Journal of Psychology and Behavioral Science
December 2017, Vol. 5, No. 2, pp. 1-13
ISSN: 2374-2380 (Print), 2374-2399 (Online)

Published by American Research Institute for Policy Development
DOI: 10.15640/jpbs.v5n2a1
URL: https://doi.org/10.15640/jpbs.v5n2a1

Investigating the Role of Episodic Gist and False Memory

Kethera Fogler[1]

Abstract

False memories can arise from a sense of familiarity due to having previously encountered similar items; however, some false memories are rich in episodic detail (phantom recollections), which cannot be explained by mere familiarity. The current study hypothesizes that episodic and semantic gist representations combine to form phantom recollections, and that this occurs due to the way information is encoded. Limited research has investigated whether phantom recollections implicate encoding or retrieval processes; however, some research suggests that they are the result of errant binding during retrieval. The role of encoding was investigated by 1) measuring response times at test and 2) examining the effects of sleep consolidation on false memory. Similar response times for true and phantom recollections were found, supporting the role of encoding in both types of memory. Furthermore, a 12 hour delay between study and test that included sleep resulted in performance that was similar to a no delay condition, as compared to a 12 hour delay condition that did not include sleep. These results suggest that sleep consolidation strengthened memory traces that were formed during encoding for both true and phantom recollections.

Keywords: false memory, encoding, memory, episodic gist

1. Introduction

Perhaps the most convincing evidence against the permanent storage of memories is found in universal and anecdotal experiences of forgetting. At one point or another, most people have experienced memory lapses, from forgetting someone's name, to forgetting whether or not one's medication has been taken. In addition to memory lapses, memories can be inaccurate. Memories are organized according to schemas, gist, and semantic categories (Alba & Hasher, 1983; Brainerd & Reyna, 2002; MacKay, 1987; Schacter, Norman, & Koutsaal, 1998) and errors can occur during encoding (MacKay & Miller, 1994; Manelis, Wheeler, Paynter, Storey, & Reder, 2011) and retrieval (Loftus & Palmer, 1974; Mazzoni, 2002) processes.

Research by Roediger and McDermott (1995) led to the development of the DRM (Deese, 1959; Roediger & McDermott) paradigm to examine the occurrence of false memory in a controlled setting. False memories tend to be based on actual events with schema-consistent or semantically related information being mistakenly reported as having occurred (Bartlett, 1932; Roediger, Watson, McDermott, & Gallo, 2001). For example, when given the words *bed*, *rest*, and *dream*, people often recall having been given the word *sleep*, which is semantically related to the presented words but was not presented (known as a *critical lure*). This error is thought to occur due to the disproportional memory trace strengths for semantic information and detailed information (Brainerd & Reyna, 1998, 2002). Detailed information is thought be more vulnerable to forgetting than semantic information. For example, the details of a class lecture may decay soon after the lecture is over, but the gist or meaning of the lecture is likely to be retained.

[1] James Madison University, Psychology Dept. MSC 7704, 91 East Grace Street Harrisonburg, foglerka@jmu.edu, VA 22807, USA, (719) 432-6980, (540) 568-3322 (fax)

Based on this reasoning, theories of false memory (e.g., Brainerd & Reyna; Jacoby, Woloshyn, & Kelley, 1989; Yonelinas, 1994, 2002) have proposed that false memory occurs when an item seems familiar due to its semantic similarity to items previously encountered. This sense of familiarity arising from an emergent semantic gist can lead participants to incorrectly designate the item as "old" (i.e., previously encountered). Detailed memory should only accompany veridical memory, as it is part of the stored representation for an actual event; thus, false memories should never be recollected in specific detail (as they were never actually part of the event).

1.2 Phantom Recollections

A memory error that challenges these views is that of phantom recollection. Phantom recollection is a phenomenon within false memory research that refers to a detailed false memory: Rather than merely recognizing something as familiar, it is 'remembered' with accompanying details of the episode in which it was supposedly encountered. Semantically similar information (gist) is accompanied by specific details (verbatim) to produce detail-rich memories of something that did not occur (e.g., not only thinking that *sleep* had been presented, but stating that it was the third word presented; Gallo & Roediger, 2003; Gardiner & Java, 1990). Errors based on gist versus verbatim details are typically measured using the Remember-Know paradigm. The Remember-Know (RK; Tulving, 1985) paradigm was designed to distinguish between judgments based on the remembrance of detailed memory (i.e., episodic memory) and a feeling of knowing (i.e., semantic memory). In this paradigm, people are typically asked to determine if an item has previously been encountered and, if so, if the item is remembered or if it is simply known that the item was previously presented. A remember (R) response reflects a memory that is accompanied by specific details regarding the experience of having seen the item: The color of the item, the psychological state of the person viewing the item, or other episodic details accompany the recollection of seeing the item. Know (K) responses indicate that an item is known to have been encountered, but the specific details (e.g., its position on the list or computer screen) of this encounter are not recollected. K responses have also been described as having a sense of familiarity with the item, but lacking specific details regarding the item. A classic example of this is the "butcher on the bus" proposed by Mandler (1980), in which seeing one's butcher out of context results in a feeling of knowing the person without the specific recollection that he is one's butcher. Using the R-K paradigm, a false R response (i.e., phantom recollection)is a detailed, specific memory for something that did not happen. Phantom recollections occur at a surprising rate, given that within most false memory explanations, they are theoretically impossible (Brainerd & Reyna, 1998, 2005; Mandler, 2008; Yonelinas, 1994, 2002). Could they arise from the combination of emergent semantic gist and emergent episodic gist?

