City Research Online

City St George's, University of London

Citation: Dewhurst, S., Anderson, R. J. & Knott, L. (2012). A gender difference in false recall of negative words: Women DRM more than men. Cognition and Emotion, 26(1), pp. 65-74. doi: 10.1080/02699931.2011.553037

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/4767/

Link to published version: https://doi.org/10.1080/02699931.2011.553037

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

Running head: GENDER AND FALSE RECALL

A Gender Difference in the False Recall of Negative Words: Women DRM more than Men.

Stephen A Dewhurst[1], Rachel J. Anderson[1], and Lauren M. Knott[2]

[1] University of Hull

[2] Lancaster University

Author note

Stephen A. Dewhurst and Rachel J. Anderson, Department of Psychology, University of Hull, England. Lauren M. Knott, Department of Psychology, Lancaster University, England.

Address for correspondence: Dr Stephen A Dewhurst, Department of Psychology, University of Hull, Cottingham Road, Hull, HU6 7RX, England. Phone +44 1482 465931. Fax +44 1482 466511. Email: s.dewhurst@hull.ac.uk

Abstract

Gender differences in susceptibility to associative memory illusions in the Deese/Roediger-McDermott paradigm were investigated using negative and neutral word lists. Women (n = 50) and men (n = 50) studied 20 lists of 12 words that were associates of a nonpresented critical lure. Ten lists were associates of negatively-valenced lures (e.g., cry, evil) and ten were associates of neutral lures (e.g., chair, slow). When asked to recall the words after each list, women falsely recalled more negative lures than men, but there was no gender difference in the false recall of neutral lures. These findings suggest that women reflect on associations within negative lists to a greater degree than men and are thereby more likely to generate the negative critical lures.

Keywords: False memory; emotion; gender

A Gender Difference in the False Recall of Negative Words: Women DRM more than Men

There is evidence from a number of studies that women have better memory than men for emotional information. For example, Davis (1999) found that women recalled more emotional childhood memories than men and recalled them more rapidly than men. Similarly, both Fujita, Diener, and Sandvik (1991) and Seidlitz and Diener (1998) found that women recalled more positive and negative autobiographical memories than men in timed retrieval tasks. Seidlitz and Diener speculated that this gender difference might occur because women encode emotional events in greater detail than men. However, Bloise and Johnson (2007) showed that a gender difference also occurs when the emotional content of the information to be remembered is controlled. They presented male and female participants with scripts containing emotional and neutral information. Consistent with findings from studies of autobiographical memory retrieval, Bloise and Johnson found that women recalled more emotional information than men but did not differ from men in the recall of neutral information. These and other studies (See Bloise & Johnson for a review) indicate that women have better recall than men for emotional information. The question addressed in the current study is whether a gender difference also exists in the false recall of emotional information.

We investigated false recall using the Deese/Roediger-McDermott (DRM) procedure, named after studies by Deese (1959) and Roediger and McDermott (1995). In this procedure, participants study lists of words that are semantic associates of a nonpresented "critical lure". For example, participants study words such as *bed*, *dream*, *wake*, and *tired* which are associates of the critical lure *sleep*. In subsequent tests of recall or recognition memory, participants often claim to have studied the critical lures, with levels of false recall and recognition often equalling or exceeding levels of correct recall and recognition (see Gallo,

2006, for a review). According to the activation-monitoring account of associative memory illusions (Roediger, Watson, McDermott, & Gallo, 2001), participants spontaneously generate the critical lures at study, which are then subject to source monitoring errors at test and erroneously judged to have been studied.

