the lure "sleep" as previously studied well above chance levels (Roediger & McDermott, 1995).

A key factor in the development of false memories in the DRM paradigm is the effect of an individual list item's associative strength to the critical lure, known as backwards associative strength (BAS). Items with larger BAS values will be more likely to activate the critical lure than items with lower BAS values (McEvoy, Nelson, & Komatsu, 1999; Roediger, Watson, McDermott, & Gallo, 2001; Hicks & Hancock, 2002). The cumulative effect of an entire list’s BAS is a strong predictor of false recall. For example, Robinson and Roediger (1997) compared lists of three, six, nine, twelve, and fifteen associative items. They found that as list length (i.e., total list BAS) increased, the probability of falsely calling the critical lure "old" increased due to greater activation of the critical lure.

Aside from inherent factors of DRM lists like cumulative activation, the current study also addressed the role source monitoring plays on false memory. Source monitoring refers to one’s ability to remember from where a piece of information was encoded; this process involves the integration and retrieval of source and content information from the original memory trace (Radvansky, 2011). Source memory can be analyzed implicitly through context reinstatement paradigms in the DRM task where the specific method used to present an item at study is used again (correctly reinstated) at test or switched for another source (incorrectly reinstated). In a related study by Arndt (2010), each DRM list was presented with a distinctive font style. During the encoding phase, all items of each list were presented in that list's font. During the retrieval phase, associative items and critical lures were presented in the font that the theme was originally seen in (matching font) or in a font not used for the theme at study (mismatching font). Presenting associative items in a matching font resulted in higher hit rates versus mismatching fonts, and false alarm rates of critical lures were also higher when the retrieval font matched the one used to present the list's associates. The reinstatement of two different sources (font style and screen location) constituted the key manipulation of the experiments in this study.

Many theories of false memory have been applied to results from the DRM paradigm. Arndt and Gould (2006) proposed the use of dual process theories of memory (activation-monitoring theory or fuzzy-trace theory) to account for error-inflating and error-reducing processes in the DRM paradigm. Activation-monitoring theory (Roediger et al., 2001) suggests that studying a theme's associates activates the critical lure and that this activation may result in false perception-like qualities of the lure during a memory test (i.e., a source monitoring error). The theory accounts for how context reinstatement paradigms affect performance on the DRM task. If a lure is strongly activated during encoding due to semantic relatedness, then the encoding context (e.g., font or location) can also become associated to the lure (Hicks & Hancock, 2002; Roediger, McDermott, Pisoni, & Gallo, 2004).

Fuzzy-trace theory claims that DRM lists are encoded into two distinct memory traces: one verbatim and one gist (e.g., Brainerd, Reyna, & Kneer, 1995). Verbatim traces contain item specific information representing the subjective experience of encoding, and as a result, retrieval of these traces supports accurate memory. On the other hand, gist traces represent semantic similarities between the encoded words. The critical lures from DRM themes are strong gist traces, and retrieval of these traces can create false memories. This theory also incorporates a mechanism called "phantom recollection," which is the illusory experience of "remembering" unstudied lures as "old" items (Gallo, 2006). This mechanism can account for source-specifying details in false memories. Thus, fuzzy-trace theory makes very similar predictions as activation-monitoring theory in terms of factors that increase and reduce false recognition. Importantly, both activation-monitoring theory and fuzzy trace theory include control processes that allow for reductions in false memory via use of systematic strategies to consider the source of a test item’s high activation.

In contrast to these dual-process theories of false memory, Arndt (2010) promoted a serious consideration of global matching models of recognition (see also Arndt & Hirshman, 1998 and Hicks & Starns, 2006), but particularly for explaining the influence of font reinstatement. These models describe how an individual employs cognitive processes to consider the entire set of memory traces during retrieval. False memories occur because critical lures are similar to a multitude of encoded memory traces (Arndt, 2010). Test items are either a better or worse match to studied information. This degree of match is evaluated by the decision maker as providing either enough or not enough evidence to call it "old." For example, in the MINERVA2 model (Hintzman, 1988), memory traces containing semantic and contextual information for each studied associate are stored. During retrieval an individual will compare the test word with all traces stored in memory and will generate an activation value for each matching trace. These values will be summed across all memory traces to give a value representative of memory activation. High activation values can result in the participant calling a study item "old" because item and context information match (a process known as interactive cueing; Clark & Gronlund, 1996; Hicks & Starns, 2006). As a result, test items with context information corresponding to encoded items will strongly activate memory. Therefore, testing items that are highly related to studied items and that are also presented in a similar context to studied items have a strong influence on the activation value. This global matching mechanism explains why hit rates for associative items increase with source reinstatement and why false alarm rates for critical lures also increase with source reinstatement (Arndt, 2010; Hicks & Starns, 2006).

The rationale for the experiments in this study was to set up a test of global matching model predictions against predictions from activation-monitoring theory and fuzzy-trace theory. All of these theories contain a gist-based and/or familiarity process, which operates on activation (i.e., a global matching process). This global matching process is essentially the first process within dual process theories. However, dual process theories incorporate retrieval mechanisms that explain how false memory can be reduced when an individual has better access to verbatim information in test cues despite what may be a high activation value (Gallo, 2004), but global matching models do not contain this error-reducing process. An example of a process that can assist in false memory reduction is disqualifying monitoring (Gallo, 2004; also known as a recall-to-reject process). This type of monitoring occurs when an individual recalls a true memory trace or aspects of that trace to reject a potentially false memory (Gallo, 2004). For example, in the DRM paradigm an individual may literally recall encoded items of a specific theme during the retrieval phase. Encountering the critical lure "sleep" at test may cue retrieval of the words that were actually studied ("bed, rest, awake, tired") allowing one to correctly reject the lure, but exhaustive recall of an entire set of studied items is necessary to best reduce false recognition (Gallo, 2004). Global matching models do not make strong predictions that false recognition can be reduced below some baseline level because these models ascertain that accumulated activation of memory traces is necessary for recollection. In other words, these models cannot account for mechanisms independent of activation that might allow for false recognition reduction to counter such activation. Thus, the primary rationale for both experiments herein was to design experimental conditions that allow for the possibility of false recognition reduction beyond the scope of an activation accumulation process.

Two primary studies were important for designing the experimental conditions that would allow for contrasts of dual process theories and global matching models within the DRM paradigm. A study by Arndt (2010) explored the impact of list length and contextual reinstatement on false recognition. This study utilized font reinstatement paradigms with varying list lengths. In his first experiment, Arndt varied the number of associates in each DRM list (eight or two) and manipulated whether the font style presented at retrieval matched or mismatched the font style used during encoding. False alarms for critical lures were highest when the list consisted of eight versus two associates and when the critical lure was presented at retrieval in the theme’s matching encoded font. However, Arndt speculated that a recall-to-reject strategy could have been utilized in the first experiment and might have interfered with the operation of font reinstatement for his shorter list lengths. In another experiment, Arndt showed that the font-match effect was stronger for longer lists even when recall-to-reject strategies could not have been used.

However, one question is whether people really were using recall-to-reject strategies in Arndt's (2010) first experiment. An unpublished data set from Hicks and Lynn (2010) is relevant to this question because in that study a series of experiments was conducted to explore the same factors that Arndt studied: list length and font reinstatement. In their Experiment 1, the researchers had people study DRM list lengths of two, four, or six words, and during the retrieval phase, test items were presented in a font that matched or mismatched the theme's items during encoding. For the critical lures, false alarm rates increased with the number of words per list, and false alarms increased when the same font was used versus a different font. Critically, the font-match effect was stronger for longer lists as compared to the shortest list length, replicating Arndt (2010). In Experiment 2, the researchers used a speeded test manipulation to eliminate recall-to-reject strategies and only used two or six item long study lists. Font reinstatement resulted in higher hit rates and false alarm rates, and the longer list also produced increased false recognition. In this case, no interaction between list length and font reinstatement occurred, suggesting that recall-to-reject strategies might have been employed in their first experiment. In their Experiment 3, the researchers attempted to analyze the cognitive processes behind the previous experiments. If an individual rejected an item as new, follow-up questions were presented to distinguish between the utilization of a recall-to-reject strategy, the utilization of a distinctiveness heuristic, and an item's lack of familiarity. Again, longer list lengths and font reinstatement were found to increase false recognition and reinstatement produced a larger effect for the longer lists. Individuals in the two item study list reported using a recall-to-reject strategy more often than individuals in the six item study list. The researchers determined that font reinstatement helps aid recall-to-reject strategies especially when theme lists contain less items compared to more items, and these results could only be explained by dual process theories;

however, in a fourth unpublished study, not enough evidence was found to claim that a recall-to-reject strategy was being applied more often in the short length conditions. Thus, one may be able to explain the false recognition results using a global matching process.

