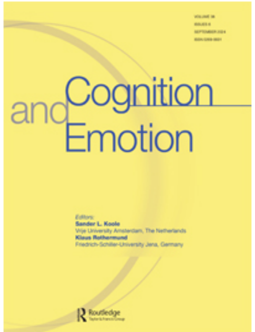

Cognition and Emotion

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/pcem20

Emotional false memories: the impact of response bias under speeded retrieval conditions

Lauren M. Cooper & Datin Shah

To cite this article: Lauren M. Cooper & Datin Shah (12 Sep 2024): Emotional false memories: the impact of response bias under speeded retrieval conditions, Cognition and Emotion, DOI: 10.1080/02699931.2024.2401611

To link to this article: https://doi.org/10.1080/02699931.2024.2401611

Published online: 12 Sep 2024.

COGNITION AND EMOTION
https://doi.org/10.1080/02699931.2024.2401611

Emotional false memories: the impact of response bias under speeded retrieval conditions

Lauren M. Cooper and Datin Shah

Department of Psychology, City, University of London, London, UK

ABSTRACT

Emotional false memory findings using the DRM paradigm have been marked by higher false alarms to negatively arousing compared to neutral critical lure items. Explanations for these findings have mainly focused on false memory-based accounts. However, here we address the question of whether a response bias for emotional stimuli can, at least in part, explain this phenomenon. Participants viewed both neutral and negative arousing DRM lists and completed a recognition test in speeded or self-paced conditions. Speeded test reduces the opportunity to adjust response bias. Analysis showed no significant difference in false recognition across critical lure types for the speeded condition, but false recognition was higher for negative compared to neutral critical lures in the self-paced condition. We argue that when retrieval does not allow for shifts in response criteria, false alarms to negative emotional critical lures appear more similar to neutral equivalents. The discussion explores memory-based and criterion-shift explanations for the enhanced emotional false memory finding.

ARTICLE HISTORY
Received 2 February 2024
Revised 24 July 2024
Accepted 2 September 2024

KEYWORDS
Emotion; false memory; DRM paradigm; response bias; speeded retrieval

False memories occur through inaccurate recollections of events or details and can manifest through a confluence of internal and external mechanisms. External factors, such as misleading post-event information, underscore the susceptibility of memory to external manipulation and distortion (e.g. Loftus & Palmer, 1974). Conversely, internal mechanisms, such as those explored through the Deese-Roediger-McDermott (DRM) paradigm, highlight the innate capability of our cognitive processes to spontaneously generate false recollections (Deese, 1959; Roediger III & McDermott, 1995). The DRM paradigm engages participants in studying word lists, each semantically associated with an unstudied "critical lure" word. In subsequent retrieval tests, this critical lure is presented as a test item, and participants frequently falsely recall and recognise it. Remarkably, false recognition rates often equate to or exceed those for genuinely studied items (Roediger III & McDermott, 1995).

The DRM paradigm has been widely employed to examine factors that influence false memory production. We have been interested in a particular focus on the emotional salience of the to-be-remembered content. Various studies, manipulating the emotional content of word lists, have sought to understand emotional false memories (Brainerd et al., 2010; Hellenthal et al., 2019; Howe et al., 2010; Knott et al., 2018; Otgaar et al., 2016). Although there are some variances, the general picture from these findings is that negative high arousing stimuli provide the "optimum" conditions for false memory production (Brainerd et al., 2010), however, this is often only evident for false recognition, not recall (Howe et al., 2010).

Several theories have been used to explain the DRM's robust findings, and in turn, consider the role of emotional salience. Activation-Monitoring Theory (AMT; Roediger III et al., 2001) attribute false

CONTACT Lauren M. Cooper Lauren.Cooper@city.ac.uk Department of Psychology, City, University of London, London EC1V 0HB, UK

memories to an implicit effect of the spreading activation (Collins & Loftus, 1975) and subsequent failures in source monitoring (Johnson et al., 1993). Words are conceptualised as "nodes" linked by their semantic similarities. When a word is studied, its corresponding node is activated. This activation then extends to the node of the critical lure, resulting in its false recall or recognition. The stronger the association between a list item and a critical lure, the more likely it is that the list item will cause activation of the critical lure. The inability to discriminate between studied items generated externally and critical lures generated internally (Johnson et al., 1993) leads to high false recognition responses. Researchers have argued that negative emotional items are embedded within dense associative networks which enhances the activation of critical lures (Otgaar et al., 2016; Talmi, 2013). A limited number of theme nodes facilitates a quick and more automatic spread of activation to the negative critical items. During retrieval, monitoring processes exhibit diminished efficacy in "editing" negative emotional critical lures due to the lack of differentiating information available.

