Psychological Research (2005) 69: 420–430
DOI 10.1007/s00426-005-0213-1

ORIGINAL ARTICLE

Bruce W. A. Whittlesea · Michael E. J. Masson
Andrea D. Hughes

False memory following rapidly presented lists: the element of surprise

Received: 18 October 2004 / Accepted: 26 November 2004 / Published online: 26 April 2005

Abstract This article examines a false memory phenomenon, the Deese-Roediger-McDermott (DRM) effect, consisting of high false alarms for a prototype word (e.g., SLEEP) following a study list consisting of its associates (NIGHT, DREAM, etc.). This false recognition is thought to occur because prototypes, although not presented within a study list, are highly activated by their semantic association with words that are in the list. The authors present an alternative explanation of the effect, based on the discrepancy-attribution hypothesis. According to that account, false (and true) familiarity results when a comparison between expectations and outcomes within a processing episode causes surprise. Experiment 1 replicates the DRM effect. Experiment 2 shows that a similar effect can occur when participants are shown lists of unrelated words and are then surprised by a recognition target. Experiments 3 and 4 show that the DRM effect itself is abolished when participants are prevented from being surprised by prototypes presented as recognition targets. It is proposed that the DRM effect is best understood through the principles of construction, evaluation, and attribution.

Introduction

This article is intended to contrast two general approaches that investigators have taken to explain the subjective quality of remembering. One is based on a direct connection between the state of representations in memory and the subjective experience of remembering. The other argues that the subjective experience results most directly from evaluative processes that interpret the quality of performance controlled by representations in memory. We compared these two approaches in an examination of a memory illusion.

The Deese-Roediger-McDermott (DRM) illusion consists of the observation that, after reading a list of words such as “bed, rest, tired, awake, dream” and so on, people often falsely recall or recognize a critical word such as “sleep” that is not presented within the list but is highly related to each member of the list (e.g., Deese, 1959; Read, 1996; Roediger & McDermott, 1995). (As Arndt & Hirshman, 1998, pointed out, these critical words are in effect the semantic prototypes of their lists, and will be referred to as such below.) False remembering of prototypes was originally observed in delayed tests, in which the participants saw all training lists before being tested on any associates or prototypes (e.g., Bredart, 2000; Hicks & Marsh, 1999; Mather, Henkel, & Johnson, 1997; McDermott, 1996; McDermott & Roediger, 1998; McEvoy, Nelson, & Komatsu, 1999; Norman & Schacter, 1997; Read, 1996; Roediger & McDermott, 1995; Schacter, Verfaellie, & Pradre, 1996). The illusion has also been observed, however, in tests immediately following rapid serial visual presentation (RSVP) presentation of the lists (e.g., Arndt & Hirshman, 1998; McDermott & Watson, 2001; Seamon, Luo, & Gallo, 1998). This procedure was used in the experiments that follow.

This false memory effect has been explained in a variety of ways, including criterion shifting during retrieval (Miller & Wolford, 1999), global matching (Arndt & Hirshman, 1998), memory for semantic gist (Brainerd, Wright, Reyna, & Mojardin, 2001), and reconstruction (Whittlesea, 2002a). Perhaps the most common form of explanation, however, in both delayed and immediate tests, is based on spreading activation and monitoring (e.g., McDermott & Watson, 2001; Nelson, McKinney, Gee, & Janczura, 1998; Roediger, Balota, & Watson, 2001; Roediger & McDermott, 2000; Roediger, Watson, McDermott, & Gallo, 2000).

B. W. A. Whittlesea (✉) · A. D. Hughes
Department of Psychology,
Simon Fraser University, Burnaby,
British Columbia, Canada, V5A 1S6
E-mail: bruce_whittlesea@sfu.ca

M. E. J. Masson
University of Victoria, British Columbia, Canada

The spreading activation explanation rests on the assumption that part of memory has a network organization in which related concepts are linked together, such that accessing one causes activation to spread to related concepts (e.g., Collins & Loftus, 1975; Nelson, Schreiber, & McEvoy, 1992). Furthermore, the idea is that experience of the list of words activates the representation of the prototype, because it is associated with every word in the list. This activation mimics the effects of actually encountering the word. In consequence, when remembering is required immediately after a rapid list, the representation of the prototype is highly active, causing false recognition or recall. False remembering in delayed tests occurs because activation of prototypes during list processing caused its semantic representation to be activated and episodically marked in much the same way as it would be if it actually had been presented. Consistent with these ideas, presenting a list of associates in the study phase causes priming for the prototypes in later indirect conceptual or meaning-based memory tests such as word association (e.g., McDermott, 1997) and lexical decision (Hancock, Hicks, Marsh, & Ritschel, 2003; Whittlesea, 2002a), and also to some degree in implicit tests emphasizing perceptual aspects, such as word-stem or fragment completion (McKone & Murphy, 2000). Furthermore, the more associates presented in a training list, the greater the probability of false recall (Robinson & Roediger, 1997). Similarly, in immediate tests, probability of false recall or recognition is greater when list words are presented for a longer time (2 s or 250 ms each) rather than for a very short interval (20 ms each; McDermott & Watson, 2001; Seamon et al., 1998, 2002).

The monitoring component of this account is described as being "a much more strategic, controlled process (or set of processes)" (McDermott & Watson, 2001, p. 160). The idea is that during the memory test, people must distinguish between ideas that come to mind because they were experienced earlier in the study session from ones that come to mind for other reasons (the issue of source monitoring: Johnson, Hashtroudi, & Lindsay, 1993). Warning people about the false-familiarity effect of prototypes before they engage in studying the lists attenuates the false-remembering effect to some degree (e.g., Gallo, Roberts, & Seamon, 1997; McDermott & Roediger, 1998). However, providing such warnings before the retrieval test but after they have read the lists has little if any effect on the size of the illusion (Gallo, Roediger, & McDermott, 2001; Whittlesea, 2002a). This result has been taken as evidence that the monitoring process cannot distinguish between traces laid down by external stimulation and traces created through spreading activation, once they are formed; strategic control processes are primarily effective when engaged during the original reading of the lists (McDermott & Watson, 2001). This reasoning is consistent with the idea that the illusion, when it occurs, is produced by encoding processes rather than decisions made during retrieval.

