Valence and the Development of Immediate and Long-Term False Memory Illusions

Mark L. Howe[1], Ingrid Candel[2], Henry Otgaar[2], Catherine Malone[1], and Marina C. Wimmer[1]
[1]Lancaster University
[2]Maastricht University

Address correspondence to: Professor Mark L. Howe
Centre for Research in Human Development
Department of Psychology
Lancaster University
Lancaster UK LA1 4YF
Phone: +44 1524 594336
e-mail: mark.howe@lancaster.ac.uk

IN PRESS: *Memory*
Do not quote without authors' permission

Abstract

Across five experiments we examined the role of valence in children's and adults' true and false memories. Using the Deese/Roediger-McDermott paradigm and either neutral or negative-emotional lists, both adults' (Experiment 1) and children's (Experiment 2) true recall and recognition was better for neutral than negative items and although false recall was also higher for neutral items, false recognition was higher for negative items. The last three experiments examined adults' (Experiment 3) and children's (Experiments 4 and 5) one-week long-term recognition of neutral and negative-emotional information. The results replicated the immediate recall and recognition findings from the first two experiments. More important, these experiments showed that although true recognition decreased over the one-week interval, false recognition of neutral items remained unchanged whereas false recognition of negative-emotional items increased. These findings are discussed in terms of theories of emotion and memory as well as their forensic implications.

Keywords: Valence and memory; False memory development; DRM paradigm

Valence and the Development of Immediate and Long-Term False Memory Illusions

Our memory is a reconstructive process in which errors often occur (for an overview see Koriat, Goldsmith, & Pansky, 2000). That means that we might falsely remember details that we in fact did not encounter. In recent years, the development of false memories has been extensively studied with the Deese/Roediger-McDermott (DRM) paradigm (Deese, 1959; Roediger & McDermott, 1995). Using this paradigm, participants are presented with word lists that consist of close semantic associates of a non-presented *critical lure*. For example, a list might consist of the words *door, glass, pane, shade, ledge, sill, house, open, curtain, frame, view, breeze, sash, screen*, and *shutter* which are all related to the non-presented word *window*. Studies have shown that participants have the tendency to falsely recall and recognize the critical lures during subsequent memory tests (for an overview see Gallo, 2006). Another robust finding obtained with the DRM paradigm is a reverse age effect. Specifically, studies have shown that the development of spontaneous false memories increases as a function of age (e.g., Brainerd, Reyna, & Brandse, 1995; Howe, 2005, 2006, 2008; Sugrue & Hayne, 2006). The low levels of spontaneous false memories found in young children may be due to a failure to automatically access associative relations between list items and the critical lure (associative-activation theory, or AAT, e.g., Howe, 2005, 2008; Howe, Gagnon, & Thouas, 2008; Howe, Wimmer, & Blease, 2009a; Howe, Wimmer, Gagnon, & Plumpton, 2009b) or to extract gist memories across words on DRM lists (fuzzy-trace theory, or FTT, e.g., Brainerd & Reyna, 2005).

Despite the robust findings produced by the DRM paradigm, researchers have questioned the ecological validity of this paradigm (Freyd & Gleaves, 1996). One concern raised by Freyd and Gleaves is that in real world scenarios false memories often involve emotional events. To address this concern, researchers have started to create negative-emotional DRM tasks to study the role of valence in the development of false memories. Although only a handful of such studies exist, findings from this line of research are important not only for their theoretical contribution but also because of their forensic implications. The purpose of this article is to establish the nature of valence effects in children's and adults' spontaneous true and false memories using the DRM paradigm. In the first part, we provide an overview of what is known about the role of valence in memory and then present two experiments in which true and false memory (recall and recognition) for negative-emotional lists is examined. In the second part, we provide an overview of what is known about the role of valence in the long-term retention of information and then present three experiments that examine the role of negative-emotional information in the persistence of true and false memory. To anticipate the outcome of this exercise, both children and adults are susceptible to spontaneous false memory illusions for negative-emotional information and these effects increase over time. Specifically, at least for recognition measures, negative-emotional information often produces false memories on immediate tests of retention and, importantly, these false memory illusions become more robust over time. That is, whereas true recognition rates decline and false recognition of neutral critical lures does not change with time, false recognition rates for negative-emotional information increase over a one-week retention interval. The theoretical and forensic consequences of these new and unique findings will be discussed in detail at the end of this article.

Part I: False Memory Illusions and Valence

Evidence concerning the facilitative role of valence and arousal in memory is mixed. It is commonly thought that emotional stimuli lead to enhanced true memory performance (e.g., Christianson, 1992). Indeed, some researchers have recently found that emotional information, when presented in well-organized and coherent lists, do result in higher levels of correct recall and recognition (e.g., Talmi & Moscovitch, 2004). However, this may be more the exception rather than the rule. As it turns out, this emotional memory advantage may be restricted to conditions in which items are presented pictorially (e.g., Talmi, Luk, McGarry, & Moscovitch, 2007; Talmi, Schimmack, Paterson, & Moscovitch, 2007) or are presented in mixed-list formats (e.g., Dewhurst & Parry, 2000; Hadley & MacKay, 2006). In fact, despite a number of studies showing an emotional memory advantage, the majority of research has shown quite the opposite. For example, there is considerable evidence indicating that it takes longer to generate associates to emotionally-arousing words (Levinger & Clark, 1961) and that emotional information is not better remembered than neutral information on immediate tests using recall measures either for adults (for a review, see Eysenck, 1977) or children (Howe, 2007). Indeed, it has been known for some time that immediate memory tests (those occurring 2 to 20 minutes following study) for negative information (e.g., words such as *rape*, *vomit*) that is arousing (as measured by galvanic skin response) is not as good as memory for low arousal, neutral information (e.g., words such as *dance*, *swim*) (Kleinsmith & Kaplan, 1963). However, following a one-week retention interval, this effect of valence/arousal is reversed. That is, high arousal negative items are better remembered than low arousal neutral items after a delay (for a review, see Eysenck, 1977).

Despite the fact that such findings are consistent with recent neurobiological evidence concerning the role of emotion in consolidation and recollection (e.g., LaBar & Cabeza, 2006; Sharot, Verfaellie, & Yonelinas, 2007), these results are difficult to replicate (for a recent review, see Mather, 2007). Indeed, when the various confounds are eliminated in studies such as these (e.g., confounding item type with serial order), there is either no effect of negative emotion on memory or the opposite effect. That is, negative emotional items are more likely to be forgotten than neutral items. Unfortunately, then, this research does not allow us to draw any firm conclusions about emotion and memory.

Interestingly, because the effects of emotion on memory are stronger and more reliable in mixed-list than pure-list designs, some have argued that these effects are similar to those found for distinctiveness (e.g., Schmidt, 2006; Schmidt & Saari, 2007). That is, instead of assuming that emotional stimuli affect attention (Christianson, 1992; Easterbrook, 1959) or encoding (e.g., Hadley & MacKay, 2006), emotional materials, being rare or unusual, are distinctive and are subject to item-specific processing in memory. For example, arousal associated with valence may lead to increased binding of information in memory, something that enhances within-object or item-specific memory (e.g., see Mather, 2007). Because of this, binding should improve recollection for the item itself but not necessarily the episodic context in which it occurred. To the extent that the effects of emotion and arousal on binding do not include between-item associations, arousal may lead to better item-specific memories. However, arousal can have a deleterious effect on memory because it does not lead to more associations between items or between items and their context. To the extent that between-item associations are reduced, false recollection rates may be similarly attenuated with emotional material. However, because DRM procedures are, by definition, pure-list manipulations, it is not clear whether these effects will extend to spontaneous false memories using this popular task.

Regardless, to the extent that emotional information promotes item-specific, distinctive processing, such material should be less susceptible to spontaneous false memory illusions. Specifically, there is solid evidence that distinctive information leads to reduced false memory rates in both children (e.g., Howe, 2008) and adults (e.g., Ghetti, Qin, & Goodman, 2002) using the DRM paradigm. In fact, false memory studies in which valence has been manipulated have revealed that participants are less likely to falsely recall and recognize negative-emotional critical lures as compared with neutral critical lures (e.g., Huang & Yeh, 2006; Kensinger & Corkin, 2004; Pesta, Murphy, & Sanders, 2001; but see Budson et al., 2006). Consistent with what we know about distinctiveness effects, the general explanation given for this finding concerns the distinctiveness of the emotional critical lures (e.g., Kensinger & Corkin, 2004). However, in all of these studies except for one (Budson et al., 2006), the list items that were presented were neutral despite their being related to either a neutral or a negative-emotional critical lure. Negative-emotional critical lures are distinct from the list items being presented and this by itself may have reduced false memory rates for the emotional lures.

