False Memories:

Evidence for Unintentional Processing of Semantically Related Words using a Stroop Variation Task

Ana Genao

Psychology Senior Thesis

Bryn Mawr College

Abstract

The present study examined false recognition of semantically related non-presented words. We were primarily interested in replicating a study by Dodd and MacLeod (2004) where false memories were produced in groups intentionally processing words (by reading or studying) and unintentionally processing words (performing a variation of the Stroop task). Seventy-five undergraduate students (30 in the unintentional group, 30 in the intentional group and 15 in the control group) completed a primary task followed by a recognition test for actually presented and non-presented words. Results revealed that false alarms to non-presented words were robust and fairly similar across all groups. These findings are consistent with an activation/monitoring theory where semantically related words may evoke similar yet non-presented words through conscious realization or through unconsciously and automatic activation.

False Memories: Evidence for Unintentional Processing of Semantically Related Words using a Stroop Variation Task

Despite the abundant research conducted on memory accuracy and distortion, false recollection or recognition of past events continues to occur. Recent interest in the field grew as researchers sought to understand the mechanisms underlying these illusory memories, spurred in part by real-life controversies such as recovered memories of childhood sexual abuse (Koriat, Goldsmith and Pansky, 2000).

One of the leading contributors to research on false memories include Elizabeth Loftus (1975) who questioned the accuracy of information content reported in eyewitness testimonies and attributed repressed memories of sexual accusation claims to a “confirmatory bias” where therapists help confirm memories rather than disconfirm them by encouraging recall of past events in a suggestive manner (Brown, Scheflin, and Hammond, 1998). Loftus claimed that memory recovery techniques focus on reconstructing memories, and therefore do not accurately reproduce aspects of information represented within memory structures (as defined in Brown, Scheflin, and Hammond, 1998).

In one of her many studies, Loftus and colleagues (1978) evaluated the idea of misguided memory and memory inaccuracies by presenting participants with certain situations and varying time between this situation and time when misleading information was given. The post-event questionnaire containing neutral, misleading or inconsistent information was administered anywhere from 20 minutes to 1 week after initial situation. After completing the questionnaire, participants were tested immediately for memory accuracy. Results showed that the percentage of participants that accurately recognized material from initial situation declined with increased interval times. Loftus and colleagues also delayed time between receipt of misleading information and retrieval of original material, and found that misinformed participants performed worse when tested immediately after completing questionnaires, but better when tested a day after.

From their findings these researchers concluded that memory may become distorted when misleading information and actual memories occurring around the same time, are subsequently followed by a long delay.

The idea of memory reconstruction dates back to 1932, when Bartlett, in his research of false memories, had participants read and recall the content of folktales. He found that people tended to produce distortions in stories consistent with their cultural expectations (as cited in McDermott, 1996).

Bartlett claimed that remembering involves a reconstruction of past events based on one's schemas or "cognitive structures that represent knowledge about a type of stimulus including its attributes and the relations among those attributes"[1]. To account for false recall and false recognitions Bartlett described an activation account of false memories where during the encoding process people expand on the information given through inferential processing and activate related information (1932). In this manner, one's schemata and associated knowledge may fill in omitted information. Along with this activation idea, monitoring memory during encoding and retrieval may help control for false memory generation by bringing information which refers to a perception of past events to consciousness and therefore helps distinguish it from information that was not presented (Johnson, 1981).

Many researchers have followed the paradigm set forth by Bartlett, using sentences, prose passages and slide sequences in order to investigate these false memory constructions (as cited in Roediger and McDermott, 1995). In 1965 Underwood, set up a technique to study false recognition of words by creating a list learning paradigm where participants had to identify whether a word had been presented in a list. In his findings, words related to previously presented words showed only a modest amount of false recognitions.

At about the same time that Underwood was using lists of words, Deese (1959) was developing a standard list learning procedure. This procedure introduced a new way of examining recollection by presenting participants with lists of words semantically associated to a non-presented critical word (or lures). Deese was interested in seeing the effect of extra-list intrusions (non-presented words) on lists of words studied. Therefore, participants were given words from 36 lists containing 12 words each and were told to study the words in preparation for a free recall test. His findings showed that for lists of semantic associates such as thread, pin, eye, sewing, pricked, point, sharp, thimble, haystack, pain, hurt, and injection, the critical word “needle” was often falsely recalled as having been presented.

Some of these semantically associated word-lists gave rise to extra-list intrusions while others did not. These findings led Deese to conclude that false recall would occur for lists that contained both forward associations (associations from the critical lures to the list words) and backward associations (associations from the list words to the critical lures). Deese found a .87 correlation between the probability that the list items elicited the critical lure on a free association test and the probability of obtaining an intrusion in recall for that list, indicating a robust effect of false recall (as cited in McDermott, 1996).