The SCAPE framework of memory (e.g., Whittlesea & Williams, 1998; Whittlesea, 2002a) provides an alternative interpretation of the DRM effect, one that does not involve spreading activation. According to this account, subjective experiences such as the feeling of familiarity are produced by an evaluation process that monitors the integrity of ongoing performance. This evaluation takes into account the apparent quality of current processing as well as those aspects of the current stimulus and context that are salient to the person; it leads the person to develop an attitude toward their performance. Of particular relevance to the DRM effect, this process can sometimes cause the person to experience a perception of discrepancy, a feeling of surprise caused by the apparent mismatch between expectations raised by some parts of a processing experience and the actual outcome. This perception motivates the person to seek an explanation, attributing the surprise to some plausible source in covert characteristics of the current stimulus or environment, the person's current state (mood or disposition), or in the past. When this perception is unconsciously ascribed to a source in the past, the person experiences a conscious feeling of familiarity.

The perception of discrepancy has been demonstrated to cause false feelings of familiarity-based remembering in a variety of paradigms, none of which involve spreading activation (Whittlesea & Williams, 2000, 2001a, 2001b; Whittlesea & Leboe, 2003; Whittlesea, 2002a, 2002b). For example, Whittlesea and Williams (1998) presented words and nonwords in a training phase; in a subsequent test phase, they presented old and new words and nonwords. The critical manipulation was that some words were presented in the test as pseudohomophones (e.g., PHRAUG). Participants were instructed to claim these items to be old if their phonological code matched that of a word shown normally in training (FROG). Presenting test words in this way created a perception of discrepancy. The unfamiliar orthography caused the participants to develop an initial impression that these items were nonwords; that expectation was violated a moment later by their realization that they were saying a natural word. That reorganization of understanding caused surprise, as evidenced by the laughter observed throughout the experiment. It also caused a large increase in false alarms (43%, compared with 19% for natural words and 14% for nonwords).[1]

[1]Not all surprises cause this misattribution, even when regarded as funny. For example, Whittlesea and Williams (2001b) presented test words as the last word of familiar phrases. They compared claims of recognition for words presented as clang associations (e.g., Row, row row your GOAT) with words presented in their original context (e.g., Row, row row your BOAT) or words simply mismatching the context (e.g., Row, row row your SHEEP). The participants laughed at many of the clangs, but that manipulation had no observable effect on claims of remembering. The difference between the two cases of surprise appears to be that when dealing with items like PHRAUG, people interpret their subsequent processing as matching well, although surprisingly so, with earlier processing, whereas presentations such as GOAT seem to be surprisingly wrong.

Not all occurrences of the perception of discrepancy involve a humorous reaction, but they all appear to involve some element of surprise. Thus, for example, regular nonwords such as HENSION produce more false alarms than either regular words such as TABLE or irregular nonwords such as STOFWUS. Whittlesea and Williams (2000, 2001a) provided a number of convergent measures to demonstrate that the regular nonwords produced surprise based on the fact that they are very word-like, but turn out to be meaningless: They thus create a sense of mismatch between early and later aspects of processing. Similarly, words that sensibly complete sentence stems cause increased false alarms when separated from constraining stems by a pause, but do not do so when there is no pause. Whittlesea (2002b) showed that such words following a pause are experienced as being unexpected relative to words without a pause, but nevertheless matched the sentence stem and so feel surprising.

Many readers of past articles that attribute illusions of memory to a perception of discrepancy (e.g., Whittlesea, 2002a, 2002b; Whittlesea & Williams, 2001b) have had trouble understanding why something that is very expectable, such as fluent processing of orthographically regular nonwords or a sensible completion of a stem after a pause, could cause a sense of surprise. Perhaps an analogy can help, based on the common expression "Waiting for the other shoe to drop." As a general truism, the expression refers to cases in which a person is in a state of extreme and conscious anticipation of a particular event in their environment. In such a case, it might be imagined that no surprise could attend the occurrence of the anticipated event. But imagine that you really are lying in bed in a hotel room, and just before falling asleep have heard the occupant of the room above drop one shoe. You know you cannot simply carry on going to sleep, because you know the second shoe will awaken you. Thus, the sound caused by the second shoe is entirely expected: And yet, when it occurs, it can be very startling. That is because you do not know *when* to expect it. Many other common experiences produce surprise for similar reasons. Thus, for example, when starting to view an adventure film, you can be completely sure that the hero will win and the villain will lose. You can also, at some moments during the film, be alert to the fact that the hero is in danger from the villain. Thus, both the final outcome and also the moment-by-moment interactions of the characters are in one sense entirely predictable. But what makes such films popular (in part) is the fact that the particular ways in which the hero wins and the villain loses are unexpected from within the local context (e.g., villain attempting to shoot hero steps on rake and knocks self out), even though they are consistent with aspects of the situation known to the viewer. The art of such films is to create suspense, based on the development of a strong but indefinite expectation, but then providing a resolution that matches the earlier expectation in an unexpected way.