Whilst the above is a memory-based account of false memory formation, researchers have also provided evidence for a criterion shift account (see Wixted & Stretch, 2000, for a review). This perspective suggests that the recognition decision is less about the experience of remembering and more about a strategic inferred judgement based on the familiarity of the critical lure. Corroborating this, researchers have employed alternative forced choice tests, limiting the role of response criterion utilised in old/new tests (Jou et al., 2018), thereby achieving a notable reduction in false memory.

Most studies highlighting the effects of emotional enhancement on false memories primarily emphasise memory-based theories. Explanations for emotional false memories centred on criterion shifts have garnered comparatively less focus. However, there is recent evidence to suggest that response bias for negative critical lures might be more lenient than response bias for neutral emotional critical lures. Such evidence suggests that because emotionally salient stimuli increase meaning-based associations, a more lenient criterion is adopted for accepting emotional stimuli as old (see Dougal & Rotello, 2007). There has been some evidence using signal detection analyses to support this suggestion (Hellenthal et al., 2019) with more liberal criterion measures reported for negative compared to neutral critical lures, even within the same recognition test (Yüvrük & Kapucu, 2022).

One would assume therefore that if a criterion shift was responsible for greater emotional false memories in the DRM paradigm, reducing the opportunity to employ strategic judgements could reduce the emotional false memory effect. One way this could be achieved is to utilise a speeded test condition to limit differences in strategic judgement, especially across neutral and emotional stimuli. Indeed, Carneiro et al. (2017) suggested that speeded testing conditions hinder editing mechanisms needed to distinguish correctly studied and non-studied items. We argue that a speeded test condition will limit the opportunity to shift criteria for negative emotional items. When this happens, we should see an elimination in the "enhanced" emotional false memory effect. We will therefore utilise a speeded versus self-paced recognition test to manipulate the propensity to form negative emotional false memories.

Method

Participants

There were 60 participants in this study (sex: 4 males, 57 females; age: 18-53, $M = 20.03$, $SD = 4.44$). All participants were undergraduate students at City, University of London. They received course credit for their participation, and all were native English speakers. A priori power analysis using G*Power indicated a total sample size minimum of 52 to detect a 2×2 interaction effect, with a medium effect size of 0.25 and Power ($1 - \beta$ err prob) of 0.95. Thirty participants were randomly assigned to the "self-paced test" condition, whilst 31 were randomly assigned to the "speeded test" condition.

Design, stimuli, and procedure

The experiment followed a 2 (list type: neutral vs negative) X 2 (test type: speeded vs self-paced recognition) mixed factorial design, with repeated measures on the first factor. The main dependent variable was the "hit" and "false alarm" rates to list items, critical lures, and distractor items.

The encoding and retrieval phases were presented using E-prime in individual testing laboratories on the University Campus. Once randomly assigned to each test type condition, participants completed the encoding phase. They were informed that they would be presented with several word lists. We used 12 DRM lists (6 negative lists and 6 neutral lists) taken from Knott et al. (2022). Each list had 12 words with the following critical lures for the negative lists: *lie, hurt, thief, dead, cry, and sick*, and the neutral lists: *window, foot, chair, mountain, shirt, and car*. They were matched for Backward Associative Strength (BAS) but differed in arousal and valence (see Table 1).[1] Each word was presented individually for 2 s, centrally on the screen, and with a 1-second interval between each word. Before the presentation of each list, an on-screen instruction would inform them of the next list (for 2s). They were reminded that they would later complete a recognition test for the words presented in the encoding phase.