Various explanations have been devised to explain these false memories. Underwood’s (1965) implicit associative response account added to Bartlett’s (1932) ideas on inferential processing and suggested that people activate semantically associated concepts/words during encoding. The activated concepts/words are said to be the implicitly associated responses. For example, when participants study a word such as “thread,” an implicit response may be to think “needle”. In cases like these, the studied words have high levels of Deese’s (1959) backward associative strength notion—the strength of the connection between list items and critical lures (as cited in Roediger, Watson, McDermott, Gallo, 2001). In addition, these associations were thought to be consciously activated and only implicit in that the response was not overtly produced, meaning that people may become aware of a word during study and rehearse it, but may only claim to recognize it due to this implicit associative response (Roediger and McDermott, 1995).

Roediger and McDermott (1995, 2000) replicated Deese’s findings of the false recall of non-presented critical items, and extended it to include a recognition test as well as auditorily presented words. They used the list that had elicited the strongest false recall tendencies. The study also included a remember-know judgment task based on Tulving’s (1985) work that examined whether participants specifically remembered some aspect during the word’s presentation or simply knew the word had been presented without being able to pinpoint the exact moment of occurrence. This judgment task was added in order to distinguish studied words remembered from non-presented words remembered, given that one would only expect studied words to be remembered.

Together with the procedure Deese had developed, their work became known as the DRM paradigm—a paradigm that set forth a standard procedure for other researchers to work with in their study of the false memory phenomena (Deese, 1959; Roediger and McDermott 1995, 2000).

Results from the Roediger and McDermott experiment showed that participants falsely recognized critical words as having been presented at a rate comparable to the hit rate for studied words. They also found that the false recognitions were often judged to be "remembered" more than "known". These findings suggested that encoding processes were critically involved in memory distortion, similarly to the manner in which researchers (e.g. Reysen and Naire, 2002; Bauml and Kuhbandner, 2003; Anderson, Bjork and Bjork, 2000) found retrieval processes to play a significant role in forgetting mechanisms. Critical non-presented words being judged as remembered suggests that some specific attribute during its occurrence was being encoded. Additionally, although a remember claim being made could indicate that the word was being processed consciously and intentionally, it could also indicate that the word was being processed unconsciously and unintentionally by means of spreading activation. Participants who falsely "remembered" items may have encoded the gist of the experience, where the interpretation of the words or their general semantic content was used rather than the verbatim or specific attributes of individual items (Brainerd and Reyna, 2002).

Following the DRM paradigm, a number of researchers have manipulated conditions such as study and test phase to help eradicate these false memory results. Seamon, Luo, Gallo (1998) varied the rate of word presentation from 2s, 250 ms 20 ms to test whether recognition of DRM list words and critical lures would vary with presentation rate. They found that although correct recognition decreased with decreased presentation rate, false recognitions still remained robust. Additionally, McDermott (1996) found that false memories persisted despite efforts to test for recall immediately after study phase. She found that these false memories were found even after multiple study and test trials. Even more striking was when Seamon et.

al (1998) warned participants of the false memory phenomena and still found that the effect although reduced was not eliminated, suggesting that complex mechanisms were involved.

Roediger and McDermott (1995) expanded on both Bartlett's (1932) and Underwood's (1965) ideas about false memories to propose that individuals can activate semantically associated words either consciously through elaborative processes or unconsciously through an automatic spreading of associations elicited by a particular word. Conscious activation of the non-presented words would involve intentionally thinking of the critical word during the study/encoding phase and confusing the source of the word during retrieval. Here, monitoring the source of memories would help reduce memory confusion, as seen in experiments where participants were advised about the false memory phenomena (e.g. Gallo, Roberts, and Seamon, 1997). In contrast, automatic or unconscious activation consists of a strongly activated word, not coming to conscious awareness (Roediger, Watson, McDermott, and Gallo, 2001). For example, one may not realize that a particular word comes to mind during the encoding phase (such as seeing the word bed and not realizing that it evokes the concept sleep), but the related word can nonetheless be processed.

Recent research, in fact, has provided evidence for both the conscious and unconscious activation of non-presented critical words (e.g. Seamon, Luo, and Gallo, 1998; Dodd and MacLeod, 2004). With the DRM paradigm in mind, Seamon et. al (1998) sought to investigate whether participants would demonstrate the false memory effect if they were unable to recognize list items. They presented participants with a list of semantically related words, either with or without a concurrent memory load (seven-digit auditory number sequence). As mentioned earlier, word presentation rates varied, such that some words were presented at unrecognizably fast rate (2s, 250 ms 20). Participants studied words for the recognition test to follow. Seamon et. al found that participants in both conditions falsely recognized semantically related, non-presented critical words. False recognitions were also found despite the variation in word presentation rate. Even though the hit rates to studied items were low for the two memory load conditions, false alarms to non-presented items were fairly high, indicating that recognition of list items was unnecessary to produce the false memory effects. This also suggests that participants must have processed and activated some aspect of the word.