Previous investigations of gender differences in susceptibility to associative memory illusions have found no difference between men and women. Seamon, Guerry, Marsh, and Tracy (2002) presented male and female participants with 16 DRM lists taken from Roediger and McDermott (1995) and found no gender difference in false recall. However, as noted by Bauste and Ferraro (2004), the lists they used were neutral with regard to gender. Bauste and Ferraro presented male and female participants with five DRM lists, including those associated with the critical lures *man* and *girl*, which they predicted might produce higher levels of false recall in men and women respectively. Contrary to their prediction, no gender differences in false memory were observed. Kreiner, Price, Gross, and Appleby (2004) presented DRM lists spoken in a male or female voice to male and female participants. They also found no significant effect of gender, even when the gender of the speaker matched that of the participants. More recently, Smeets, Jelicic, and Merckelbach (2006) investigated the effects of acute stress on false memories in men and women using the Trier Social Stress Test (Kirschbaum, Pirke, & Hellhammer, 1993). They found no effects of stress or gender in false recall or false recognition. These findings suggest that gender does not influence susceptibility to associative memory illusions, even when lists are used that might have particular salience for one gender. However, given the findings discussed above that gender influences the correct recall of emotional information, it is possible that a gender difference in false recall will emerge if emotional DRM lists are presented.

A number of previous studies have extended the DRM procedure to emotional lists. For

example, Pesta, Murphy, and Sanders (2001) found lower levels of false recall for emotional relative to neutral lures and attributed this to the greater distinctiveness of emotional lists. Similar reductions in false memory with emotional lists were reported by Kensinger and Corkin (2004) and by Palmer and Dodson (2009). In contrast, Budson, Todman, Chong, Adams, Kensinger, Krangel, and Wright (2006) found no difference between emotional and neutral DRM lists in elderly adults. More recently, El Sharkawy, Groth, Vetter, Beraldi, and Fast (2008) found higher levels of false recognition with negatively-valenced lists than with neutral lists but no difference in levels of false recall. Brainerd, Stein, Silveira, Rohenkohl, and Reyna (2008) extended previous research by manipulating the valence of the emotional lists. Consistent with El Sharkawy et al., they found that negatively-valenced lists led to higher levels of false recognition than neutral lists. In contrast, positively-valenced lists led to lower levels of false recognition relative to neutral lists. More recently, Howe, Candel, Otgaar, Malone, and Wimmer (2010) found that negative DRM lists produced higher levels of false recognition relative to neutral lists but lower levels of false recall, indicating that the effect of emotionality on false memory depends on how memory is tested.

Although there are discrepancies regarding the direction of the effect of emotion on the DRM illusion, the studies discussed above show that associative memory illusions can be observed with emotional lists, particularly if they are negatively-valenced. The aim of the current study was to investigate whether susceptibility to negative false memories is influenced by the gender of the participants. Howe et al. (2010) reported no gender difference in levels of false recall or false recognition (though gender was not the focus of their study) and collapsed their data across gender. However, their groups included fewer than 50 participants of each gender, which Seamon et al. (2002) suggested was the minimum number necessary for a gender comparison.

In order to investigate whether gender influences the false recall of negative DRM lists, we presented 50 female and 50 male participants with lists of words that were associates of either negatively-valenced critical lures or neutral critical lures. Participants studied ten lists of each type and were given tests of free recall after each list. We chose to investigate recall rather than recognition in order to be consistent with the previous investigations of gender differences discussed above, which found no gender differences in false recall using neutral lists. It has also been suggested that recall is a more sensitive measure than recognition in terms of understanding the processes that underlie the DRM illusion (see, for example, Hege & Dodson, 2004). An additional reason is that Bloise and Johnson (2007) found a gender difference for emotional information in correct recall but not in correct recognition. Based on previous findings that women focus on emotional information to a greater degree than men (e.g., Bauer, Stennes, & Haight, 2003; Bloise & Johnson, 2007; Hess, Senecal, Kirouac, Herrera, Philippot, & Kleck, 2000), particularly negative emotional information (e.g., Bradley, Codispoti, Sabatinelli, & Lang, 2001; Krohne & Hock, 2008; Sethi & Nolen-Hoeksema, 1997), our prediction was that women would falsely recall more negative lures than men.

Method

Participants

One hundred undergraduate students (50 women and 50 men in the age range 18-25) from Lancaster University took part in the study. They were tested in quiet research labs in groups of between two and six. All were native English speakers.