After the presentation of all lists, participants were instructed to complete a 10-minute set of math problems on paper as a filler task. The recognition test then followed. The test consisted of 72 items. There were 12 critical lures (6 neutral, 6 negative), 36 list items (18 neutral, 18 negative), 12 unrelated distractor items (6 neutral, 6 negative), and 12 weak related distractor items (6, neutral, 6 negative). Neutral and negative distractors matched the valence and arousal levels of the list items. Participants were told that they would be presented with a word, and they needed to indicate whether that word was "old" (they had seen it during the encoding phase), or "new" (they had not seen it during the encoding phase), using a keyed response. In the self-paced condition, no time limit was set for participants to provide their responses. For the speeded test condition, participants were instructed to give their responses quickly, not exceeding the time limit of 750 ms. After the time limit, the next word appeared. A lack of response within the time limit was recorded as a miss. Participants in the speeded condition were given four practice trials with unseen words presented at the pace of the speeded condition to familiarise themselves with the time frame for a response. If participants did not select an answer within the time limit, a message appeared on-screen reading, "Please respond faster", encouraging them to be quicker in selecting a response.

All stimuli, data and codes are available at https://osf.io/73j46/

Table 1. Mean (and Standard Deviations) values, with independent *t*-test mean comparisons and Bayes Factor (BF) analysis for valence, arousal and Backward Associative Strength (BAS) by list emotion.

	Negative Lists	Neutral Lists	*t* value	*p* value	BF_{10}
List item Valence	3.14 (.65)	5.27 (.29)	7.35	<.001	[a]567.50
Critical lure Valence	2.13 (.38)	6.06 (1.13)	7.36	<.001	[a]200.66
List item Arousal	5.47 (.38)	4.32 (.38)	5.27	<.001	[b]62.12
Critical lure Arousal	5.74 (.93)	4.20 (1.06)	2.45	=.040	[c]2.12
BAS	.25 (.07)	.23(.05)	0.45	=.664	[d]0.50

Notes: $BF_{10} > 1$ supports the alternate hypothesis, and a $BF_{10} < 1$ supports the null hypothesis. Thus, the interpretation of the BF values in the table are as follows:
[a]Extreme evidence in favour of a difference between Negative and Neutral lists.
[b]Very strong evidence in favour of a difference between Negative and Neutral lists.
[c]Anecdotal evidence in favour of a difference between Negative and Neutral lists.
[d]Anecdotal evidence in favour of no difference between Negative and Neutral lists.

Results

One participant from the speeded condition was removed from the analysis as they responded new to 92% of the recognition test items. Separate 2 (emotion: negative vs neutral) X 2 (test condition: speeded vs self-paced) mixed factors ANOVAs were conducted to examine *old response rates* (hits to list items and false alarm rates for critical lures, weak related distractors, and unrelated distractors). The same analysis was conducted for discrimination sensitivity (A') and bias measures (B'_H).[2] Higher values of discrimination sensitivity (A') indicate better discrimination of hits from false alarms, whilst values of zero indicates minimal response bias (B'_H). Lower values signify more liberal responding, prioritising not missing true signals but at the cost of increased false positives. We conducted two discrimination analyses. First, discrimination of critical lures from unrelated distractors ($A'_{\_CL}$) and second, discrimination of studied list words from unrelated distractors ($A'_{\_List\ Item}$). In addition, two bias measures were calculated; $B'_{H\ \_CL}$ to indicate bias used to discriminate critical lures from distractors, and $B'_{H\ \_List\ Item}$ to indicate bias used to discriminate list words from distractor items. Bonferroni corrected multiple comparisons were used for all significant main effects and interactions. Alpha was set at 0.05.

Old response rates

For old responses to list items (hits), there was a significant main effect of emotion with greater hits to negative compared to neutral list items, $F(1, 57) = 28.82$, $p < .001$, $\eta_p^2 = .34$, and greater number of hits in the self-paced compared to speeded test condition, $F(1, 57) = 19.43$, $p < .001$, $\eta_p^2 = .25$, but there was no significant interaction, $F(1, 57) = 1.91$, $p = .173$, $\eta_p^2 = .30$. For old responses to critical lures (false recognition), there were significant main effects of emotion, $F(1, 57) = 7.37$, $p = .01$, $\eta_p^2 = .12$ and test condition, $F(1, 57) = 6.42$, $p = .014$, $\eta_p^2 = .10$, and a significant interaction, $F(1, 57) = 5.59$, $p = .02$, $\eta_p^2 = .09$. Bonferroni pairwise comparisons showed that old responses to negative and neutral critical lures did not differ significantly in the speeded condition ($p = .81$). However, for the self-paced condition, there was a significant difference, ($p < .001$) with more false recognition responses to negative compared to neutral critical lures, ($p < .001$; see Figure 1). As expected, for weak related distractors, there was no main effect of emotion, $F(1, 57) = 2.16$, $p = .15$, $\eta_p^2 = .04$ nor interaction, $F(1, 57) = 0.93$, $p = .338$, $\eta_p^2 = .10$. However, false alarms were greater in the speeded compared to self-paced condition, $F(1, 57) = 11.95$, $p < .001$, $\eta_p^2 = .17$. For unrelated distractors, there were no significant results (all $Fs < 1.10$, see Table 2).