To resolve the inconsistencies within Hicks and Lynn's (2010) study and Arndt's (2010) study, optimal conditions to elicit recall-to-reject strategies were developed for both experiments, and rationales for old/new judgments were gathered for Experiment 2. In order to examine the potential for contextual features to effectively cue studied items, DRM themes of short and long list lengths were used in both experiments of this study. However, a more extreme version of a short list condition was used, as people studied only one list item in this condition. Participants learned one word for some associative themes and six words for longer associative themes. Each theme was presented in a unique font style. The former condition should provide an optimal opportunity for disqualifying monitoring processes to be utilized because the likelihood of recalling and rejecting a single item is much greater than exhaustively recalling all six items of the long lists. Additionally, the amount of contextual cues available to aid in rejecting critical lures was analyzed by reinstating the font style of studied words, the location of studied words, or both during the test phase. Adding location as a second context feature extends prior research and allows for an additional potential recall-to-reject mechanism.

Based on these prior studies, critical lures presented in the same context as their studied theme items were predicted to increase false recognition for the longer list themes. More specifically, false recognition should increase from the completely switched condition, to the conditions in which one or the other contextual feature is reinstated, to the condition in which both features are reinstated. An interaction between list length and font reinstatement is expected based on the prior work by Arndt (2010) and Hicks and Lynn (2010). The nature of the results in the short length condition is of primary interest for the experiments in this study. When critical lures from these themes are presented in the same font and location as their studied associates, rejection-based processes are most likely to occur (i.e., the test font and location may cue people to recall the specific single item studied and call the critical lure "new"). This would result in an interaction between list length and switched versus reinstated feature conditions as in prior work (e.g., Arndt, 2010), but because rejection processes help reduce false recognition for reinstated conditions. Reinstating one or both features should increase critical lure activation but at a lower rate than the long list condition. The continued increase in activation for the short lists as context reinstatement increases would be indicative of a global matching model (Arndt, 2010). However, if disqualifying monitoring processes are being utilized to counteract this activation, false recognition may be *lower* when one or both contextual features are reinstated as opposed to when both features are switched. In other words, by giving more contextual support through reinstatement, which will increase activation in general as well as false recognition for the long lists, participants should be able to recall the single studied item related to the critical lure in the matching font style and/or location thereby allowing them to reject the lure. The strongest version of this hypothesis would be that false recognition for short lists in the completely reinstated conditions should be as low as the conditions in which one or both contexts are switched. This potential outcome is not compatible with global matching models of recognition and would require an explanation based on a second retrieval monitoring process, like activation-monitoring theory (Roediger et al., 2001) or fuzzy-trace theory (Brainerd et al., 2005).

To gain further evidence for the possibility of disqualifying monitoring in Experiment 2, after each test item was answered, follow-up questions were asked to gain insight into participants’ cognitive processes. Similar to Hicks and Lynn (2010), participants were asked to provide supplementary information when any test item was rejected (called "new"). Rejection could be based on the participant's ability to correctly recall one or more studied items because: a) he or she realized the critical item was semantically related to the encoded associate(s), but he or she remembered that it was not in the original set, b) the critical lure's reinstated font cued the studied item(s) in that font, c) the critical lure's reinstated location cued the studied item(s) in that location, or d) the test item itself simply did not seem familiar. If the participant chose any of options a) through c) as the basis for their rejection, they were prompted to recall as many of the studied items related to his or her rationale (e.g., which words *were* studied in this font, if not the test item?).

Experiment 1

List length, font reinstatement, and location reinstatement were manipulated in Experiment 1 to investigate their singular and combined effects on false recognition. Participants were expected to falsely recognize critical lures from long lists more often than from short lists because of increased activation of the lure during study. Font and location reinstatement were hypothesized to have a combined effect such that false recognition will occur at a higher rate when both contextual features are reinstated compared to only one or neither. However, based on pilot testing the font manipulation is expected to be more salient than the location manipulation, and as a result, reinstating font should cause a greater increase in false recognition compared to location reinstatement. If people utilize controlled rejection strategies, list length, font reinstatement, and location reinstatement are expected to interact in such a way that false alarm rates are exceedingly depressed in the short list condition when both contextual features are reinstated compared to the long list condition with full reinstatement.

Method

Participants

Participants (*N* = 102) were undergraduate students enrolled in psychology courses at Louisiana State University. They received extra credit or fulfilled a course research learning requirement for participating in the experiment.

Materials

The materials consisted of 24 associative lists compiled from the University of South Florida Free Association Norms Website (Nelson, McEvoy, & Scheiber, 1998). Each associative list contained the critical lure and the six related words with the highest backwards associative strength (BAS) values. See the Appendix for these lists. Sixteen fonts were obtained from the website 1001 Free Fonts (Nolan, 1998). The 16 associative lists presented during encoding were randomly matched to these fonts, and 8 of these fonts were randomly selected for use during testing for presentation of new unstudied items. The paired lists and fonts did not resemble one another; for example, the *fire* list items were not presented in a font with flames. The only constraint was that all words must be legible in their assigned fonts. E*Prime version 2.0 was used for all phases of the experiment.

During the encoding phase, 16 of the 24 associative lists were presented. Half of the 16 lists consisted of all six of the associative items (long study list condition). The other half were composed of only the associate with the highest BAS value (short study list). Therefore, 56 stimuli were presented in the encoding phase.

Each list theme was matched to a specific location on the computer screen. A 3 × 3 matrix was used to create the eight locations; no words were presented directly in the center. One long list and one short list were presented in each of the locations. Whenever the list themes were presented, they were shown in only one location in 30-point font size of the assigned font style. As a result, all associative items of the same theme were shown in the same font and location. For long lists, the 6 associates related to a given theme were presented sequentially (i.e., they were blocked by theme); otherwise, list items were presented in a pseudo-random manner.

At test each of the 24 themes' associate with the highest BAS to the critical lure and the critical lure were presented individually. Sixteen of the test items were the highest-BAS list associates (eight from the long study list condition, eight from the short study list) studied by the participants, and eight of the associates were new words from 8 nonpresented DRM themes. Of the eight studied associates from the long study list condition, two were presented in the same font and location that they were studied in, two in the same font but in a switched location, two in the same location but in a switched font, and two in a switched location and a switched font. These test conditions also applied to the highest-BAS associates from the short lists and to the critical lures from each theme such that 2 critical lures from each list length were presented in each of the four test conditions. Test fonts and locations were always familiar. Moreover, in switched test conditions, the font and/or location were switched between list themes in a given length condition rather than across length conditions.

For the eight unstudied themes presented at test, 16 test items were presented including the highest-BAS associate and the critical lure from that associate's theme. The encoding fonts and locations were used to present these items as well. Half were shown in a previously studied font in the location paired with the font, and half were shown in an old font not in that font's original location. These conditions served to identify baseline rates of false recognition for unfamiliar themes.

The associative lists were split into six counterbalanced sets across participants. The counterbalanced sets ensured that each associative list was presented in all of the reinstatement conditions (same font and location, same font but switched location, same location but switched font, and switched font and switched location for studied items; matched font and location as well as mismatched font and location for new items). The counterbalanced sets also ensured that a list was presented twice in the long condition during encoding, twice in the short condition during encoding, and twice at test as unstudied items.

