Cognition and Emotion

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

Generative processing and emotional false memories: a generation "cost" for negative false memory formation but only after delay

Lauren Knott, Samantha Wilkinson, Maria Hellenthal, Datin Shah & Mark L. Howe

COGNITION AND EMOTION
https://doi.org/10.1080/02699931.2022.2128063

Routledge
Taylor & Francis Group

BRIEF ARTICLE

OPEN ACCESS   Check for updates

Generative processing and emotional false memories: a generation "cost" for negative false memory formation but only after delay

Lauren Knott, Samantha Wilkinson, Maria Hellenthal, Datin Shah and Mark L. Howe

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

ABSTRACT

Previous research shows that manipulations (e.g. levels-of-processing) that facilitate true memory often increase susceptibility to false memory. An exception is the generation effect. Using the Deese/Roediger–McDermott (DRM) paradigm, Soraci et al. found that generating rather than reading list items led to an increase in true but not false memories. They argued that generation led to enhanced item-distinctiveness that drove down false memory production. In the current study, we investigated the effects of generative processing on valenced stimuli and after a delayed retention interval to examine factors that may lead to a generation effect that increases false memories. At the immediate test, false recognition rates for both negative and neutral valanced critical lures were similar across read and generate conditions. However, after a one-week delay, we saw a valence differentiation, with a generation effect for false recognition but only for negative stimuli. The roles of item-specific and relational processing during encoding and their interaction with long-term retention are discussed.

ARTICLE HISTORY
Received 17 May 2022
Revised 16 September 2022
Accepted 19 September 2022

KEYWORDS
DRM paradigm; false memories; emotion; generation effect

Encoding situations that facilitate retention often produce greater levels of illusory memories (Toglia et al., 1999). This phenomenon, often referred to as a *more is less effect*, has been studied using a well-known list learning paradigm (Deese-Roediger-McDermott [DRM]; Deese, 1959; Roediger & McDermott, 1995). Here, participants are presented with a list of words (e.g. *table*, *sit*, *chair*), which are all semantically associated with a non-presented critical item (*chair*). At recall or recognition, participants often report seeing or hearing this critical item as part of the original study list. Roediger and McDermott reported that participants recalled approximately 50% of the critical items, with false memories at a similar level to true memory in a recognition test.

Two widely cited theoretical explanations include associative-activation theories (AAT; e.g. Howe, 2005) and the dual-process fuzzy-trace theory (FTT; Brainerd & Reyna, 2001). AAT argues that false recall and recognition in the DRM paradigm occur due to the spread of semantic activation from associative words presented at the study. Activated items might later be incorrectly retrieved as items that were seen or heard. FTT argues for two parallel processes. Gist traces represent the core meaning of the memory but not its specific details whereas verbatim traces capture the specific attributes of the memory (e.g. visual features). Retrieving verbatim traces results in accurate recollection of list items. Retrieving gist traces can lead to correct recognition or false recognition. List items cue the semantic relation to the critical lure. The more cuing, the stronger the gist representation to the corresponding critical lure.

DRM encoding conditions are often manipulated to better understand the role of gist processing or semantic activation on veridical and false memory

CONTACT Lauren Knott lauren.knott.1@city.ac.uk Department of Psychology, City, University of London, Northampton Square, London EC1V 0HB, UK

rates. Toglia et al. (1999) referred to factors that increased both veridical and false memory as the *more is less* effect. So called because conditions that facilitate recall and recognition induce higher levels of false memory which leads to less accurate memory performance. For example, Toglia et al. varied levels-of-processing at encoding by using semantic and non-semantic judgements of list items. Semantic processing increased both accurate recall of list items and false recall of critical lures. Thapar and McDermott (2001) extended these findings to recognition memory. However, not all manipulations that enhance accurate memory, increase false intrusions. For example, the *generation effect* (e.g. Slamecka & Graf, 1978; Soraci et al., 1994) shows a robust memory enhancement effect but does not also increase false recollection (Soraci et al., 2003). Using the DRM paradigm, Soraci et al. found that generating list items from word-stems increased correct recognition with no comparable increase in critical lures. They argued that generative processing increased item distinctiveness. Unlike other levels-of-processing manipulations, Soraci et al. argued that the generation condition enhances item-specific processing relative to relational or meaning-based information for list items (e.g. Gardiner & Hampton, 1988). Such focus on the surface form disrupts the use of meaning-based information (e.g. Hirshman & Bjork, 1988). In relation to false memory formation, increasing item-distinctiveness likely increases reliance on verbatim processing and disrupting the list item's relational information hinders gist trace formation and spreading activation.

It appears that classic levels-of-processing manipulations that produce the *more is less* effect enhance semantic and relational meaning during acquisition. This strengthens associative activation or gist processing that leads to an increase in both true and false memory. Generative processes enhance verbatim processing in study impeding the use of gist representations needed for false memories (Soraci et al., 2003). However, can we manipulate generation under conditions that "dilute" its' verbatim booster, thus transforming it into a more is less manipulation that leads to an increase in false recognition?

There are two conditions that come to mind. First is to measure generation effects after a delay of several days. We know that item-specific details decay more quickly than meaning-based or relational information (Colbert & McBride, 2007). If generative processing enhances item-specific processing and the acquisition of such information is used at the test to facilitate item differentiation, then a longer retention interval that leads to the decline of such information may reduce its role in driving down false memories associated with the generated list items. After a longer delay, recognition will have to rely more on relational information which, as we know, is what false memories rely on. However, if generative processing suppresses the reliance of relational details, will there now be sufficient information to falsely identify the critical lure? This is where our next factor may be key.