Associative or semantic theories of memory organization (e.g., Collins & Loftus, 1975; MacKay, 1987) posit that information is stored and organized according to semantic, lexical, and phonological systems. Within this structure, nodes or units of information that are associated are more closely connected, and activation of one closely related concept will activate (or prime, meaning to prepare for activation) another closely related concept (i.e., spreading activation). This associative organization explains the role of semantic gist in false memory formation: A critical lure (e.g., *sleep*) is primed and/or activated by exposure to semantically related words (e.g., *bed*, *rest*, *dream*). These items would be associated semantically and temporally, making it possible for episodic details present at the time to also become incorporated in what is being encoded. The spreading activation through the semantic system could result in episodic details being linked to a non-presented, but semantically associated, concept. For example, if the associated structure linking *bed*, *rest*, *dream* includes details of the study episode, other semantically associated words (e.g., *sleep*) could also then be linked to those episodic details. In short, when *bed, rest,* and *dream* are presented, the context in which they were presented is linked to these items, as are semantically related associates.

1.3 Encoding or Retrieval Error?

Another key question is which process, encoding or retrieval, is primarily involved in the errant binding of semantic and episodic gist representations. Lampien, Meier, Arnal, and Leding (2005) and Lampien, Ryals, and Smith (2008) argue that familiar items presented at test cause a people to search their memories for corroborating evidence as to why the item seems familiar. This search process can lead to content from a particular episode or event being "borrowed" and bound to the familiar item. Thus, according to content borrowing theory, errant binding is the result of a retrieval process. The procedure they used to test this required participants to speak aloud as they determined whether an item had been previously encountered, and these instructions may have prompted the observed reasoning process and memory search for corroborating evidence. This leaves open the possibility that their finding is an artifact of their experimental design.

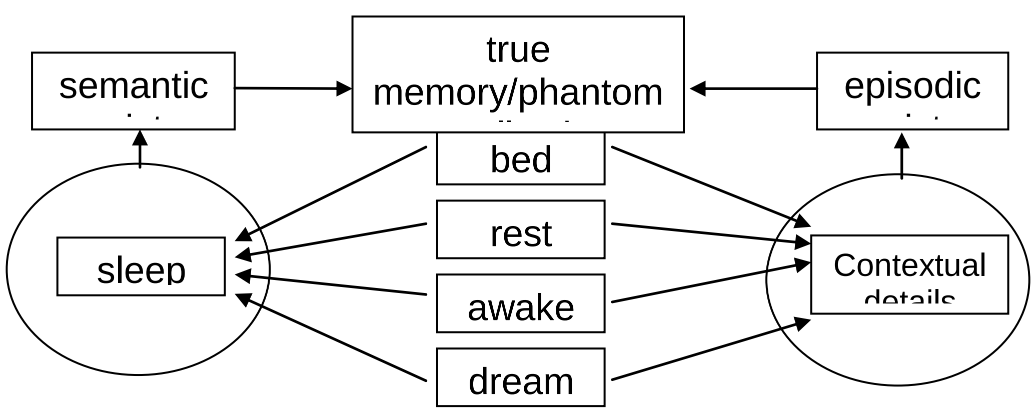
Lampinen et al. (2005, 2008) and other theories that posit a decision-making process at retrieval (see Yonelinas, 2002) predict slower response times for phantom recollections because a memory search for corroborating evidence or metacognitive decision processes must occur prior to the false R judgment (which is obvious since these decision processes are what lead to the false R judgment). Results for reaction times for false memories are mixed, with some studies showing slower reaction times for false memories (e.g., Hintzman, Caulton, & Levitin, 1998) and others showing that some false memory judgments are as fast as true memory judgments (e.g., Stretch & Wixted, 1998). One possible reason for the disparity in reaction time results is that Stretch and Wixted specifically measured true and false R responses, in comparison to K responses, whereas other studies have only measured the time it takes to produce an 'old' response (thus, combining R and K response times). In contrast to errors driven by retrieval processes, Dewhurst et al. (2009) manipulated test conditions to favor associations made at encoding versus those made at retrieval and found that false memory was influenced more by associations during the study episode than at test, supporting the view that errant binding can occur at encoding. Reaction times were not measured; however, it would be predicted that if the errant association or binding has occurred at encoding, reaction times for false R judgments should be similar to those of true R judgments.

1.4 Episodic gist.

The best option for accounting for phantom recollection may be one that incorporates an interactive contribution of *episodic gist* that is tied to a specific semantic context. Fuzzy trace theory (Brainerd & Reyna, 2002) allows for semantic gist to emerge from semantic associations, but this trace is separate from perceptual details that would link it back to a specific context. According to fuzzy trace theory, perceptual consistency between items could become part of the semantic gist representation; however, they would be akin to a sense of strong familiarity and could not include specific details regarding the episode in which the critical lure may have been activated (Brainerd & Reyna, 2005). A detailed memory trace that would tie an item back to its specific context would be a veridical trace, which can only lead to a correct rejection of the critical lure.

The proposed episodic gist model (Figure 1) hypothesizes that the process involved in forming a detailed true memory is similar to that which results in a detailed, false memory. For both types of recollection, detailed memory traces and semantic memory traces can be bound together during an episode in which they are both encountered. Thus, the memory representation would consist of an integration of episodic and semantic gist tied to the specific episode in which they emerge. In other words, episodic gist would not be perceptual familiarity with the entire study episode, but would be specific to contextual details present when studied items are present and critical lures are associatively activated.

Figure 1. Proposed episodic and semantic gist model.



Allowing for episodic gist to be present at encoding would eliminate the need of a corroborative or global memory search at retrieval in order to explain a detailed memory. If the integration of episodic gist and semantic gist occurs at encoding, it would explain the fast (and similar) reaction times for true and false R responses in Stretch and Wixted (1998).

Contrary to theories that rely on a retrieval based memory search and decision process, the proposed episodic and semantic gist model proposes mental representations can be formed at encoding for any item activated at encoding; therefore, for both true and phantom recollection all that is required at retrieval is for the mental representation to be reactivated. The model proposed in Figure 1 allows for phantom recollection to be the product of information bound (or learned) at encoding. Information, both episodic and semantic, that is present during an event (e.g., word list study session) could become encoded together, creating a memory trace for the *combined* association. This would result in a memory pattern or representation that includes both episodic and semantic information. This approach is different from the theories discussed, as it shifts the focus of phantom recollection to an encoding, rather than retrieval, error. That this process can occur for studied items is not controversial or disputed; the present model simply extends this process to items internally generated during study. This approach differs from source monitoring theories because the feature integration does not, necessarily, occur at retrieval, nor require cognitive deliberation. It is predicted that at retrieval, the memory representation is reactivated, but that this representation already includes episodic and semantic details, eliminating the need for a memory search as to why the item seem familiar, or a search for details to corroborate a strong semantic gist trace.