Design and Procedure

The experimental design of this study is a 2 (study list length) × 2 (font condition) × 2 (location condition) fully repeated measures factorial. Study list length represents the number of associates studied (long or short). The font condition represents whether or not the appropriate studied font was reinstated at test (same versus switched). Similarly, the location condition represents whether the location was reinstated or switched between study and test.

Before beginning the experiment, participants' informed consent was obtained. During the study phase, participants were told that they would be shown a list of words in different fonts and locations on the screen. They were instructed to do their best to learn the words, fonts, and locations for a later test. Each word was viewed one at a time for four seconds. When long lists were presented, all items in that list were shown sequentially in a given font and location before the next list theme was presented. All individuals studied 56 items. After the study phase, participants completed simple arithmetic problems for one minute.

During the test phase, the 24 associates and 24 critical lures were presented individually as described previously. A theme's critical lure was always tested prior to its related associate to ensure that any potential test priming did not affect estimates of false recognition. Participants were told to decide whether the presented word was studied (old) or unstudied (new). Participants were informed that the font styles and locations may be switched between test themes and that some test words may be similar in meaning to studied words. This information was provided so that participants could make their old/new decisions carefully and technically without regard to the test context. The test phase was self-paced, but the experimenter recorded the participant's answers to help ensure that participants were thinking carefully about their decisions. The entire experiment took participants approximately twenty minutes to complete. At the conclusion of testing, participants were fully debriefed.

Results and Discussion

An initial consideration of baseline false alarm rates was made prior to examining the hit rates for studied items and the false alarm rates for critical items. For the brand new items, some of the fonts and locations used were previously paired for studied items; others were recombined. An example of a previously-paired combination would be if the word "bed" was presented in the **Bauhaus 93** font style and in the upper-left corner of the screen. A brand new item unrelated to any studied themes might be shown at test in this font and in this location. In contrast, a recombined font-location test context would be to show a test item in the **Bauhaus 93** font but in a different location. Both contexts were experienced during the study phase, but for different studied items as opposed to the same item. Moreover, some of these font styles were experienced six times each (for long DRM themes) and others were experienced only once. Therefore, seeing brand new test items in fonts that were experienced six times could increase baseline false recognition because of the increased activation due solely to font experience.

In fact, such differences did occur. For new items shown in previously-paired fonts and locations seen only once, false recognition was 13%. For recombined fonts and locations seen only once, false recognition was 11%. However, these rates were larger for fonts seen six times each: 18% for previously-paired fonts and locations and 17% for recombined fonts and locations. Baseline false recognition was higher when items were shown in fonts and locations that were experienced more often.

These baseline false alarm rates were therefore subtracted from the hit rates for studied items and from false recognition rates for critical lures. This correction accounts for the baseline probability that items seem familiar solely because of the fonts and locations used to present studied items. Therefore, the analyses reported below for hit rates and false recognition rates were performed on these corrected values. An alpha level of .05 was used as the significance criterion for hypothesis tests. For post hoc comparisons, Bonferonni corrections were applied to determine significance.

Studied list items. The corrected means for all test conditions are reported in Figure 2. A fully repeated measures 2 (study list length) × 2 (font manipulation) × 2 (location manipulation) ANOVA was conducted on corrected list item recognition. The effect of location reinstatement was not significant, $F(1, 101) = 0.23$, $MSE = 0.11$, $p \geq .63$, partial eta squared < .01. The main effect of list length was only marginal, $F(1, 101) = 3.31$, MSE = 0.25, $p = .07$, partial eta-squared = .03. Recognition was nominally better when longer lists were studied. The font by length interaction was also not significant, $F(1, 101) = 1.65$, MSE = 0.12, $p \geq .20$, partial eta-squared = .02. The three-way interaction was not significant, $F(1, 101) = 1.65$, $MSE = 0.12$, $p \geq .20$, partial eta squared = .02.

A significant main effect of font reinstatement was produced, $F(1, 101) = 32.51$, $MSE = 0.13$, $p < .05$, partial eta squared = .24. Hit rates were higher for studied items when they were presented in the same font rather than a different font. A significant interaction between length and location was found, $F(1, 101) = 10.70$, $MSE = 0.08$, $p < .05$, partial eta squared = .10. Pairwise-comparisons revealed that hit rates were greater for the short study lists when the location was correctly reinstated compared to when it was incorrectly reinstated, $t(101) = 2.38$, $p < .025$. The opposite trend was present for long lists but did not reach significance, $t(101) = 1.86$, $p \geq .06$.

A significant interaction between font and location was also found, $F(1, 101) = 11.03$, $MSE = 0.12$, $p < .05$, partial eta squared = .10. Pairwise-comparisons revealed that switching the font had an inordinately greater effect when location was reinstated, $t(101) = 5.63$, $p < .025$, compared to when location was different, $t(101) = 2.26$, $p < .025$, although both comparisons were significant.

Critical lure false recognition. Mean corrected false alarm rates for critical items related to studied themes are presented in Figure 3. A fully repeated measures 2 (study list length) × 2 (font manipulation) × 2 (location manipulation) ANOVA was conducted on these data. The effect of location reinstatement was not significant, $F(1, 101) = 0.18$, $MSE = 0.11$, $p \geq .65$, partial eta squared < .01. The interaction between length and location was not significant, $F(1, 101) = 0.09$, $MSE = 0.12$, $p \geq .76$, partial eta squared < .01. The interaction between font and location was also not significant, $F(1, 101) = 0.26$, $MSE = 0.12$, $p \geq .60$, partial eta squared = .01. In addition, the three-way interaction was not significant, $F(1, 101) = 0.35$, $MSE = 0.13$, $p \geq .55$, partial eta squared < .01.

A significant main effect of list length was found, $F(1, 101) = 67.12$, $MSE = 0.27$, $p < .01$, partial eta squared = .40. Participants were more likely to call critical items "old" from long study lists than from short study lists. A significant main effect of font reinstatement was also found, $F(1, 101) = 30.42$, $MSE = 0.13$, $p < .01$, partial eta squared = .23. False recognition was greater when the font used to present a list's associates was reinstated for that list's critical lure compared to when the font was switched for one not paired with the list during encoding.

Of primary interest to the current study is the interaction between length and font reinstatement, and this interaction was significant, $F(1, 101) = 30.58$, $MSE = 0.11$, $p < .05$, partial eta squared = .23. Pairwise-comparisons revealed that font reinstatement had little effect on false recognition for short list items, $t(101) = 0.21$, $p \geq .82$, with reinstated and non-reinstated conditions producing about 11% corrected false recognition. But reinstating the studied fonts at test resulted in significantly higher false alarm rates with the long lists (54%) compared to when the font was incorrectly reinstated (28%), $t(101) = 7.62$, $p < .01$.

Summary. For both studied associates and critical lures, the font reinstatement manipulation had the greatest influence on performance. Providing participants with a greater amount of contextual information by reinstating the font at test resulted in better recognition discriminability as compared to switched fonts. This was seen most drastically for long list themes, especially for critical lures from those themes. Short list themes produced much less false recognition, consistent with the notion that activation for critical lures is lower when fewer list associates are presented (cf. Robinson & Roediger, 1997). The lower false recognition rates may also be caused by participants using strategies to reject critical lures of short lists even when the contextual information should theoretically activate false memories (cf. Experiment 1 of Arndt, 2010). Because there is only one list item for the short study lists, it may be far easier compared to the long study lists for participants to reject the critical lure based on the test context. Experiment 2 was conducted to investigate this interaction further and to understand the cognitive processes responsible for this interaction.