Interestingly, Budson et al. (2006) found no differences in rates of false recognition for emotional over neutral lists. This raises the possibility that there is no advantage for negative-emotional items in false recollection either. In a recent study, Howe (2007) did find higher false recognition (but not recall) rates for negative-emotional items than neutral items in a DRM task with children. Unlike Budson et al. (2006), Howe (2007) included measures of recall and, for recognition tests, included emotional control items (items that were not presented but were related to items on the emotional lists). This critical control was missing from the Budson et al. (2006; a critical limitation of this study, as the authors pointed out on p. 73) study and may be one of the reasons that their results differed from those of Howe (2007). (Of course, other differences between these studies may have contributed to the different outcomes including the use of a recall measure by Howe and the fact that Howe's participants were 8- and 12-year-old children and Budson et al.'s were younger and older adults.)

In order to clarify the effects of negative valence, both in children's and adults' false memory illusions, we first present two experiments that compare neutral and negative-emotional DRM lists. In the first experiment, we examine these effects in adults when both recall and recognition tests are used and when recognition tests also contain foils that are emotional. Following this experiment, we present a second experiment in which the effects of emotion are examined in children's true and false recall and recognition.

Experiment 1: Valence and False Memory Illusions in Adults

The purpose to this first experiment was to obtain baseline evidence concerning adults' false memory illusions when the studied material was negatively, emotionally charged. We were interested in studying these effects using both recall and recognition measures. As well, for recognition, non-presented emotional foils were added to the recognition lists. We did this in order to measure any general tendency to false alarm to emotional items that were consistent with the lists presented originally.

Method

Participants

The 40 participants were undergraduate students (20 males and 20 females) who participated for course credit.

Materials and Procedure

There were 12 DRM lists, 6 negative-emotional and 6 neutral. The critical lures for the emotional lists (randomly sampled from the lists used by Budson et al., 2006) were *rape, danger,*

anger, thief, hungry, and *hell* and for the neutral lists (randomly sampled from both Budson et al., 2006, and Stadler, Roediger, & McDermott, 1999) were *chair fruit, sweet, sleep, lion,* and *needle*. Unlike Budson et al.'s lists, which consisted of only 8 items, we added 4 items to each list making each list 12 items long. We did this in part to have lists of optimal length and to have enough words so we could include emotionally-charged lures on the recognition tests. The added items were drawn either from Stadler et al. (1999) or by sampling related terms from the Nelson, McEvoy, and Schreiber (1999) norms. These additional items all exhibited rates of backward associative strength to the critical lure that were similar to those of exiting list members. Ten items from each list were presented to participants, with the 11th and 12th items not being presented and instead, were held in reserve for use as related distractors on recognition tests. The negative-emotional and neutral lists were equated not only on word length, but also on mean backward associative strength ($t[10] = 1.72$, $p = .12$) and word frequency ($t[10] = 1.35$, $p = .20$) (Leech, Rayson, & Wilson, 2001).

The 12 lists were recorded on tape and played to the participants. Items were presented at a 3-second rate. Half of the participants were presented with the emotional lists first followed by the neutral lists and the other half were given the neutral lists first followed by the emotional lists.[1] Within each list type (emotional, neutral), lists were presented randomly one at a time. Following the presentation of a list, participants immediately performed a distractor task (circling random letter pairs) for 30 seconds, and were then instructed to write down as many of the words from the list as they could remember. Participants had a maximum of 3 minutes for recall and then the next study-distractor-recall cycle began. This procedure was repeated until all 12 lists had been presented.

Immediately after the 12th list was complete, participants were given a recognition test. The list of 72 items was presented via a tape recorder and participants simply had to mark on their sheet (by ticking beside "yes" or "no") whether they recognized the words from the lists or not. The 72-item recognition test consisted of 36 targets (3 randomly sampled from each of the 12 lists), all 12 critical lures, 12 words that were semantically related to the presented lists (6 of which were neutral and 6 of which were emotional; either the 11th or 12th unpresented list word) but were not presented, and 12 words which were unrelated and not presented (6 of them were neutral and 6 of them were emotional words). Several versions of this recognition test were created and each test was constructed so that the words appeared in random order.[2]

Results[3]

Recall

The proportion of true and false recall scores were analyzed using separate within-subject analyses of variance (ANOVAs). For true recall, there was a main effect for valence, $F(1, 39) = 36.99$, $p < .001$, ($_p^2 = .487$, where neutral lists ($M = .64$, $SD = .09$) were better recalled than negative-emotional lists ($M = .55$, $SD = .11$). The ANOVA for false recall also produced a main effect for valence, $F(1, 39) = 11.29$, $p < .01$, ($_p^2 = .225$, where false recall was higher for neutral lists ($M = .40$, $SD = .10$) than for negative-emotional lists ($M = .27$, $SD = .08$).

Recognition[3]

Like recall, the proportion of true and false recognition scores were analyzed using separate ANOVAs. The results for true recognition indicated no significant effects for valence where true recognition of negative-emotional items ($M = .83$, $SD = .15$) was no different than that observed for neutral items ($M = .82$, $SD = .11$). However, for false recognition, there was a main effect for valence, $F(1, 39) = 15.38$, $p < .001$, ($_p^2 = .283$, where negative-emotional items were more frequently falsely recognized ($M = .67$, $SD = .13$) than neutral items ($M = .51$, $SD = .12$).

Interestingly, when examining false alarm rates to related and unrelated foils on the recognition test, although no valence effects were observed for unrelated foils, significantly higher false alarm rates were found for related negative-emotional foils (M = .16, SD = .08) than for related neutral foils (M = .10, SD = .05), $F(1, 39) = 6.04$, $p < .02$, ($_p^2$ = .134.

Discussion

The findings from this experiment are similar to those obtained by Howe (2007) with children but are not consistent with those of Budson et al.'s (2006) study with adults. That is, like Howe (2007), but not Budson et al.'s (2006) findings, adults in this study were more likely to false alarm to negative-emotional critical lures than neutral ones during the recognition test. Moreover, when semantically related distractors were presented for both neutral and negative-emotional lists, adults were more likely to false alarm to these distractors if they were related to negative-emotional information than to neutral information. Thus, despite being more likely to correctly and falsely recall neutral information, false recognition rates are higher for negative-emotional information than neutral information.

Of particular interest was the finding that adults were more likely to false alarm to weakly related, but not presented, negative-emotional items than neutral items. This may be attributable to the density of negative-emotional lists and that for adults, whose associative networks are well-developed for negative-emotional information, false alarms occurred not just for strongly activated items (the critical lures) but also for the weaker, related distractors. Recall that measures of backward associative strength were equivalent across negative-emotional and neutral lists meaning that the advantage of negative-emotional information in eliciting false memory illusions must be due to the greater intralist richness of associations (Talmi et al., 2007). It may be that the effects of emotionality extend to even weakly related items when recognition is used to measure the role of emotion in false memory illusions. Thus, consistent with the idea that negative-emotional information is more densely integrated and more richly interconnected (e.g., Talmi et al., 2007), it would seem that once such information becomes activated in memory there exists an increased likelihood that most, if not all, related negative-emotional information will be primed and become available for false recollection regardless of backward (or forward) associative strength. Before examining this idea further, we present data from a similar experiment with children.

Experiment 2: Valence and False Memory Illusions in Children

Next we examined the effects of negative valence in children's true and false recall and recognition. In particular, we wanted to see whether the effects obtained by Howe (2007) were generally replicable and whether these effects paralleled those for adults. Recall that like the adults in Experiment 1, Howe (2007) found that 8- and 12-year-olds recalled more neutral words than negative-emotional words after studying neutral and negative-emotional DRM lists. In addition, and again like the adults in Experiment 1, children were more likely to falsely recall neutral critical lures than emotional critical lures. When recognition measures were used, children recognized neutral items more often than negative-emotional items and false recognition was higher for negative-emotional than for neutral critical lures.