Similarly, Dodd and MacLeod (2004) sought to investigate whether the false memory effect required intentional and therefore conscious word processing or whether these false memories could be produced by the mere presentation of DRM lists without any intention to process. In order to achieve this, they adapted a Stroop-like task, given that distracter words in a color identification task are often processed unintentionally. In the standard Stroop task (1935) color words (such as red, blue, green and yellow) are printed in either compatible or incompatible colors. When word and word color match, relative to when word and word color do not match, facilitation is said to occur and participants are faster to identify the color that a word is presented in.

In their experiment, Dodd and MacLeod presented participants with a list of semantically related words, one after the other, in varying print colors. Testing instructions were manipulated such that in one of two experiments, (unintentional group) participants were told to ignore the words presented and simply identify the color of the word, as in the original Stroop task. In a second experiment, participants were instructed to intentionally read the words presented but ignore the print color. Neither group was aware of the recognition test to follow. Dodd and MacLeod did this in order to validate their unintentional task, in contrast to previous experiments which administered weak tests of unintentional processing. These previous studies, as in Seamon et. at (1998), had informed participants of a memory test to follow, thus measuring intentional reading and processing of list items instead of unintentional processing.

Dodd and MacLeod predicted that if false recognition is solely due to intentional processing, then the false memory effects should be eliminated when DRM words are presented within a color-identification task where words are processed automatically albeit unintentionally.

Results for the unintentional condition indicated that hits to actually presented words were much greater than false alarms to unrelated/non-presented or weakly-related/non-presented words. In the intentional condition (experiment two) memory for actually presented words was better than false alarms to those words not presented, as expected given that participants were aware of the impending memory test. Importantly, false alarms to critical lures in both conditions were greater than hit rates.

Although, false memories were found under both intentional and unintentional conditions, questions arose concerning the procedural methods involved. The study used only 36 words, all semantically related to one of three DRM lists. Such a limited number of words all related may have aroused suspicion in participants. That is, instructing university students to simply read or ignore blatantly related words may not have been sufficient to make the test very believable and prevent suspicions of a memory test to follow. Additionally, researchers did not follow-up by asking whether participants believed the cover story. Due to such limitations, a more stringent test is necessary to avoid having participants speculate the true nature of the experiment and thus affect results.

The present study sought to replicate the DRM paradigm's false memory findings under unintentional processing as well as make the experimental procedures more believable to participants. For this reason, the same 36 semantically related DRM list items were used, but also embedded within a larger list of words containing color words (e.g. red, blue, green and yellow) and unrelated or weakly-related non-color words (e.g. banner, sickness and elastic), which were randomly chosen from other DRM lists. In the unintentional group participants were given a variation of the standard Stroop task (a color bar presented with either a color-word or non-color word) and told that the purpose of the experiment was to evaluate color bar naming speed. This deception was necessary and unavoidable in order to show unintentional processing of words. There were two types of conditions in addition to the unintentional group: an intentional group and a control group.

The Stroop variation in the current study was based on Kahneman and Chajczyk's (1983) experiment where word and color were spatially separated in order to observe the effects of attention allocation on color-identification tasks and to show that reading was automatic. We adapted this version of the Stroop Task in our experiment in order to validate our test of unintentional processing (since words did not have to be read directly in order to be attended to or processed), but more importantly to make the test more believable to participants.

We predicted that the unintentional group as well as the intentional group would show these false memory effects due to the activation/monitoring theory. Additionally, we expected the mean for actually presented words to be less in the unintentional condition compared to the control condition, given that participants in the former condition would not be processing words directly.

Methods

Participants.

The sample was comprised of 75 students from Haverford and Bryn Mawr College ages 18-24, who either volunteered or received course credit for participation. Each participant completed an informed consent form prior to beginning the experiment. Sixty participants were randomly assigned to either the unintentional condition or the intentional condition. The remaining 15 students served as a control group.

Materials.

All stimuli were created using an E-prime version 1.1 program and displayed on an IBM-compatible computer. Stimuli consisted of a bar in the color red, blue, yellow or green and a word above and below the bar. Participants identified colors using four response keys (z, c, b, and m) representing the four colors.

The 120 non-color words and the recognition test items were adapted from the DRM lists of semantically associated words (Roediger, Watson, McDermott and Gallo, 2001). Three 12 word lists, each related to a non-presented critical word or target (sleep, cold and window) were used to test for false memories. These three lists were used given that they induced high levels of false recognition in previous studies (Stadler, Roediger, and McDermott, 1999). The presentation of each list block during the color-identification task was counterbalanced and embedded within the rest of the words. Within each list, words were always presented in the same order (from strongest to weakest associate) as is typically done with this procedure. The rest of the words used included randomly chosen items from each of the remaining 52 DRM lists and 120 color words (30 presentations of the words red, blue, yellow and green).