Stimuli

Ten emotionally neutral DRM lists of 12 words each were taken from Stadler, Roediger, and McDermott (1999) and consisted of associates of the following critical lures:

sleep, chair, foot, high, rough, king, fruit, sweet, mountain, slow. Ten negatively-valenced lists of 12 items each were constructed for the purposes of the study and consisted of associates of the following critical lures: *sick, lie, anger, fear, evil, cry, pain, hate, alone, danger*. The negative lists were constructed by choosing negatively-valenced words that had at least 12 associates listed on the University of South Florida free associations norms (Nelson, McEvoy, & Schreiber, 1998). The negative and neutral lists were matched for backwards associative strength (BAS) according to the Nelson et al. norms. Mean BAS values were .19 for both the negative and neutral lists. Half the participants studied the lists in the order shown above with negative and neutral lists alternating. This order was reversed for the remaining participants. Each participant was given a response booklet containing the list number and 12 lined spaces in which they were to recall the items.

Mean valence and arousal ratings for the list items and critical lures were taken from ANEW (Bradley & Lang, 1999) and included ratings for all subjects plus separate ratings for male and female subjects. The means for all negative and all neutral study items are shown in Table 1, along with ratings for concreteness, frequency, and word length. Concreteness and frequency ratings were taken from the MRC Psycholinguistic Database (Coltheart, 1981). The full set of stimuli is shown in the Appendix.

Independent samples t-tests showed that negative list items had significantly lower (i.e., more negative) ratings of valence and significantly higher ratings of arousal than neutral list items, both in terms of overall ratings and separate ratings for males and females, all $p<.001$. For the negative list items, male and female ratings of valence did not differ significantly from another, $t(104)=1.14$, $p=.26$, but female ratings of arousal were significantly higher than male ratings, $t(104)=2.08$, $p<.05$. For neutral list items, there were no significant differences between male and female ratings for valence, $t(104)=.26$, $p=.79$, or arousal, $t(104)=.08$,

p=.94. There were no significant differences between negative and neutral list items in terms of frequency, *t*(215)=.55, *p*=.58, but emotional list items had significantly lower ratings of concreteness, *t*(144)=10.44, *p*<.001 and higher ratings of word length, *t*(238)=2.89, *p*<.005, than neutral list items.

For the critical lures, independent samples t-tests showed that emotional lures had significantly lower ratings of valence and significantly higher ratings of arousal than neutral lures, both in terms of overall ratings and separate ratings for each gender, all *p*<.05. The female ratings for the negative lures were significantly more negative than the male ratings, *t*(16)=5.05, *p*<001, but there was no reliable difference between male and female ratings of arousal, *t*(16)=.88, *p*=.39. For the neutral lists, there were no reliable differences between the male and female ratings for valence, *t*(12)=.03, *p*=.98, or arousal, *t*(12)=.28, *p*=.79. Critical lures of negative and neutral lists did not differ in terms of frequency, *t*(18)=1.05, *p*=.31, or word length, *t*(18)=1.47, *p*=.16, but emotional critical lures had significantly lower ratings of concreteness than neutral critical lures, *t*(17)=3.36, *p*<.005.

Procedure

The 20 lists were presented on Apple Macintosh computers using custom-written software. Each list was preceded by the instruction *List 1, List 2*, etc., which was shown for 2 seconds, after which the 12 associates appeared one at a time for 1 second each, separated by a 1 second interval. At the end of each list a message instructed the participants to attempt to recall the list they had just studied. The recall test was self-paced. Participants were instructed to press the space bar when they could recall no further items and the next list was then displayed.

At the end of the testing session, participants were asked to complete the Rumination Scale (Scott & McIntosh, 1999) and the 12-item Emotional and Interpersonal Sensitivity

Scale (Bloise & Johnson, 2007). The only significant gender difference to emerge from these questionnaires was that the male participants scored higher than the female participants on the motivation subscale of the Rumination Scale. Given the largely nonsignificant results, we do not discuss these findings further.

Results

Table 2 shows the mean proportions of correct and false recall of negative and neutral items by men and women. Mean correct and false recall scores were analyzed in separate 2 (gender) x 2 (list type) mixed ANOVAs. The analysis of correct recall showed a significant main effect of gender, whereby women recalled more words than men, $F(1,98 = 12.79$, $MSE=.02$, $p<.005$, $\eta_p^2=.12$. A main effect of list type also showed that participants recalled more neutral than negative words, $F(1,98)=361.04$, $MSE=.002$, $p<.001$, $\eta_p^2=.79$. The interaction between gender and list type was not statistically significant, $F(1,98)=1.76$, $MSE=.02$, $p=.19$, $\eta_p^2=.02$.