Memory sensitivity and response bias

Similar 2(emotion: negative vs. neutral) X 2(test condition: speeded vs. self-paced) mixed ANOVAs were conducted for each sensitivity measure ($A'_{\_CL}$) and ($A'_{\_List\ Item}$). For $A'_{\_CL}$, there was a significant main effect of test condition which suggests greater sensitivity (greater distinction in false recognition to critical lures compared to unrelated distractors) in the self-paced compared to speeded condition $F(1, 57) = 10.22$, $p = .002$, $\eta_p^2 = .15$. The was no significant main effect of emotion, $F(1, 57) = .96$, $p = .33$, $\eta_p^2 = .02$ or interaction, $F(1, 57) = .15$, $p = .70$, $\eta_p^2 = .003$. For $A'_{\_List\ Item,}$ there was a significant main effect of test item, $F(1, 57) = 27.68$, $p < .001$, $\eta_p^2 = .33$ but there was no main effect of emotion, $F(1, 57) = 3.45$, $p = .07$, $\eta_p^2 = .06$, or interaction, $F(1, 57) = 0.07$, $p = .79$, $\eta_p^2 = .001$ (see Table 2).

For the analysis of response bias for critical lures ($B'_{H\ \_CL}$), there was no significant main effect of test condition, $F(1, 57) = .10$ $p = .76$, $\eta_p^2 = .002$ or emotion, $F(1, 57) = 2.35$, $p = .13$, $\eta_p^2 = .04$, but there was a significant interaction, $F(1, 57) = 7.00$, $p = .01$, $\eta_p^2 = .11$, indicating that the effects of emotion on response bias may depend on the test condition. Bonferroni-adjusted pairwise comparisons showed that response bias between negative and neutral critical lures did not differ significantly in the speeded condition ($p = .44$). For the self-paced condition, there was a significant difference, ($p = .004$) and response bias was more liberal for negative compared to neutral critical lures (see Figure 2). For list item response bias ($B'_{H\ \_List\ Item}$), with no significant main effect of test condition $F(1, 57) = .30$, $p = .59$, $\eta_p^2 = .01$, but significant a main effects of emotion, $F(1, 57) = 5.76$, $p < .02$, $\eta_p^2 = .09$ and a significant interaction, $F(1, 57) = 9.74$, $p = .003$, $\eta_p^2 = .15$. Pairwise comparisons showed that response bias between negative and neutral critical lures did not differ significantly in the speeded condition ($p = .62$). For the self-paced condition, there was a significant difference, ($p < .001$) and response bias was more liberal for negative compared to neutral critical list items.

Discussion

This study primarily aimed to investigate whether limiting processing time at retrieval through speeded test conditions would impact the formation of false memories, particularly those associated with negative emotions. Previous research has suggested that negative-emotional words often produce more false memories than neutral words (Brainerd et al., 2010; Howe et al., 2010), but the reasons for the increase in emotional over neutral false memories have been less clear. Some have argued for a memory-based account whereby emotional items activate a dense associative network and increase gist/theme processing (Bookbinder & Brainerd, 2016; Otgaar et al., 2016), while others have suggested response bias for emotional stimuli (Hellenthal et al., 2019; Yüvrük & Kapucu, 2022). Our study examined the production of both emotional and neutral false memories when the retrieval phase allows for little to no opportunity to make criterion shifts. Our findings support previous results, with significantly higher false alarms to negative compared to neutral critical lures, but only in the self-paced condition. In the speeded condition, analysis revealed no significant difference in false recognition responses between emotion types. Similarly, there was no significant difference in response bias for negative and neutral critical lures in the speeded condition, but there was a significant difference for the self-paced condition with more liberal responding for negative critical lures compared to neutral critical lures.