In the same way, we suggest that the illusions of familiarity in the three paradigms discussed earlier (PHRAUG, HENSION, and sensible words after a pause) involve surprise (a perception of discrepancy between expectation and outcome) in one way or another. Normatively, the experience of surprise in these cases is wrong: In an ideal world, people would understand the causes of variations in their behavior and in each case attribute it to its actual source (the odd spelling, the orthographic regularity, or the pause). However, the participants in such studies are either unaware of or do not understand the effects of such factors. Because they are unable to understand the cause of variations in their performance, they consciously experience surprise. That surprise can, in principle, be attributed to a source in the past, in the covert properties of the stimulus or some characteristic of the person (e.g., mood or skill). Within the context of a remembering experiment, it is likely to be attributed to an unknown source in the past. That attribution occurs spontaneously and unconsciously, based on those aspects of the task and situation that are salient to the person: Such nonspecific unconscious attribution to the past causes the person to experience a conscious sense of familiarity.

The surprise that people experience in those paradigms appears to mimic in some way what the participant would ordinarily expect to feel when presented with an item they had encountered previously. That is, it suggests that appropriate feelings of familiarity also derive from a sense of surprise. The experiments in this article examine this idea by investigating the consequences of inducing and preventing surprise during a remembering test.

Returning to the DRM, this discussion suggests that the proximal source of the illusion is that prototypes following lists of their associates are experienced as being surprising. That surprise could come about in one of several ways. One is that prototypes following related lists are processed more swiftly than they would be otherwise (e.g., Hancock et al., 2003; McDermott, 1997; Whittlesea, 2002a), thereby violating a fluency expectation. Another is that, by virtue of being a prototype, that item is more related to all of the words in the list than any other word is (Arndt & Hirshman, 1998). It could thus cause surprise by providing an unexpectedly clear summary or central meaning relating to the words in the list.

The purpose of the experiments in this article is to investigate the idea that it is the surprise caused by the prototype at the time it is evaluated as a candidate during a memory test, rather than some other aspect of its processing, that is the proximal cause of the DRM effect. The alternatives considered include the ideas that it is the greater semantic relatedness or the greater fluency of the prototypes compared with other items that directly causes the effect. We will suggest that those factors are part of what makes the effect occur, but that surprise is the crucial aspect.

Experiment 1

Our first study was a simple demonstration that the DRM effect occurs robustly with the particular stimuli and means of presentation that we used. As mentioned above, the procedure of presenting word lists in RSVP format has been used by others (e.g., Arndt & Hirshman, 1998; McDermott & Watson, 2001; Seamon et al., 1998) and has been shown to generate robust DRM effects.

Method

Participants

All of the participants tested in the experiments reported here were students at Simon Fraser University, participating for extra credit in an undergraduate course or for a chance to win a Canadian $400 lottery. Nineteen participants were tested in Experiment 1.

Materials

Stimuli for these experiments were taken from the Appendix presented by Stadler, Roediger, and McDermott (1999). Their materials consist of 36 semantic prototypes, together with lists of 15 high-frequency associates of each prototype. For example, one prototype word is ANGER, with associates (in descending order of associative frequency) MAD, FEAR, HATE, RAGE, TEMPER, FURY, IRE, WRATH, HAPPY, FIGHT, HATRED, MEAN, CALM, EMOTION, and ENRAGE. We split each list into two six-item lists, one consisting of the odd-numbered items from the Stadler et al. lists, beginning with the third word (e.g., HATE, TEMPER, IRE, etc.), the other consisting of the even-numbered items, beginning with the fourth word (e.g., RAGE, FURY, WRATH, etc.), the 15th word being dropped. The prototype and the two highest associates were reserved for special treatment, as described below. This arrangement resulted in 72 lists of six words each. Eighteen of these lists were assigned to each of the four test conditions, at random except that the two lists belonging to any particular prototype were assigned to different conditions. Assignment of lists to conditions was independently randomized for each participant.

Procedure

In each trial of each experimental condition, the participant was shown a list of six words, always in descending order of associative frequency. Each trial began with a READY prompt. On a key press by the participant, the screen blanked for 250 ms. Following that, each word was presented in uppercase letters at the center of the screen for 120 ms, followed immediately by the next word. These six words will be referred to as the "RSVP list.". In addition, each list was pre- and post-masked by a row of symbols (e.g., #$%, etc.), presented for 250 ms. Immediately following the RSVP presentation, a probe word was displayed on the screen, with the instruction "Did you see that word in the list?"

There were four conditions, each involving 18 trials. In the first, a prototype was presented as a recognition target immediately after one of its six-item lists of associates. In the second, a prototype was presented immediately after one of the lists for a different prototype. The prototypes were never presented in the lists, whether related to the list or not. In the third condition, a highest associate was presented as a recognition probe following a list associated with its prototype. In that condition, the target associate was inserted into the list, replacing one of the original list items at random in locations 2–5, and was consequently old. The placing of the associate in the list in those locations ensured that it was always pre- and post-masked by at least one other word. In the fourth condition, a highest associate was presented as a recognition probe following a list associated with a different prototype; in this case, the highest associate was not present in the list. Thus, the test item was a prototype or highest associate, was related or unrelated to the foregoing list, and was old only if it was a highest associate and related to its list. Assignment of lists to conditions and order of presentation of trials within and between conditions was freshly randomized for each participant. Type I error probability was set at .05 for all tests reported in this article.

Results and discussion

The participants were quite accurate in recognizing highest associates, claiming old items to be old 73% more often than new items (see Table 1). The prototypes were generally judged to be new, but were claimed to be old 12% more often when following a related than an unrelated list, $F(1, 18) = 21.03$, MSE $= .01$, $\eta^2 = .54$. They were also claimed to be old 14% more than novel highest associates, $F(1, 18) = 43.66$, MSE $= .01$, $\eta^2 = .72$. That is, we observed a DRM effect—an illusion of familiarity—for prototypes prepared by their own lists of associates. We then attempted to investigate how that effect came about.