The analysis of false recall produced nonsignificant main effects of gender, $F(1,98)=2.09$, $MSE=0.05$, $p=.15$, $\eta_p^2=.04$, and list type, $F(1,98)=1.74$, $MSE=.02$, $p=.19$, $\eta_p^2=.02$. There was, however, a significant interaction between gender and list type, $F(1,98)=7.27$, $MSE=.02$, $p<.01$, $\eta_p^2=.07$. Planned comparisons showed that women falsely recalled more negative critical lures than men[1], $t(98)=2.39$, $p<.05$, but there was no significant gender difference in the false recall of neutral critical lures, $t(98)=.06$, $p=.95$. Women also falsely recalled more negative lures than neutral lures, $t(49)=2.57$, $p<.05$, but levels of false recall of negative versus neutral lures did not differ significantly for men, $t(49)=1.11$, $p=.27$. Unrelated intrusions were low and not significantly affected by gender or list type.

We also calculated recall accuracy by dividing correct recall scores by correct plus

false recall scores (see Table 2). Note that the proportions are high due to the fact that each list had 12 studied items but only one critical lure. A 2 (gender) x 2 (list type) ANOVA showed that accuracy was significantly greater for neutral relative to negative lists, $F(1,98)=8.68$, $MSE=.00$, $p<.01$, $\eta_p^2=.08$. The main effect of gender was not significant, $F<1$, but there was a significant gender x list type interaction, $F(1,98)=5.97$, $MSE=.00$, $p<.01$, $\eta_p^2=.06$. Planned comparisons showed that, with a one-tailed hypothesis, men were more accurate than women in recalling the negative lists, $t(98)=1.78$, $p=.08$, but there was no reliable gender difference in recall accuracy for the neutral lists, $t(98)=.72$, $p=.48$. Women were more accurate in their recall of neutral lists relative to negative lists, $t(49)=3.61$, $p<.05$, but accuracy for negative versus neutral lists did not differ significantly for men, $t(49)=.38$, $p=.71$.

Discussion

The present study investigated gender differences in associative memory illusions by presenting lists of associates of negative and neutral critical lures to male and female participants. Consistent with previous research (Bauste & Ferraro, 2004; Kreiner et al., 2004; Seamon et al., 2002; Smeets et al., 2006) there was no gender difference in the false recall of neutral critical lures. In contrast, a significant gender difference was found in the false recall of negative critical lures, whereby women falsely recalled more negative lures than men. This gender difference was present both in overall levels of false recall and in recall accuracy. Previous studies have shown that women typically recall more emotional information than men (Bloise & Johnson, 2007; Davis, 1999; Fujita et al., 1991; Seidlitz & Diener, 1998). The present findings show that, under some conditions, women's memory for negative emotional information may be less accurate than that of men.

Bloise and Johnson (2007) suggested that women are more likely than men to discover

and reflect on connections within the study material, thereby forming richer associative connections that serve as retrieval cues in subsequent recall tests. Although this account was presented to explain why they found a gender difference in recall but not in recognition memory, it can be extended to explain the gender difference in false recall found in the current study. According to the activation-monitoring account of associative memory illusions (Roediger et al., 2001), such errors occur when participants generate associates of the words presented in the study lists. The present findings suggest that women reflect on associations within negatively-valenced lists to a greater degree than men and are thereby more likely to generate the negative critical lures.