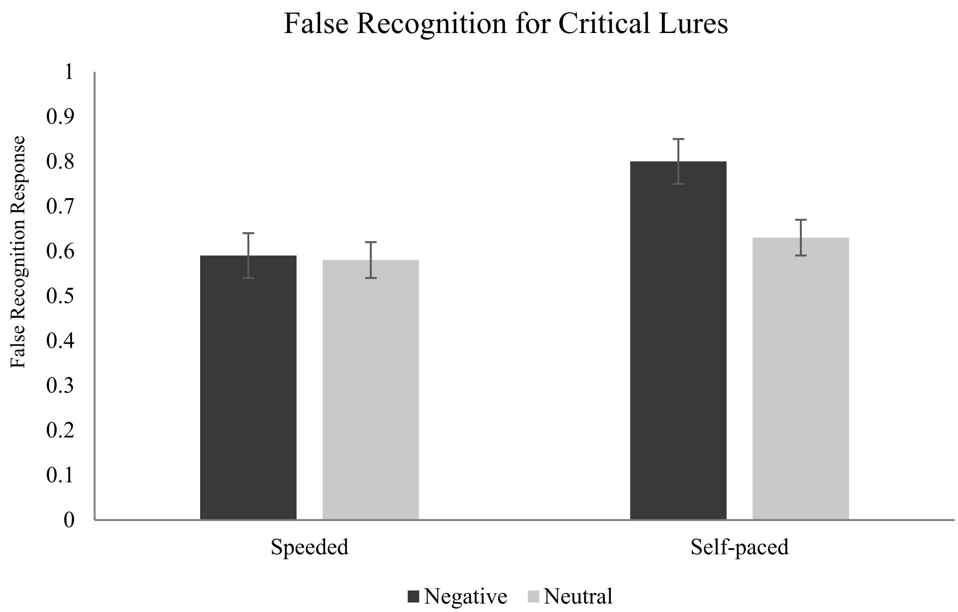

Figure 1. Proportion of false alarms to critical lures as a function of emotion and test type (error bars represent standard error).

Differences across emotion conditions could also reflect variations in monitoring strategies during retrieval (see Gallo et al., 2001). Implicit spreading activation, which is the source of false memory, is automatic and fast (Underwood, 1965), but monitoring, which is the basis of a correct rejection, is an effortful, slow, controlled process (Roediger III et al., 2001). Johnson et al. (1993) argued sufficient information for old-new discrimination becomes available before information for source-monitoring indicating the additional time needed to reject the familiar distractor after additional differentiating information is provided. Whether or not the correct rejection takes place will depend on whether the familiarity of the lure provides enough evidence to prevent the onset of additional monitoring. If negative critical lures feel particularly familiar, they may not evoke source-monitoring strategies because there is no sufficient differentiating evidence for a rejection decision. When monitoring, individuals focus on the quality of evidence related to the questioned event itself and whether it meets certain expected criteria (Gallo et al., 2001). This type of diagnostic monitoring seems to be compatible with the response bias of signal detection theories as both refer to expected criteria to make a recognition decision. High levels of familiarity for negative critical lure items convince us to accept less evidence to decide whether an item is old or new. In the self-paced condition, where participants can call upon source-monitoring strategies, the criterion for accepting negative items appears to rely more on feeling of familiarity as opposed to more differentiating diagnostic information that is available for neutral critical lures, thus increasing negative emotional false responses. With the speeded condition, there is no opportunity to consider response criteria, thus response biases cannot shift, and consequently we see similar false alarm responses to negative and neutral critical lures.

Table 2. Mean "old" response rate and measures of sensitivity and response bias (and standard deviation), as a function of test type and emotion.

