The role of spurious feature familiarity in recognition memory

RACHEL A. DIANA, **MARGARET J. PETERSON**, and **LYNNE M. REDER**
Carnegie Mellon University, Pittsburgh, Pennsylvania

Abstract

In two experiments, we investigated the role of perceptual information in spurious recognition judgments, Participants viewed lists of words in various unusual fonts. The frequency with which each font was presented was manipulated at study: Each font was presented with 1 or 12 different words in Experiment 1 and with 1 or 20 words in Experiment 2. Although the participants were instructed in a word recognition test to judge only on the basis of the word, regardless of font, there were significantly more false alarms for new words seen in a previously presented font than for new words presented in a novel (not seen at study) font in Experiment 1. In Experiment 2, the participants were significantly more likely to make a false alarm to a new word seen in a font that had been used to present 20 words during study than to a font that had been used to present only 1 word during study. The data show a mirror effect, in which words tested in low-frequency fonts produced more hits and fewer false alarms than did words tested in high-frequency fonts. These results show that irrelevant perceptual information plays a role in recognition judgments by providing spurious sources of familiarity and, thus, provide evidence that perceptual information is represented and processed in the same way as semantic information.

The phenomenon of false memories has received considerable attention in recent years. Roediger and McDermott's (1995) reexamination of Deese's (1959) paradigm has been seminal in this exploration of the properties conducive to spurious recollection. In this article, we explore factors that might exacerbate human susceptibility to spurious recognition. Roediger and McDermott (1995), like Deese, used sets of semantically related words to produce spurious recall and recognition of a critical nonpresented item (CNI) that was essentially the prototype of the set of studied elements (e.g., *sleep* is the CNI for *bed, night, dream,* and *yawn)* but was not, in fact, studied. This effect has primarily been attributed to semantic similarities between words on the list and the CNI (e.g., Robinson & Roediger, 1997; Roediger & McDermott, 1995, 2000). Even in (Loftus's 1975, 1979) real-world studies, the misinformation involved the insertion of a semantically related term (e.g., *yield sign* for *stop sign*).

Indeed, the term semantic memory has often been used interchangeably with long-term memory, suggesting that nonsemantic aspects of our memory are somehow in a different state or memory system. Some researchers have speculated that use of perceptual information may be automatic, whereas conceptual information requires conscious application (Craik, Moscovitch, & McDowd, 1994). Roediger (1990) has stated that one of the major dissociations between explicit and implicit memory is that explicit tests, although sensitive to semantic elaboration, are relatively insensitive to changes in perceptual features. In contrast, research has shown that perceptual information does influence performance on implicit tasks (e.g., Jacoby, 1983; Rajaram & Roediger, 1993; Roediger & Blaxton, 1987). Jacoby and Whitehouse (1989) found that a subliminal flash of a word just prior to its visible presentation increased

participants' tendency to respond *old*. They interpreted this finding as showing that the brief flash leads to more fluent processing of the word and that participants attribute perceptual fluency to recognition-based familiarity. This effect is modulated by the perceptual characteristics of the word (Jacoby & Hayman, 1987). That is, words that look more similar to their first presentation are more easily identified than those that look different. However, their result does not rule out the possibility that the word was semantically primed by its earlier presentation, The goal of this article is to demonstrate that familiarity judgments can be based on perceptual information alone; therefore, in the present experiments, we used a manipulation that isolated the effects of perceptual information (divorced from semantic information) on spurious recognition in explicit tasks.

Arndt and Reder (2003) found that when each word in a list of semantically related items is presented in a unique font, the rate of false recognition for the CNI is significantly lower than if it is presented in a font that was seen with 12 other words, whether or not these 12 words were from the same semantic list. These results demonstrate that perceptual information can influence recognition judgments when it is linked to semantic information.

However, perceptual information can also affect explicit recognition performance without any semantic similarity. For example, false alarms has been found for semantically unrelated but phonologically and orthographically similar words (Pesta, Murphy, & Sanders, 2001; Sommers & Lewis, 1999) Israel and Schacter (1997) have shown that distinctive perceptual information can influence participants to produce fewer false memories. Word recognition was affected by changes in the voice for auditory tasks (Sheffert, 1998), which the author concluded was evidence that spoken words are represented in episodic memory traces containing both conceptual and perceptual information Graf and Ryan (1990) found effects of perceptual characteristics when the encoding task focused on perceptual processing.