As described in the *Stimuli* section, the negative lists were of lower valence and higher arousal than the neutral lists, both in terms of overall ratings and in terms of the separate male and female ratings provided in the ANEW database (Bradley & Lang, 1999). There were, however, some differences between the male and female ratings. Specifically, females rated the negative critical lures as significantly more arousing, and the negative study items as significantly more negative, than did males. This pattern is consistent with research showing that women experience greater affective intensity than men (e.g., Fujita et al., 1991) and rate emotional stimuli as more emotional than do men (e.g., Ferree & Cahill, 2009). There were no gender differences in the ratings of the neutral items. However, we do not believe that these differences compromise our findings. Indeed, the fact that females tend to give more extreme ratings for negative items than do males is consistent with the suggestion above that women reflect on the negatively-valenced lists more so than men. Crucially, as noted above, the negative study items and critical lures are rated as significantly more negative and more arousing than their neutral counterparts by both males and females in the ANEW database.

The fact that the negative lists differed from the neutral lists in terms of both valence

and arousal raises the question as to which of these factors is responsible for the gender difference in false recall. Brainerd et al. (2008) found higher levels of false recognition for negative lists relative to neutral and positive lists when the targets and critical lures were matched for arousal, and concluded that negative valence increases the familiarity of the critical lures, relative to neutral lures. Although Brainerd et al. gave their participants tests of recognition memory rather than recall, their findings suggest that valence is the crucial factor underlying the effect of emotion on false memory. However, Brainerd et al. did not separate the data from male and female participants, therefore the possibility remains that arousal may have contributed to the effect of emotion in the female participants. The data from the current study do not allow us to determine whether the gender difference we observed was due to valence or arousal. Given that the female ratings for the negative items in the ANEW database were significantly more negative and more arousing than the male ratings, it is possible that both factors contributed to the observed gender difference.

Although the negative and neutral items in the current study were matched for mean BAS according to the Nelson et al. (1998) word association norms, there were differences in terms of other variables such as concreteness and word frequency. This is due largely to the difficulty of finding sufficient numbers of emotional words in the Nelson et al. norms, which makes it difficult, if not impossible, to match emotional and neutral lists for other variables (see Palmer & Dodson, 2009, for a similar dilemma). Again, however, we do not believe that our findings are compromised by differences in these variables. Roediger et al. (2001) reported a multiple regression analysis of the stimulus factors that might affect false recall, including the concreteness and log frequency of the critical lures. The only significant predictors of false recall were BAS (the higher the BAS the higher the false recall rate) and correct recall (the more list items correctly recalled the lower the false recall rate). Neither

the concreteness nor the log frequency of the critical lures accounted for any unique variance in the level of false recall. The current findings are, therefore, unlikely to be due to differences in terms of these variables. In addition, to the best of our knowledge there is no evidence that any of these variables interact with gender, whereas there is abundant evidence for an interaction between gender and emotion (see Bloise & Johsnon, 2007, for a review).

In contrast to the effects observed in false recall, the analysis of correct recall showed significant main effects of gender and list type, but these factors did not interact. The main effect of gender showed that women recalled more studied items than men, but, unlike the difference observed in false recall, the advantage occurred with both negative and neutral lists. This is consistent with previous findings that women often outperform men on tests of episodic memory for verbal material (e.g., Kramer et al., 1988), though effect sizes are typically small (Herlitz, Nilsson, & Bäckman, 1997; see Herlitz & Rehnman, 2008, for a review of gender differences in episodic memory). The main effect of list type showed that correct recall was higher for neutral than for negative words. This is consistent with a number of previous studies showing that levels of correct recall are lower for emotional than for neutral words when testing occurs immediately after study (e.g., Kleinsmith & Kaplan, 1964; Parkin, Lewinsohn, & Folkard, 1982). This finding is also consistent with the more recent study by Howe et al. (2010) which found reduced recall of negative relative to neutral words with DRM lists.

If, as suggested above, women are more likely than men to reflect on associations within emotional lists, one might expect the female participants to show higher levels of correct recall for negative list items as well as higher levels of false recall for negative critical lures. One possible explanation for the pattern we observed is that reflecting on the negative lists led women to encode the "emotional gist" to a greater degree than men (see Brainerd et

al., 2008). According to fuzzy trace theory (FTT, e.g., Reyna & Brainerd, 1998) participants encode a gist trace that reflects the general semantic properties of the studied items. In subsequent memory tests, the critical lures are falsely recalled because of their overlap with the gist trace. Brainerd et al. suggested that participants also form an emotional gist trace when studying lists of words that share the same emotional theme. If, as stated by FTT, gist traces underlie false recall but not correct recall, which is supported more by verbatim traces of the studied items, then reflecting on the gist of emotional lists will lead to a selective increase in false but not correct recall.