The 54 recognition items were divided into presented words (old) and non-presented words (new). The actually presented words included: a total of 12 words (studied items) from the three DRM lists (Words 2, 5, 8 and 11 from each list) and 12 randomly chosen words from the remaining DRM lists (random studied items). Non-presented words included: the 3 target words, 21 randomly chosen unrelated words from the other DRM lists, 6 weakly-related words from the other DRM lists (Words, 13 and 14). Recognition items were presented in random order. (See Appendices for words used.)

Design.

The study used a between-subjects design: the unintentional condition was set up to evaluate unintentional processing of semantically related words and the intentional condition was set up to evaluate intentional processing. The control condition served to validate the findings in the unintentional condition where one would expect recall for presented items to be greater here than in the control group. Independent variables were the condition assigned to, while dependent variables were percentage of veridical and false memory produced.

Procedure.

Participants in each condition were presented with 240 words (120 color and 120 non-color) above and below a color bar. Those in the intentional condition were simply instructed to study all words (e.g. banner, sickness and elastic) other than color words (e.g. red, blue, yellow and green) in preparation for a memory test to follow. Each bar/word was displayed on the screen for 1000 milliseconds.

Participants in the unintentional condition were unaware of the memory test to follow and were simply instructed to identify the color of the bar presented as quickly and accurately as possible, by pressing the corresponding keys labeled in one of the four colors. They were told to ignore the words above and below the bar. Similarly, participants in the control group were not aware of the memory test to follow and were instructed to say each word aloud and then press the spacebar key. They were told to work as quickly as possible. None of the groups were aware of the semantically relatedness of the words presented.

During the recognition test, participants were told to identify a subset of words from the study as either old or new (by entering “o” or “n”, respectively). Upon completion, participants were fully debriefed. In addition, participants in the unintentional condition were questioned concerning the believability of the primary task.

Results

Unintentional Group and Stroop Effects

Two paired samples t-tests evaluating mean color identification time and mean accuracy, indicated that reaction times in the congruent conditions were significantly faster (M=832.23, SD=124.47) than reaction times in the incongruent conditions (M=864.59, SD=146.35), $t(29)=3.11$, $p<.05$ and that mean accuracy scores in the congruent conditions (M=.99, SD=.015) were significantly greater than mean scores in the incongruent conditions (M=.98, SD=.024), $t(29)=3.093$, $p<.05$. These findings, which demonstrate facilitation and inhibition, suggest that the study was effectively designed in a manner to replicate the Stroop effect.

Recognition Performance using Uncorrected Scores

Three separate one-way repeated measures ANOVAs were performed for all experimental conditions in order to test for differences in mean recognition according to word type. Using uncorrected or raw scores for the unintentional groups, a significant difference in recognition was found, ($F(4, 116) = 28.35$, $p<.05$). Follow-up analyses using protected t-tests revealed that almost all word-types were significantly different from one another, such that hits to actually presented words (both studied and random studied) were significantly greater than false alarms to weakly-related/non-presented words or to unrelated/non presented words, but significantly less than false alarms to non-presented critical lures. Of great importance was the fact that the false alarms to these critical lures were very high. As in the findings of Dodd and MacLeod, no statistically significant difference was found between false alarms to weakly-related/non-presented and false alarms to unrelated/non-presented words. There was also no significant difference between random studied words and studied DRM list items, and between false alarms to critical lures and weakly-related/non-presented words or unrelated/non-presented words. Table 1 lists all the means, standard deviations, and t-values.

Significant recognition differences in word-types were also found for the intentional group, (F (4, 116) =158.51, p<.05), such that almost all word-types were significantly different from each other (see Table 2 for means, standard deviations and t-values). The hit rates for both presented words were significantly greater than either of the false alarms to weakly-relate/non-presented and unrelated/non-presented items. The hit rate of random studied words was significantly greater than the false alarm to critical lures. However, with the protected t-tests and adjusted alpha levels, no difference was found between recognition for studied items and critical lures. As expected, no difference was found between studied and random studied items and between weakly-related/non-presented items and unrelated/non-presented items. False alarms to critical lures were found to be significantly greater than false alarms to weakly-related/non-presented items and unrelated/non-presented items.

The ANOVA for the control group indicated that significant differences existed between word-types, (F (4, 46) =31.90, p<.05). Post hoc tests revealed that the mean of presented words (both studied and random studied) was significantly greater than either means of weakly-related/non-presented items and unrelated/non-presented items, indicating that the hit rate was greater than false alarms to these two word types in the control condition. Again, false alarms to critical lures were found to be significantly greater than false alarms to weakly-related/non-presented items and unrelated/non-presented items. False alarms to unrelated/non-presented words were significantly greater than false alarms to weakly-related/non-presented items. No difference was found between studied items and critical lures and between random studied items and critical lures. No significant differences were found in recognition accuracy between random studied and studied DRM list items. Table 3 lists all the means, standard deviations, and t-values.