	Speeded		Self-paced	
	Negative	Neutral	Negative	Neutral
Old responses				
List item	60 (.19)	49 (.20)	80 (.16)	62 (.15)
Critical lure	59 (.25)	58 (.25)	80 (.20)	63 (.23)
Weak related distractor	34 (.22)	33 (.23)	22 (.24)	14 (.17)
Unrelated distractor	30 (.25)	29 (.22)	26 (.20)	19 (.22)
Sensitivity & Response Bias A' **and** B'_H				
$A'_{\_CL}$	68 (.17)	67 (.18)	80 (.16)	77 (.16)
$B'_{H\ \_CL}$	17 (.48)	08 (.41)	-.07 (.52)	27 (.41)
$A'_{\_List\ Item}$	68 (.20)	62 (.17)	82 (.11)	78 (.11)
$B'_{H\ \_List\ Item}$	20 (.53)	15 (.33)	-.08 (.50)	34 (.37)

Our key explanation for our findings is that criterion can shift on a trial-by-trial basis for critical lures that were either negatively arousing or neutral in nature. This occurred in a self-paced recognition test but not

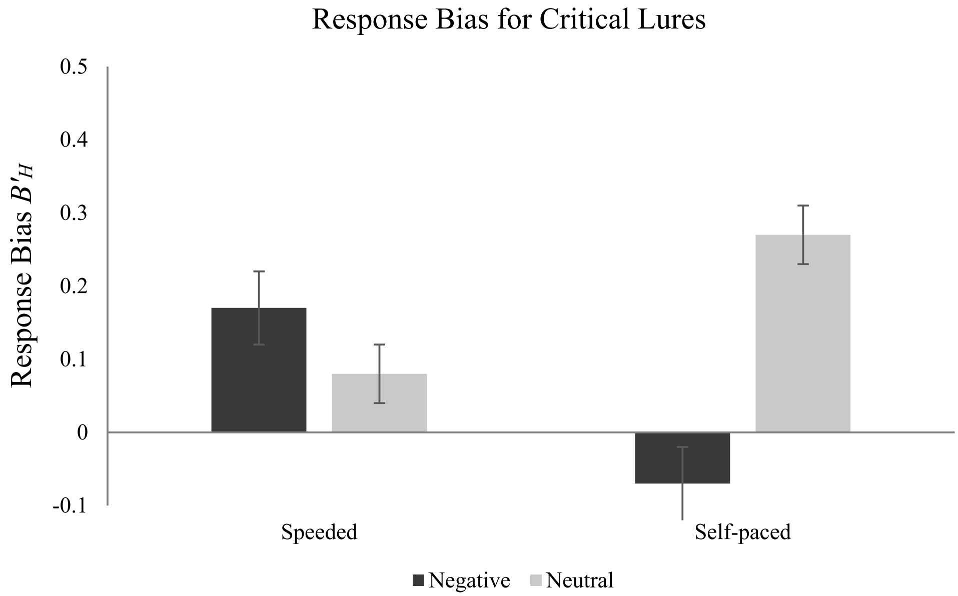

Figure 2. Response Bias B'_H, for critical lures as a function of emotion and test type (error bars represent standard error). Lower values signify more liberal responding.

in a test where response time was severely limited. Stretch and Wixted (1998) have argued that trial-by-trial criterion shifts for items differentially encoded but tested together do not occur because shifting criteria is an onerous, resource-demanding process. Starns and Olchowski (2015) argued that tailoring the response criterion to expected strength is a natural part of the recognition process. They found that participants can shift criterion trial-by-trial based on changes in expected strength. One would assume that the availability of differentiating information for neutral but not negative stimuli would cause a change in expected strength. This is what we found in the self-paced condition (similar to Yüvrük & Kapucu, 2022). However, changing criteria if resource-demanding, may be more difficult in the speeded test condition. Our analysis provides support for this with no significant difference in response bias across emotion types in the speeded test condition. One question for further study might be to understand, at what point, response time becomes too fast to utilise differentiating information, or indeed, make trial-by-trial criterion shifts.

We note that our findings differ from those of Shah and Knott (2018). They too examined reduced attention for emotional false formation by using a random number generation task during retrieval and found that false recognition was higher for negative compared to neutral critical lures for both full and divided attention conditions (but only for remember responses, they were similarly matched for old responses). Although divided attention task reduces attentional resources for recognition decisions, the recognition test was self-paced and this can give one opportunity to engage monitoring strategies and response biases. Shah and Knott did not report criterion response measures, although they did report higher levels of negative distractors compared to neutral distractors. This is a sign that response bias may be more liberal for negative items.