Experiment 2

In order to ascertain the various strategies that participants may be using when identifying critical lures as "new," they were asked follow-up questions in Experiment 2. These follow-up questions presented various rationales for the participants' decisions and served as secondary evidence for specific cognitive processes or strategies used by the participant (e.g., recall-to-reject strategies). Of particular interest is whether the use of different strategies correlates with the various context reinstatement conditions at test. For example, presenting a critical lure at test with the same font style and location as its studied associates may provide the participant with retrieval cues allowing for the rejection of that critical lure. Although reinstating font and location in Experiment 1 produced the highest rates of false recognition, these rates were substantially lower for short lists as compared to long lists. Thus, these test conditions for short lists may motivate the utilization of rejection-based strategies. On the other hand, test conditions in which the font style and/or location are switched may serve as poor retrieval cues for studied associates. The patterns of false alarm rates, hit rates, and corrected recognition scores are expected to be approximately the same as those obtained in Experiment 1.

Method

Participants

Participants (N = 112) were undergraduate students enrolled in psychology courses at Louisiana State University. They received extra credit or fulfilled a course research learning requirement for participating in the experiment. Individuals who participated in Experiment 2 did not take part in Experiment 1.

Materials, Design, and Procedure

The materials and the study phase for Experiment 2 were identical to those used in Experiment 1. The experimental design remained a 2 (study list length) × 2 (font condition) × 2 (location condition) fully repeated measures factorial.

The test phase in Experiment 2 differed slightly from the test phase in Experiment 1. The only difference was that in Experiment 2 participants were asked follow-up questions in order to understand their rationales for calling test items "old" or "new." Participants answered the follow-up question for each item before the next test item was presented. The test phase was self paced, but the experimenter recorded the participant's answers to the old/new recognition task and to all follow-up questions. The entire experiment took participants approximately 30 minutes to complete.

When a participant called an item "old" at test, he or she was immediately asked why the word was called old. Six options were presented: a) the participant knew the word was old because of its generic familiarity, b) the specific word had distinct visual imagery, c) the word elicited memorable associations, d) the font was memorable, e) the location was memorable, or f) the word's position in the study list was memorable. The first response for this follow-up question represents a "know" judgment in the traditional "Remember/Know" paradigm for false memory. The other responses represent different variations of "remember" judgments. Participants were allowed to select as few or as many of the rationales as they wanted.

When a participant called an item "new" at test, he or she was immediately asked why he or she was confident the word was new. Four options were presented: a) other similar items could be recalled *in that font* that do not include the presented word, b) other similar items could be recalled *in that location* that do not include the presented word, c) other similar, semantically related items could be recalled that did not include the presented word, and d) the word was just not familiar enough. If the participant chose one of the first three options, he or she was then instructed to recall as many similar studied words (based on font, location, or semantics) as possible. These words were recorded by the experimenter.

Results and Discussion

As in Experiment 1, hypothesis tests were conducted separately for list and critical items. Baseline false alarm rates for new unstudied items were on average around 27%, but differed dramatically across the various font/location conditions as in Experiment 1. For new items shown in previously-paired fonts and locations seen only once during encoding, false recognition was 29%. For recombined fonts and locations seen only once during encoding, false recognition was much lower at 17%. For previously-paired fonts and locations seen six times each during encoding, this rate was 27%. Baseline false recognition was 30% for recombined fonts and locations seen only once before. These rates were subtracted from the corresponding hit rates for studied items and false recognition rates for critical lures to create corrected measures. Interestingly, these baseline false alarm rates are greater than those found in Experiment 1. Having participants think about their rationales for calling an item "old" or "new" may inherently cause participants to doubt themselves and incorrectly call new items "old." Moreover, 14 participants in this experiment were extreme "old" responders; these individuals called 75% or more of all items "old." The results described below did not change when these participants were excluded from analyses, and therefore, they were retained.

An alpha level of .05 was used in all hypothesis tests. For post hoc and pairwise-comparisons, a Bonferonni correction to the experimentwise Type I error probability was applied to determine significance.

Studied list items. The corrected means for all test conditions are reported in Figure 5. A fully repeated measures 2 (study list length) × 2 (font manipulation) × 2 (location manipulation) ANOVA was conducted on studied list item true recognition. The main effects of length and location reinstatement were not significant: for length, $F(1, 111) = 0.84$, MSE = .30, $p \geq .36$, partial eta squared < .01; and for location, $F(1, 111) = 1.77$, MSE = .13, $p \geq .18$, partial eta squared = .01. The two-way interaction between length and location was not significant, $F(1, 111) = 2.29$, MSE = .14, $p \geq .13$, partial eta squared = .02. Finally, the three-way interaction was not significant, $F(1, 111) = 2.54$, $MSE = 0.17$, $p \geq .14$, partial eta squared = .02.

A significant main effect of font reinstatement was found, $F(1, 111) = 76.45$, $MSE = 0.09$, $p < .05$, partial eta squared = .41. Two interactions were significant, the length × font interaction, $F(1, 111) = 11.66$, MSE = .15, $p < .01$, partial eta squared = .10, and the font × location interaction, $F(1, 111) = 5.52$, MSE = .11, $p < .05$, partial eta squared = .05. List length interacted with font because font reinstatement produced a larger benefit for long lists, $t(111) = 8.90$, $p < .01$, as compared to short lists, $t(111) = 2.63$, $p < .025$. Font reinstatement interacted with location reinstatement because reinstating font improved recognition more when location was also reinstated, $t(111) = 7.45$, $p < .01$, as compared to when location was not reinstated, $t(111) = 4.38$, $p < .01$.

For the most part, these results replicate those found in Experiment 1. In both experiments, font reinstatement had a powerful impact on corrected true recognition. In addition, font reinstatement interacted with location reinstatement, wherein changing or reinstating font had a bigger influence when location was also reinstated. The primary difference between the two experiments was that length and font reinstatement interacted in Experiment 2, but did not in Experiment 1, although the same trend was present in Experiment 1.

Critical lure false recognition. Mean corrected false alarm rates for critical items related to studied themes are presented in Figure 6. A fully repeated measures 2 (study list length) × 2 (font manipulation) × 2 (location manipulation) ANOVA was conducted on critical lure false recognition. The interaction between length and location was not significant, $F(1, 111) = 0.09$, $MSE = 0.13$, $p \geq .76$, partial eta squared < .01. The interaction between font and location was also not significant, $F(1, 111) = 0.29$, $MSE = 0.15$, $p \geq .59$, partial eta squared = .01. Finally, the three-way interaction was also not significant, $F(1, 111) = 1.22$, MSE = 0.13, $p \geq .27$, partial-eta squared = .01.

All three main effects were significant. Longer lists promoted higher rates of false recognition, $F(1, 111) = 36.56$, $MSE = 0.36$, $p < .05$, partial eta squared = .25. Reinstating the correct font style at test also promoted higher rates of false recognition, $F(1, 111) = 38.43$, $MSE = 0.15$, $p < .05$, partial eta squared = .26, as did reinstating the correct location at test, $F(1, 111) = 6.00$, $MSE = 0.12$, $p < .05$, partial eta squared = .05.

The significant main effects of font and length were qualified by an interaction between length and font reinstatement, $F(1, 111) = 30.56$, $MSE = 0.15$, $p < .05$, partial eta squared = .22. Pairwise-comparisons revealed that changing the font for short study lists had no significant effect, $t(111) = 0.53$, $p > .50$, but this effect was significant for long study lists, $t(111) = 7.79$, $p < .025$. As in Experiment 1, font reinstatement had a larger impact on false recognition for longer lists. The lone difference between the patterns in Experiments 1 and 2 for critical lures was the significant effect of location reinstatement. It is likely that having people carefully consider font style and location information in their recognition decisions made the use of location information more prevalent.

Examination of rejection strategies. Although follow-up questions were asked when participants called items "old" and "new," only the rationales for items called "new" will be discussed. Participants were allowed to select more than one rationale for each test item, and for items called "old" the pattern of rationales was not easily interpretable because most responded with the familiarity option alongside the other choices. Also, participants selected a wide range of the choices, which made interpretation of the primary rationale for calling items "old" difficult. Identifying trends in the responses made following "new" responses was clearer. If a participant called an item "new" and selected a rationale where items could be recalled, the individual was asked to recall the items that aided selection. Recall was rarely exhaustive, and like Hicks and Lynn (2010), recall rarely correlated with the participant's rationales.