Table 1 Experiment 1: Prototypes not in related list

Probe	Related List	Unrelated List
Associate	.82	.09
Prototype	.23	.11

Recognition: *p* (claim old)

Experiment 2

Our suggestion is that the DRM effect occurs because something about the prototype causes surprise when it is presented after a list of its associates, and that that surprise is attributed to prior experience. We propose that a primary reason that the prototype is surprising is because, unlike other recognition probes, it matches especially well any of the items that the participant can remember from the list—much more so than even the highest associates from the list.

Our next experiment involved a radical departure from the procedures of the DRM effect, but with the intention of showing the critical role of surprise. We elected to do that by presenting jokes. Critical probes appeared as part of the punch line for a joke, thereby creating a feeling of surprise, even though the probes were not list prototypes. By creating a feeling of surprise, we expected that an elevated false alarm rate would result, relative to a case in which no surprise was generated.

Method

Participants

There were 15 participants in Experiment 2A and 17 in Experiment 2B.

Procedure

In Experiment 2A, in each trial, participants were shown six words in RSVP format, just as in the previous study. These words were taken from the stock used in the first study, but were randomly sampled from the stock, subject to the condition that no two list items came from the same original category, so that there was no relationship among words within any six-word list. These lists were presented with the same temporal and display characteristics as in Experiment 1.

Following each list, participants were given a question to read, and then given an answer to the question. The cover story for this procedure was that they would later be asked about the answers to these questions (which they were not). The majority of these questions were about simple trivia, either well known to the participants or ones that they were likely to figure out, such as "What are the colors in the Canadian flag?" (red and white) or "What is the international shape of a stop sign?" (octagon). Immediately after they had read the answer, one word from the list (e.g., WHITE) was presented as a recognition probe, with the question "Did you see that word in the list?" Half of those words had indeed been in the list (being substituted for a word that would otherwise have been in that list, at random in location 2, 3, or 4 of the list). The other half of the probe words had not been in the RSVP list. Assignment of questions to old vs. new conditions was freshly randomized for each participant. Twenty-two questions were shown in each of these conditions. Questions remained on the screen for 3 s, were automatically replaced by the answer, which was shown for 2 s, which in turn was automatically replaced by the probe (shown in capital letters). The probe word remained on the screen until the participant responded by pressing one of two keys.

In addition, in 22 other trials, the question and answer made a joke. The participants were warned that this would happen occasionally. An example is "What kind of nail does a carpenter hate to hit?" (a finger nail)". Again, one word (FINGER) would then be presented as a recognition probe. In parallel to the way that prototypes were treated in the last study, probes in these trials were always novel. Trials of all three conditions were presented in a random order determined independently for each participant.

Because the words used as recognition targets following jokes were presented only in that condition, a difference in performance between this condition and the nonjoke control could conceivably be due to some *in sui* characteristic of those words, rather than their participation in the jokes. To examine that possibility, we conducted a control study (Experiment 2B), presenting the same RSVP lists and recognition target words as in Experiment 2A, but without questions or answers between those two presentations.

Results and discussion

As shown in Table 2, in Experiment 2A the participants discriminated well (about 37%) between old and new probes following trivia questions, $F(1, 14) = 36.74$, $MSE = .03$, $\eta^2 = .73$. In the critical comparison of trivia vs. jokes, participants were 15% more likely to commit a false alarm following a joke than a trivia item, $F(1, 14) = 15.27$, $MSE = .01$, $\eta^2 = .53$. In contrast, in the control study (Experiment 2B), the probe words that had been used in the joke condition of the main study produced only 2% more false alarms than words used in the nonjoke condition, the difference not being reliable, $F(1, 16) = 1.74$, $MSE = .01$, $\eta^2 = .10$, although the power of this experiment to detect a 15% effect was .99. In a between-groups comparison, the difference in false alarms between those two conditions was 13% greater in the main experiment than in the control, $F(1, 30) = 10.18$, $MSE = .01$, $\eta^2 = .25$.

Table 2 Experiment 2: Jokes and false familiarity

Question type	Old	New
Trivia	.61 (.57)	.24 (.12)
Joke		.39 (.14)

Recognition: *p* (claim old)
Note: Values in parentheses show claims for the same words when presented after RSVP lists without questions or answers in the control study

The critical element of the main study (Experiment 2A) compared with the control (Experiment 2B) was that some answers would be surprising and others not. Although knowledge and tastes vary, it could generally be supposed that our joke answers were on average more surprising than our trivia questions. Although the participants undoubtedly realized which were the jokes, that knowledge by itself probably did not affect the responses. None of the words in the answers to the jokes were contained within any RSVP list: Had the participants realized that, they could have said "new" whenever a joke appeared, quite the opposite of what we observed. We therefore concluded that surprise caused by joke endings had caused an illusion of familiarity for the target words.

The false alarm effect caused by the jokes was clearly not produced by spreading activation, either from the list (because none of the joke answers was presented within the lists) or from the question (otherwise the punch line would be anticipated and not funny). Also, it was probably not caused by facilitating the processing of the answer, so that the person could read it with enhanced fluency. Instead, we suggest that the key element is the surprise caused by the joke answers, and that that surprise was caused by the re-interpretation of understanding, from that apparently posed by the question to that caused by the answer. That surprise would not normally be attributed to a source in the past (which would lead people to say "I've heard that before" whenever they heard a joke). However, within the context of a recognition test, the surprise was often falsely attributed to a prior experience of the answer within the RSVP list, as though such a prior experience could make the joke funny. (In fact, of course, the opposite would happen: Hearing the punch line earlier spoils the joke.) Despite this doubly counternormative thinking, this phenomenon illustrates our point: Surprise resulting from a violation of expectation, which results in a new understanding, causes a perception of discrepancy and an attribution to a plausible source outside of the current event. Put more simply, surprise caused a feeling of familiarity.