To conclude, the enhanced subjective vividness and familiarity for all negative test items may generate a more liberal bias to classify them as "old" during a recognition test compared to neutral items. Neutral list items and critical lures provide more differentiating information compared to their negative counterparts, making them less reliant on familiarity. Source monitoring is more effective, and we expect a higher strength of memory. Only when we remove the opportunity to monitor source and evaluate strength, does the false recognition of neutral and negative critical lures in the same test become similar, presumably because criterion shifts cannot occur under such fast response conditions. There is, of course, an adaptive explanation. Yüvrük and Kapucu (2022) suggested that negative items carry potential threats, and the liberal bias serves to facilitate memory by increasing the sense of familiarity, ensuring the potential threat is not overlooked, even when those memories were never actually experienced. This finding certainly underscores the importance of considering both the emotional content of memories and the conditions under which they are retrieved, to fully understand the mechanisms underlying memory recognition and the potential biases that may arise.

Notes

1. Although emotional and neutral lists were developed to match for associative strength whilst manipulating valence and arousal, we acknowledge that there are other list variables that may contribute to accurate and false recognition rates. Neutral lists are typically more concrete than emotional lists, and this was the case in our stimuli, where values were available ($M = 405.28$, $SD = 80.72$ vs $M = 565.21$, $SD = 30.84$, $p < .001$; Coltheart, 1981), although analysis of differences in word frequency did not reach significance ($M = 32.06$, $SD = 17.68$ vs $M = 59.67$, $SD = 33.02$, $p = .10$; Kucera & Francis, 1967). Previous research has demonstrated that when comparing neutral concrete, neutral abstract, and emotional lists, false memories are similar for concrete and abstract neutral lists, but still higher for emotional lists (Bauer et al., 2009). This would certainly be something to consider, if stimuli could be designed in such a way to control for these list variables (and BAS) whilst still maintain the key manipulations of valence and arousal.
2. Sensitivity measure A' (Pollack & Norman, 1964) and bias measure B'_H, equivalent to B' (Hodos, 1970) are used here at the recommendation of one of our reviewers. We acknowledge that without a confidence rating to construct ROC curves, we do not have a measure of sensitivity uncontaminated by bias, however our speeded test condition meant that it was impractical to collect confidence ratings. Although McNicol (1972) notes that if there is bias in either direction, A' will underestimate sensitivity, our main reason for this analysis is to directly compare criterion shifts between the two test conditions. To note, we also calculated and analysed d' and Criterion C and found that the pattern of results were similar to discrimination and response bias measures reported here.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The author(s) reported there is no funding associated with the work featured in this article.

References

Bauer, L. M., Olheiser, E. L., Altarriba, J., & Landi, N. (2009). Word type effects in false recall: Concrete, abstract, and emotion word critical lures. *The American Journal of Psychology*, *122*(4), 469–481. https://doi.org/10.2307/27784422

Bookbinder, S. H., & Brainerd, C. J. (2016). Emotion and false memory: The context–content paradox. *Psychological Bulletin*, *142*(12), 1315–1351. https://doi.org/10.1037/bul0000077

Brainerd, C. J., Holliday, R. E., Reyna, V. F., Yang, Y., & Toglia, M. P. (2010). Developmental reversals in false memory: Effects of emotional valence and arousal. *Journal of Experimental Child Psychology*, *107*(2), 137–154. https://doi.org/10.1016/j.jecp.2010.04.013

Carneiro, P., Garcia-Marques, L., Lapa, A., & Fernandez, A. (2017). Explaining the persistence of false memories: A proposal based on associative activation and thematic extraction. *Memory (Hove, England)*, *25*(8), 986–998. https://doi.org/10.1080/09658211.2016.1239742

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428. https://doi.org/10.1037/0033-295X.82.6.407

Coltheart, M. (1981). MRC psycholinguistic database user manual: Version 1.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*(1), 17–22. https://doi.org/10.1037/h0046671

Dougal, S., & Rotello, C. M. (2007). "Remembering" emotional words is based on response bias, not recollection. *Psychonomic Bulletin & Review*, *14*(3), 423–429. https://doi.org/10.3758/BF03194083

Gallo, D. A., Roediger, H. L., & McDermott, K. B. (2001). Associative false recognition occurs without strategic criterion shifts. *Psychonomic Bulletin & Review*, *8*(3), 579–586. https://doi.org/10.3758/BF03196194