Method

Participants

Sixty (30 males, 30 females) children, 30 7-year-olds (M_{age} = 7 years, 5 months; SD = 3 months) and 30 11-year-olds (M_{age} = 11 years, 4 months; SD = 2 months), participated in this experiment. The children, predominantly White and middle class, participated with parental consent and themselves assented to the procedure on the day of testing.

Materials and Procedure

True and false recall and recognition were tested using a 2(age: 7- vs. 11-year-olds) x 2(valence: neutral vs. negative-emotional) mixed design, where age was between-subjects and valence was within-subject (as in the first experiment). For Experiments 2 through 4, lists were sampled from the pool of DRM lists provided in the Appendix. The critical lures for the emotional lists (randomly sampled from the lists used by Budson et al., 2006) were *cry, lie, anger, thief, alone,* and *hell* and for the neutral lists (randomly sampled from both Budson et al., 2006, and Stadler et al., 1999) were *sweet, lion, fruit, sleep, chair,* and *needle*. Again, as in the previous experiment, we added additional items to each lists to achieve a length of 12 items, 10 of which were presented and the remaining 2 used as distractors or foils in the recognition tests. The added items were drawn either from Stadler et al. (1999) or by sampling related terms from the Nelson et al. (1999) norms (see Appendix). The negative-emotional and neutral lists were equated not only on word length, but also on mean backward associative strength and word frequency (*t*s(10) (1.30, *p*s > .10).

For Experiment 2, 8 DRM lists were drawn from this pool, four negative-emotional (*cry, lie, anger*, and *thief*) and four neutral (*chair, sweet, lion*, and *sleep*). Like Experiment 1, only the first 10 words of each list were recorded on tape and presented to the children, with the critical lure and the 11^{th} and 12^{th} words not presented but used on recognition tests. The lists were chosen in part to once again control for levels of backward associative strength across valence and in part to insure their appropriateness of use with children (equated on word length and frequency).

Each child was tested individually in a quiet room in their school. The lists were presented on audiotape at a 3-second rate per item. Following presentation of the list, there was a 30-second distractor task (circling pairs of random letters), followed by oral recall (due to age differences in writing speed and skill, children were not required to write their responses). This study-distractor-recall procedure continued until all eight lists had been presented and recalled. For half of the children, the negative lists were presented first followed by the neutral lists, and for the remaining half the neutral words were presented first followed by the negative lists (again, no differences were observed as a consequence of blocking valence – see Footnote 1).

When all the lists had been presented and recalled the child was given a recognition test. The recognition test consisted of a total of 48 words: 3 (randomly selected) words from each presented list, all 8 critical lures, 8 related but not presented list words (4 negative-emotional and 4 neutral; either the 11^{th} or 12^{th} non-presented list word), and 8 unrelated but not presented words (4 negative-emotional and 4 neutral). The child was instructed to listen to each word on the tape and report whether it had been on the lists read out earlier (saying “yes”) or not (saying “no”).

Results

Recall

The proportion of true and false recall was analyzed separately using 2(age: 7- vs. 11-year-olds) x 2(valence: neutral vs. negative-emotional) mixed ANOVAs. For true recall, there was a main effect for age, $F(1, 58) = 30.53$, $p < .001$, ($_p^2$ = .345, where 7-year-olds (*M* = .32, *SD* = .14) recalled fewer items than 11-year-olds (*M* = .48, *SD* = .15), and a main effect for valence, $F(1, 58) = 86.85$, $p < .001$, ($_p^2$ = .600, where neutral lists (*M* = .46, *SD* = .16) were better recalled than negative-emotional lists (*M* = .33, *SD* = .12). Similar effects were observed for false recall. Specifically, there was a main effect for age, $F(1, 58) = 5.88$, $p < .02$, (^{2} = .092, where 7-year-olds (*M* = .23, *SD* = .10) falsely recalled fewer items than 11-year-olds (*M* = .33, *SD* = .13), and a main effect for valence, $F(1, 58) = 11.01$, $p < .01$, (${}_p^2 = .159$, where neutral lists ($M = .34$, $SD = .16$) produced more false recall than negative-emotional lists ($M = .23$, $SD = .14$).

Recognition[5]

Like recall, the proportion of true and false recognition scores were analyzed using separate ANOVAs. For true recognition, there was a main effect age, $F(1, 58) = 19.49$, $p < .001$, (${}_p^2 = .251$, where 7-year-olds ($M = .70$, $SD = .20$) recognized fewer items than 11-year-olds ($M = .83$, $SD = .23$), and a main effect for valence, $F(1, 58) = 49.93$, $p < .001$, (${}_p^2 = .463$, where neutral lists ($M = .85$, $SD = .26$) were better recognized than negative-emotional lists ($M = .68$, $SD = .22$). For false recognition, there was a main effect for age, $F(1, 58) = 8.25$, $p < .01$, (${}_p^2 = .125$, where 7-year-olds ($M = .58$, $SD = .23$) falsely recognized fewer items than 11-year-olds ($M = .70$, $SD = .17$), and a main effect for valence, $F(1, 58) = 7.85$, $p < .01$, ($^2 = .119$, where negative-emotional lists ($M = .68$, $SD = .22$) produced more false recognition responses than neutral lists ($M = .59$, $SD = 18$). Finally, analyses of false alarm rates to related and unrelated foils on the recognition test produced no reliable differences for children, regardless of age.

Discussion

The results of Experiment 2 are consistent with those of Experiment 1 with adults as well as previous research with children (Howe, 2007). As is usual in DRM experiments, younger children recalled and recognized fewer true and false items and this held regardless of the emotional nature of that information (negative or neutral). More important, children, regardless of age, were more likely to falsely recognize negative-emotional critical lures than neutral ones, despite being more likely to correctly recall and recognize, and falsely recall, neutral items. Unlike adults, children's false recognition of weakly related items did not differ as a function of valence. Apparently, regardless of any presumed differences in semantic density across lists, children's processing did not extend to these weaker associates. This is consistent with AAT's assumptions concerning age differences in the structure and relative automaticity of associative activation in children's and adults' knowledge base. However, before considering possible interpretations of this developmental difference, we present Experiments 3 through 5.

Part II: The Persistence of False Memory Illusions and Valence

The first two experiments give a different picture of true and false recollection of negative-emotional information than that provided by Budson et al. (2006) but one that is consistent with previous research with children (Howe, 2007). Together, these findings suggest that negatively valenced information is not better remembered than neutral information. Indeed, neutral information was better recalled (both for adults and children) and recognized (for children, but no difference for adults) than negative-emotional information. However, regardless of age, children (critical lures only) and adults (critical lures and weakly related items) were more likely to falsely recognize negatively valenced than neutral information.

Perhaps these different trends for true and false recollection may be the result of using immediate testing conditions only. Recall that earlier research has produced mixed results for negatively valenced material – sometimes it was better recalled later than earlier and sometimes the reverse was true (Eysenck, 1977; Mather, 2007). Recently, LaBar and Cabeza (2006) have indicated that memory for emotional information is better than memory for neutral information when tested days rather than minutes following study. That is, consistent with recent neurobiological studies that show increases in recollection of emotional material over time (e.g., Sharot et al., 2007), perhaps children's and adults' true memory for emotional items will exceed that for neutral items only after a delay. Although overall rates of true memory decline with time, perhaps the rate of this decline is attenuated for emotional material.

What about false recollection rates over time? Although valence has not been manipulated in studies of false memory rates over a delay, there are several studies that have compared true and false recollection rates over time for neutral material. For example, McDermott (1996) found that over a two-day delay, false recall rates exceeded true recall. That false memories persist over time while true memories decline is a quite robust phenomenon, having been obtained for both recall and recognition measures (e.g., Seamon et al., 2002). More generally, false memories are more persistent than true memories, showing lower rates of decline between immediate and delayed tests, than correct recollection (McDermott, 1996; Seamon et al., 2002; Thapar & McDermott, 2001; Toglia, Neuschatz, & Goodwin, 1999). These same false memory persistence effects have also been obtained with children using neutral material (Brainerd et al., 1995).