Despite the radical differences between this study and the procedures of the DRM paradigm, we believe that there is one element in common: Surprise is sufficient to cause a feeling of familiarity, at least within the context of a recognition test. We earlier proposed two reasons why prototypes might be surprising following lists of their associates. One is at least analogous to the way that punch lines are surprising, through providing a new way of organizing or understanding that which has gone before. People do not find prototypes funny, but they may find them surprising in their relationship to what has gone before.

Experiment 3

The demonstration that surprise can cause a false alarm effect similar to the DRM illusion does not, of course, establish that the DRM illusion itself results from surprise. In Experiments 3 and 4, we reversed the logic, examining whether the DRM illusion occurs when participants could be expected not to feel surprise. If the illusion is due to surprise, then a procedure that eliminates surprise should also abolish the effect.

It is notorious in the literature on the prototype illusion that it can be somewhat reduced but not eliminated by warning participants about the existence and nature of prototypes (e.g., Gallo et al., 1997; McDermott & Roediger, 1998). Even when participants are coached about the illusory effect of prototypes and asked to identify a recognition target as a prototype or nonprototype item prior to judging it for recognition, the effect still occurs (Whittlesea, 2002a). We suspected, however, that being told about prototypes or guessing that they are prototypes is pallid information that is unlikely to have an impact on the experience of surprising relatedness to the list and the consequent feeling of familiarity. In Experiment 3, we attempted to cause participants to become aware of the true prototypical nature of the stimuli by generating the prototypes for themselves before attempting to recognize them. The idea was that when people use their knowledge of the RSVP list to generate the most typical or representative word that they can, they should not then be surprised that that word matches the list well. This lack of surprise should then eliminate the feeling of familiarity typically produced by presentation of a prototype and with it, the memory illusion.

Method

Seventeen participants were tested. As indicated in the Method section of Experiment 1, each prototype had two six-item lists of associates. In Experiment 3, the prototype was inserted into one or the other of its two lists, at random, replacing the associate in a random location from positions 2–5. All 72 lists were presented in the same RSVP format as before, in a random sequence, each list being initiated by a key press by the participant. Immediately after each RSVP presentation, the participants were shown the list again, minus the prototype if it had been presented. This list was presented as a "reminder" of the earlier experience: It was shown in clear view and remained on the screen until the trial was over. The participants were also told that one word shown in the earlier list was missing from the reminder list. They were instructed to use the reminder list to produce a "theme word"—a word that best represented all of the words in the reminder list. In doing so, they were not permitted to produce a word already in the reminder list. Following that, they were asked to decide whether the word they generated had been in the RSVP list.

Results and discussion

The data are presented in Table 3. Participants generated the prototype in about 42% of trials. Considering

Table 3 Experiment 3: Immediate generation and recognition

Prototype	p (generate prototype)	p (claim old/generate prototype)	p (claim old/generate other word)
Old	.45	.46	.16
New	.39	.05	.17

that the original lists consist of 15 associated words, and the "reminding" lists contained only five or six words, so that in principle the participants had many choices for generating a "theme word," this consistency suggests that the prototypes truly are special words.

The generation task also serves as an implicit test of memory for the experience of the prototype within the list. Participants were about 6% more likely to generate the prototype when it had been presented in the list than when it had not, $F(1, 16) = 8.24$, MSE = .01, $\eta^2 = .25$, indicating that the experience of seeing it within the list established an effective representation of that word in only a small number of trials. This result suggests that presenting prototypes within RSVP lists does not provide much support for re-processing them later. Certainly, this factor was much less influential in causing participants to generate the prototype as a "theme word" than its special status as prototype of the list, as shown by the observation that participants generated that word in 39% of trials when it was not in the list.

Analyses of the recognition decisions were conditioned on whether the participant produced a prototype or nonprototype word as the "theme word." In trials in which the prototype was generated, hits exceeded false alarms by 41%, $F(1, 16) = 38.63$, MSE = .03, $\eta^2 = .71$. Hits on the prototype also exceeded claims of recognition for generated nonprototype words (that, because of the requirement to generate a word not on the reminder list, were all false alarms). The critical finding, however, was that false alarms occurred in only 5% of trials in which participants generated and made a recognition decision about a prototype that had not been in the list. That rate was reliably less than false alarms to generated nonprototype words (17%), $F(1, 16) = 9.05$, MSE = .01, $\eta^2 = .38$.[2] It was also substantially smaller than the false alarm rate for the same condition in Experiment 1, in which we observed false alarms in 23% of trials, $F(1, 34) = 16.50$, MSE = .03, $\eta^2 = .33$, in comparing false alarm rates from the two experiments. Thus, the DRM illusion appears to be substantially reduced or eliminated. The only major difference in procedure between Experiments 1 and 3 was that the participants in the earlier study received prototypes passively as recognition targets, whereas participants in this study generated them. In consequence, participants in Experiment 3 would not be surprised by how well the prototypes matched the list, because they deliberately generated them to match the remainder of the list. In turn, this outcome suggests that the illusion occurred in Experiment 1 because the participants, not anticipating any particular word, were surprised that the prototypes matched their lists so well.

[2]The rate of false alarms for prototypes might be expected to be equal in this case to those for nonprototoype items, but it is actually less. We suspect that the difference occurred because the prototypes actually did summarize the lists satisfactorily, whereas the nonprototoype words the participant generated were much less representative. Generating a prototype would thus give the participant a sense of having produced a word totally under control of the attempt to summarize, and so would be attributed only to that source. In contrast, because generated nonprototype words were less representative of the list, they would not so clearly appear to be a product of a summarization process, so that the participants could not as definitely attribute their occurrence in mind to that source.