Some previous studies have shown that negative lists produce lower levels of false recall than neutral lists (Howe et al., 2010; Palmer & Dodson, 2009; but see El Sharkawy et al., 2008, who reported no difference in false recall between negative and neutral lists). In contrast, the present study found that negative lists produced higher levels of false recall than neutral lists, at least with female participants. There are a number of differences that make it difficult to compare the current findings with those of the previous studies mentioned above, the most telling of which is the use of different emotional lists. For example, Palmer and Dodson (2009) created their own positive and negative lists via a word norming study. Other studies (e.g., Budsen et al., 2006) have included emotional words that do not appear in the Nelson et al. (1998) word association norms. Given the importance of BAS as the main predictor of false recall, we felt it important to construct our lists entirely from the Nelson et al. norms rather than borrow lists from previous studies. It is clear, however, that a number of issues remain to be resolved before a more complete understanding of the effect of emotion on the DRM illusion is achieved. Previous research has produced inconsistent findings in terms of the presence and direction of an emotionality effect, and comparison between studies is hindered by their use of different lists and different methods of testing memory.

The findings from the current study suggest that the gender of the participants must also be taken into account when assessing the effect of emotion on false memory.

Notwithstanding the discrepancies with previous studies, the current findings are consistent with the suggestion by Bloise and Johnson (2007) that women reflect on associations within emotional lists to a greater degree then men. Although these associative processes may enable women to outperform men in the correct recall of emotional information, they are also the very processes that give rise to associative memory illusions, thereby rendering women's recall of emotional words potentially less accurate than that of men.

[1]Footnote

A reviewer of an earlier draft of this article suggested that the gender difference for negative lists might be influenced by changes in mood across the study phase. In order to investigate this, we compared men's and women's false recall of critical lures from the first half of the lists with those from the second half. There was no sign of false recall increasing over time with either list type, as levels of false recall for the second half of the lists were either identical to or slightly lower than those for the first half. The same reviewer also suggested that correct recall may have been influenced by gender differences in serial position effects. In order to test this, we partitioned the study items from each list into three blocks of four (items from positions 1-4, 5-8, and 9-12) and conducted statistical analyses to determine whether serial position interacted with gender or list type. There was a significant interaction between list type and serial position, which was due to a higher primacy effect for the negative lists, but this was the case for both male and female participants. There were no significant interactions involving gender.

References

Bauer, P. J., Stennes, L., & Haight, J.C. (2003). Representation of the inner self in autobiography: Women's and men's use of internal states language in personal narratives. *Memory, 11,* 27-42.

Bauste, G., & Ferraro, F.R. (2004). Gender differences in false memory production. *Current Psychology, 23*, 238-244.

Bloise, S.M., & Johnson, M.K. (2007). Memory for emotional and neutral information: Gender and individual differences in emotional sensitivity. *Memory, 15*, 192-204.

Bradley, M.M, Codispoti, M, Sabatinelli, D, & Lang, P.J. (2001) Emotion and motivation II: sex differences in picture processing. *Emotion, 1*, 300–319

Bradley, M.M., & Lang, P.J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings*. Technical Report C-1, The Center for Research in Psychophysiology, University of Florida.

Brainerd, C. J., Stein, L. M., Silveira, R. A., Rohenkohl, G., & Reyna, V. F. (2008). How does negative emotion cause false memories? *Psychological Science, 19*, 919-925.

Budson, A. E., Todman, R. W., Chong, H., Adams, E. H., Kensinger, E. A., Krangel, T. S., & Wright, C. I. (2006). False recognition of emotional word lists in aging and Alzheimer disease. *Cognitive and Behavioral Neurology, 19*, 71-78.

Coltheart, M. (1981). The MRC Psycholinguistic Database, *Quarterly Journal of Experimental Psychology, 33A*, 497-505.

Davis, P. J. (1999). Gender differences in autobiographical memory for childhood emotional experiences. *Journal of Personality & Social Psychology, 76*, 498-510.