Whether negatively valenced information exhibits similar false memory persistence effects is unknown. However, if true memories for emotional information increase over a delay interval (or at least the rate of decline is less than for neutral information) then it might also be the case that false memory rates change across retention. This is because if activation levels increase over time (or remain relatively constant) for emotional stimuli, whereas such activation declines for neutral stimuli, there is a greater likelihood that false recollection of negatively valenced information should increase over that same retention interval given increased associative activation. Thus, for measures of false recollection, neutral information should persist in memory whereas negative-emotional information might increase.

Experiment 3: Valence and False Memory Persistence in Adults

In Experiment 3, we examined the role of valence in adults' true and false recollection over a one-week retention interval. To our knowledge, this is the first time emotionally-laden stimuli have been used over a significant delay interval to examine false memory rates. If activation of emotional material increases over time (e.g., due to some time-dependent neurobiological process such as consolidation – see LaBar & Cabeza, 2006; Sharot et al., 2007), then emotional false memories should be more evident over extended delays.

Method

Participants

The 60 participants were undergraduate students (30 males and 30 females) who participated for course credit.

Materials and Procedure

There were 12 DRM lists, 6 negative-emotional and 6 neutral (see Appendix). The critical lures for the emotional lists were *cry, lie, anger, thief, alone,* and *hell* and for the neutral lists were *sweet, lion, fruit, sleep, chair,* and *needle*. Again, as in the previous experiments, 10 of the 12 items were presented and the remaining 2 were used as distractors or foils in the recognition tests.

The 12 lists were recorded on tape and played to the participants at a rate of 3-seconds per item. Half of the participants were presented with the emotional lists first followed by the neutral lists and the other half were given the neutral lists first followed by the emotional lists. Within each list type (emotional, neutral), lists were presented randomly one at a time. Following the presentation of a list, participants immediately performed a distractor task (circling random letter pairs) for 30 seconds, and were then instructed to write down as many of the words from the list as they could remember. Participants had a maximum of 3 minutes for recall and then the next study-distractor-recall cycle began. This procedure was repeated until all 12 lists had been presented.

Immediately after the 12th list was complete, half of the participants (N = 30) were given a recognition test. The remaining half of the participants (N = 30) were given the recognition test one-week later. Regardless of when the recognition test was administered, the list of 72 items was presented via a tape recorder and participants simply had to mark on their sheet (by ticking beside “yes” or “no”) whether they recognized the words from the lists or not. As in Experiment 1, the 72-item recognition test consisted of 36 targets (3 randomly sampled from each of the 12 lists), all 12 critical lures, 12 words that were semantically related to the presented lists (6 of which were neutral and 6 of which were emotional) but were not presented, and 12 words which were unrelated and not presented (6 of them were neutral and 6 of them were emotional words). Several versions of this recognition test were created and each test was constructed so that the words appeared in random order.

Results

Recall

The proportion of true and false recall was analyzed separately using two within-subject ANOVAs. For true recall, there was a main effect for valence, $F(1, 58) = 15.14$, $p < .01$, ($_p^2$ = .231, where neutral lists ($M = .62$, $SD = .13$) were better recalled than negative-emotional lists ($M = .54$, $SD = .16$). Similar effects were observed for false recall. Specifically, there was a main effect for valence, $F(1, 58) = 11.01$, $p < .01$, ($_p^2$ = .159, where neutral lists ($M = .30$, $SD = .11$) produced more false recall than negative-emotional lists ($M = .20$, $SD = .10$).

Recognition

Recognition data were analyzed using a 2(item: true vs. false) x 2(valence: neutral vs. negative-emotional) x 2(time of test: immediate vs. 1-week) ANOVA where the first two variables were within-subject and the latter variable between-subjects. The results revealed a main effect for item, $F(1, 36) = 10.41$, $p < .01$, ($_p^2$ = .224, where true recognition ($M = .80$, $SD = .18$) exceeded false recognition ($M = .66$, $SD = .14$). Although like Experiment 1, immediate false recognition rates were higher for negative-emotional than neutral critical lures, there was an Item x Valence x Time of test interaction, $F(1, 36) = 9.87$, $p < .01$, ($_p^2$ = .163. As can be seen in Figure 1, and was confirmed in a series of post-hoc tests ($p < .05$), for neutral lists, true recognition rates fell and false recognition rates did not change significantly, across the retention interval (Figure 1, Panel A). However, for negative-emotional lists, although true recognition rates fell (and did so more dramatically than for neutral items), false recognition rates increased across the retention interval (Figure 1, Panel B). Thus, as predicted, although true recognition rates declined over time, false memories either persisted (neutral items) or increased (emotional items).

Discussion

The results of this experiment are generally consistent with the previous two experiments. That is, neutral information was better recalled than emotional information and the same effects were observed for false recall. Although there were marginal valence effects at immediate recognition (negative items better recognized than neutral items) as in the previous experiments, there were no valence differences in true recognition rates at one week and false recognition rates either remained the same (for neutral items), replicating previous research on false memory persistence effects, or increased (for emotional items) representing a hitherto unreported effect. These findings have important theoretical as well as forensic implications. That negative-emotional true memories were no better recognized than neutral ones and that both declined following a one-week retention interval, but false emotional memories increased over this same interval whereas neutral false memories remained unchanged, suggests that activation levels change differently for presented and unpresented information as a function of valence and time. Before discussing the implications of these effects, we present the results for these same manipulations with children.

Experiments 4 and 5: Valence and False Memory Persistence in Children

The aim of these final two experiments was to examine the effect of valence on children's long-term true and false memory. In both experiments, children were presented with neutral and negative-emotional DRM lists. All children were involved in an immediate free recall task and one recognition test, either immediately or one-week later. Experiment 4 tested young (5- and 8-year-olds) children's emotional recollection and Experiment 5 tested older (7- and 11-year-olds) children's emotional recollection.

Experiment 4: Method

Participants

Thirty (16 male, 14 female) children took part in the experiment, 15 5-year-olds (*M* = 5.67 years, *SD* = .5) and 15 8-year-olds (*M* = 7.54 years, *SD* = .45). All of the children were fluent speakers of English, were predominately White, middle-class and all participated following parental consent and their own assent on the day of testing.

Materials and Procedure

Four neutral (*chair, lion, sleep,* and *sweet*) and four negative-emotional (*anger, cry, lie,* and *thief*) word lists (see Appendix) were used in this experiment. Each list contained 12 items, the top 10 associates being used at study (with the remaining 2 used in the recognition tests – see below). Again, all of the words were matched across lists for length and frequency and were equated for mean backward associative strength (all *t*s < 1).

Each child was tested individually and was told that they were going to hear a series of words, which they should try to remember. The lists were presented on audiotape at a rate of one word every 3 seconds. For half of the children, the emotional lists were presented first and for the other half, the neutral lists were presented first. After each list of words was presented, the child was given a distractor task (circling random pairs of letters) for 30 seconds prior to recall. This study-distractor-recall procedure was repeated until all 8 lists had been presented.

After all of the lists had been presented and recalled by the children, half of the participants were given an immediate recognition test and the remaining children were given the recognition test one week later. The recognition test was audio recorded and children simply indicated "yes" if the word had been presented earlier and "no" if the word had not appeared earlier. The 48-item recognition test consisted of the 8 critical lures, 24 targets (3 randomly selected from each of the 8 lists), 8 semantically-related distractors (i.e., one unpresented item from each list that occurred in the 11th or 12th position), 4 unrelated neutral distractors, 4 unrelated negative-emotional distractors.

Results and Discussion

Recall

The proportion of true and false recall was analyzed separately using two 2(age: 5- vs. 8-year-olds) x 2(valence: neutral vs. negative-emotional) ANOVAs where the first factor was between-subjects and the last factor was within-subject. For true recall, there was a main effect for age, $F(1, 28) = 15.75$, $p < .001$, $(_p^2 = .335$, where older children correctly recalled more items (*M* = . 34, *SD* = .21) than younger children (*M* = .23, *SD* = .14), and a main effect for valence, *F*(1, 28) =21.95, *p* < .001, ($_p^2$ = .453, where neutral lists (*M* = .32, *SD* = .18) were better recalled than negative-emotional lists (*M* = .25, *SD* = .13). Similar effects were observed for false recall. Specifically, there was a main effect for age, *F*(1, 28) = 20.86, *p* < .001, ($_p^2$ = .425, where older children falsely recalled more critical lures (*M* = .28, *SD* = .10) than younger children (*M* = .12, *SD* = .09), and a main effect for valence, *F*(1, 28) = 28.85, *p* < .001, ($_p^2$ = .497, where neutral lists (*M* = .25, *SD* = .10) produced more false recall of critical lures than negative-emotional lists (*M* = .10, *SD* = .06).