Strategy selection can be influenced by perceptual information without semantic priming, so that spurious feelings of familiarity are produced. For example, Reder and Ritter (1992) required participants to make a splitsecond decision as to whether or not they would be able to retrieve the answer to an arithmetic problem or would need to calculate the answer. The participants' judgment of whether the answer was known was influenced by the number of times the problem operands had been presented together. When the participants saw a problem with frequently paired operands but a different operator, they had a spurious feeling of knowing the answer. However, the participants judged that they did *not* know the answer to problems that they had studied when the spatial relations of the operands were inverted. In this case, the answer was known, but the problem looked unfamiliar, and the participants responded accordingly. These results suggest that recognition errors may depend, at least in part, on the familiarity of superficial aspects and, in particular, perceptual features of the stimulus.

The evidence presented above indicates that perceptual information can influence explicit memory tasks in the same way as semantic information We believe that this finding supports the proposal that perceptual information and semantic information are represented and processed within the same system. In order to further test our hypothesis, we will examine the effects of the frequency of the perceptual information on the production of false memories. Reder, Donavos, and Erickson (2002) demonstrated that perceptual cues *extraneous* to the recognition judgment might affect accuracy. In that study, words were presented in one of 150 unusual fonts. At test, words could be re-presented in the font seen at study (original), a font seen at encoding with a different word (used), or an unstudied font (novel). Memory performance was better when the font matched between study and test; however, this perceptual match effect was greater when the font was presented at a lower frequency. This effect occurred even though the participants were explicitly told to base their judgments on the word, regardless of the font.

Our explanation for the perceptual match effect's being modulated by font frequency (the number of words studied with a font) is analogous to the account of the word frequency mirror effect[1] (e.g., Glanzer & Adams, 1985; Glanzer, Adams, Iverson, & Kim, 1993) offered by Reder et al. (2000). That account posits that recognition can be based on one of two processes, familiarity or recollection, and that prior experience with information affects these processes differently. As words, fonts, and other types of information are seen more often, the number of associations with the concept increases, thereby decreasing the likelihood of retrieving any *specific* memory trace associated with it. This decrease in recollection occurs because the amount of activation that can spread to any contextual association depends on the number of competitors sharing the activation. The less the amount of activation that arrives at a contextual trace, the less likely it is that it will surpass the retrieval threshold. However, when a concept is experienced more frequently, its base activation is raised, making it more prone to elicit a familiarity response and, thereby, raising the number of hits and false alarms. This account of the perceptual match effect proposes that the effects observed are based on differences in retrieval of information, rather than on differences at time of encoding.

This dual-process account of the mirror effect for word frequency predicts changes based on the familiarity of semantic information. Because we propose that perceptual and semantic information have the same properties in memory, we wish to test this account's predictions under this assumption. Therefore, we would expect to find similar effects based on perceptual similarity alone Some evidence has already been provided by Reder et al. (2002), who demonstrated that the effect of manipulated font frequency (or fan, as they call it)[2] on hits and d' and found better recognition for studied words that were tested in the original font. Recognition was further improved when the font had not been studied with other words than when it had been studied with many words. We reanalyzed Reder et al.'s (2002) Experiment 2 data to determine whether there would be a spurious perceptual familiarity effect on false alarms and found a significant effect [$F(2,52) = 4.53, p < .05$], such that the percentage of false alarms was least when words were presented in novel fonts and greatest for words tested in high-frequency fonts. Their Experiment 3 results did not reach significance on the same contrasts, but the trend was replicated.

The goal of the present study was to further explore (1) how false recognition can be influenced by nonsemantic aspects of test stimuli and (2) whether perceptual information behaves analogously to semantic information. Specifically, we investigated the effects of extraneous perceptual cues on the false alarm rate and the mirror effect, as modulated by font frequency. Therefore, we hypothesized that these experiments would show that false alarms can be made to perceptually familiar stimuli, modulated by the frequency of the perceptual characteristics. That is, false alarms are made on the basis of the font information stored in memory, as well as on the basis of semantic information.