The current study investigates whether encoding or retrieval is primarily responsible for phantom recollections using two different approaches. In Experiment 1, response times were measured during testing. If errant binding requires a memory search and deliberation process, false R responses should have longer response times than true R responses. If false R responses are the result of errant binding at encoding, this deliberation process at retrieval is unnecessary; thus, there should be similar response times for true and false R responses.In Experiment 2, the effect of sleep consolidation on false memory was investigated, again with purpose of investigating encoding versus retrieval processes in phantom recollections. False memories, such as those produced by the DRM (Deese, 1959; Roediger & McDermott, 1995) involve the medial temporal lobe (Cabeza, Rao, Wagner, Mayer, & Schacter, 2001; Schacter et al., 1996), a brain region that is also implicated in memory consolidation during sleep (Wixted, 2004). Sleep consolidation refers to the strengthening of associations and new memory representations during sleep, and it has been shown to be most beneficial when sleep is obtained after encoding and prior to retrieval (Diekelmann & Born, 2010). Hippocampus-dependent memory representations, especially those that are associative or require binding of conceptual and perceptual details (Walker & Stickgold, 2010), should be strengthened during a period of sleep that occurs after information has been encoded and prior to retrieval. This would include information that was errantly bound during encoding; in other words, both true and false detailed memory should be enhanced by sleep consolidation.

Another potential benefit of sleep consolidation is that memories are protected from associative interference (Ellenbogen, Hu, Payne, Titone, & Walker, 2006). If false R responses occur due to a corroborative memory search and content borrowing that occurs at retrieval, false R responses should decrease after sleep. Sleep consolidation should protect veridical memory traces from associative interference, thereby decreasing the likelihood that content borrowing would occur at retrieval. This would then result in decreased false R responses after sleep. In Experiment 2, two competing predictions are tested: 1) if errant binding occurs at encoding, sleep consolidation should result in increased false R responses by strengthening encoded associations (both true and false associations); and 2) if errant binding occurs at retrieval, sleep consolidation should decrease false R responses by protecting veridical, encoded associations from interference.

2. Experiment 1: Method

2.1 Participants

Fifty-seven under-graduate students from Saint Louis University (SLU) participated. Seven of the participants did not follow directions (e.g., consistently using only two response types, designating over half the filler items as "old", etc.) and were dropped. The remaining 50 participants (*M* age = 20.04 years, *SD* = 1.27) were mostly female (78%), Caucasian, (74%), and right-handed (82%). None of the participants were color-blind. Participants were recruited via the Psychology Department's on-line research participant site (SONA systems) and received extra-credit for participation.

2.2 Materials

The primary materials for this research included: (1) the Ishihara color plates color blind test, delivered electronically via E-Prime to ensure that only data from participants who do not have color blindness were analyzed; (2) a demographics questionnaire (see Appendix); (3) an E-Prime presentation of 12 lists, each containing 12 semantically related words and presented in either green or blue font, with the percentage of items in each color (strong context = 80:20 ratio; weak context = 60:40) varied in a counterbalanced fashion across participants; and (4) an E-prime recognition test consisting of a mix of 84 studied items (e.g., *bed, rest, awake*), 12 unstudied critical lure items (items associated with studied items, e.g., *sleep*), and 96 unstudied filler items (e.g., *chalk*). All items were presented in black font. Each word appeared on the screen for as long as it took participants to respond. Each word was presented on two consecutive screens, with the first screen presenting following response options: Remember, Know, Don't Know, or New. Participants responded by selecting the appropriate keys on the keyboard assigned to each response. The keys assigned to responses were counter-balanced across participants. After the judgment response was made, the word appeared on a second screen in which participants were asked to rate how confident they were that the word was "old" (a word encountered during study) or "new" (a word that had not previously been studied) using the following scale: 1 = sure old; 2 = probably old; 3 = maybe old; 4 = maybe new; 5 = probably new; 6 = sure new).

2.3 Procedure

Participants were tested at various times throughout the day, with some participants tested in the morning (8:00 a.m.) and some in the evening (8:00 p.m.) for the purpose of serving as a control group for Experiment 2. Participants were randomly assigned to either the strong context ratio (n = 26) or the weak context ratio (n = 24). The E-Prime experiment began with a color-blind test in which participants used the keyboard to indicate what number was depicted in a slide. This was followed by the word list presentation, with each word presented on-screen for 1 s, with a 500 ms blank screen between individual list words, and a 1 min delay between lists during a study phase. Following the word list presentation, participantswere verbally instructed that they would be tested on the words they had seen during the study phase. They were told they would have four options for responding to the items on the screen: Remember (explained as a situation in which they can recollect specific details about having seen that item), Know (explained as knowing the item was studied, but specific details about having seen the item can not be recollected), Don't Know (explained as not knowing whether the word was previously studied), or New (explained as know the word was not previously studied). Participants were also told that they would be asked how confident they are that the word is old (i.e., previously studied) or new. They were told that these instructions (see Appendix) would also be presented on the computer screen as they continued to the test phase of the experiment (i.e., the verbal instructions were repeated in written form prior to the start of the recognition test). Finally, they were instructed to respond as quickly as possible, while still being accurate in their responses. These instructions were followed by a practice test that used the same procedure as the word list presentation, except 20 numbers were used instead of words. Participants were tested on these numbers using the recognition test procedure in order to ensure they were comfortable using the keyboard to make their responses and to familiarize themselves with the response options. This practice test lasted approximately 5 min and was immediately followed by the word list recognition test. Upon completion of the recognition test, participants filled out a demographics questionnaire and were debriefed.