Recognition Performance using Corrected Scores

The next set of analyses used the same procedure as Dodd and MacLeod (2004) to account for the fact that weakly-related and unrelated words had low, albeit non-negligible false alarm rates. This procedure involved correcting false recognition scores for both critical lures and presented items. The new corrected scores were determined by subtracting the false alarm rate for the average of weakly-related and unrelated words from the false alarm rate for critical lures. These values are presented in Table 4.

One-way repeated measures ANOVAs were also calculated for the three conditions (unintentional, intentional and control) using these corrected scores.

In the unintentional group, a significant difference in recognition was found, ($F(2, 58) = 18.93$, $p<.05$). Follow-up analyses using protected t-tests revealed that the false alarms to critical lures were significantly greater than the hit rates to both random studied and studied words, $t(29)=4.88$, $p<.01$ and $t(29)=4.51$, $p<.01$. This was also the case using the uncorrected scores. Again, as with the uncorrected mean scores, no difference was found between means for random studied and studied items, t (29) =.83, p>.01.

In the intentional group, a significant difference in recognition was also found, (F (2, 58) = 7.53, p<.05). Post hoc testing showed that false alarms were significantly greater than the hit rates to random studied and studied words, t(29)=2.75, p<.01 and t(29)=3.89, p<.01. This was different from the findings using the uncorrected scores, where no difference was found between the studied and critical lures and where the hit rates to the random studied were greater, not less than the false alarms to critical lures. As expected again, no difference was found between the means for random studied and studied items, t (29) =1.19, p>.01.

Finally, in the control condition, the one-way repeated measures ANOVA showed no statistically significant difference between any of the word-type means, unlike the differences found with the uncorrected scores, (F (2,28)= .190, p>.05).

Between Groups Analyses

In order to examine differences between hit rates and false alarms in the unintentional, intentional and control groups, three one-way between subjects ANOVAs were performed using the corrected scores only. A significant effect of group type on recognition performance was found, for the studied items (F (2, 74) =29.82, p<.05), random studied items (F (2, 74) =29.54, p<.05), and for the critical lures (F (2, 74) =7.41, p<.05).

Post-hoc comparisons using the Tukey revealed that for the studied items, a significant difference was found between intentional and unintentional group means, and control group and unintentional group means only, such that the hits to studied items were greater in control group compared to the unintentional group. This finding demonstrates that participants did in fact ignore the words as instructed and therefore found the task very believable. Hits to studied items were also greater in the intentional group compared to the unintentional group, as expected since participants in the intentional group were aware of a memory test to follow.

Post-hoc comparisons for the corrected random studied items, revealed a significant difference between unintentional and intentional as well as unintentional and control groups, such that hits to these presented items were greater in intentional compared to the unintentional group and greater in the control group compared to the unintentional. Again, the findings between the unintentional and control demonstrate that participants in the unintentional group did not suspect a memory test and therefore ignored the words. No other significant differences were found.

Finally, post-hoc comparisons for the corrected critical lures revealed a significant difference between unintentional and intentional groups and between intentional and control groups only, such that false alarms to critical lures were greater in the intentional condition. Nonetheless, false alarms were found in all groups. Table 5 lists all the means, standard deviations, and significant values for all post hoc tests.

Of the 30 participants in the unintentional condition, 28 or 93.3 % were completely surprised when a memory recognition test followed the color identification task. Upon follow-up questioning, the two remaining participants said that although they did not have any memory test suspicions they were curious as to why the list items were so long.

Discussion

The primary goal of the present study was to replicate false memory findings using conditions where words had to be processed both intentionally and unintentionally (as done by Dodd and MacLeod, 2004). We were interested in supporting the activation/monitoring theory where non-presented words are said to be activated in memory both consciously (with the realization that the word comes to mind) and unconsciously (without the awareness that the word is evoked) (Roediger and McDermott, 2000). A second goal was to modify the unintentional condition procedure used by Dodd and MacLeod in order to make it more believable. We did this in order to be certain that participants were fully engaged in the primary task of identifying the color of the bar and therefore actually ignoring words. This would show that participants were indeed unintentionally processing DRM list items.

Recognition Performance Within-Groups

As predicted, participants in both the unintentional group and the intentional group developed false memories. False memories were even found in the control condition. The fact that our robust measure of unintentional processing was able to elicit false memories indicates that as Seamon, Luo, and Gallo (1998) found, related non-presented words (critical lures) do not need to be consciously activated or rehearsed during the encoding phase in order to be recognized. The activation/monitoring theory holds that associative activation (as with these critical lures) can occur automatically without the realization that one is extracting information from a presented word to in turn conjure up a related word (Roediger and McDermott, 1995). As Bartlett noted, extracting information from a word during encoding or constructing networks of associations without awareness may lead to the false recognition of words that were never presented. Similarly, the theory holds that the associative activation may also occur through a more controlled process where participants remember encoding some aspect of the word during study. As Roediger, Watson, McDermott and Gallo (2001) described, the failure of source monitoring or distinguishing perceptions of past from reality of past, can lead to these false memories.