EXPERIMENT 1

Method

Participants. Thirty-three undergraduates at Carnegie Mellon University participated for partial fulfillment of a research experience requirement. The 15 participants who were randomly assigned to the delay condition were also paid $5 for the second session.

Design and Materials. The design consisted of four factors: delay (immediate test or I-week-delayed test), font frequency (high-frequency font, low-frequency font, or novel font), word status (studied or not studied), and font-matching condition (original, used, or novel). Test

---

1. The mirror effect refers to the phenomenon that two distinct classes of items, such as high- and low-frequency words, produce opposite orderings in likelihood to respo
2. We called the manipulation *font fan* rather than *font frequency* in deference to the underlying theoretical construct that motivated our predictions— the number of comp

delay was the only between-subjects factor. We manipulated delay in case participants' memory performance was at ceiling in the immediate condition. The original-font-matching condition was undefined for nonstudied test words, and thus the design was not fully factorial.

The stimulus materials were 240 words and 150 unusual fonts (see Figure 1 for examples of the fonts used), The assignment of words to fonts and each word-font pair to condition was randomly determined for each participant, For the 120 words presented at study, half were seen with a high-frequency font (seen with 12 words at encoding), and half with a low-frequency font (seen with only 1 word). Half of the words studied in each font frequency condition were tested in their encoding font (original), one quarter in a font seen with a different word (used), and one quarter with a font not seen during encoding (novel). For used fonts, the level of font frequency was maintained from study to test (e.g., words studied in a high-frequency font were shown in another high-frequency font). Of the 120 new words presented at test, three quarters were presented in a font that had been seen during the encoding phase (a used font), and one quarter were presented in a novel font. New words in the *used* condition were divided evenly between high- and low-frequency fonts.

Procedure. The participants were tested individually either at two sessions separated by a week or all in one session (40 min total). All the participants were led to believe that they would come back for a second session. The experiment was presented on a Macintosh computer running PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993).

During the study phase, the participants were instructed to judge the appropriateness of a font for word meaning. They were not informed of the subsequent memory test. During each trial, a word appeared in the center of the screen for 2 sec before being replaced with a 4-point Likert scale in which I indicated *very inappropriate* and 4 indicated *very appropriate*. After this study phase, the participants assigned to the delay condition were excused and instructed to return a week later to complete the experiment. The other group immediately began the test phase of the experiment. The instructions given to both groups were otherwise identical.

In the test phase, the participants were instructed to silently read a word that was displayed in the center of the screen and to respond by pressing one of three keys: "N" (new) if they felt the word was not presented during the study phase, "R" (remember) if they had a conscious recollection of experiencing the word in the study phase, or "K" (know) if they believed that the word had been on the original list but did not consciously remember experiencing it.[3] It was emphasized to the participants that when making a judgment as to whether a word had been seen, they were to take into account only the word itself, regardless of the font. Words were presented one at a time and remained on the screen until the participants chose one of the three possible responses.

Results and Discussion

Table 1 presents the hits, false alarms, and d' scores for the immediate test condition as a function of font frequency and font match. Table 2 presents the analogous data for the delay condition. The data were analyzed according to the frequency of the font used at test. For the *original* and the *used* conditions, the retrieval and encoding font frequencies were the same. Half of the words presented at test in a novel font were encoded with a high-frequency font, and the other half with a low-frequency font. Our account of font frequency effects posits that the manipulation affects retrieval, rather than encoding. Therefore, we expected that there would be no difference in performance for studied words encoded in a high-frequency font but tested in a novel font versus studied words encoded in a low-frequency font but tested in a

---

3. We collected remember-know judgments for both Experiments 1 and 2; however, due to space constraints, please refer to Reder et al. (2002) for a discussion of the data

novel font. Consistent with our account, these conditions did not produce a significant difference in hits [$t(32) = 1.16$] or in d' [$t(32) = -1.09$]. On the basis of these findings, all subsequent analyses will collapse across encoding condition to create one *novel* condition.