Tables 1, 2, 3, and 4 show the rate of rationale selection for critical items from long study lists, critical items from short study lists, and unstudied associate items shown in fonts presented one time or six times during encoding, respectively. Note that the total number of items within each row of these tables often differs. Therefore, the focus is on the percentages of each type of rejection strategy within a given condition (i.e., the rows).

Focus first on Table 1 for critical lures related to long study lists. In the first row where both font and location were combined as during encoding (SFSL), participants called critical items "new" using rationales that support recall-to-reject strategies (font, location, or semantics rationales) 77.1% of the time. In this same condition, participants stated that the item was unfamiliar only 22.9% of the time. Whenever at least one piece of contextual information was not reinstated at test, participants were more likely to select the unfamiliar rationale (46.6% for same font different location, 59.8% for different font same location, and 49.2% for different font different location).

For critical lures related to short study lists in Table 2, this trend reemerged, except participants selected the unfamiliar rationale much more often in general as compared to the long study lists. When both pieces of contextual information were reinstated, participants called critical items "new" using the font, location, and semantic rationales 46.4% of the time. Participants used the unfamiliar rationale 53.7% of the time when both font and location were the same, but this number increased drastically when the contexts were not correctly reinstated (73.4% for same font different location, 65.9% for different font same location, and 70.6% for different font different location).

Interestingly, participants very rarely declared that other similar items in the same font could be recalled when selecting their rationales, especially when the font and location were reinstated at test. When both sources were reinstated at test (SFSL conditions), participants chose the location rationale (41.7% for long list items and 28.7% for short list items) more often than the font rationale (8.3% for long list items and 5.9% for short list items). Yet the font manipulation had a strong influence on calling both studied and critical items "old." Because of this strong influence, one would expect to see a similar influence of font manipulation for rejecting items as "new" because font information is much more salient than location information. Participants may be selecting the location rationale more often than the font rationale because there are more fonts utilized in this study than there are locations. Each of the study lists is paired with a single font (resulting in 16 distinctive fonts used during encoding), but only 8 locations are used during encoding such that two associative lists are presented in each location. As a result, participants may be recalling both of the associative lists that they studied in a specific location when they consider the location rationale. More contextual and semantic information would be available in this situation leading to the greater preponderance of location rationales.

Focusing now on Tables 3 and 4, the participants called a majority of the unstudied items "new" by using the unfamiliar rationale. Table 3 reports correct rejections for brand new lists themes shown in fonts presented once during encoding; Table 4 reports data for items shown in fonts presented six times during encoding. In Table 3, for items shown in previously studied fonts that were matched to the fonts' original locations, participants chose the unfamiliar option 57.4% of the time and either the font, location, or semantic rationales 42.6% of the time. For new items shown in previously studied fonts that were incorrectly paired with a studied location, participants selected the unfamiliar rationale 82.1% of the time and either the font, location, or semantic rationales 17.9% of the time. Clearly, new items were more often called "new" because the participant felt the items were unfamiliar, but when previously studied fonts and locations were paired for new items, participants were more likely to utilize recall-to-reject strategies compared to when the previously studied fonts and locations were incorrectly paired at test. As in Tables 1 and 2, location was identified as the basis for rejection much more often when the test item’s font was previously shown in that same test location. (22.4% vs. 11.5%).

Interestingly, font was also identified more often in the top row of Table 3 as a basis for rejection. Participants may have been able to recall the studied items that were actually presented in the paired fonts and locations for new items in order to reject these unstudied items as "new." However, when the new items were presented in mismatched studied fonts and locations, participants were less likely to use this strategy. Table 4 also deviates from Tables 1 and 2 but in a different way. When fonts and locations were paired as before, location was identified more often as a basis for rejection (30.4% vs. 19.2%). Oddly, the patterns for font-based rejection was opposite, with the bottom row of that table producing 19.8% of rejections as compared to the top row's 6.4%.

Summary. In almost every case, the results from Experiment 1 were validated in Experiment 2. Of primary interest was the length × font interaction in true and false recognition. This interaction revealed that for true recognition, source activation was greater in the long study lists compared to the short study lists leading to increased hit rates in the former condition when the font style was reinstated (although this trend was present but not significant for true recognition in Experiment 1). This interaction was also found to be significant in false recognition for critical lures, replicating prior work by Arndt (2010) and Lynn and Hicks (2010). In this situation, source activation for reinstated font styles resulted in an increased likelihood to falsely recognize the critical items, especially in longer lists when the same font style for the critical item's associative theme had been presented six times during encoding. For short lists in which only one item was studied, font reinstatement had a negligible impact on false recognition.

The condition that best supports the use of a recall-to-reject strategy is the one where both font and location were reinstated, and the participants' rationales in Experiment 2 revealed modest evidence for this claim. Providing participants with the contextual information used to study the associates related to the critical item potentially allowed them to recall those list items in order to reject the critical item. However, participants were more likely to select the font and/or location rationale for long lists than they were for short lists. This lends credence to the notion that contextual information is more strongly activated when items are studied six times in the long study lists as compared to only one time in the short study lists. As a result, there is little evidence for the hypothesis that recall-to-reject strategies are employed more often by participants when rejecting items from short study lists as compared to long study lists. In hindsight, this makes sense because a given font is viewed six times in long lists in a given location, giving people a good sense that new test items sometimes do not match well with that retrieval environment. This process may be best described as a strong sense of "not belonging" because of distinctiveness as opposed to a disqualifying rejection strategy, because no one could exhaustively recall all 6 items that were presented in that encoding context. It is in the long list conditions where greater activation for reinstated fonts and locations led to higher false recognition for critical lures and to the preponderance of recall-to-reject strategies based on contextual information. This overall pattern is consistent with the speculation made by Arndt 2010) that context reinstatement may not only play a role in error inflation (increased activation) but also error reduction (recall-to-reject strategies).

General Discussion

Nearly all significant main effects and interactions were replicated in Experiment 2 from Experiment 1. The goal of these experiments was to provide participants with optimal conditions for utilizing recall-to-reject strategies by manipulating list length and multidimensional context reinstatement. Clearly, there is a large effect of list length on false recognition (Figures 3 and 6), which replicates Arndt's (2010) study as well as the study by Hicks and Lynn (2010). True recognition and false recognition increase as the number of studied items increases (Figures 2, 3, 5, and 6), although this pattern was more noticeable for critical false recognition as opposed to true recognition, after correcting for baseline false alarm rates.

Like Arndt (2010) and Hicks and Lynn (2010), context reinstatement affected false recognition for critical lures. Presenting critical items in the same font style used to study the theme's associates greatly increased false recognition. The font manipulation had a greater effect on performance than location reinstatement, especially for long lists as evidenced by the length ×

font interaction on false recognition for critical lures in both Experiments 1 and 2 (Figures 3 and 6). For long lists, reinstating the font drastically increased rates of false recognition. For short lists, reinstating the font only increased false recognition when the location was the same as compared to when the location was different. The main effects of length and font reinstatement as well as the length × font interaction seem consistent with global matching models. Increased global activation for the critical item as a result of list length, font reinstatement, and the combined effects of these variables resulted in greater false recognition for critical items (Arndt, 2010).

The location reinstatement did not appear to have as strong of an impact on performance as font reinstatement. However, this trend was not uniformly present as evidenced by the significant two-way interactions in both Experiments (Figures 2 & 5) for font and location. Location reinstatement did have an impact in some cases but never as strongly as font reinstatement. Also, location was identified as a basis for rejection most often when people rejected critical lures (Tables 1 and 2). In Experiment 2, the rate of false recognition increased as the amount of context reinstatement increased but only for long lists. Although there is evidence for increased activation in the same font, same location condition (as evidenced by the length × font interaction), evidence for rejection based on recollection does exist. In the same font, same location condition, participants were more likely to select rejection rationales based on recollection than they were to reject an item because of its unfamiliarity. Participants appear to latch onto location information but only when the font information is present as well. So for the same font, same location condition, activation of the critical item by context reinstatement occurs in some cases, but in other circumstances, the increased activation is overridden by rejection processes when contextual information is recombined at test. This dissociation hints at dual process theories of false recognition. However, rationale rates were only analyzed descriptively and these trends were present for both long and short lists.