As expected in all three conditions, memory performance for actually presented words was greater than the false alarms to the unrelated and weakly-related non-presented words. However, the presence of the false alarms suggests that while actual memory appears to be intact, false memories are not eliminated. This is consistent with the idea that false memories persist despite accuracy in veridical memory (McDermott, 1996).

More interestingly, was the finding for the unintentional and intentional groups using corrected scores which showed that participants were more likely to false alarm to critical lures than they were to correctly recognize actually presented words. This also demonstrates the robustness of the false memory effect found by Dodd and MacLeod (2004).

Our results differ from the findings of Dodd and MacLeod in that both uncorrected and corrected scores for the unintentional group only, showed false alarms to critical lures to be much greater than hits to actually presented words. Not only does this serve as strong evidence that false alarms can be elicited non-consciously, but also that these illusory memories can be as high as veridical memory. Roediger and McDermott (1995) attributed the robust finding in false alarms to the lists of related words used, which according to Underwood (1965) can make seeing associated words (such as hot, weather and ice) contribute to implicitly thinking of a non-presented associate, (cold). The low levels of false recognition observed in Underwood's study may be accredited to his word lists which consisted of single words modestly related to subsequent words. Additionally, Koriat et al. (1998) argue that false alarm rates only tend to be as high as hit rates for presented words in contrived situations as in these studies and are not representative of memory in general. According to such researchers, our findings along with many involving the DRM paradigm are inconsistent with claims that in reality a recalled/recognized item is more likely to be correct rather than false. Although, we do not claim that false alarms tend to be as high as hit rates in general, the findings in the intentional and unintentional condition show this for semantically-related word lists.

The insignificant differences in memory performance between random studied (e.g. sickness, banner and summer) and studied semantically related list items (e.g. snooze, snore, blanket) were expected given that both items were actually presented ("old") items. Although one might think that grouping items under a similar category (as in the studied words) would help with retrieving related words and therefore improve memory performance, studies have shown that this is not the case given that retrieval can impair remembering of related items (Anderson, Bjork and Bjork, 2000). However, according to Deese's (1959) interitem associative strength, the more that items within a list are associated, the better memory performance for those items.

In general, the findings in all conditions where false alarms to non-presented critical lures were greater than false alarms to weakly-related or unrelated non-presented words show that inferential processing is occurring and semantically related words are in fact spreading activation to the lexical system as Bartlett (1932) suggested. Because of their association to the list items, the critical lures are evoked more strongly than the other words which are not as strongly related. This is consistent with Deese's (1959) idea of connectivity where, the stronger the backward association (or tendency for words in a study list to elicit the critical items on a free association test), the greater the false recall or recognition.

R*ecognition Performance Between-Groups*

For actually presented words, the finding that the intentional group performed better than the unintentional group is consistent with the idea that the former group expected a memory test and therefore would be making an effort to memorize words. Similarly, it makes intuitive sense that the control group, while not expecting a memory test, performed better than the unintentional group due to the fact that they were reading words aloud (and therefore directly processing items) rather than ignoring these words. The unintentional groups low veridical memory performance supports the notion that words were being processed unintentionally.

The finding that false alarms to critical non-presented words were greater in the intentional group compared to the other two groups also makes sense given that participants who intentionally look at and study words may notice that the items are semantically-related and in turn activate/evoke other related but non-presented words.

Although only a minor issue, one limitation of the present study was the cumbersome number of trials presented. Participants in each of the three conditions processed 240 words. Although only 120 of them were actual words of interest, participants appeared more fully engaged towards the beginning of the experiment where they verbally showed interest in the study. Towards the end, participants seemed fidgety and commented on the length of the words, which irritated their eyes. They therefore, may not have followed task instructions carefully towards the end.

One way to improve this for future research is to simply reduce the number of trials, perhaps using 200 words (100 color and 100 non-color words) instead of 240. This small reduction could lead to task improvements without compromising the need for more than 36 words. Another possibility to reduce eyesight irritation and therefore allow full engagement in the task would be to display a key for corresponding color keys directly in line with participants view, rather than have them look up and down from key board to screen, as was done in the unintentional condition.

In sum, the activation/monitoring theory (Roediger & McDermott, 2001) holds valid in the present study and provides a reasonable explanation of why false recognition occurs in the DRM paradigm, given the robust false memories that occurred in the absence of conscious elaborative processes.

False memories are a phenomena that illustrate the concern we should have about the authenticity of our memories, even those we believe with great certainty. Equally disturbing, is the notion that awareness of these memories does not eliminate them. The fact that our memory is made up information that we interpret based on our past experiences, limits are ability to process new information or words without relating them to that which we are familiar with. For this reason, it remains important to evaluate our memories more carefully and understand that although we may be convinced that we have experienced a particular event/concept/word we may in fact be incorrect. What remains to be evaluated is how we can trust our memories knowing that we do not have control over false albeit, related memories given that they can arise without our realization.