Replicating Reder et al. (2002), we found that in the immediate testing condition, d'[4] was better when the test font had only been studied with one word, rather than with many words [$F(l,17) = 26.74, p < .0001$] and when the encoding and the test fonts matched [$F(1,17) = 9.34, p < .01$] . In order to test the effect of delay, and because the experimental design was not fully factorial (we make no distinction between high and low frequency for the *novel* condition), we treated the various font conditions as five levels of one factor that was crossed with delay. There was a main effect of delay on d' [$F(1,31) = 30.785, p < .001$] and an interaction between font condition and delay [$F(4,28) = 4.377, p < .01$]. By dropping the *novel* condition, we could perform an analysis of variance (ANOVA) on font frequency, font match (between study and test), and delay. There was a significant interaction between font frequency and delay [$F(1,31) = 10.963, p < .01$] and between match condition and delay [$F(1,31) = 4.20, p < .05$], but no three-way interaction ($F < 1$).[5] These effects on d' indicate that the pattern of the data cannot simply be attributed to a response bias.

Our main concern, however, was the effect of font frequency on false alarms. Figure 2 shows the false alarm rates as a function of font frequency and delay. There was a main effect of font frequency [$F(2,62) = 15.37, p < .001$], so that more false alarms were elicited when foils were presented in high-frequency fonts than in novel (unstudied) fonts. Also, a mirror pattern was found in the immediate testing condition: The participants produced more hits for words presented in low-frequency fonts and more false alarms for words presented in high-frequency fonts. An ANOVA performed on the *used* condition for no-delay participants showed a marginally significant interaction between hits/false alarms and font frequency [$F(l,17) = 3.95, p < .06$]. When word frequency is manipulated, the same pattern is reported, indicating that perceptual features (font) affect memory in a fashion similar to that for aspects of lexical/conceptual memory. The pattern in the delay condition was different: The participants produced both more hits and more false alarms when viewing words presented in high-frequency fonts. This concordant pattern is similar to that seen in word frequency experiments when participants' opportunities to recollect are reduced by manipulating the study phase (e.g., Joordens & Hockley, 2000) or when the population studied is made up of amnesic individuals (e.g., Huppert & Piercy, 1976).

An ANOVA on false alarms showed a main effect of delay [$F(1,31) = 8.44, p < .01$], so that participants made more false alarms when the retention interval increased. The interaction of delay with font frequency was not reliable [$F(2,62) = 1.20, p > .05$]; however, there was a suggestion that the effects of font frequency varied with delay, and we decided to examine these effects, using t tests.[6] For the participants tested immediately after the encoding phase, the contrast between high- and low-frequency fonts was only marginally significant [$t(17) = 1.88, p = .07$], but new words presented in a novel font produced fewer false alarms than did words presented in a low-frequency font [$t(17) = 2.53, p < .05$]. On the other hand, at a 1-week delay, the difference in false alarm rates for words presented in a high-frequency font was reliably larger than that for words presented in a low-frequency font [$t(14) = 2,83, p < .05$], whereas the difference between low-frequency and novel fonts was not reliable [$t(14) = 1.07, p > .05$].

---

4. The d' value for the *original* condition was calculated using the false alarm late of the used condition, due to the fact that the design was not fully factorial. It is not possi

5. The interaction between font frequency and match showed the same pattern as that found in Reder et al. (2002), although it was not statistically significant ($F < 1$). How

6. Although t tests are not strictly endorsed when the higher level interaction is not reliable, we felt that these tests would be informative.

We suspect that the relative familiarity of the fonts in the conditions changed over time, resulting in a reliable difference between the novel and the low-frequency font in the immediate condition, but not after the delay. Perhaps after a week's delay, the participants found fonts that had been seen 12 times familiar, whereas fonts that had been seen only once were not appreciably more familiar than those not seen at all. Conceivably, the high-versus low-frequency difference could be obtained in an immediate test if the frequency difference was accentuated. This was our motivation for Experiment 2.