2.4 Design

A mixed design was used in which context (strong vs. weak ratio) was manipulated between-groups and responses to items (studied, critical lure, and filler) was manipulated within-groups. Filler items (unrelated, non-presented items) were only included in an initial analysis of the overall proportion of "old" responses per context ratio to establish a false memory effect (i.e., higher proportion of "old" responses to critical lure vs. filler items; Gallo, 2010) and to measure response bias. Furthermore, descriptive statistics are reported forall response types (*remember*, *know*, *don't know*, and *new*); however,*don't know* and *new* responses were not included in 5nferential analyses because they are not of primary interest; they serve to balance response options between "old" and "not old," and to guard against guess-based responses for *remember*® and *know* (K) judgments.

3. Results

3.1 Overall false memory effect

The proportion of "old" responses as a function of context ratio was analyzed using a 2 (context ratio: strong vs. weak) x 3 (item type: studied, critical lure, filler) mixed factorial analysis of variance (ANOVA). Remember (R) and Know (K) responses were combined for this analysis and constitute "old" responses. A main effect of item type was observed, F (2, 48) = 177.05, $p< .001$, $\eta_p^2 = .79$. Bonferroni post hoc comparisons revealed that there were fewer "old" responses to filler items, $p< .001$, than to studied items and critical lures, which did not significantly differ. There was no main effect of context, nor was there an interaction between context and item type, *F*s < 1.0.

3.2 Context and R-K judgments

To address differences in proportions of R and K responses as a function of context ratio, a 2 (response type: R vs. K) x 2 (item type: studied vs. critical lure) x 2 (context ratio: strong vs. weak) mixed factorial ANOVA was conducted. A main effect of response type was found, F (1, 48) = 22.16, $p< .001$, $\eta_p^2 = .32$, such that proportions of R responses were higher than those of K responses (see Table 1 for *M*s and *SD*s). There was no main effect of item type or context, *p*s > .05, nor were any interactions found (i.e., item type by context, response by context, item type by response, and item type by response by context were all non-significant, $p> .05$). The proportions of "old" responses for all responses and item types are displayed in Table 1.

Table 1. Mean Proportions of Responses per Item Type.

Item Type	Response (*SD*)			
	Remember	Know	Don't Know	New
Studied	.41 (.18)	.20 (.10)	.28 (.15)	.11 (.08)
Critical Lure	.40 (.25)	.25 (.15)	.27 (.22)	.09 (.11)
Filler	.08 (.08)	.16 (.09)	.49 (.20)	.29 (.17)

3.3 Response latencies

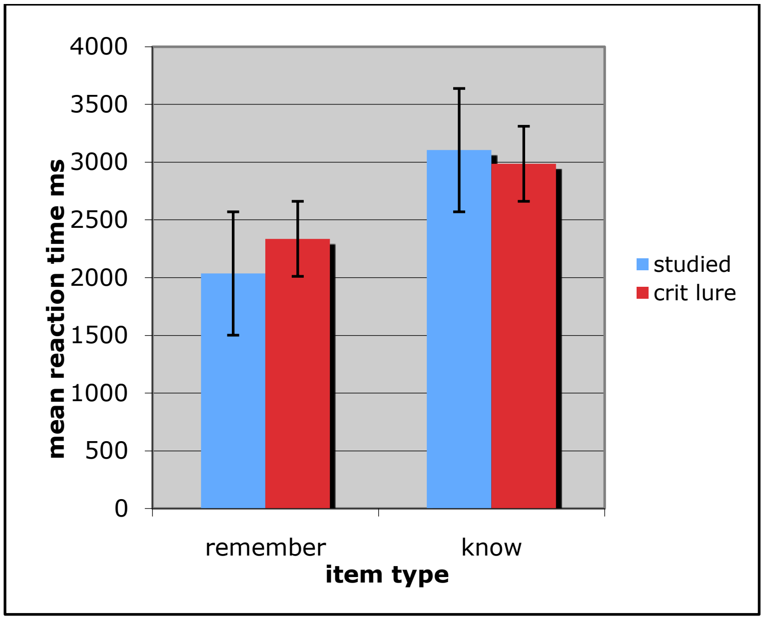
In order to guard against extreme values pulling the mean one way or another, median response times were used. Median response times per item type and response can be found in Table 2. Descriptive statistics for Don't Know and New items are included in Table 2; however, these response types were not included in the main analyses because they are not of primary interest, nor were predictions made regarding response latencies for these items.

Table 2. Median Response Time in Milliseconds per Item Type and Response.

Item Type	Response (*SD*)			
	Remember	Know	Don't Know	New
Studied	1720 (484)	2685 (663)	2741 (840)	2460 (4026)
Critical Lure	1787 (956)	2424 (1400)	2503 (1300)	2971 (2672)
Filler	2130 (679)	2800 (793)	2482 (825)	2329 (771)

Response latencies were analyzed using a 2 (item type: studied vs. critical lure) x 2 (response: R vs. K) within groups ANOVA. A main effect of response was found, such that R responses (for studied and critical lure items) were faster than K judgments (for studied and critical lure items), F (1, 49) = 38.27, $p< .001$, $\eta_p^2 = .44$ (see Figure 2). Neither the main effect of item type, nor the interaction with response was significant, *F*s< 1.0.

Figure 2. Median reaction times for R and K responses per item type.



3.4 Confidence ratings

Mean confidence ratings for all item types and responses can be found in Table 3. Confidence ratings were analyzed using a 2 (item type: studied vs. critical lure) x 2 (response: R vs. K) within groups ANOVA. There was a main effect of response, F (1, 43) = 125.34, p< .001,η_p^2 = .75, such that confidence ratings for R responses were higher than those for K responses. There was no main effect for item type, nor was there an interaction, p> .05.

Table 3. Mean Confidence Ratings per Item Type and Response.