Recognition

Recognition data was analyzed using a 2(age: 5- vs. 8-year-olds) x 2(time of test: immediate vs. 1-week) x 2(item: true vs. false) x 2(valence: neutral vs. negative-emotional) ANOVA where the first two variables were between-subjects and the last two variables were within-subject. The results revealed a main effect for age, *F*(1, 28) = 15.87, *p* < .001, ($_p^2$ = .324, where older children (*M* = .72, *SD* = .23) were better at recognition than younger children (*M* = .54, *SD* = .17). Although there were no differences as a function of valence for immediate false recognition rates, there was an Item x Valence x Time of test interaction, *F*(1, 28) = 14.13, *p* < .01, ($_p^2$ = .200. As can be seen in Figure 2, and was confirmed in a series of post-hoc tests (*p* < .05), for neutral lists, true recognition rates fell and false recognition rates did not change significantly, across the retention interval (Figure 2, Panel A). However, for negative-emotional lists, although true recognition rates fell (and unlike adults, did so at the same rate as true neutral information), false recognition rates increased across the retention interval (Figure 2, Panel B). These effects did not vary with age and, like the findings for adults in Experiment 3, show that although true recognition rates decline across time, false recognition rates remain stable (neutral information) or actually increase (emotional information) over a 1-week retention interval. Before elaborating on the theoretical and forensic implications of these findings, we describe the effects of these same manipulations with older children in Experiment 5.

Experiment 5: Method

Participants

This study involved 80 (38 females, 42 males) children from schools in The Netherlands, 40 7-year-olds (*M* = 7.47 years, *SD* = .57) and 40 11-year-olds (*M* = 11.72 years, *SD* = .58). Children participated after parents had given their written informed consent and each child's assent at the time of testing. Children were tested individually and were given a small present in return for their participation.

Materials and Procedure

For the neutral word lists, five critical lures (*bread, window, sweet, smoke,* and *foot*) were selected from Peters, Jelicic, and Merckelbach's (2007) norming study. For the emotional word lists, five critical lures (*murder, pain, punishment, death,* and *cry*) were chosen and lists generated. List items were selected from the Dutch word association norms (Van Loon-Vervoorn & Van Bekkum, 1991). All lists contained 12 items, but only the top 10 associates were presented at study (the others were used as distractors in the recognition tests – see below). All items were presented in order of associative strength, from strongest to weakest. Using the Celex lexical database (Centre for Lexical Information, 1995), it was established that the mean word frequency of neutral and emotional critical lures did not differ (*t*(8) = 0.22, n.s.). Likewise, the mean backward associative strength between the neutral list items and their critical lures and the mean backward associative strength between the emotional list items and their critical lures did not differ ($t(8) = 1.69$, n.s.).

Children were tested individually. The experiment consisted of two phases. In the first phase, children were exposed to each word list followed by a distractor task followed by recall. During the second phase, children engaged in a recognition task. Children were instructed to listen to the words carefully and to recall as many words as possible after each list. Lists were presented on an audio recording at a rate of one word every 3 seconds. List order was counterbalanced so that half of the children were presented with the neutral lists first whereas the other half listened to the emotional lists first. After each list and before free recall, children were involved in a 30 second distractor task. Following presentation of the 10th list, half of the children of each age group (N = 40) were administered a yes-no recognition task (immediate recognition group). The remaining half of the children in each age group (N = 40) completed the recognition task following a one-week delay (delayed recognition group). The 78-item recognition task consisted of the 10 critical lures, 40 studied items (i.e., 4 items from each list), 10 semantically-related distractors (i.e., one unpresented 11th or 12th word from each list), 6 unrelated neutral distractors (e.g., *castle*) taken from unstudied lists from Peters et al. (2007), 6 unrelated negative distractors (e.g., *enemy*), and 6 unrelated neutral distractors (e.g., *fun*). Recognition items were presented randomly.

Results and Discussion

Recall

The first step in the analysis was to examine children's proportions of true and false recall. True and false recall data were analyzed separately using two 2(Age: 7- vs. 11-year-olds) x 2 (Valence: neutral vs. negative-emotional) ANOVAs where the first factor was between-subjects and the latter factor within-subject. For true recall, the analyses revealed a significant main effect for age, $F(1, 78) = 119.09$, $p < .001$, ($_p^2 = .604$, where 7-year-olds ($M = .34$, $SD = .15$) recalled fewer items than 11-year-olds ($M = .54$, $SD = .22$), and a significant main effect for valence, $F(1, 78) = 19.73$, $p < .001$, ($_p^2 = .202$, where neutral words ($M = .46$, $SD = .16$) were better recalled than negative-emotional words ($M = .42$, $SD = .15$).

For false recall the analysis also showed significant main effects of age, $F(1, 78) = 28.80$, $p < .001$, ($_p^2 = .270$, where 7-year-olds ($M = .17$, $SD = .09$) falsely recalled fewer critical lures than 11-year-olds ($M = .35$, $SD = .15$), and list valence, $F(1, 78) = 11.67$, $p = .001$, ($_p^2 = .130$, where neutral lists elicited false recall of more critical lures ($M = .31$, $SD = .17$) than did negative-emotional lists ($M = .20$, $SD = .11$).

Recognition

Recognition data were analyzed using a 2(age: 7- vs. 11-year-olds) x 2(time of test: immediate vs. 1-week) x 2(item: true vs. false) x 2(valence: neutral vs. negative-emotional) ANOVA where the first two variables were between-subjects and the last two variables were within-subject. The results revealed a main effect for age, $F(1, 76) = 93.30$, $p < .001$, ($_p^2 = .551$, where younger, 7-year-old children correctly recognized fewer items ($M = .56$, $SD = .16$) than older, 11-year-old children ($M = .72$, $SD = .21$). The only other significant effect was an Item x Time of test interaction, $F(1, 76) = 7.76$, $p < .01$, ($_p^2 = .093$, where post-hoc tests ($p < .05$) showed that true recognition ($M = .69$, $SD = .17$) was higher than false recognition ($M = .62$, $SD = .15$) on immediate tests, but false recognition ($M = .69$, $SD = .24$) was higher than true recognition ($M = .60$, $SD = .14$) on 1-week tests. Although the 3-way Item x Valence x Time of test interaction was not significant in the overall ANOVA, Bonferroni tests ($p < .05$) did reveal significant differences in true and false recognition rates due to valence and time. As shown in Figure 3, true recognition rates did decrease over time for both neutral and negative-emotional items and, like Experiment 4, did so at similar rates (Panel A). For false recognition (Figure 3, Panel B), performance on neutral items remained the same across time whereas for negative-emotional items, false recognition of critical lures increased over the 1-week retention interval. Thus, like the children in Experiment 4 and the adults in Experiment 3, 7- and 11-year-olds' true memory declined with time but their false recollection rates either remained unchanged (neutral information) or increased (negative-emotional information) over a 1-week delay. The theoretical and forensic ramifications of these findings will be discussed next in the context of the other four experiments.

General Discussion

The aim of these experiments was to examine the role of valence in the development and retention of children's and adults' true and false memories. Several key findings emerged, some of which replicate developmental trends found previously for neutral materials and extend these findings to emotional material. More important, several new and hitherto unreported effects emerged, ones having to do with the effects of emotional material on long-term retention of true and false memories.

First, our findings replicate previous developmental DRM studies inasmuch as (a) younger children's true and false recollection was generally poorer than older children's, (b) all children correctly and falsely recalled more neutral items than negative-emotional items, and (c) all children tended to correctly recognize more neutral than negative-emotional items, but, on average, falsely recognized more negative-emotional than neutral items.