Experiment 4

Another way to eliminate surprise for a prototype might be to present it out of context, in isolation from its list. We made use of that idea in the next study. Experiment 4 was identical to Experiment 1 in every way, except that the words presented after lists in Experiment 1 were instead presented before their lists. We reasoned that, when not prepared by lists of associates, prototypes would not feel surprising. If participants are surprised at the how well the prototypes match when they have been prepared by a list of their associates, and if that surprise is the proximal cause of the DRM effect, then that effect should be substantially reduced or eliminated by this change of procedure.

By presenting the probe item in advance of the word list, it might be claimed that we have changed a memory task into a search task. We argue, however, that if spreading activation from associates to the prototype is responsible for false alarms to the prototype, then that spreading activation should still occur in the present experiment and should still lead to a high rate of false alarms to prototype probes. Furthermore, if presenting probes in advance of the lists somehow protects participants against false alarms in general, then we should see an overall drop in false alarms relative to the rates found in Experiment 1. Our expectation, however, was that advance presentation of the probes would have one selective consequence: The prototypes would no longer seem surprising when considered as probes, so that the false alarm rate for these items would be no larger than for other probes. Performance in other conditions, including false alarms to nonprototype probes, should be similar to that found in Experiment 1.

Method

Nineteen participants were tested. As just indicated, the procedure was identical to Experiment 1, except that probes were presented in advance of the lists. These probes were presented for 2 s, following which the screen blanked for 500 ms. The RSVP list then ran off automatically, and was followed by the question "Did you see that word in the list?" As before, half of the probes were prototypes and half highest associates; crossed with that factor, half of each type of probe was associated with its own list, the remainder with a list to which they did not correspond. Subject to those conditions, assignment of lists to conditions was independently randomized for each participant.

Results and discussion

Presenting targets in advance of the lists is a clear departure from the usual method of testing recognition. However, as can be seen by comparing Tables 1 and 4, that change in procedure had little effect on performance for highest associates presented with their own lists and for both prototypes and highest associates presented with their unrelated lists. In contrast, the pattern for prototypes presented with related lists changed between the studies. In Experiment 1, which tested participants from the same population as the current experiment, the false-familiarity effect (the difference in claims about prototypes presented with their own list vs. with a list associated with some other prototype) was 12% and highly reliable; in Experiment 4, the false-familiarity effect was reduced to 2%, and was not significant, $F(1, 18) = 1.16$, MSE $= .01$, $\eta^2 = .06$, although the power of this experiment to detect a 12% effect was .99. Moreover, the 10% reduction in the size of the false-familiarity effect between the studies was highly reliable, $F(1, 37) = 10.07$, MSE $= .01$, $\eta^2 = .22$. That is, once again, a manipulation intended to make the prototypes unsurprising also eliminated the false-familiarity effect.

General discussion

In the experiments of this article, we have presented a simple argument. In Experiment 1, we observed the DRM illusion, such that prototypes presented as probes after RSVP lists of their associates were falsely judged to be old 12% more than prototypes presented after unrelated lists. We then showed that surprise (caused by jokes) produces a similar illusion of familiarity (15%), although the probe words following jokes were in no way related to the RSVP lists. In Experiments 3 and 4, we demonstrated that the DRM illusion is substantially reduced or eliminated when procedures designed to preclude surprise were introduced, although in those cases the prototypes were as related to their lists of associates as they were in Experiment 1. We thus concluded that the proximal mechanism producing the DRM illusion is surprise associated with how well the prototype fits with list items.

Table 4 Experiment 4: Search task (prototypes not in related list)

Probe	Related list	Unrelated list
Associate	.83	.06
Prototype	.10	.08

Recognition: *p* (claim old)

Our conclusion contradicts a straightforward spreading-activation account of the illusion, by which the degree of activation of representations in memory directly produces feelings of remembering. However, that account could be modified to incorporate a mediating role for spreading activation in causing surprise. The idea would be that pre-activation of the prototype by its list of associates could cause the activation level of its representation to be as great, or even greater than, representations of words that actually had been seen. That pre-activation could cause the prototype, when presented after the list, to be experienced with a strength of feeling or a degree of fluency exceeding some of the words actually presented, resulted in a feeling of surprise.

Even that modification, however, would not permit the spreading-activation account to explain the similar illusion caused by presenting jokes in Experiment 2, because probes in that case were not related to the RSVP lists. Thus, a different proximal mechanism would have to be invoked for that experiment. Moreover, if pre-activation of prototypes causing surprise were the direct source of the DRM illusion, then that effect should still have been observed in Experiment 3, in which prototypes were judged immediately after related lists, and could also have been expected in Experiment 4 because presentation of the lists should further activate representations of prototypes presented before the lists.

Instead, we suggest that the experiments require some other mechanism that causes prototypes to be surprising in Experiment 1 but not in Experiments 3 and 4. We believe that that mechanism is an evaluation function similar to that assumed by the SCAPE framework. The evaluation function is much more extensive than the monitoring component of the spreading-activation account, which, as described above, is limited to strategic (possibly conscious) decisions about whether to withhold a response to a stimulus that already feels familiar for reasons other than its appearance on the study list. In contrast, under the assumptions of the SCAPE framework, the stimulus comes to feel familiar or unfamiliar due to the operation of the evaluation function itself, as it attempts to make sense of the quality of current processing within salient aspects of the task and the context. These different views of what monitoring may entail can be contrasted by thinking about Experiment 3. That fact that participants generated the prototype rather than had it presented to them should make no difference to the plausibility that such a word was or was not presented within the list, so that the simple monitoring function posited by the spreading-activation account should not be influenced by that change in procedure. In contrast, under the assumptions of the

SCAPE framework, the fact that the participants became aware of the prototype in the act of generating it as a word that best matched the theme of the list would certainly be one of the salient factors to be taken into account in attributing the current processing of that word to a plausible source.