Item Type	Response (*SD*)			
	Remember	Know	Don't Know	New
Studied	1.24 (.45)	2.24 (.47)	3.56 (.40)	5.03 (.74)
Critical Lure	1.32 (.50)	2.28 (.65)	3.49 (70)	5.21 (97)
Filler	1.44 (.40)	2.50 (.41)	3.79 (.40)	5.10 (.63)

Note: Confidence scale: 1 = certain item is old; 2 = probably old; 3 = maybe old; 4 = maybe new; 5 = probably new; 6 = certain item is new.

3.5 Discussion

As predicted, response times for R responses were similar for critical lures and studied items, both of which were faster than K responses. This result is consistent with those of Stretch and Wixted (1998) who found that reaction times for R responses were faster than those of K responses and did not differ across item type (i.e., between studied items and critical lures). Furthermore, confidence ratings were highest for R responses and did not differ between studied items and critical lures: Participants were just as confident in their phantom recollection as they were in their true recollection. These results support the predictions of the model in Figure 1. Critical lures that are bound to episodic details appear to be treated just like studied items that are bound to episodic details, reflected by similar reaction times and confidence levels. Episodic context did not increase R responses; however, proportions of R responses were higher than K responses, indicating details about the study episode were being encoded, even if they did not pertain to font color.

4. Experiment 2: Method

4.1 Participants

Participants consisted of 35 under-graduate students from Saint Louis University. Three of the participants were excluded for not following directions (e.g., consistently using only a couple response options) and one participant in the no sleep group was excluded for taking a nap prior to testing. Thirty-one participants were included in the analyses (no sleep group n = 15, sleep group n = 16; M age = 19.47 years, SD = 1.25). Most of the participants were female (77%), Caucasian, (68%), and right-handed (86%). Participants reported an average of 6.34 hours sleep the night before participation (no sleep group M = 6.30 hours, SD = 1.58; sleep group M = 6.38 hours, SD = 1.38). Average time to bed was 1:00 a.m., and did not differ between the no sleep and the sleep groups, nor did wake time, which was 7:45 a.m., per group. There were no differences between groups on the morning-, evening-type questionnaire (no sleep group: M = 45, SD = 7.73; sleep group: M = 43, SD = 6.72), and scores reflect intermediate types. Participants were recruited via the Psychology Department's on-line research participant site (SONA systems) and received extra-credit for participation.

4.2 Materials

The primary materials for this study consisted of all those used in Experiment 1, with the addition of a sleep log consisting of questions regarding sleep behavior immediately prior to participation (e.g., bedtime, waketime, total hours sleep) and the Morningness-Eveningness Questionnaire (MEQ; Horne & Ostberg, 1976). MEQ scores range between 16-96, with scores below 30 indicating evening types, scores above 54 indicating morning types, and scores between 30-54 indicating intermediate or "neither" types.

4.3 Procedure

Participants were randomly assigned to either the no sleep group or the sleep group. The procedure was identical between the two groups, except for the study and test times. Participants in the no sleep group reported for the study session at 8:00 a.m. and returned for testing 12 hours later at 8:00 p.m. The sleep group reported for study at 8:00 p.m. and returned for testing 12 hours later at 8:00 a.m. the next morning, thus incorporating a night's sleep. The study and testing procedure followed the same protocol as Experiment 1, except that following the word list study presentation, participants were either told not to take any naps and to return in the evening for testing (no sleep group) or to get a good night's sleep and return in the morning for testing (sleep group). When they returned for testing, they were given the same instructions, practice test, and recognition test as was described in Experiment 1.

4.4 Design

A mixed design was used in which group (sleep vs. no sleep) was manipulated between-groups and responses to items (studied, critical lure, and filler) was manipulated within-groups. As in Experiment 1, filler items were only included in an initial analysis of the overall proportion of "old" responses, and only *remember* and*know*responses were used in inferential analyses.

5. Results

5.1 Overall false memory effect

The significance level for all analyses was .05 unless otherwise stated. The proportion of "old" responses was analyzed using a 2 (group: sleep vs. no sleep) x 3 (item type: studied, critical lure, filler) mixed factorial ANOVA. Remember (R) and Know (K) responses were combined for this analyses and constitute "old" responses. A main effect of item type was observed, F (2, 28) = 44.24, p< .001, η_p^2 = .62.Bonferroni post hoc comparisons revealed that there were fewer "old" responses to filler items, p< .001, than to studied items and critical lures, which did not significantly differ (see Figure 3). There was no main effect of group, F< 1.0; however a marginal interaction between item type and group was found, F(2, 28) = 3.06, p = .055, η_p^2 = .10, such that the no sleep group had a higher proportion of "old" responses to filler items, t(28) = 1.94, p = .06, two-tailed.

5.2 R-K judgments

The proportion of R versus K responses was analyzed using a 2 (group: sleep vs. no sleep) x 2 (item type: studied vs. critical lure) x 2 (response: R vs. K) mixed factorial ANOVA.

There was no main effect of item type, response, or group, *p*s > .05, nor were any interactions found, *F*s < 1.0. Proportions of "old" responses for all item types and responses can be found in Table 4.

Table 4. Mean Proportions of Responses per Item Type and Group.