Second, the current findings extend these outcomes to adults. That is, adults correctly and falsely recalled more neutral than negative-emotional items, tended to correctly recognize more neutral than negative items, but falsely recognized more negative-emotional than neutral items. These findings are also consistent with a recent study in which negative items were more frequently falsely recognized than neutral or positive items even when arousal levels were controlled (Brainerd, Stein, Silveira, Rohenkohl, & Reyna, 2008). Together, these outcomes suggest that valence, independent of arousal, contributes to false memory production and that the role of valence in false memory may be developmentally invariant. Although adults in the current experiments tended to false alarm to both strong (critical lures) and weak associates in the recognition task whereas children false alarmed only to the critical lure, this difference more quantitative rather than qualitative.

This developmentally invariant pattern was also observed for our third and perhaps most important set of findings, namely, the presence of false memory persistence for neutral items but the unmistakable "growth" of false memory for negative-emotional information. That is, for both children and adults, false memories for neutral critical lures remained stable over a one-week retention interval whereas false memories for critical lures increased for negative-emotional material. This despite the fact that true recognition rates declined for both neutral and negative-emotional information. For children, these rates diminished at the same rate for both neutral and negative information whereas for adults, true recognition declined more rapidly for negative-emotional than neutral material.

These new findings have important implications for understanding the effects of negative-emotional information on children's and adults' true and false memories, immediately and in the long-term. In contrast to what would have been anticipated on the basis of the research literature concerning memory advantages for emotional stimuli (e.g., Kensinger, 2007; Talmi & Moscovitch, 2004; Talmi et al., 2007), both children and adults were better at recalling neutral than emotional items. However, consistent with that literature, both children and adults were more likely to falsely recall neutral critical lures than negative-emotional critical lures. The finding that fewer false memories were produced for negative-emotional information than for neutral information using recall measures also corresponds with previous work that has shown that negative emotion leads to fewer reconstructive-memory errors than positive emotion (e.g., Kensinger & Schacter, 2006; Levin & Bluck, 2004).

However, lower false recollection rates for negatively valenced information appears to be limited to recall measures. That is, although true recognition rates were frequently higher for neutral than negative-emotional information, unlike recall, false recognition tended to be higher for negative-emotional information and this effect increased over a one-week retention interval. This finding is difficult to reconcile with the idea that emotional material increases memory accuracy.

Of course, we are not the first to report differences in false recollection rates that depend on how memory is being measured, recall tests versus recognition tests (e.g., see Dewhurst, Barry, Swannell, Holmes, & Bathurst, 2007). Moreover, although some studies have found that emotion enhances correct recognition accuracy (e.g., Ochsner, 2000), a number of others found no advantage for emotional material (Comblain, D'Argembeau, Van der Linden, & Aldenhoff, 2004; Pesta et al., 2001; Windmann & Kutas, 2001), and still others have found lower recognition accuracy for negative than neutral words (Maratos, Allen, & Rugg, 2000). Our findings add to the growing consensus that when pure-list designs are used, negative-emotional material does not enjoy an advantage in true recall or recognition relative to neutral words. Indeed, neutral words are better recalled and recognized than negative-emotional ones.

Interestingly, these same trends hold for false recall – that is, false recall is more likely for neutral than negative items. However, the reverse was observed for false recognition – that is, false recognition was more likely for negative than neutral items. More important, although false recognition of neutral items shows the classic persistence effect across a one-week retention interval for both children and adults, false recognition for negative information actually increased over that same interval.

This pattern of findings might be consistent with predictions from fuzzy-trace theory (Brainerd & Reyna, 2005). Specifically, it could be argued that because valenced (either positive or negative) information produces relatively more gist than neutral information (see Rivers, Reyna, & Mills, 2008), true recall and recognition should be better for neutral than negative lists. This is because true memories are often based on verbatim information whereas false memories are almost always gist-based (see Brainerd & Reyna, 2005). As well, according to FTT, because gist processing may be more likely with negative items, such items would be expected to more readily invoke false memories than neutral items. This prediction is partially confirmed in our results. However, when there were more false memories for negative than neutral items, this outcome was restricted to measures of recognition not recall. Because gist processing is more likely on recall tests because they involve more reconstruction than recognition measures, FTT can only provide a partial explanation of the current findings. Moreover, because true and false recollective processes are dissociated in FTT, it is not clear how lower true and false recall for negative items, along with higher true and false recall for neutral items, can be reconciled within this model.

However, recently, Brainerd et al. (2008) showed that when it comes to negatively valenced information, meaning-based processing might play more of a role than previously expected even on recognition tests. Specifically, they found that participants were more likely to accept critical lures for negative than neutral or positive information and that these errors were based on familiarity (meaning) judgements. Not only might this explain our immediate recognition findings but might also provide an account of our long-term retention findings. Because verbatim traces decay more rapidly than gist traces, FTT anticipates that false memories for negative items should be more robust than those for neutral items. Given that FTT can accommodate the false memory persistence effect with neutral items, it may not be too difficult to see how it can also accommodate for the increase in false recognition rates over time. That is, gist-rich material (negative items) may also give rise to more stable gist, one that continues to foment over time. If so, then false recognition rates for negatively valenced information should increase with time. Thus, according to FTT, for negative information, it is the meaning content of critical lures that becomes increasingly familiar to participants while at the same time they are less able to use verbatim information to suppress false acceptance of critical lures. This may not only explain the differences we observed between recall and recognition measures in our experiments, but also account for the differences in "growth" rates over time for false recollections of negative but not neutral information.

Alternatively, associative-activation theory (e.g., Howe, 2006; Howe et al., 2009a, 2009b) can also provide an account of these findings. Specifically, although neutral and negative-emotional lists were equated on backward associative strength, as noted earlier, these lists may have differed in semantic density such that emotional lists contained more intralist cohesion than neutral lists (e.g., Talmi & Moscovitch, 2004). In fact, when measured across the lists used in the five experiments in this article, semantic density was twice as large, and significantly higher, in the emotional (M = .15) than neutral (M = .07) lists ($t[10] = 2.48$, $p < .05$). Higher levels of semantic density may confer greater confusability between what was and what was not presented because items are more interrelated, a phenomenon that was well captured in a network model depicting the influence of emotion on associative processing a number of decades ago (Bower, 1981). However, such confusion may be limited to measures of recognition where items are already generated for participants, but not on tasks (e.g., free recall) where participants must generate information themselves. Indeed, as already mentioned, it is well known that false memory production can depend critically on how memory is tested, particularly on whether one uses recall or recognition tests (e.g., Dewhurst et al., 2007). This would explain why such effects were restricted to recognition measures of false memory and not recall measures.

In addition, because semantically dense lists are more cohesive, they are more likely to maintain their stability in memory over time. Indeed, over time, negative-emotional lists might give rise to higher rates of false recognition because consolidation of emotional material is a more protracted process (e.g., Sharot et al., 2007) or because participants adopt a more lenient response criterion on recognition tasks, both initially as well as after a long-term retention interval. Indeed, semantically dense lists (e.g., ones containing emotional stimuli) do not enhance recognition accuracy, but rather, engender a more liberal response criterion (Dougal & Rotello, 2007), one that might lead to greater false recognition responses over time.

Conclusion

Regardless of which theory provides the best explanation, the current experiments have produced a series of consistent and hitherto unreported findings concerning children's and adults' processing of emotional material. Indeed, there were three key, developmentally invariant findings: (1) more neutral than negative-emotional information was correctly recalled and recognized, a result that is consistent with previous research contrasting the use of pure- rather than mixed-list designs; (2) more neutral than negative-emotional information was falsely recalled but more negative-emotional than neutral information was falsely recognized; and (3) regardless of valence, true recognition rates declined over time but false recognition rates either remained unchanged (neutral lists) or actually increased over time (negative-emotional lists). Findings 2 and 3 are consistent with the literature showing that emotional information may be semantically denser than neutral information and that this may promote greater longevity of emotional memories and a more liberal response bias on recognition tests.

There are at least two important take-home messages. The first is that whether emotional information is more likely to produce false memory illusions in children and adults seems to depend on whether one administers a recall or recognition test, a finding not unheard of in the false memory literature (e.g., Dewhurst et al., 2007) and one that replicates earlier findings concerning children's spontaneous false memory production (Howe, 2007). The second is that unlike false memory illusions for neutral material, memory illusions that tend to persist over a one-week retention interval, false memory illusions for negative-emotional information tend to increase over that same interval.