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

El Sharkawy, J., Groth, K., Vetter, C., Beraldi, A, & Fast, K. (2008). False memories of emotional and neutral words. *Behavioural Neurology, 19,* 7-11.

Ferree, N. and Cahill, L. (2009). Post-event spontaneous intrusive recollections and strength of memory for emotional events in men and women *Consciousness and Cognition, 18*, 126-134.

Fujita, F., Diener, E., & Sandvik, E. (1991). Gender differences in negative affect and well-being: The case for emotional intensity. *Journal of Personality and Social Psychology, 61*, 427-434.

Gallo, D.A. (2006). *Associative illusions of memory*. New York: Psychology Press.

Hege, A. C. G., & Dodson, C. S. (2004). Why distinctive information reduces false memories: Evidence for both impoverished relational encoding and distinctiveness heuristic accounts. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 30*, 787-795.

Herlitz, A., Nilsson, L.-G., & Bäckman, L. (1997). Gender differences in episodic memory. *Memory & Cognition, 25*, 801–811.

Herlitz, A., & Rehnman, J. (2008). Sex differences in episodic memory. *Current Directions in Psychological Science*, *17*, 52-56.

Hess, U., Senecal, S., Kirouac, G., Herrera, P., Philippot, P., & Kleck, R. E. (2000). Emotional expressivity in men and women: Stereotypes and self-perceptions. *Cognition & Emotion, 14*, 609-642.

Howe, M.L., Candel, I., Otgaar, H., Malone, C., and Wimmer , M.C. (2010). Valence and the development of immediate and long-term false memory illusions. *Memory, 18,* 58-75.

Hunt, R. R., & Einstein, G. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning & Verbal Behavior, 20*, 497–514.

Kensinger, E. A., & Corkin, S. (2004). The effects of emotional content and aging on false memories. *Cognitive, Affective, & Behavioural Neuroscience, 4*, 1-9.

Kirschbaum, C., Pirke, K.-M., Hellhammer, D.H., 1993. The 'trier social stress test' — a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology 28*, 76–81.

Kleinsmith, L.J., & Kaplan, S. (1964). Interaction of arousal and recall in nonsense syllable paired-associate learning. *Journal of Experimental Psychology, 67*, 124-126.

Kreiner, D.S., Price, R.Z., Gross, A.M., & Appleby, K.L. (2004). False recall does not increase when words are presented in a gender-congruent voice. *Journal of Articles in Support of the Null Hypothesis, 3*, 1-18.

Krohne, H.W., & Hock, M. (2008). Cognitive avoidance, positive affect, and gender as predictors of the processing of aversive information. J*ournal of Research in Personality, 42*, 1572-1584.

McCabe, D.P., Presmanes, A.G., Robertson, C.L., & Smith, A.D. (2004). Item-specific processing reduces false memories. *Psychonomic Bulletin & Review, 11*, 1074-1079.

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

Palmer, J.E., & Dodson, C.S. (2009). Investigating the mechanisms fuelling false recall of emotional material. *Cognition & Emotion, 23*, 238-259.

Parkin, A.J., Lewinsohn, J., & Folkard, S. (1982). The influence of emotion on immediate and delayed retention: Levinger and Clark reconsidered. *British Journal of Psychology, 73*, 389-393.

Pesta, B. J., Murphy, M. D., & Sanders, R. E. (2001). Are emotionally charged lures immune

to false memory? *Journal of Experimental Psychology: Learning,Memory, & Cognition, 27,* 328-338.

Reyna, V.F., & Brainerd, C.J. (1998). Fuzzy-trace theory and false memory: New frontiers. *Journal of Experimental Child Psychology, 71*, 194-209.

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

Seamon, J.G., Guerry, J. D., Marsh, G. P., & Tracy, M. C. (2002). Accurate and false recall in the Deese/Roediger and McDermott procedure: A methodological note on sex of participant. *Psychological Reports, 91*, 423-427.

Scott, V.B., Jr, & McIntosh, W.D. (1999). The development of a trait measure of ruminative thought. *Personality & Individual Differences, 26*, 1045-1056.