Item type	Sleep				No Sleep			
	Response (*SD*)				Response (*SD*)			
	R	K	DK	N	R	K	DK	N
Studied	.38 (.15)	.22 (.06)	.26 (.14)	.14 (.10)	.33 (.16)	.24 (.14)	.27 (.17)	.16 (.14)
Critical lure	.38 (.28)	.32 (.20)	.20 (.12)	.11 (.12)	.33 (.23)	.31 (.23)	.23 (.21)	.14 (.14)
Filler	.08 (.07)	.19 (.11)	.44 (.17)	.30 (.21)	.16 (.08)	.22 (.15)	.40 (.19)	.22 (.18)

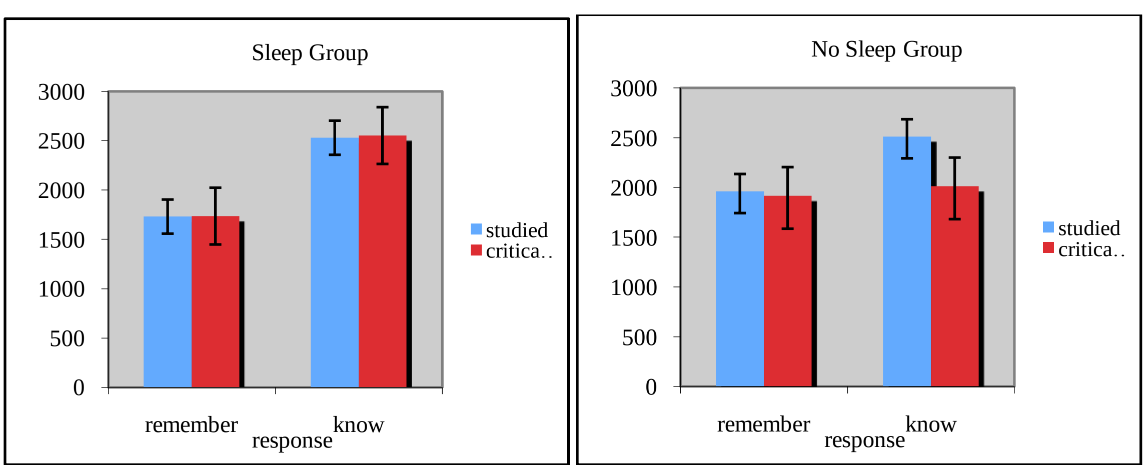
Note: R = remember; K = know; DK = don't know; N = new

Due to the high proportion of "old" responses to filler items for the no sleep group (see Figure 3), corrected recognition rates were also computed for both groups by subtracting hits (R or K responses) to filler items from hits to studied items and critical lures, thus controlling for base-rate false alarm rates (i.e., studied R proportion – filler R proportion; studied K proportion – filler K proportion; critical lure R proportion – filler R proportion; critical lure K proportion – filler K proportion). A 2 (group: sleep vs. no sleep) x 2 (item type: studied vs. critical lure) x 2 (response: R vs. K) mixed factorial ANOVA was conducted using these corrected proportions. A main effect of response type was found, $F(1, 29) = 12.30$, $p < .01$, $\eta_p^2 = .30$, such that R responses were higher than K responses. No interactions were found, however, there was a main effect of study, $F(1, 29) = 6.16$, $p = .019$, $\eta_p^2 = .18$, indicating that sleep group had a higher proportion of R responses for studied items, $t(29) = 2.52$, $p < .05$, as compared to the no sleep group. Mean differences in R responses for critical lures was not significant between the sleep and no sleep groups, although the effect size was medium to large ($d = .71$), indicating that the sample size was too small to detect an effect given the higher variance for critical lures, as compared to studied items.

5.3 Response latencies

There were no direct predictions made regarding response latencies for the sleep versus no sleep groups; however, response time instructions for the recognition test in Experiment 2 were identical to that of Experiment 1 (i.e., respond as quickly as possible while still being accurate). Therefore, reaction times were also analyzed to determine if the pattern found in Experiment 1 had been replicated. Median response times were analyzed using a 2 (group: sleep vs. no sleep) x 2 (item type: studied vs. critical lure) x 2 (judgment: R vs. K response) ANOVA. A main effect of response was found, such that R judgments (for studied and critical lure items) were faster than K judgments (for studied and critical lure items), $F(1, 28) = 8.90$, $p < .01$, $\eta_p^2 = .24$ (see Figures 5a and 5b). Neither a main effect of item type or group were found, *F*s< 1.0, nor were any interactions found, *p*s > .05.

Figure 3.Median reaction times in ms for R and K responses for the sleep and no sleep groups.



5.4 Confidence ratings

Confidence ratings for all response types and items can be found in Table 5. Confidence ratings were analyzed using a 2 (group: sleep vs. no sleep) x 2 (item type: studied vs. critical lure) x 2 (response: R vs. K) ANOVA. There was a main effect of response, $F(1, 24) = 26.31$, $p < .001$, $\eta_p^2 = .52$, such that confidence ratings for R responses were significantly higher than those of K responses. There was no main effect of item type or group, $F < 1.0$. There were no interactions found, ps < .05.

Table 5. Mean Confidence Ratings per Response, Item Type, and Group.

Item type	Sleep				No Sleep			
	Response (*SD*)				Response (*SD*)			
	R	K	DK	N	R	K	DK	N
Studied	1.34 (.36)	2.26 (.42)	3.82 (.52)	4.93 (.62)	1.58 (.59)	2.09 (.71)	3.69 (.55)	4.69 (.64)
Critical lure	1.44 (.53)	2.35 (.59)	3.75 (.82)	4.95 (.75)	1.36 (.44)	2.16 (.73)	3.90 (.93)	4.98 (1.15)
Filler	1.53 (.64)	2.32 (.44)	3.96 (.47)	5.03 (.57)	1.69 (.65)	2.36 (.56)	3.93 (.68)	475 (.61)

Note: R = remember; K = know; DK = don't know; N = new;
Confidence scale: 1 = certain item is old; 2 = probably old; 3 = maybe old; 4 = maybe new; 5 = probably new; 6 = certain item is new.

5.5 Discussion

It was predicted that sleep would result in an increase in the proportion of R responses for both studied items and critical lures due to consolidation of encoded associations. This prediction was partially supported. The proportion of studied items that were recollected (R response) increased after sleep; although the proportion of critical lures recollected after sleep was not statistically significant for critical lures, potentially due to a power problem, the large effect size for the mean difference in corrected phantom recollection rates does suggest that sleep consolidation strengthens learned associations. The opposing prediction that phantom recollections would be reduced after sleep due to consolidation protecting studied information from associative interference, was not supported. Additionally, the pattern for reaction times for R and K responses found in Experiment 1 was replicated. Reaction times were similar for critical lures and studied items that were "remembered" and were faster than those for "know" responses. Confidence ratings for studied items and critical lures were also similar and higher for R versus K responses.