Both of these messages may have forensic implications. First, consistent with claims by some (e.g., Freyd & Gleaves, 1996), it is important to examine emotional information when developing a more complete theory of false memory illusions in children and adults. In fact, the current research has produced some findings that may be potentially worrisome, ones that have some rather daunting implications. For example, consider the findings that (a) true recollection of negative events is not always better than that for neutral events and (b) false memories increase for negative-emotional information over time. Taken together, the results suggest that memories for negative experiences may not always be well recollected and, in fact, may be more, not less, susceptible to increases in false memory production over time.

Of course, there are concerns with whether findings from research using word lists generalizes to more typical forensic situations in which entire events are being (mis)remembered (e.g., Pezdek & Lam, 2007). Moreover, to the extent that recollection of "real-life" autobiographical events is more frequently recall-based than recognition-based, the concerns expressed here about growth of false recollection over time may be moot. However, if autobiographical recollection is (a) cued, as it often is when looking at old photographs or talking to others about the past or (b) subsequent introspection about past events involves cuing or recognition-like prompts (e.g., answering specific yes/no questions about the remembered event), there may be some legitimacy to findings from word-list studies such as these that show greater false recollection over time when recognition-like memory prompts are used. Moreover, it should be pointed out that the current results correspond well with other recent findings in which entire events are implanted. For example, there is some evidence that valenced events are more readily implanted in children's memory than more mundane or neutral events (Otgaar, Candel, & Merckelbach, 2008). Thus, although procedures used to investigate *spontaneous* false memories differ from those used to examine *implanted* false memories, it does seem that the results are similar – that is, negative-emotional material, regardless of whether it consists of a word list (the current research; Brainerd et al., 2008; Howe, 2007) or an entire event narrative (e.g., Otgaar et al., 2008; for an overview, see Loftus, 2004), is more vulnerable to false memory illusions than neutral material. Indeed, it may be that the same underlying mechanisms are responsible for false memory illusions regardless of one's paradigm of preference (also see Wade et al., 2007).

Finally, it is important to note that not only might the same mechanism underlie the construction of spontaneous and implanted false memories, but this same mechanism may be responsible for false memory production regardless of age. Recall that despite age-related increases in both true and false memory rates, a finding typical in this area of research (see Brainerd & Reyna, 2005), increased susceptibility to false memory illusions in recognition but not recall of negative-emotional information was developmentally invariant. That is, regardless of age, children and adults were more likely to falsely recognize negatively valenced items than neutral ones and this effect increased over the long-term retention interval. That children and adults differ quantitatively but not qualitatively is important and argues that similar mechanisms are involved in false memory illusions throughout development.

References

Bower, G. H. (1981). Mood and memory. *American Psychologist, 36*, 129-148.

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

Brainerd, C. J., Reyna, V. F., & Brandse, E. (1995). Are children's false memories more persistent than their true memories? *Psychological Science, 6*, 359-364.

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.

Christianson, S.-A. (1992). Emotional stress and eyewitness memory: A critical review. *Psychological Bulletin*, *112*, 284-309.

Comblain, C., D'Argembeau, A., Van der Linden, M., & Aldenhoff, L. (2004). The effect of ageing on the recollection of emotional and neutral pictures. *Memory*, *12*, 673-684.

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., Barry, C., Swannell, E. R., Holmes, S. J., & Bathurst, G. L. (2007). The effect of divided attention on false memory depends on how memory is tested. *Memory & Cognition*, *35*, 660-667.

Dewhurst, S. A., & Parry, L. A. (2000). Emotionality, distinctiveness, and recollective experience. *European Journal of Cognitive Psychology*, *12*, 541-551.

Dougal, S., & Rotello, C. M. (2007). "Remembering" emotional words is based on response bias, not recollection. *Psychonomic Bulletin & Review*, *14*, 423-429.

Easterbrook, J. A. (1959). The effects of emotion on cue utilization and the organization of behavior. *Psychological Review, 66*, 183-201.

Eysenck, M. W. (1977). *Human memory: Theory, research, and individual differences*. Oxford: Pergamon.

Freyd, J. J., & Gleaves, D. H. (1996). "Remembering" words not presented in lists: Relevance to the current recovered/false memory controversy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 811-813.

Gallo, D. A. (2006). *Associative illusions of memory: False memory research in DRM and related tasks.* New York: Psychology Press.

Ghetti, S., Qin, J., & Goodman, G. S. (2002). False memories in children and adults: Age, distinctiveness, and subjective experience. *Developmental Psychology*, *38*, 705-718.

Hadley, C. B., & MacKay, D. G. (2006). Does emotion help or hinder immediate memory? Arousal versus priority-binding mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 79-88.

Howe, M. L. (2005). Children (but not adults) can inhibit false memories. *Psychological Science, 16*, 927-931.

Howe, M. L. (2006). Developmentally invariant dissociations in children's true and false memories: Not all relatedness is created equal. *Child Development, 77*, 1112-1123.

Howe, M. L. (2007). Children's emotional false memories. *Psychological Science*, *18*, 856-860.

Howe, M. L. (2008). Visual distinctiveness and the development of children's false memories. *Child Development*, *79*, 65-79.

Howe, M. L., Gagnon, N., & Thouas, L. (2008). Development of false memories in bilingual

children and adults. *Journal of Memory and Language*, *58*, 669-681.
Howe, M. L., Wimmer, M. C., & Blease, K. (2009a). The role of associative strength in children's false memory illusions. *Memory*, *17*, 8-16.
Howe, M. L., Wimmer, M. C., Gagnon, N., & Plumpton, S. (2009b). An associative-activation theory of children's and adults' memory illusions. *Journal of Memory and Language, 60*, 229-251.
Huang, Y. M., & Yeh, Y. Y. (2006). Memory errors on emotional lures: Is it possible to mistake a positive stimulus for a negative one? *Cognition and Emotion*, *20*, 646-670.
Kensinger, E. A. (2007). Negative emotion enhances memory accuracy: Behavioral and neuroimaging evidence. *Current Directions in Psychological Science, 16*, 213-218.
Kensinger, E. A., & Corkin, S. (2004). The effects of emotional content and aging on false memories. *Cognitive, Affective, & Behavioural Neuroscience, 4*, 1-9.
Kensinger, E. A., & Schacter, D. L. (2006). Reality monitoring and memory distortion: Effects of negative, arousing content. *Memory & Cognition*, *34*, 251-260.
Kleinsmith, L. J., & Kaplan, S. (1963). Paired-associate learning as a function of arousal and interpolated interval. *Journal of Experimental Psychology*, *65*, 190-193.
Koriat, A., Goldsmith, M., & Pansky, A. (2000). Toward a psychology of memory accuracy. *Annual Review of Psychology, 51*, 481-537.
LaBar, K. S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews: Neuroscience, 7*, 54-64.
Leech, G., Rayson, P., & Wilson, A. (2001). *Word frequencies in written and spoken English*. Harlow, England: Longman.
Levine, L. J., & Bluck, S. (2004). Painting with broad strokes: Happiness and the malleability of event memory. *Cognition & Emotion*, *18*, 559-574.
Levinger, G., & Clark, J. (1961). Emotional factors in the forgetting of word associations. *Journal of Abnormal and Social Psychology*, *62,* 99-105.
Loftus, E. F. (2004). Memories of things unseen. *Current Directions in Psychological Science*, *13*, 145-147.
Maratos, E. J., Allen, K., & Rugg, M. D. (2000). Recognition memory for emotionally negative and neutral words: An ERP study. *Neuropsychologia*, *38*, 1452-1465.
Marche, T. A., Brainerd, C. J., Lane, D. G., & Loehr, J. D. (2005). Item method directed forgetting diminishes false memory. *Memory*, *13*, 749-758.
Mather, M. (2007). Emotional arousal and memory binding: An object-based framework. *Perspectives on Psychological Science*, *2*, 33-52.
McDermott, K. B. (1996). The persistence of false memories in list recall. *Journal of Memory and Language, 35*, 2212-230.
Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1999). *The University of South Florida word association, rhyme, and word fragment norms*. Unpublished manuscript, University of South Florida, Tampa.
Ochsner, K. N. (2000). Are affective events richly recollected or simply familiar? The experience and process of recognizing feelings past. *Journal of Experimental Psychology: General, 129*, 242-261.
Otgaar, H., Candel, I., & Merckelbach, H. (2008). Children's false memories: Easier to elicit for a negative than for a neutral event. *Acta Psychologica*, *128*, 350-354.
Pastore, R. E., Crawley, E. J., Berens, M. S., & Skelly, M. A. (2003). "Nonparametric" A(and other modern misconceptions about signal detection theory. *Psychonomic Bulletin &*