One way to resolve this conflict of accounts might be to assume an evaluative function as complex as that posited by the SCAPE framework along with the assumption of spreading activation. In that case, however, the spreading-activation assumption itself adds nothing to the explanation of the effects observed; the surprise causing false claims of familiarity in Experiment 1 could as easily be caused by a retrospective comparison of the prototype with those associates the participant can remember, and the surprise occurring in Experiment 2 could not be due to spreading activation.

We suggest that a more general explanation can be put forward without assuming spreading activation. The surprise causing false feelings of familiarity in Experiment 1 resulted from a perception of discrepancy, occurring because the prototype (related to all words in the list) was experienced as matching its list exceptionally well, compared with the match experienced when shown a nonprototype target. This effect is analogous to one described earlier, in which words that completed high-constraint stems sensibly were experienced as being surprisingly related to their stems when separated from them by a pause (Whittlesea, 2002b). In both cases, the participants were led into a feeling of surprise because they did not take into account one aspect of the target display: The effect of the pause in causing suspense, in the one instance, and the greater semantic matching of prototypes with their lists, in the other.

The same explanation works for the other three studies. Surprise was engendered by jokes in Experiment 2 by causing a re-interpretation of the meanings given by the questions. In Experiment 3, surprise was prevented by making the prototype's special status obvious to the participants by requiring them to use that dimension in generating recognition candidates. In Experiment 4, surprise was prevented by initially exposing the prototypes in isolation. Like presenting a punch line before a joke, this procedure prevented the participants from being surprised by the relationship between prototypes and their subsequent lists.

There are many other ways of inducing surprise and consequent false remembering, as illustrated by the PHRAUG, HENSION, and sentence-stem effects described in the Introduction, in which spreading activation has no critical role. We conclude that spreading activation is an idea of very little generality as an explanation of remembering. Moreover, despite its widespread appeal as a psychological construct, it is of questionable utility in explaining other phenomena of the mind. For example, spreading activation was the original and still most frequently invoked explanation of semantic priming. Although many investigators have argued for some version of a spreading activation account of that effect (e.g., Collins & Loftus, 1975; McRae, de Sa, & Seidenberg, 1997; Neely & Keefe, 1989; Plaut & Booth, 2000), others have suggested that the effect is retrospective, occurring because the probe recruits relevant information from the prime experience, rather than because the prime pre-activates the probe (e.g., Whittlesea & Jacoby, 1990). The strongest evidence in favor of a spreading-activation interpretation of semantic priming is the common observation that it occurs only if the probe is presented within about 2 s of the prime, and with at most one item intervening between the prime and the probe (e.g., Joordens & Besner, 1992; McNamara, 1992; Neely, 1977, 1991). More recent investigations, however, have demonstrated that when the work required on the prime and the probe is more elaborate than the usual naming and lexical decision tasks, the effect can last much longer (up to half an hour) and with many intervening words (Becker, Moscovitch, Behrmann, & Joordens, 1997; Hughes & Whittlesea, 2003; Joordens & Becker; 1997). That evidence at least makes it questionable whether the usual short duration of the priming effect is due to decaying activation or instead to the insensitivity of the measurement task. It further suggests that semantic priming is not directly due to presentation of the prime, but instead due to presentation of the probe. That is, the prime does not pre-activate the probe, but instead creates a latent resource that can be utilized when the probe is shown.

Similarly, our view of the DRM memory illusion is that the presentation of a list of associates of a prototype creates a resource that influences how a prototype is evaluated when presented in a recognition memory test or when it comes to mind in the course of free recall. It is that evaluation that leads to elevated false alarms and false recall, rather than pre-activation of the prototype during presentation of a list of its associates.

As a final conclusion, we suggest that accounts of subjective experience that are based on automatic flow of activation within semantic networks underrate the complexity of the processes by which people come to experience subjective reactions. Such accounts treat familiarity as a direct product of the fluency of stimulus processing. In contrast, the SCAPE framework suggests an intervening layer of inferential and attributional processes that interpret the significance of salient aspects of processing within the salient context. As demonstrated by these experiments, those processes mediate between the quality of processing and the subjective reaction experienced by the person. The quality of stimulus processing constrains the range of possible subjective reactions, but the evaluative processes are the direct precursor of phenomenology.

Acknowledgements This research was supported by grants from the Natural Sciences and Engineering Research Council of Canada to B.W.A.W. and M.E.J.M.

References

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

Becker, S., Moscovitch, M., Behrmann, M., & Joordens, S. (1997). Long-term semantic priming: A computational account and empirical evidence. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *23*, 1059–1082.

Brainerd, C. J., Wright, R., Reyna, V. F., & Mojardin, A. H. (2001). Conjoint recognition and phantom recollection. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 307–327.

Bredart, S. (2000) When false memories do not occur: Not thinking of the lure or remembering that it was not heard? *Memory*, *8*, 123–128.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*, 407–428.

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

Gallo, D. A., Roberts, M. J., & Seamon, J. G. (1997). Remembering words not presented in lists: Can we avoid creating false memories? *Psychonomic Bulletin & Review*, *4*, 271–276.

Gallo, D. A., Roediger, H. L. III, & McDermott, K. B. (2001). Associative false recognition occurs without strategic criterion shifts. *Psychonomic Bulletin & Review*, *8*, 579–586.

Hancock, T. W., Hicks, J. L., Marsh, R. L., & Ritschel, L. (2003). Measuring activation level of critical lures in the Deese–Roediger–McDermott paradigm. *American Journal of Psychology*, *116*, 1–14.

Hicks, J. L., & Marsh, R. L. (1999). Attempts to reduce the incidence of false recall with source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1195–1209.

Hughes, A. D., & Whittlesea, B. W. A. (2003). Long-term semantic transfer: An overlapping-operations account. *Memory & Cognition*, *31*, 401–411.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3–28.

Joordens, S., & Becker, S. (1997). The long and short of semantic priming effects in lexical decision. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *23*, 1083–1105.