6. General Discussion

The primary research question addressed in this research was whether phantom recollection is produced in the same way that true recollection is produced: Could details from the event's episode become bound to items externally present (e.g., studied items), as well as those internally generated (e.g., critical lures)? It was predicted that semantic and episodic details are bound at encoding to items externally and internally activated (see Figure 1), and that this would be reflected by reaction times that were equivalent for true and phantom recollection. This prediction was supported: There were no differences in reaction times for R responses between studied items and critical lures, both of which produced significantly faster reaction times than did K responses.

These results are similar to those found by Stretch and Wixted (1998) who found response times were fastest for both true and false R responses, which did not differ from each other. Stretch and Wixted did not have an explanation for how or why critical lures would be given R responses (i.e., the memory trace should not be strong enough to meet the criteria for R responses), but predicted that in such cases, the R response should be just as fast for studied items as for critical lures. The present research also predicted fast R response times, whether they were for studied items or critical lures, and posits that this is due to the reactivation of a memory representation that was formed during encoding. Reactivation of an existing memory pattern would be faster than a deliberate memory search for information that then becomes bound together during retrieval. That is, memory representations formed at encoding that include both episodic and semantic information would already exist and would be readily reactivated at retrieval. These types of memories (whether for studied items or critical lures) would be faster than those formed during the retrieval process because they would not require a deliberate memory search for corroborating details or metacognitive decision processes (e.g., Brainerd & Reyna, 2002; Johnson et al., 1993; Lampien et al., 2005).

Lampinen et al. (2005) and theories that posit a criteria judgment, including "process-pure" dual process theories and fuzzy trace theory, would argue for slower response times for phantom recollection because a memory search for corroborating evidence or metacognitive decisions must occur prior to the false R judgment (which is obvious since these decision processes are what lead to the false R judgment, according to these theories). Content borrowing, metacognitive decision processes, and memory searches prompted by familiarity may occur in some situations; however, they do not account for the present study's results, nor do they account for those of Stretch and Wixted (1998). If episodic and semantic information is bound at encoding, there would be no need of a corroborative memory search. Thus, true recollection and phantom recollection should produce response times that are similar, as was found in the present study. In addition to faster reaction times for both true and phantom recollection, high confidence ratings were also found for R responses. These ratings did not differ between items that were actually presented and critical lures. This suggests that critical lures were experienced in a similar fashion as studied items. They were 'recollected' as quickly and as confidently as studied items. Fuzzy trace theory would predict high confidence levels for a phantom recollection due to the premise that the gist trace would have to be very strong for those items in order for an R response to be given; however, as noted previously, it does not adequately explain how details associated with particular study episodes become part of that memory trace. Yonelinas (2007) designates phantom recollections as guesses resulting from a strong sense of familiarity. Responses that involve guessing, even those stemming from a strong sense of familiarity, would not result in confidence levels depicting certainty that the item is "old." Additionally, the confidence ratings for true and phantom recollection did not differ; thus, if guessing is implicated in phantom recollection, confidence levels should have been lower for those items as compared to studied items (for which guessing is not posited to be involved). The present research demonstrates that this is not the case.

Episodic context did not increase R responses, as was predicted; however, proportions of R responses were higher than K responses, indicating details about the study episode were being encoded, even if they did not pertain to font color. Arndt (2010, 2006) found an effect of both font color and font type on the proportion of false memory, such that matching the font at test increased false memory. Arndt did not differentiate between R and K responses; however, LaVoie et al. (2009) found that word lists presented with strong contextual details (i.e., 80 % of the words in one font, 20% in another) were more likely to result in phantom recollection that was tied to the dominant font color. That is, when participants designated an item as "old" they were more likely to assign the dominant font color to the item. That there was no effect of font ratio in the present study does not alter the proposed model of episodic and semantic gist because the model does not require strong contextual manipulations (which would then fail to account for high R response rates in typical DRM tasks). The model simply posits that episodic detail can get bound to semantic information during encoding due to those details being present while studied items are present (externally) and critical lures are activated (internally). Results of the present research indicate a higher proportion of R versus K responses, indicating that details of some sort were being encoded.

It was predicted that sleep consolidation would increase the proportion of R responses by strengthening associations made during encoding. This prediction was partially supported; while the corrected mean difference was not statistically significant for critical lures, potentially due to a power problem, the large effect size for the mean difference in corrected phantom recollection rates does suggest that sleep consolidation strengthens learned associations. However, the contrary prediction supporting the role of retrieval processes in phantom recollection was not supported. This prediction would have been supported had phantom recollection been reduced after sleep. The results from Experiment 2 also support the role of encoding processes in the formation of phantom recollection because even after a 12-hour delay in that reaction times were faster for R responses, as compared to K responses, and these reaction times were equivalent for critical lures as compared to studied items. Furthermore, similar confidence ratings for studied items and critical lures were also replicated. The procedures for Experiment 1 and 2 were exactly the same, apart from the differences in delay between study and test. Therefore, a subset of participants from Experiment 1 who were tested at 8:00 a.m. and 8:00 p.m. can be used as a comparison for the sleep and no sleep groups in Experiment 2. There were no differences in reaction times for the morning and evening participants in Experiment 1, which eliminates a potential time of day testing effect. However, a comparison between these participants and the sleep and no sleep groups in Experiment 2 revealed a significant interaction between response and study, $p < .01$, such that the no sleep group's reaction times for R responses were slightly slower (M = 2090, SE = 175) as compared to the sleep group (M = 1861, SD = 194).