Seidlitz, L., & Diener, E. (1998). Sex differences in the recall of affective experiences. *Journal of Personality and Social Psychology, 74*, 262-271.

Sethi, S., & Nolen-Hoeksema, S. (1997). Gender differences in internal and external focusing among adolescents. *Sex Roles, 37*, 687-700.

Smeets, T., Jelicic, M., & Merckelbach, H. (2006). Stress-induced cortisol responses sex differences, and false recollections in the DRM paradigm. *Biological Psychology, 72*, 164-172.

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

Appendix

Emotional word lists

SICK	LIE	ANGER	FEAR	EVIL
ill	fib	rage	terror	wicked
flu	deceive	hostility	doubt	devil
nausea	untrue	fury	panic	villain
virus	bluff	frustration	fright	demon
hospital	truth	temper	anxiety	satan
fever	rumour	tantrum	afraid	witch
disease	deception	violent	scared	corrupt
healthy	dishonest	fight	monster	sin
vomit	betray	annoy	horror	taboo
germ	deny	argument	coward	cruel
well	cheat	irritate	scream	good
cough	false	conflict	hide	hell

CRY	PAIN	HATE	ALONE	DANGER
weep	agony	despise	isolated	hazard
sob	hurt	dislike	solo	risk
tears	discomfort	love	secluded	beware
emotional	headache	revenge	lonely	caution
upset	pleasure	prejudice	single	warning
sorrow	sore	enemy	private	safe
sensitive	relief	disgust	individual	harm
grief	dentist	war	withdrawn	help
sad	sharp	kill	bored	threat
misery	soothe	insult	empty	trouble
laugh	cut	condemn	together	jeopardy
tissue	cramp	shun	without	deadly

Neutral words lists

SLEEP	CHAIR	FOOT	HIGH	ROUGH
bed	table	shoe	low	smooth
rest	sit	hand	clouds	bumpy
awake	legs	toe	up	road
tired	seat	kick	tall	tough
dream	couch	sandals	tower	sandpaper
mattress	desk	walk	jump	jagged
blanket	recliner	ankle	above	ready
doze	sofa	arm	building	coarse
slumber	cushion	boot	cliff	uneven
snore	stool	sock	sky	gravel
pillow	sitting	knee	over	rugged

quiet	bench	mouth	elevate	crude
KING	FRUIT	SWEET	MOUNTAIN	SLOW
queen	apple	sour	hill	fast
castle	vegetable	candy	valley	lethargic
crown	orange	sugar	climb	stop
prince	kiwi	bitter	summit	snail
palace	citrus	taste	top	cautious
throne	bowl	tooth	peak	delay
chess	pear	honey	glacier	traffic
rule	banana	soda	bike	hesitant
monarch	basket	chocolate	climber	speed
royal	orchard	cake	range	tortoise
emperor	strawberry	tart	steep	crawl
lord	grape	pie	ski	motion

Table 1. Mean ratings of valence, arousal, concreteness, frequency, and word length for list items and critical lures.

	List items		Critical lures	
	Negative	Neutral	Negative	Neutral
Valence (all)	3.45	5.78	2.51	5.69
Valence (female)	3.27	5.80	1.95	5.70
Valence (male)	3.70	5.73	3.68	5.68
Arousal (all)	5.69	4.31	6.31	4.13
Arousal (female)	5.90	4.33	6.51	4.05
Arousal (male)	5.42	4.31	6.01	4.25
Concreteness	363.17	527.86	369.67	501.40
Frequency	63.10	76.92	80.00	135.10
Word length	5.87	5.26	4.20	4.90

Table 2. Mean proportions (with standard errors) of correct recall, false recall, and recall accuracy as a function of list type (emotional versus neutral) and gender.

	Correct recall		False recall		Accuracy (proportion correct)	
	Emotional	Neutral	Emotional	Neutral	Emotional	Neutral
Women	.58 (.01)	.70 (.01)	.27 (.03)	.20 (.03)	.96 (.01)	.98 (.01)
Men	.52 (.01)	.63 (.02)	.18 (.02)	.21 (.02)	.97 (.01)	.97 (.01)