One possible conclusion for the evidence of error-inflating and error-reducing processes is that recall-to-reject strategies only occur in best-case scenarios. For example, the experiments in this study promoted the utilization of these strategies by providing participants with multiple contextual cues at test in the short lists where recalling a single item is far easier than exhaustively recalling all six items in the long study list condition. One can speculate that recall-to-reject strategies are only employed in similar best-case scenarios such that these strategies are not always available to individuals. The implications of this speculation are that error-inflating processes override the error-reducing processes when participants do not have optimal conditions to exhaustively recall all studied items of a given theme. Yet these conditions do sometimes create retrieval environments that make new items seem different enough from what was studied to cause rejection. Thus, error-reducing processes may not actually be fully employed by participants in most DRM studies.

As proposed by Arndt and Gould (2006), future research on the DRM paradigm should attempt to manipulate factors that affect error-inflating processes and error-reduction processes, separately and in conjunction, to identify when and how false memories can be created or diminished. The present study could be improved upon by intermingling factors that are known to affect error-inflating processes (e.g., mean BAS, study list length) with factors that are known to affect error-reducing processes (e.g., study time, number of study item presentation), like in Arndt and Gould (2006).

In addition, the location manipulation could be made more salient by limiting the number of list themes presented in each location. Instead of presenting multiple list themes per each location (as in this study), each list theme could be presented in a distinctive location. Alternatively, one could use a different contextual feature that could be applied uniquely to each list theme, such as different spoken voices accompanying each theme.

Also, limiting the number of rationales a participant can select would aid in interpretation. Requiring participants to follow their first instinct and/or to only choose one item may yield more interpretable results for why items are called "old," and the results for why items were called "new" could change drastically.

Lastly, one should carefully consider the rationale findings in Experiment 2. On the one hand, asking people for rationales reduced false recognition in the short study list conditions to near zero when one or more contexts switched. This finding shows that changing retrieval context may be sufficient to sometimes eliminate false recognition. On the other hand, asking for rationales caused baseline false alarm rates to almost double in some conditions. Clearly, the number of fonts and locations used can help in some cases to create distinctive encoding environments, but other times, contextual information can cause interference and confusion.

In summary, global matching models can account for the significant main effects and interactions that were identified in this study. Overall, the findings, especially the font by length interactions in false recognition, are consistent with a global activation mechanism. However, such a mechanism cannot explain why recollection-based rejections were greater for both long and short study lists when multidimensional context reinstatement took place (i.e., same font, same location conditions). These findings together reinforce Arndt’s (2010) speculation that conditions influencing activation may also influence distinctiveness-based rejection mechanisms. Dual process theories of false memory, such as the activation-monitoring theory and fuzzy trace theory, can account for both increased activation (error-inflating processes) and the utilization of recall-to-reject strategies (error-reducing processes).

References

Arndt, J. (2006). Distinctive information and false recognition: The contribution of encoding and retrieval factors. *Journal of Memory & Language, 54,* 113-130.

Arndt, J. (2010). The role of memory activation in creating false memories and encoding context. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 36 (1),* 66-79.

Arndt, J., & Gould, C. (2006). An examination of two-process theories of false recognition. *Memory, 14,* 814-833.

Arndt, J., & Hirshman, E. (1998). True and false recognition in MINERVA2: Explanations from a global matching perspective. *Journal of Memory & Language, 39,* 371-391.

Brainerd, C. J., Reyna, V. F., & Kneer, R. (1995). False-recognition reversal: When similarity is distinctive. *Journal of Memory & Language*, *34,* 157-185.

Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin & Review, 3,* 37-60.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58,* 17-22.

Gallo, D. A. (2004). Using recall to reduce false recognition: Diagnostic and disqualifying monitoring. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 30,* 120-128.

Gallo, D. A. (2006). *Associative illusions of memory: False memory research in DRM and related tasks.* New York, NY: Psychology Press.

Hicks, J. L., & Hancock, T. W. (2002). Backward associative strength determines source attributions given to false memories. *Psychonomic Bulletin & Review, 9 (4),* 807-815.

Hicks, J. L., & Lynn, S. D. (2010). Activation and recall-to-reject processes in false recognition:

An interaction between font style and list length. Unpublished manuscript, Louisiana State University, Baton Rouge, LA.

Hicks, J. L., & Starns, J. J. (2006). Remembering source evidence from associatively related items: Explanations from a global matching model. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 32,* 1164-1173.

Hintzman, D. (1988). Judgments of frequency and recognition memory in a multiple trace memory model. *Psychological Review, 95,* 528-551.

McEvoy, C. L., Nelson, D. L., & Komatsu, T. (1999). What is the connection between true and false memories? The differential roles of interitem associations in recall and recognition. *Journal of Experimental Psychology: Learning Memory, & Cognition, 25,* 1177-1194.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. http://w3.usf.edu/FreeAssociation/.

Nolan, J. (1998). *1001 Free Fonts*. Retrieved March 25, 2012, from http://www.1001freefonts.com/.

Radvansky, G. (2011). *Human memory* (2nd ed.). Boston, MA: Allyn & Bacon.

Robinson, K. J., & Roediger, H. L. III. (1997). Associative processes in false recall and false recognition. *Psychological Science, 8,* 231-237.

Roediger, H. L. III, & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning Memory, & Cognition, 21*, 803-814.

Roediger, H. L. III, McDermott, K. B., Pisoni, D. B., & Gallo, D. A. (2004). Illusory recollection of voices. *Memory, 12,* 586-602.

Roediger, H. L. III, Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review, 8,* 385-407.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychologist, 26,* 1-12.

Appendix. Associative word lists with critical items.

bee	hive	chair	table	box	cardboard	back	front
	bumble		recliner		carton		spine
	sting		seat		package		pack
	buzz		stool		fuse		rear
	hornet		wicker		mail		hatch
	wasp		desk		storage		yard
answer	question	animal	zoo	dinner	supper	doctor	physician
	reply		beast		lunch		nurse
	response		farm		banquet		stethoscope
	solution		cage		meal		medical
	problem		vet		breakfast		surgeon
	explanation		creature		reservation		patient
fight	brawl	fire	blaze	fix	repair	pants	trousers
	quarrel		flame		mend		slacks
	feud		inferno		adjustment		zipper
	argument		torch		broke		jeans
	struggle		smoke		restore		belt
	fist		log		replace		shorts
pen	ink	picture	frame	plane	airport	run	jog
	quill		portrait		pilot		flee
	pencil		photo		jet		walk
	paper		camera		aircraft		track
	marker		draw		helicopter		chase
	write		album		passenger		dash
rock	boulder	rope	knot	old	new	shoe	sneaker
	stone		jump		elderly		sock
	granite		string		antique		foot
	pebble		noose		ancient		heel
	geology		twine		young		cobbler
	solid		loop		age		clog
sweet	honey	trash	garbage	smart	intelligent	soft	hard
	bitter		dump		clever		velvet
	sugar		litter		brilliant		loud
	sour		junk		gifted		tender
	candy		waste		genius		fluffy
	tart		pile		wise		pillow

Table 1

Frequencies and Rates of the Various Rationales for Long Study List Items

	Rationales			
Condition	Font	Location	Semantic	Unfamiliar
SFSL				
Frequency	4	20	13	11
%	8.3	41.7	27.1	22.9
SFDL				
Frequency	4	9	18	27
%	6.9	15.5	31.0	46.6
DFSL				
Frequency	12	17	16	67
%	10.7	15.2	14.3	59.8
DFDL				
Frequency	20	23	21	62
%	15.9	18.3	16.7	49.2

Note: SFSL = items at test were presented in the same font and location as presentation during encoding, SFDL = items at test were presented in the same font as presentation during encoding but the location was switched, DFSL = items at test were presented in the same location as presentation during encoding but the font was switched, DFDL = items at test were presented in switched fonts and locations.