EXPERIMENT 2

Experiment 2 was similar to Experiment 1, with the following exceptions. Having demonstrated that participants can make false alarms on the basis of perceptual information, we wanted to focus on differences that are based on the frequency of the font. The delayed test was dropped because there was no evidence that ceiling effects had been a problem in Experiment 1. Experiment 2 focused on high-versus low-frequency fonts exclusively. In an effort to maximize the chance of getting a reliable difference in the immediate test condition, we increased the number of presentations in the high-frequency font condition from 12 to 20 and dropped the novel font condition. The contrast between novel and studied fonts on false alarms had already been replicated in Experiment 2 of Reder et al. (2002) and in the present Experiment 1, so we felt it was not needed. Also, without the novel font condition, the participants might be more sensitive to the familiarity difference between the high- and the low-frequency fonts. The novel font condition provided the participants with response choices that had no familiarity, making the other two levels of familiarity seem more similar. Furthermore, given the tacit desire to keep the old/new responding at around 50%, the participants might be more inclined to judge words as *new* with a rarely exposed font when there were no choices that provided less familiarity. The notion that people adjust their criterion for familiarity-based responding has been proposed elsewhere and has explained response patterns in word and pseudoword frequency effects (Reder et al., 2000).

Method

Participants The participants were 28 undergraduates at Carnegie Mellon University, who participated for partial fulfillment of a research experience requirement.

Design and Materials. Experiment 2 was identical to Experiment 1, except that the number of presentations for high-frequency fonts was increased and the novel font and delay conditions were dropped. Because we increased the number of presentations of high-frequency fonts, the total number of words used also increased from 240 to 400 Of the words that had been seen during study, half were presented at test in their original font, and half in a used font (a different font of the same frequency). For new words, half were shown in high-frequency fonts, and half in low-frequency fonts.

Procedure. The procedure of Experiment 2 was identical to that of the no-delay conditin in Experiment 1, except that the length of the session was slightly increased (approximately 50 min).

Results and Discussion

Table 3 presents hits, false alarms, and d' as a function of font frequency and font match (whether the font at test was the original encoding font or one that had been used with another word). A 2 (font frequency) $\times$ 2 (font match) ANOVA on d' once again showed a main effect of font frequency [$F(1,27) = 22.345$, $p < .0011$]. That is, words presented in low-frequency fonts had higher d' scores than did those presented in high-frequency fonts. A main effect of matching was also found [$F(1,27) = 37.904$, $p < .001$], so that words tested in the original

encoding font were more easily recognized than those tested in used fonts. There was no significant interaction between font frequency and match condition $[F(1,27) = 2.860, p > .05]$. These d' effects were driven by the hit rates. The same patterns were found for the hit rates for both font frequency $[F(1,27) = 14.771, p < .001]$ and font matching condition $[F(1,27) = 28.931, p < .01]$, with no interaction $(F < 1)$.

Of central interest was the effect of font frequency on false alarms. Because new words could not be tested in their original fonts (undefined), the design was not a complete factorial, so *t* tests were used instead of an ANOVA to examine the effects of font frequency on false alarm rates. As was predicted, unstudied words presented in high-frequency fonts were erroneously judged as old significantly more often than foils presented in low-frequency fonts $[t(27) = 2.60, p < .05]$.

Finally, it is interesting to note that font frequency has the opposite effect on hits and false alarms, with words tested in low-frequency fonts producing more hits and fewer false alarms than did words tested in high-frequency fonts. Indeed, an ANOVA examining the relationship between hits/false alarms and the font frequency for the *used* condition found a significant interaction $[F(1,27) = 22.282, p < .001]$. This finding replicates the borderline significant mirror effect found in Experiment 1. As was discussed earlier, this result is reminiscent of the mirror effect observed for low-versus high-frequency words.