Review, 10, 556-569.
Pesta, B. J., Murphy, M. D., & Sanders, R. E. (2001). Are emotionally charged lures immune to false memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 328-338.
Peters, M. J. V., Jelicic, M., & Merckelbach, H. (2007). Inducing false memories: A Dutch version of the Deese-Roediger-McDermott paradigm. *Manuscript submitted for publication*.
Pezdek, K., & Lam, S. (2007). What research paradigms have cognitive psychologists used to study “false memory” and what are the implications of their choices? *Consciousness and Cognition*, *16*, 2-17.
Rivers, S. E., Reyna, V. F., & Mills, B. (2008). Risk taking under the influence: A fuzzy-trace theory of emotion in adolescence. *Developmental Review*, *28*, 107-144.
Roediger, H. L., III, & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 803-814.
Schmidt, S. R. (2006). Emotion, significance, distinctiveness, and memory. In R. R. Hunt & J. Worthen (Eds.), *Distinctiveness and memory* (pp. 47-64). New York: Oxford University Press.
Schmidt, S. R., & Saari, B. (2007). The emotional memory effect: Differential processing or item distinctiveness. *Memory & Cognition*, *35*, 1905-1916.
Seamon, J. G., Luo, C. R., Kopecky, J. J., Price, C. A., Rothschild, L., Fung, N. S., & Schwartz, M. A. (2002). Are false memories more difficult to forget than accurate memories? The effect of retention interval on recall and recognition. *Memory & Cognition*, *30*, 1054-1064.
Sharot, T., Verfaellie, M., & Yonelinas, A. P. (2007). How emotional strengthens the recollective experience: A time-dependent hippocampal process. *PLoS ONE*, *2(10)*, e1068. Doi:10.1371.
Stadler, M. A., Roediger, H. L., III, & McDermott, K. B. (1999). Norms for word lists that create false memories. *Memory & Cognition*, *27*, 494-500.
Sugrue, K., & Hayne, H. (2006). False memories produced by children and adults in the DRM paradigm. *Applied Cognitive Psychology, 20*, 625-631.
Talmi, D., & Moscovitch, M. (2004). Can semantic relatedness explain the enhancement of memory for emotional words? *Memory & Cognition*, *32*, 742-751.
Talmi, D., Luk, B. T. C., McGarry, L. M., & Moscovitch, M. (2007). The contribution of relatedness and distinctiveness to emotionally-enhanced memory. *Journal of Memory and Language, 56*, 555-574.
Talmi, D., Schimmack, U., Paterson, T. , & Moscovitch, M. (2007). The role of attention and relatedness in emotionally enhanced memory. *Emotion*, *7*, 89-102.
Thapar, A., & McDermott, K. B. (2001). False recall and recognition induced by presentation of associated words: Effects of retention interval and level of processing. *Memory & Cognition*, *29*, 424-432.
Toglia, M. P., Neuschatz, J. S., & Goodwin, K. A. (1999). Recall accuracy and illusory memories: When more is less. *Memory, 7,* 233-256.
Van Loon-Vervoorn, W. A., & Bekkum van, I. J. (1991). *Woord associatie lexicon (Word association lexicon)*. Amsterdam: Swets & Zeitlinger.
Wade, K., Sharman, S., Garry, M., Memon, A., Mazzoni, G., Merckelbach, H., & Loftus, E. (2007). False claims about false memory research. *Conscousness and Cognition*, *16*, 18-

28.
Windmann, S., & Kutas, M. (2001). Electrophysiological correlates of emotion-induced recognition bias. *Journal of Cognitive Neuroscience*, *13,* 577-592.

APPENDIX

List Pool for English Words used in Experiments 2-4 (Critical lures in CAPs)

Negative-Emotional Lists

CRY	LIE	ANGER	HELL	THIEF	ALONE
Tears	Fib	Mad	Devil	Steal	Single
Sad	Cheat	Fear	Satan	Robber	Isolated
Tissue	Truth	Hate	Evil	Crook	Abandoned
Sorrow	False	Rage	Damned	Burglar	Solitary
Eyes	Mislead	Temper	Sin	Money	Apart
Weep	Trick	Fury	Lucifer	Cop	Lonesome
Sob	Fake	Wrath	Demon	Bad	Separate
Bawl	Sneak	Fight	Heaven	Rob	Quiet
Frown	Pretend	Hatred	Soul	Jail	Detached
Unhappy	Deceive	Mean	Judgment	Gun	Solo
Upset	Secret	Calm	Beast	Bank	Self
Down	Honest	Enrage	Fire	Bandit	Unaided

Neutral Lists

FRUIT	CHAIR	SWEET	LION	NEEDLE	SLEEP
Apple	Table	Sour	Tiger	Thread	Bed
Vegetable	Sit	Candy	Circus	Pin	Rest
Orange	Leg	Sugar	Jungle	Eye	Awake
Pear	Seat	Bitter	Tamer	Sewing	Tired
Banana	Couch	Good	Den	Sharp	Dream
Berry	Desk	Taste	Cub	Point	Wake
Cherry	Sofa	Tooth	Africa	Prick	Snooze
Basket	Cushion	Nice	Mane	Thimble	Blanket
Juice	Sitting	Honey	Cage	Haystack	Doze
Salad	Stool	Chocolate	Feline	Thorn	Slumber
Bowl	Bench	Pie	Hunt	Cloth	Yawn
Cocktail	Rocking	Heart	Pride	Knitting	Drowsy

Authors' Note

Mark L. Howe, Centre for Research in Human Development, Department of Psychology, Lancaster University, Lancaster, United Kingdom LA1 4YF. Ingrid Candel and Henry Otgaar, Faculty of Psychology, Maastricht University, The Netherlands. Catherine Malone and Marina C. Wimmer, Department of Psychology, Lancaster University, Lancaster, United Kingdom LA1 4YF.

This research was supported by a grant to MLH from the Economic and Social Research Council of Great Britain (RES-062-23-0452) and by a grant to IC from the Netherlands Organisation for Scientific Research (NWO 451-03-013).

Correspondence concerning this research should be addressed to Prof. Mark L. Howe, Centre for Research in Human Development, Lancaster University, Lancaster, UK LA1 4YF.

Footnotes

[1]No differences were found as a consequence of whether emotional lists appeared first or second. Because different orders were constructed for methodological not theoretical reasons, the data were collapsed across order for all of the analyses reported in this article.

[2]No differences were found across the various versions of the recognition tests. Because these different versions were constructed for methodological not theoretical reasons, the data were collapsed across test version for all of the analyses reported in this article.

[3] Across all five experiments in this article, there were no effects of blocking or gender. As there were no theoretical hypotheses concerning these variables in this article, and because neither variable affected performance, they were not included in subsequent analyses for any of the experiments.

[4]Throughout this article, it is important to note that the recognition findings were unaffected by the prior recall test. That is, when recognition probabilities were conditionalized on whether an item had been recalled earlier, the patterns of findings were no different than when the unconditional recognition probabilities were analyzed (for similar findings, see Marche, Brainerd, Lane, & Loehr, 2005). In addition, using a separate but smaller sample of 5-, 7-, and 11-year-olds and adults, we examined recognition performance without an intervening recall test. Analyses of these data showed the same pattern of results as the ones reported here that were preceded by a recall test.

[5]For Experiments 2, 4, and 5 involving children, the recognition results were the same regardless of whether raw scores or A(was used in the analyses (for a discussion of the use of signal detection analyses, see Pastore, Crawley, Berens, & Skelly, 2003). Because there were no differences in the outcome of the analyses, and because we report the raw score recognition findings for adults in Experiments 1 and 3, we report the untransformed, raw score analyses for the recognition data from all participants across the five experiments in this article.

Figure Captions

Figure 1. Adults' true and false recognition rates as a function of valence and time of test in Experiment 3.

Figure 2. Children's true and false recognition rates as a function of valence and time of test in Experiment 4.

Figure 3. Children's true and false recognition rates as a function of valence and time of test in Experiment 5.

Figure 1.

Figure 2.

Figure 3.