Table 2

Frequencies and Rates of the Various Rationales for Short Study List Items

	Rationales			
Condition	Font	Location	Semantic	Unfamiliar
SFSL				
Frequency	8	39	16	73
%	5.9	28.7	11.8	53.7
SFDL				
Frequency	3	27	17	130
%	1.7	15.3	9.6	73.4
DFSL				
Frequency	18	30	10	112
%	10.6	17.6	5.9	65.9
DFDL				
Frequency	12	29	11	125
%	6.8	16.4	6.2	70.6

Note: SFSL = items at test were presented in the same font and location as presentation during encoding, SFDL = items at test were presented in the same font as presentation during encoding but the location was switched, DFSL = items at test were presented in the same location as presentation during encoding but the font was switched, DFDL = items at test were presented in switched fonts and locations.

Table 3

Frequencies and Rates of the Various Rationales for New Unstudied Items Shown in Fonts Seen One Time During Encoding

	Rationales			
Condition	Font	Location	Semantic	Unfamiliar
SFL				
Frequency	50	71	14	182
%	15.8	22.4	4.4	57.4
DFL				
Frequency	5	21	7	150
%	2.7	11.5	3.8	82.0

Note: SFL = items were presented in previously studied font-location pairings, DFL = items were presented in studied fonts and locations not previously paired.

Table 4

Frequencies and Rates of the Various Rationales for New Unstudied Items Shown in Fonts Seen Six Times During Encoding

	Rationales			
Condition	Font	Location	Semantic	Unfamiliar
SFL				
Frequency	21	99	16	190
%	6.4	30.4	4.9	58.3
DFL				
Frequency	92	89	31	252
%	19.8	19.2	6.7	54.3

Note: SFL = items were presented in previously studied font-location pairings, DFL = items were presented in studied fonts and locations not previously paired.

Figure Captions

Figure 1. Mean false recognition for new unstudied items in Experiment 1.

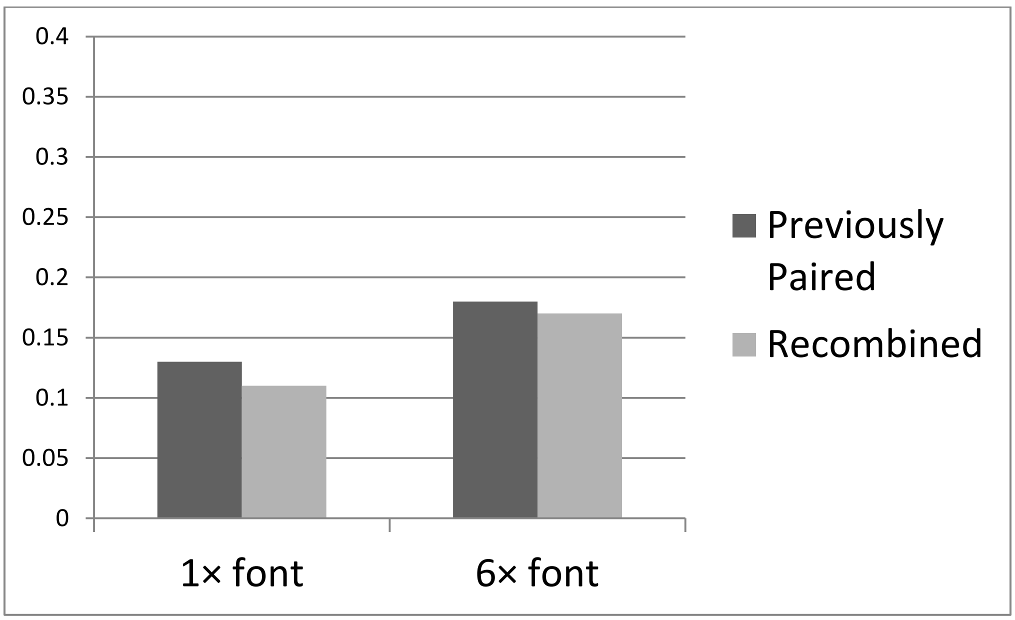
Figure 2. Mean corrected true recognition for studied list items in Experiment 1.

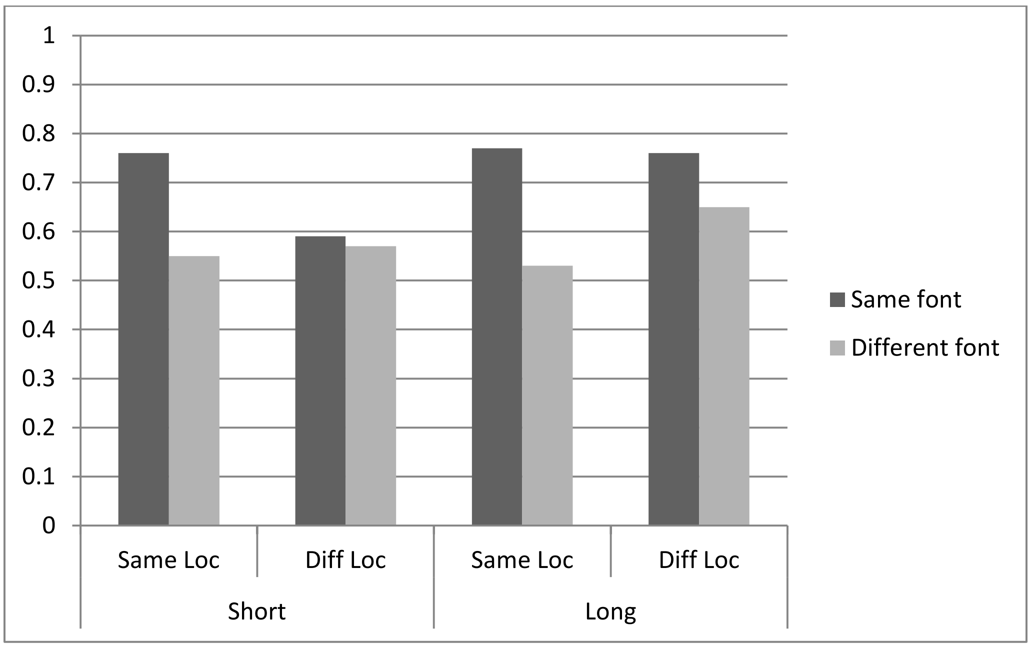
Figure 3. Mean corrected false recognition for critical lures in Experiment 1.

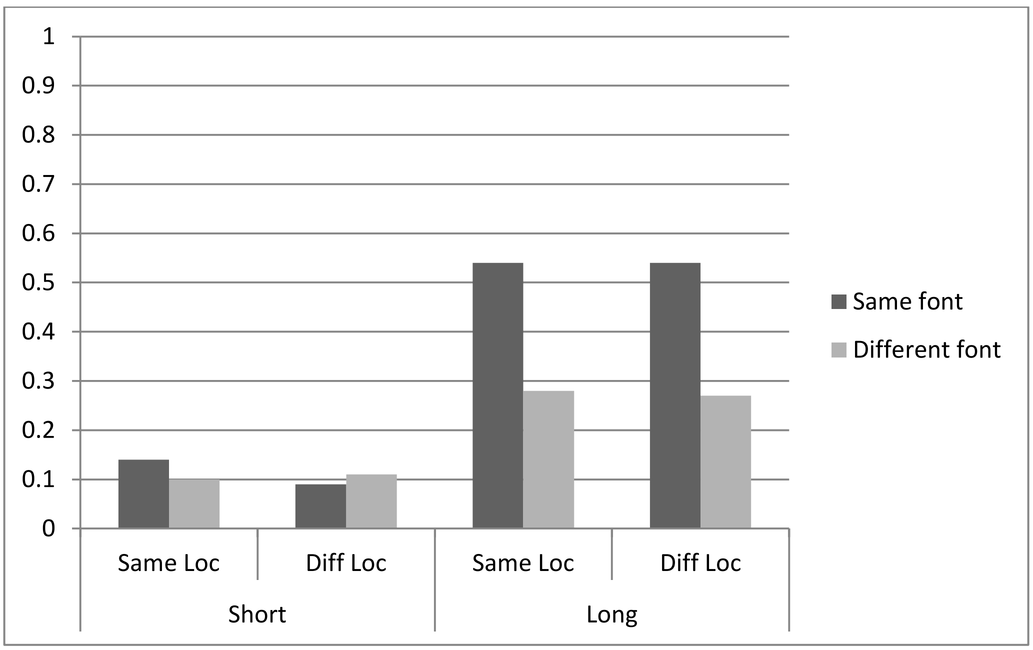
Figure 4. Mean false recognition for new unstudied items in Experiment 2.