GENERAL DISCUSSION

In two experiments, we have demonstrated a greater tendency for false alarms to unstudied words in familiarized fonts than for words in novel fonts. In Experiment 1, with delay, low-frequency fonts were treated like novel fonts. In Experiment 2, unstudied words presented in high-frequency fonts (seen with 20 other words) were spuriously recognized significantly more often than foils presented in a font that had been seen with only I other word. Although the participants were instructed to make their recognition judgments on the word alone, these perceptual cues clearly affected their judgments. The data show a mirror pattern in which low-frequency fonts result in more hit responses but fewer false alarm responses, Here, we have also replicated Reder et al.'s (2002) result that these superficial, perceptual features facilitate veridical recognition (measured in d' to tease out effects due to response bias). This facilitation of retesting in the original encoding font was modulated by the number of other words encoded in that font.

These data challenge single-process models of recognition memory. The mirror effect of font frequency is especially difficult for such models to explain, given that they propose that memory consists of a familiarity process only. Straightforward interpretations of single-process models, such as MINERVA 2 (e.g., Hintzman, 1984, 1988), would predict more false alarms and more hits to a word that is highly familiar due to a high-frequency font, as compared with words in unfamiliar fonts. According to MINERVA 2 (Hintzman, 1984, 1988), the tendency to identify a probe as old increases with the number of list items that partially match the probe for both targets and foils. Probe words that are presented in a high-frequency font can be described as partially matching (in their perceptual features) a greater number of items on the study list than do probe words that are presented in a low-frequency font. Thus, this interpretation of MINERVA 2 would predict a concordant pattern of responses.

Our explanation for these results is based on a dual-process model of recognition memory (SAC) in which recognition can be based either on retrieving the episodic memory trace or on a generalized sense of familiarity of the stimulus. The episodic trace is postulated to be associated both to the word and to the contextual features of the encoding event. At test, the context, the font, and the word send activation out to all of their associations. The amount of activation that accrues at the relevant episode node depends on the number of competing

associations. If a word is studied and tested in the same unusual font that is not studied with other words, all activation from that font is sent to the episode node, boosting the chance of a recollection. In SAC simulations, familiarity judgments are based on the activation level of the concept node. High-frequency words have a higher base-level activation (due to more prior exposures) and, thus, generate more false alarms (e.g., Reder et al., 2000).

It has become apparent from this study that the familiarity of the font affects familiarity judgments in an analogous fashion to familiarity of the concept. A font seen with many words provides less activation to the episodic trace, resulting in a diminished hit rate. Conversely, a font at test that has been seen many times at study will elicit a feeling of familiarity, regardless of whether or not the word presented in that font had been studied. This research provides compelling evidence that semantic information and perceptual information affect memory in the same way. In particular, we have demonstrated, in two experiments, a mirror effect, based on frequency of perceptual characteristics, that is reminiscent of mirror effects based on frequency of semantic information (Glanzer & Adams, 1985; Glanzer et al., 1993). Overall, it is clear that human memory is vulnerable to false recognition based on a number of different features. In-creased familiarity with any portion of a new stimulus, whether relevant or spurious, conceptual or perceptual, can lead to false recognition.

REFERENCES

ARNDT J, REDER LM. The effect of distinctive visual information on false recognition. Journal of Memory & Language 2003;48:1–15.

COHEN JD, MACWHINNEY B, FLATT M, PROVOST J. PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. Behavior Research Methods. Instruments, & Computers 1993;25:257–271.

CRAIK FLM, MOSCOVITCH M, McDowD JM. Contributions of surface and conceptual information to performance on implicit and explicit memory tasks. Journal of Experimental Psvchologv: Learning, Memory, & Cognition 1994;20:864–875.

DEESE J. On the prediction of occurrence of particular verbal intrusions in immediate recall. Journal of Experimental Psvchology 1959;58:17–22.

GLANZER M, ADAMS JK. The mirror effect in recognition memory. Memory & Cognition 1985;13:8–20.

GLANZER M, ADAMSs JK, IVERSON GJ, KiM K. The regularities of recognition memory. Psychological Review 1993;100:546–567. [PubMed: 8356189]

GRAF P, RYAN L. Transfer-appropriate processing for implicit and explicit memory. Journal of Experimental PsYchology. Learning, Memory, & Cognition 1990;16:978–992.

HINTZMAN DL. MINERVA 2: A simulation model of human memory. Behavior Research Methods, Instruments, & Computers 1984;16:96–101.