Joordens, S., & Besner, D. (1992). Priming effects that span an intervening unrelated word: Implications for models of memory representation and retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 483–491.

Mather, M., Henkel, L., & Johnson, M. K. (1997). Evaluating characteristics of false memories: Remember/know judgments and memory characteristics questionnaire compared. *Memory & Cognition*, *25*, 826–837.

McDermott, K. B. (1996). The persistence of false memories in list recall. *Journal of Memory and Language*, *35*, 212–230.

McDermott, K. B. (1997). Priming on perceptual implicit memory tests can be achieved through presentation of associates. *Psychonomic Bulletin & Review*, *4*, 582–586.

McDermott, K. B., & Roediger, H. L. III (1998). False recognition of associates can be resistant to an explicit warning to subjects and an immediate recognition probe. *Journal of Memory and Language*, *39*, 508–520.

McDermott, K. B., & Watson, J. M. (2001). The rise and fall of false recall: The impact of presentation duration. *Journal of Memory and Language*, *45*, 160–176.

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, and Cognition*, *25*, 1177–1194.

McKone, E., & Murphy, B. (2000). Implicit false memory: Effects of modality and multiple study presentations on long-lived semantic priming. *Journal of Memory and Language*, *43*, 89–109.

McNamara, T. P. (1992). Theories of priming. I. Associative distance and lag. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *18*, 1173–1190.

McRae, K., de Sa, V., & Seidenberg, M. S. (1997). On the nature and scope of featural representations in word meaning. *Journal of Experimental Psychology: General*, *126*, 99–130.

Miller, M. B., & Wolford, G. L. (1999). Theoretical commentary: The role of criterion shift in false memory. *Psychological Review*, *106*, 398–405.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited capacity attention. *Journal of Experimental Psychology: General*, *106*, 226–254.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Hillsdale, NJ: Erlbaum.

Neely, J. H., & Keefe, D. E. (1989). Semantic context effects on visual word processing: A hybrid prospective/retrospective processing theory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Vol 24* (pp. 207–248). New York: Academic.

Nelson, D. L., Schreiber, T. A., & McEvoy, C. L (1992). Processing implicit and explicit representations. *Psychological Review*, *99*, 322–348.

Nelson, D. L., McKinney, V. M., Gee, N. R., & Janczura, G. A. (1998). Interpreting the influence of implicitly activated memories on recall and recognition. *Psychological Review*, *105*, 299–324.

Norman, K. A., & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory & Cognition*, *25*, 838–848.

Plaut, D. C., & Booth, J. R. (2000). Individual and developmental differences in semantic priming: Empirical and computational support for a single-mechanism account of lexical processing. *Psychological Review*, *107*, 786–823.

Read, J. D. (1996). From a passing thought to a false memory in 2 minutes: Confusing real and illusory events. *Psychonomic Bulletin and Review*, *3*, 105–111.

Robinson, K., & 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, and Cognition*, *21*, 803–814.

Roediger, H. L. III, & McDermott, K. B. (2000). Tricks of memory. *Current Directions in Psychological Science*, *9*, 123–127.

Roediger, H. L. III, Watson, J. M., McDermott, K. B., & Gallo, D. A. (2000). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review*, *3*, 385–407.

Roediger, H. L. III, Balota, D. A., & Watson, J. M. (2001). Spreading activation and arousal of false memories. In H. L. Roediger III, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *The nature of remembering: Essays in honour of Robert G. Crowder* (pp. 95–115). Washington, DC: American Psychological Association

Schacter, D. L., Verfaellie, M., & Pradre, D. (1996). The neuropsychology of memory illusions: False recall and recognition in amnesic patients. *Journal of Memory and Language*, *35*, 319–334.

Seamon, J. G., Luo, C. R., & Gallo, D. A. (1998). Creating false memories of words with or without recognition of list items: Evidence for nonconscious processes. *Psychological Science*, *9*, 20–26.

Seamon, J. G., Luo, C. R., Schwartz, M. A., Jones, K. J., Lee, D. M., & Jones, S. J. (2002). Repetition can have similar or different effects on accurate and false recognition. *Journal of Memory and Language*, *46*, 323–340.

Stadler, M. A., Roediger, H. L. III, & McDermott, K. B. (1999). Norms for words that create false memories. *Memory & Cognition*, *27*, 494–500.

Whittlesea, B. W. A. (2002a). False memory and the discrepancy-attribution hypothesis: The prototype-familiarity illusion. *Journal of Experimental Psychology: General*, *131*, 96–115.

Whittlesea, B. W. A. (2002b). Two routes to remembering (and another to remembering not). *Journal of Experimental Psychology: General*, 131, 325–348.

Whittlesea, B. W. A., & Jacoby, L. L. (1990). Interaction of prime repetition with visual degradation: Is priming a retrieval phenomenon? *Journal of Memory and Language*, *29*, 546–565.

Whittlesea, B. W. A., & Leboe, J. P. (2003). Two fluency heuristics (and how to tell them apart). *Journal of Memory and Language*, *49*, 62–79.

Whittlesea, B. W. A., & Williams, L. D. (1998). Why do strangers feel familiar, but friends don't? The unexpected basis of feelings of familiarity. *Acta Psychologica*, *96*, 141–165.

Whittlesea, B. W. A., & Williams, L. D. (2000). The source of feelings of familiarity: The discrepancy-attribution hypothesis. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *26*, 547–565.

Whittlesea, B. W. A., & Williams, L. D. (2001a). The discrepancy-attribution hypothesis. I. The heuristic basis of feelings of familiarity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *27*, 1–13.

Whittlesea, B. W. A., & Williams, L. D. (2001b). The discrepancy-attribution hypothesis. II. Expectation, uncertainty, surprise and feelings of familiarity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *27*, 14–33.