References

Anderson, M.C., Bjork, E.L., & Bjork, R.A. (2000). Retrieval-induced forgetting: Evidence for a recall-specific mechanism. *Psychonomic Bulletin & Review*, 7, 522-530

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

Bauml, K.H., & Kuhbandner, C. (2003). Retrieval-induced forgetting and part-list cuing in associatively structured lists. *Memory and Cognition*, 31(8), 1188-1197.

Brainerd, C.J., & Reyna, V.F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science*, 11, (5), 164-169.

Brown, D., Scheflin, A., & Hammond, D. (1998). *Memory, Trauma Treatment, and the Law*. W.W. Norton & Company: New York, London.

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

Dodd, Michael D. MacLeod, Colin M. (2004). False recognition without intentional learning. *Psychonomic Society, Inc*. 11(1), 137-142

Johnson, M.K., & Raye, C.L. (1981). Reality Monitoring. *Psychological Review*, 88, 67-85.

Kahneman, D., & Chajczyk, D. (1983). Tests of the Automaticity of Reading: Dilution of Stroop Effects by Color-Irrelevant Stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, 9(4), 497-509.

Koriat, A., Goldsmith, M., Pansky, A. (2000). Toward a Psychology of Memory Accuracy. *Annual Reviews in Psychology*, 51, 481-537

Loftus, E. (1975). Leading questions and the eyewitness report. *Cognitive Psychology*, 7, 560-72

Loftus, E., Miller, D.G., Burns, H.J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory,* 4,19-31.

McDermott, K.B. (1996). The Persistence of False Memories in List Recall. *Journal of Memory and Language,* 35, 212-230.

Reysen, M.B., & Nairne, J.S. (2002). Part-set cuing of false memories. *Psychonomic Bulletin & Review*, 9, 389-393.

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

Roediger, H.L., III, & McDermott, K.B. (2000). Tricks of Memory. *American Psychological Society*, 9(4), 123-127.

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

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.

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643-662.

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

Underwood, B. J. (1965). False recognition produced by implicit verbal responses. *Journal of Experimental Psychology*, 70, 122-129.

Author's Note

I would like to thank my advisor Anjali Thapar, for her assistance and support in doing a thesis that I found very challenging although extremely valuable. I would also like to thank my friends for their constant confidence and belief in my abilities and of course God for helping me feel proud of my efforts to complete the thesis.

Footnotes

[1] Fiske and Taylor, 1991 refer to schemas as “cognitive structures that represent knowledge about a type of stimulus including its attributes and the relations among those attributes”. This idea basically means that the accumulation of information that individuals acquire help them to organize and fill in information missing from memory in order to understand the world. In this manner, schemas may help to store, retrieve, and reconstruct memories. For example, when trying to retrieve a word that is very similar or associated to another word, memory may become distorted if schemas are used.

Table 1

Recognition Performance for Unintentional Condition using Uncorrected Scores

Word-type Pair	Mean	SD	t(df=29)
Random studied	.45	.22	
Studied	.47	.23	.83
Weakly-related	.28	.24	
Unrelated	.29	.13	.042
Studied	.47	.23	
Critical lure	.70	.28	4.51*
Random studied	.45	.22	
Critical lure	.70	.28	4.88*
Studied	.47	.22	
Weakly-related	.28	.24	4.18*
Studied	.47	.22	
Unrelated	.29	.13	5.20*
Random studied	.45	.22	
Weakly-related	.28	.24	3.43*
Random studied	.45	.22	
Unrelated	.29	.13	4.55*
Critical lure	.70	.28	
Weakly-related	.28	.24	7.19*
Critical lure	.70	.28	
Unrelated	.29	.13	7.72*

Note: *$p<.005$ (with new alpha level)

Table 2

Recognition Performance for Intentional Condition using Uncorrected Scores

Word-type Pair	Mean	SD	t(df=29)
Random studied	.74	.22	
Studied	.79	.20	1.19
Weakly-related	.17	.15	
Unrelated	.20	.089	1.13
Studied	.79	.20	
Critical lure	.87	.23	2.75
Random studied	.74	.22	
Critical lure	.87	.23	3.89*
Studied	.79	.20	
Weakly-related	.17	.15	14.94*
Studied	.79	.20	
Unrelated	.20	.089	14.82*
Random studied	.74	.22	
Weakly-related	.17	.15	13.75*
Random studied	.74	.22	
Unrelated	.20	.089	12.95*
Critical lure	.87	.23	
Weakly-related	.17	.15	16.093*
Critical lure	.87	.23	
Unrelated	.20	.089	14.47*

Note: *p<.005 (with new alpha level)