Who did not differ from the subset of participants from Experiment 1 (*M* = 1866, *SD* = 109). These results cannot be explained by interference between study and test for the no sleep group because, after study, the sleep group was awake an average of 5 hours before sleeping. These results suggest that sleep may have strengthened associations made at encoding, resulting in faster reaction times for R responses, rather than a higher proportion of R responses. In fact, if sleep consolidation strengthens learned associations, it may not necessarily increase the number of associations, but could facilitate responding to these prior associations. The no sleep group maintained the same pattern of reaction times, R responses for both studied items and critical lures were faster than K responses; however, the responses were slower overall. The sample size for Experiment 2 was small (observed power = .30), which may have obscured this effect in the main analyses for Experiment 2, so these results should be interpreted cautiously; however, they are consistent with Scullin and McDaniel (2010) who found that sleep consolidation returned performance on prospective memory tasks to that of a no-delay group, as compared to a wake delay group.

Without allowing the integration of episodic gist or emergent episodic details with semantic gist at encoding (e.g., Brainerd & Reyna, 2002; Johnson et al., 1993; Yonelinas, 2002), the results of Experiments 1 and 2 are difficult to explain. Theories that have developed in response to phantom recollection, such as content borrowing, do not account for the results reported here, because they rely on an extended memory search to corroborate the feeling of familiarity generated by a critical lure, which would result in longer reaction times for phantom recollections. The integration of episodic and semantic gist representations at encoding can occur due to associative activation: A critical lure (e.g., *sleep*) is primed and/or activated by exposure to associated words (e.g., *bed*, *rest*, *dream*) and episodic details that are present when the semantic associates are activated (i.e., at study) can become part of the prepositional information for that semantic network. The spreading activation through the semantic system could result in episodic details being linked to a non-presented, but internally activated, concept. This new explanation for phantom recollection shifts the primary process responsible for the error from retrieval to encoding. During retrieval, the detailed memory for a critical lure would already have been formed, required only a reactivation of the memory representation.The proposed model, which allows for the integration of episodic and semantic gist representations at encoding, provides an answer to many of the questions that have otherwise been left unanswered.

Acknowledgement: The author would like to that Dr. Donna LaVoie for her valuable input on this project.

References

Alba, J. W., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin, 93,* 203- 231.

Bartlett, F. C. (1932). *Remembering: a study in experimental and social psychology.* Cambridge, England: Cambridge University Press.

Brainerd, C. J., & Reyna, V. F. (1998). When things that never happened are easier to remember than things that did. *Psychological Science, 9,* 484-489.

Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Psychological Science, 11,* 164-169.

Brainerd, C. J., & Reyna, V. F. (2005). *The science of false memory*. New York, NY: Oxford University Press.

Cabeza, R., Rao, S. M., Wagner, A. D., Mayer, A., & Schacter, D. L. (2001). Can medial temporal lobe regions distinguish true from false memories? An event related fMRI study of veridical and illusory recognition memory. *PNAS, 98,* 4805-4810. doi: 10.1073ypnas.081082698

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

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

Dewhurst, S. A., Bould, E., Knott, L. M., & Thorley, C. (2009). The roles of encoding and retrieval processes in associative and categorical memory illusions. *Journal of Memory and Language, 60,* 154-164. doi: 10.1016/j.jml.2008.09.002

Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature, 11*, 114-127. doi:10.1038/nrn2762

Ellenbogen, J. M., Hu, P. T., Payne, J. D., Titone, D., & Walker, M. P. (2007). Human relational memory requires time and sleep. *Proceedings of the National Academy of Sciences, USA, 104,* 7723-7728.

Gardiner, J. M., & Java, R. I. (1990). Recollective experience in word and nonword recognition. *Memory & Cognition, 18,* 23-30.

Hintzman, D. L., Caulton, D. A., & Levitin, D. L. (1998). Retrieval dynamics in recognition and list discrimination: Further evidence of separate processes of familiarity and recall. *Memory & Cognition, 26,* 449-462.

Horne, J. A., & Ostberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International Journal of Chonobiology, 4,* 97-110.

Jacoby, L. L., Woloshyn, V., & Kelley, C. M. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General, 118,* 115-125.

Lampinen, J. M., Meier, C. R., Arnal, J. D., & Leding, J. K. (2005). Compelling untruths: Content borrowing and vivid false memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 954-963.

Lampinen, J. M., Ryals, D. B., & Smith, K. (2008). Compelling untruths: The effect of retention interval on content borrowing and vivid false memories. *Memory, 16,* 149-156. doi: 10.1080/09658210701839277

LaVoie, D. J., Fogler, K. A., & Besse, A. (2009, May). *Episodic context bias in false memory.* Poster session presented at the American Psychological Science conference, San Francisco, CA.

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior, 13,* 585-589.

MacKay, D. G. (1987). *The organization of perception and action: A theory for language and other cognitive skills*. New York, NY: Springer-Verlag.

MacKay, D. G., & Miller, M. D. (1994). Semantic blindness: Repeated concepts are difficult to encode and recall under time pressure. *Psychological Science, 5,* 52- 55.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87,* 252-271.

Mandler, G. (2008). Familiarity breeds attempts: A critical review of dual-process theories of recognition. *Perspectives on Psychological Science, 3*, 390-399.

Manelis, A., Wheeler, M. E., Paynter, C. A., & Storey, L. (2011). Opposing patterns of neural priming in same-exemplar vs. different exemplar repetition predict subsequent memory. NeuroImage, doi: 10.1016/j.neuroimage. 2010.12.034

Mazzoni, G. (2002). Naturally occurring and suggestion-dependent memory distortions: The convergence of disparage research traditions. *European Psychologist, 7,* 17- 30.

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

Schacter, D. L., Norman, K. A., & Koutstaal, W. (1998). The cognitive neuroscience of constructive memory. *Annual Review of Psychology, 49,* 289-318.

Stretch, V., & Wixted, J. T. (1998). On the difference between strength-based and frequency-based mirror effects in recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 24,* 1379-1396.

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

Walker, M. P., & Stickgold, R. (2010). Overnight alchemy: Sleep-dependent memory evolution. *Nature Reviews Neuroscience, 11,* 218. doi:10.1038/nrn2762-c1

Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annual Review of Psychology, 55,* 235-269.

Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 20,* 1341-1354.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46,* 441-517.