Hellenthal, M. V., Knott, L., Howe, M. L., Wilkinson, S., & Shah, D. (2019). The effects of arousal and attention on emotional false memory formation. *Journal of Memory and Language*, *107*, 54–68. https://doi.org/10.1016/j.jml.2019.03.010

Hodos, W. (1970). Nonparametric index of response bias for use in detection and recognition experiments. *Psychological Bulletin*, *74*(5), 351–354. https://doi.org/10.1037/h0029933

Howe, M. L., Candel, I., Otgaar, H., Malone, C., & Wimmer, M. C. (2010). Valence and the development of immediate and long-term false memory illusions. *Memory (Hove, England)*, *18*(1), 58–75. https://doi.org/10.1080/09658210903476514

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*(1), 3–28. https://doi.org/10.1037/0033-2909.114.1.3

Jou, J., Escamilla, E. E., Arredondo, M. L., Pena, L., Zuniga, R., Perez, M., & Garcia, C. (2018). The role of decision criterion in the Deese–Roediger–McDermott (DRM) false recognition memory: False memory falls and rises as a function of restriction on criterion setting. *Quarterly Journal of Experimental Psychology*, *71*(2), 499–521. https://doi.org/10.1080/17470218.2016.1256416

Knott, L. M., Howe, M. L., Toffalini, E., Shah, D., & Humphreys, L. (2018). The role of attention in immediate emotional false memory enhancement. *Emotion*, *18*(8), 1063–1077. https://doi.org/10.1037/emo0000407

Knott, L. M., Wilkinson, S., Hellenthal, M., Shah, D., & Howe, M. L. (2022). Generative processing and emotional false memories: A generation 'Cost' for negative false memory formation but only after delay. *Cognition and Emotion*, *36*(7), 1448–1457. https://doi.org/10.1080/02699931.2022.2128063

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Brown University Press.

Loftus, E. F., & Palmer, J. C. (1974). Recollection of auto-mobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, *13*(1), 585–589. https://doi.org/10.1016/S0022-5371(74)80011-3

McNicol, D. (1972). *A primer of signal detection theory*. George Allen & Unwin.

Otgaar, H., Howe, M. L., Brackmann, N., & Smeets, T. (2016). The malleability of developmental trends in neutral and negative memory illusions. *Journal of Experimental Psychology: General*, *145*(1), 31–55. https://doi.org/10.1037/xge0000127

Pollack, I., & Norman, D. A. (1964). A non-parametric analysis of recognition experiments. *Psychonomic Science*, *1*(1-12), 125–126. https://doi.org/10.3758/BF03342823

Roediger III, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 803–814. https://doi.org/10.1037/0278-7393.21.4.803

Roediger III, H. L, Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review*, *8*(3), 385–407. https://doi.org/10.3758/BF03196177

Shah, D., & Knott, L. M. (2018). The role of attention at retrieval on the false recognition of negative emotional DRM lists. *Memory (Hove, England)*, *26*(2), 269–276. https://doi.org/10.1080/09658211.2017.1349803

Starns, J. J., & Olchowski, J. E. (2015). Shifting the criterion is not the difficult part of trial-by-trial criterion shifts in recognition memory. *Memory & Cognition*, *43*(1), 49–59. https://doi.org/10.3758/s13421-014-0433-y

Stretch, V., & Wixted, I. T. (1998). On the difference between strength- based and frequency-based mirror effects in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(6), 1379–1396. https://doi.org/10.1037/0278-7393.24.6.1379

Talmi, D. (2013). Enhanced emotional memory: Cognitive and neural mechanisms. *Current Directions in Psychological Science*, *22*(6), 430–436. https://doi.org/10.1177/0963721413498893

Underwood, B. J. (1965). False recognition produced by implicit verbal responses. *Journal of Experimental Psychology*, *70*(1), 122–129. https://doi.org/10.1037/h0022014

Wixted, J. T., & Stretch, V. (2000). The case against a criterion-shift account of false memory. *Psychological Review*, *107*(2), 368–376. https://doi.org/10.1037/0033-295X.107.2.368

Yüvrük, E., & Kapucu, A. (2022). False (or biased) memory: Emotion and working memory capacity effects in the DRM paradigm. *Memory & Cognition*, *50*(7), 1443–1463. https://doi.org/10.3758/s13421-022-01298-y