HINTZMAN DL. Judgments of frequency and recognition memory in a multiple-trace memory model. Psychological Review 1988;95:528–551.

HUPPERT FA, PIERCY M. Recognition memory in amnesic patients: Effect of temporal context and familiarity of material. Cortex 1976;12:3–20. [PubMed: 1261280]

ISRAEL I, SCHACTER DL. Pictorial encoding reduces false recognition of semantic associates. Psychonomic Bulletin & Review 1997;4:577–581.

JACOBY, LL. Remembering the data: Analyzing interactive processes in reading. Journal of Verbal Learning & Verbal Behavior 1983;22:485–508.

JACOBY LL, HAYMAN CA. Specific visual transfer in word identification. Journal of Experimental Psychology. Learning, Memory, & Cognition 1987;13:456–463.

JACOBY LL, WHITEHOUSE K. An illusion of memory: False recognition influenced by unconscious perception. Journal of Experimental Psychology: General 1989;118:126–135.

JOORDENS S, HOCKLEY WE. Recollection and familiarity through the looking glass: When old does not mirror new. Journal of Experimental Psychology: Learning, Memory, & Cognition 2000;26:1534–1555.

LOFTUS EF. Leading questions and the eyewitness report. Cognitive Psychology 1975;7:560–572.

LOFTUS EF. Reactions to blatantly contradictory information. Memory & Cognition 1979;7:368–374.

PESTA BJ, MURPHY MD, SANDERS RE. Are emotionally charged lures immune to false memory? Journal of Experimental Psychology: Learning, Memory, & Cognition 2001;27:328–338.

RAJARAM S, ROEDIGER HL III. Direct comparison of four implicit memory tasks. Journal of Experimental Psychology: Learning Memory & Cognition 1993;19:765–776.

REDER LM, DONAVOS DK, ERICKSON MA. Perceptual match effects in direct tests of memory: The role of contextual fan. Memory & Cognition 2002;30:312–323.

REDER LM, NHOUYVANISVONG A, SCHUNN CD, AYERS MS, ANGSTADT P, HIRAKI K. A mechanistic account of the mirror effect for word frequency: A computational model of remember knowjudgements in a continuous recognition paradigm. Journal of Experimental Psychology: Learning, Memory, & Cognition 2000;26:294–320.

REDER LM, RITTER FE. What determines initial feeling of knowing? Familiarity with question terms, not with the answer. Journal of Experimental Psychology: Learning, Memory, & Cognition 1992;18:435–451.

ROBINSON KJ, ROEDIGER HL III. Associative processes in false recall and false recognition. Psychological Science 1997;8:231–237.

ROEDIGER HL III. Implicit memory: Retention without remembering. American Psychologist 1990;45:1043–1056. [PubMed: 2221571]

ROEDIGER HL III, BLAXTON TA. Effects of varying modality, surface features, and retention interval on priming in wordfragment completion. Memory & Cognition 1987;15:379–388.

ROEDIGER HL III, McDERMoTT KB. Creating false memories: Remembering words not presented in lists. Journal of Experimental Psychology. Learning, Memory, & Cognition 1995;21:803–814.

ROEDIGER HL III, McDERMoTT KB. Tricks of memory. Current Directions in Psychological Science 2000;9:123–127.

SHEFFERT SM. Contributions of surface and conceptual information to recognition memory. Perception & Psychoplhysics 1998;60:1141–1152.