Table 3

Recognition Performance for Control Condition using Uncorrected Scores

Word-type Pair	Mean	SD	t(df=14)
Random studied	.63	.13	
Studied	.67	.25	.48
Weakly-related	.078	.12	
Unrelated	.26	.11	4.19*
Studied	.67	.25	
Critical lure	.62	.35	.77
Random studied	.63	.13	
Critical lure	.62	.35	.12
Studied	.67	.25	
Weakly-related	.078	.12	9.35*
Studied	.67	.25	
Unrelated	.26	.11	6.62*
Random studied	.63	.13	
Weakly-related	.078	.12	13.054*
Random studied	.63	.13	
Unrelated	.26	.11	8.44*
Critical lure	.62	.35	
Weakly-related	.078	.12	6.61*
Critical lure	.62	.35	
Unrelated	.26	.11	3.93*

Note: *p<.005 (with new alpha level)

Table 4

Recognition Performance for All Conditions using Corrected Scores

Word-type Pair	Mean	SD	t	d
Unintentional				
Random studied	.17	.21		
Studied	.19	.20	.83	29
Studied	.19	.20		
Critical lure	.42	.29	4.51*	29
Random studied	.17	.21		
Critical lure	.42	.29	4.88*	29
Intentional				
Random studied	.56	.22		
Studied	.60	.21	1.19	29
Studied	.60	.21		
Critical lure	.68	.23	2.75*	29
Random studied	.56	.22		
Critical lure	.68	.23	3.89*	29
Control				
Random studied	.47	.15		
Studied	.50	.23	.48	14
Studied	.50	.23		
Critical lure	.46	.33	.77	14
Random studied	.47	.15		
Critical lure	.46	.33	.12	14

Note: *p<.01 (with new alpha level)

Table 5

Between Groups Memory Performance

Word-type	Group	Mean	SD
Random Studied	Unintentional	.17	.21
	Intentional	.56	.22
	Control	.47	.15
Studied	Unintentional	.19	.20
	Intentional	.60	.21
	Control	.50	.23
Critical lure	Unintentional	.42	.29
	Intentional	.68	.23
	Control	.46	.33

Word-type	Group Pairs	Significance
Random Studied	Intentional-Unintentional	.000*
	Intentional-Control	.349
	Unintentional-Control	.000*
Studied	Intentional-Unintentional	.000*
	Intentional-Control	.298
	Unintentional-Control	.000*
Critical Lures	Intentional-Unintentional	.001*
	Intentional-Control	.034*
	Unintentional-Control	.893

*. The mean difference is significant at the .05 level

Appendix A

Non-Color Words used and Associated Critical Items

Item	**Word**	**Word**
Anger	mad	wrath
Army	rifle	draft
Beautiful	ugly	picture
Bitter	sweet	rice
Black	dark	funeral
Bread	butter	jelly
Butterfly	moth	summer
Cabbage	head	slaw
Car	truck	Ford
Carpet	rug	magic
Chair	table	stool
Citizen	alien	member
City	town	New York
Command	order	soldier
Cottage	house	ivy
Cup	mug	-
Doctor	nurse	-
Flag	banner	-
Foot	shoe	-
Fruit	apple	-
Girl	boy	-
Health	sickness	-
High	low	-
Justice	peace	-
King	queen	-
Lamp	light	-
Lion	tiger	-
Long	short	-
Man	woman	-
Mountain	hill	-
Music	note	-
Mutton	lamb	-
Needle	thread	-
Pen	pencil	-
River	stream	-
Rough	smooth	rugged
Rubber	elastic	latex
Shirt	blouse	button
Slow	fast	turtle
Smell	nose	sniff
Smoke	cigarette	tobacco

Soft	hard	loud
Spider	web	feelers
Stove	oven	kitchen
Sweet	sour	heart
Swift	water	stream
Thief	steal	rob
Trash	garbage	scraps
Trouble	bad	fear
Whiskey	drink	wine
Whistle	stop	dog
Wish	want	star

Appendix B

The Three DRM Critical Lures and Semantically Related Word-Lists

Sleep	**Cold**	**Window**
bed	hot	door
rest	snow	glass
awake	warm	pane
tired	winter	shade
dream	ice	ledge
wake	wet	sill
snooze	frigid	house
blanket	chilly	open
doze	heat	curtain
slumber	weather	frame
snore	freeze	view
nap	air	breeze

Appendix C

The Fifty-Four Recognition Items Used
(consisting of actually presented words and non-presented words)

rest	dream	blanket	snore	sleep	snow	ice
chilly	freeze	cold	glass	ledge	open	view
window	uniform	pilot	bee	stomach	evil	bar
sip	cure	mouth	sister	robust	elevate	sky
stripes	tell	bug	bulb	instrument	palace	marker
thin	hunt	iron	wash	army	doctor	slow
summer	dog	rifle	butter	heart	rug	order
banner	sickness	loud	turtle	rice		