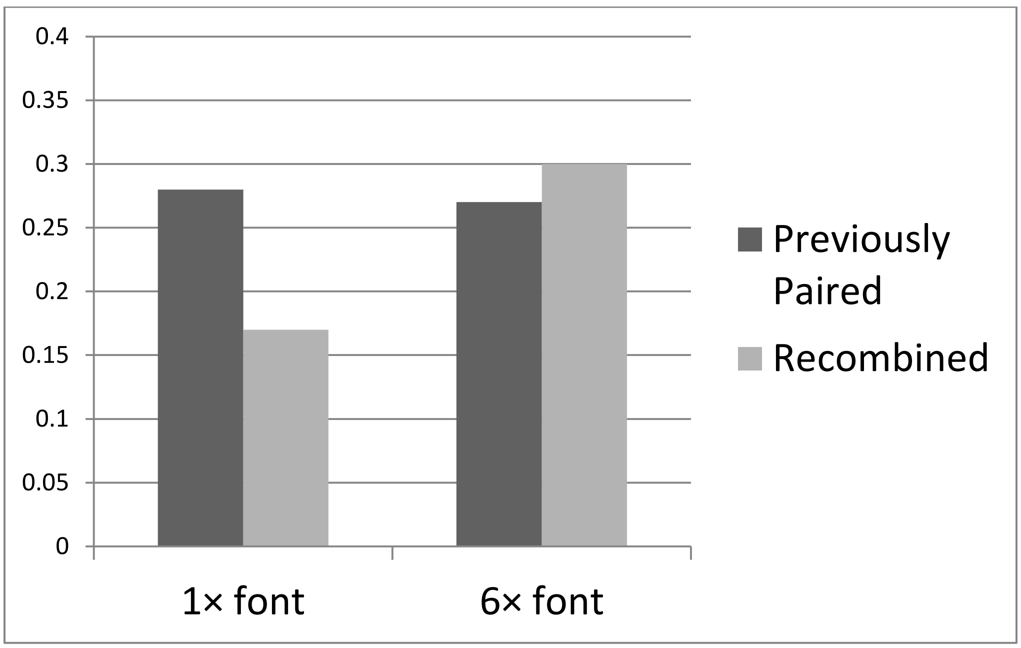
Figure 5. Mean corrected true recognition for studied list items in Experiment 2.

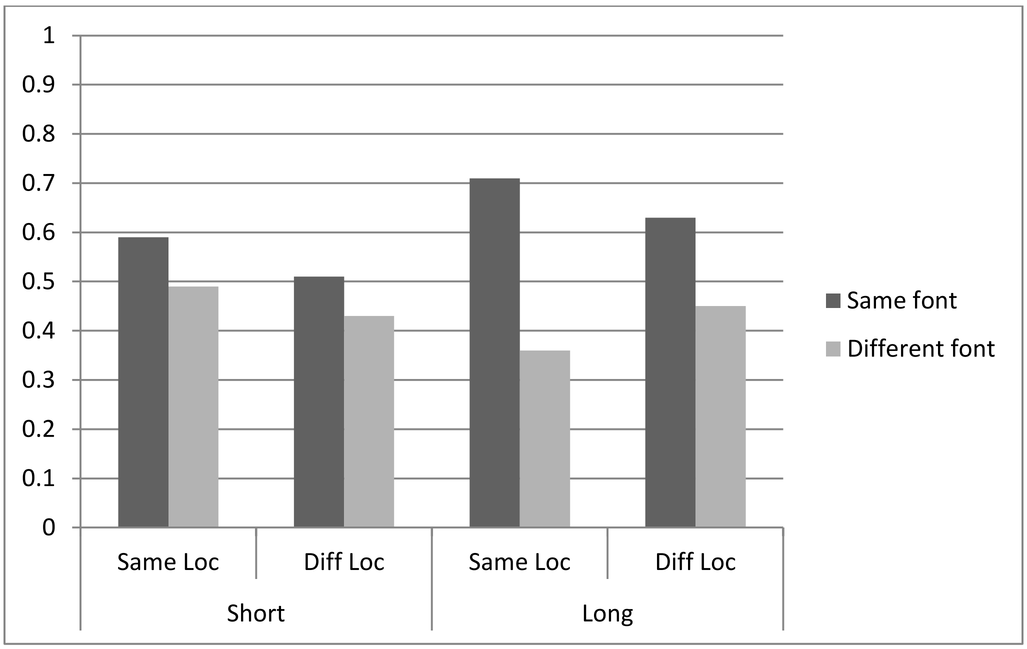
Figure 6. Mean corrected false recognition for critical lures in Experiment 2.

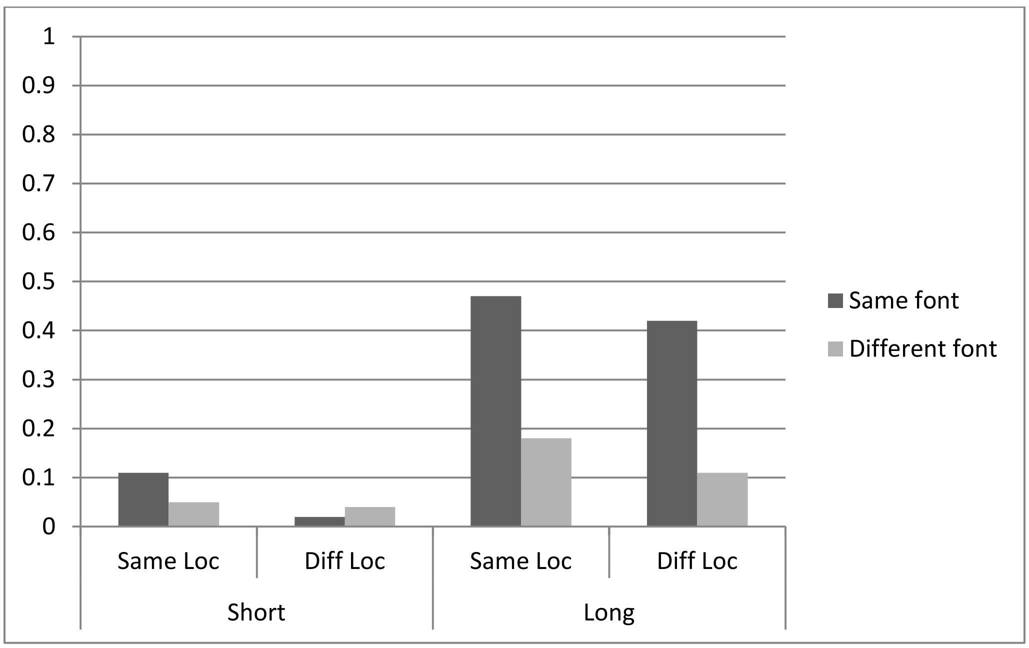

0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
1× font
6× font
Previously Paired
Recombined


1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Same Loc
Diff Loc
Same Loc
Diff Loc
Short
Long
Same font
Different font


1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Same Loc
Diff Loc
Same Loc
Diff Loc
Short
Long
Same font
Different font


0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
1× font
6× font
Previously Paired
Recombined


1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Same Loc
Diff Loc
Same Loc
Diff Loc
Short
Long
Same font
Different font


1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Same Loc
Diff Loc
Same Loc
Diff Loc
Short
Long
Same font
Different font