SOMMERS MS, LEWIS BP. Who really lives next door: Creating false memories with phonological neighbors. Journal of Memory & Language 1999;40:83–108.
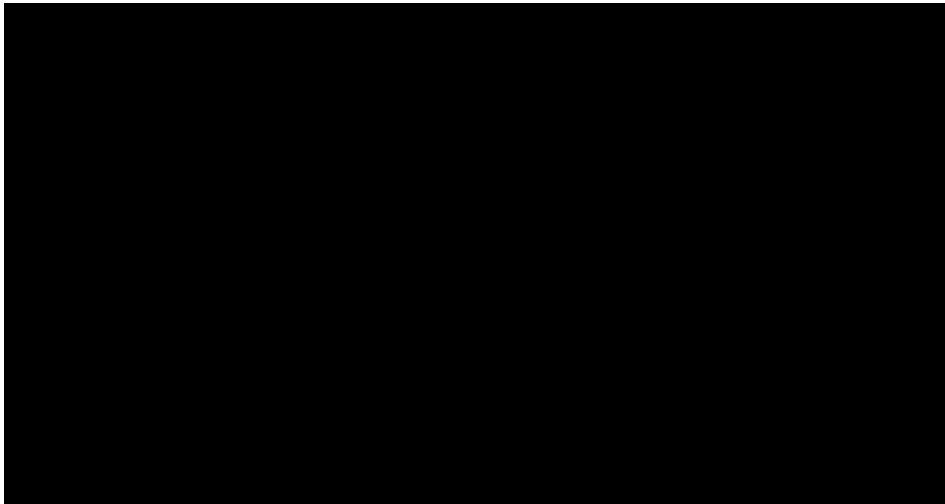
Figure 1.
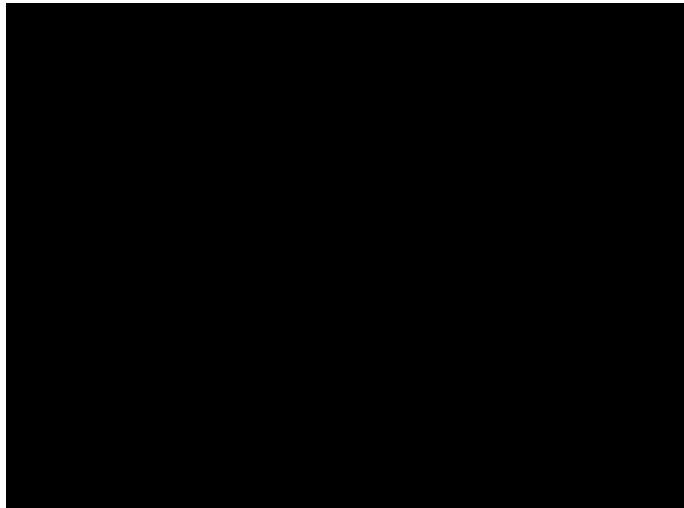Examples of the fonts used in the experiment.

Figure 2.
False alarm rate as a function of frequency of font use in the experiment and delay of test for Experiment 1.

Table 1

Proportions of Hits and False Alarms (FAs) and Mean d' Scores as a Function of Font Match and Font Frequency Conditions for the No-Delay Condition in Experiment I

Font Match Condition	Frequency Condition								
	High			Low			Novel(Collapsed)		
	Hit	FA	d′	Hit	FA	d′	Hit	FA	d′
Original	.87	NA	1.45	.91	NA	1.86			
Used	.79	.42	1.11	.81	.37	1.32			
Novel	.76[*]			.79[*]			.78	.33	1.34

Note—See the Results section of Experiment I for an explanation of the reason for examining only the collapsed data in the novel condition.

[*] Frequency condition reflects the font used at study

Table 2

Proportions of Hits and False Alarms (FAs) and Mean d' Scores as a Function of Font Match and Font Frequency Conditions for the Delay Condition in Experiment I

Font Match Condition	Frequency Condition								
	High			Low			Novel(Collapsed)		
	Hit	FA	d'	Hit	FA	d'	Hit	FA	d'
Original	.79	NA	0.64	.69	NA	0.51			
Used	.76	.64	0.44	.64	.53	0.33			
Novel	.57[*]			.59[*]			.58	.48	0.28

Note—See the Results section of Experiment I for an explanation of the reason for examining only the collapsed data in the novel condition

[*] Frequency condition reflects the font used at study.

Table 3

Proportions of Hits and False Alarms (FAs) and Mean d' Scores as a Function of Font Match and Font Frequency Conditions in Experiment 2

| | Frequency Condition | | | | | |
| | High | | | Low | | |
Font Match Condition	Hit	FA	d′	Hit	FA	d′
Original	.81	NA	1.18	.86	NA	1.53
Used	.75	.43	0.95	.78	.40	1.16