Second, we manipulate valence. Previous research that demonstrated the suppressing effect of generative processing on false memories only used neutral valenced stimuli. However, recent research has shown an emotionally enhanced false memory effect. That is, encoding of high arousing negative valenced associative lists produces greater false recognition rates compared to equivalent neutral lists (see Bookbinder & Brainerd, 2016; Howe et al., 2010). Memory-based accounts attribute this to a greater reliance on relational information or gist representations for valenced stimuli and more overlapping theme nodes (Otgaar et al., 2016) leading to the easier spread of activation and greater difficulty distinguishing externally presented and internally generated items. Either manipulation on its own may not be sufficient to prevent the generation effect from reducing false recognition, however, given the findings by Howe et al. (2010), together, there may be sufficient decay of item-specific processing and enhanced meaning extraction for negative valenced stimuli to cause a *more is less* generation encoding manipulation. We suggest this because Howe et al. found that over a one-week retention interval, although true recognition declined, false recognition for neutral valence items stayed consistent and false recognition for negative valence items increased. They similarly argued that item-specific details fade more quickly, but that gist-rich material (negative items) may also give rise to more stable gist over time.

Therefore, the aim of the experiment was to examine the effects of generative processing on false recognition rates for neutral and negative valenced critical lures with immediate versus (one-week) delayed testing. Based on the previous generation effect research, we predicted that a word-fragment generation task (Chechile & Soraci, 1999) would facilitate correct recognition compared to a read condition but there would be no loss of accuracy. That is, false recognition rates in the generation condition would be similar to those produced in the read condition (Soraci et al., 2003). However, after one week, and decay of the enhanced item-specific processing, we expected to see recognition that will rely more on meaning-based information. As we know from previous research (e.g. Howe et al., 2010), this appears to be better consolidated over time and likely richer for negative valence stimuli (Bookbinder & Brainerd, 2016). We predict this will lead to a generation effect for false recognition of negative but not neutral critical lures following a delayed retention interval.

Method

Participants

Eighty-seven undergraduate and postgraduate students from City, University of London participated in the current experiment ($M_{age} = 23.24$, $SD_{\text{age}} = 8.58$, 26% male) for either course credits or a small remuneration fee. All participants gave written informed consent and were debriefed at the end of the experiment. Ethics approval was obtained from the Psychology Department's Ethics Committee. The sample size was determined a prior using G*Power, with effect size $f = .23$, alpha = .05, power = .95. Effect size was chosen based on a prior study with a repeated measures interaction examining effects of encoding presentation on false recognition of negative valence critical lures, $\eta p^2 = 0.05$ (Hellenthal et al., 2019).

Design

The experiment followed at 2(Valence: neutral vs. negative) × 2(Format: read vs. generate) × 2(Time of Test: immediate vs. one-week delay) mixed design with repeated measures on all but the last factor. Participants were randomly assigned to either the immediate or delayed testing group. The dependent variables for the recognition data were the old response rates to list items and critical lure items, and related and unrelated distractors (although distractors could not be compared across generation and read condition).

Materials

A set of twelve DRM lists were used in this experiment. Six were neutral and were taken from Stadler et al. (1999). They consisted of the top ten associate items (measured using Backward Associate Strength [BAS]) to the following critical lures: *chair*, *city*, *mountain*, *pen*, *shirt*, and *window*. Six were emotional-negative lists and were taken from Brainerd et al. (2010) and Howe et al. (2010). BAS values were taken from South Florida free association norms database (Nelson et al., 1998). They consisted of the top ten negatively-valenced associate items (measured using BAS) to the following critical lures: *alone*, *anger*, *dead*, *gun*, *sick*, and *thief*. Neutral and negative lists were matched for arousal but differed in valence. This was achieved using mean valence and arousal ratings for list items and critical lures from the Affective Norms for English Words (ANEW; Bradley & Lang, 1999) database. Paired samples *t*-tests showed that the emotional-negative list items ($M = 3.61$, $SD = .66$) and critical lures ($M = 2.37$, $SD = .58$) had significantly lower ratings of valence than neutral list items ($M = 5.26$, $SD = .55$) and critical lures ($M = 5.90$, $SD = .62$, $p = .002$ and $p < .001$ respectively) but no significant differences in arousal ratings (list items [$M^{\text{negative}} = 5.32$, $SD = 1.20$ and $M^{\text{neutral}} = 4.30$, $SD = 1.33$] and critical lures [$M^{\text{negative}} = 6.07$, $SD = 1.33$ and $M^{\text{neutral}} = 4.46$, $SD = 1.10$] $p = .21$ and $p = .08$, respectively). Negative and neutral lists were also matched for BAS ($M = .27$, $SD = .10$ and .20, $SD = .06$, respectively, $p = .19$).

Valence was blocked so that half of the participants studied six neutral lists followed by six negative lists and vice versa. In each valence condition, half of the word lists were presented intact (read condition) and the remaining half were presented with a letter (excluding the first or the last) missing (generate condition). The format was also blocked and counterbalanced within each valence condition. Each neutral and negative list served an equal number of times in the read or generate condition. The order of lists within each valence condition and within each format condition was randomised for each participant.

A recognition test was constructed using 72 items: 6 neutral critical lures, 6 negative critical lures, 24 presented items (two were taken from each list; one high associate [i.e. from positions 1 to 5] and one low associate [i.e. from positions 6 to 10]), and 36 distractor items (12 critical distractor items [one unseen item from each DRM list; this was typically the 14th or 15th item from the original DRM list] and 12 neutral and 12 emotional-negative non-critical distractor items [unrelated but matched for valence ratings to corresponding negative and neutral critical lure items using the ANEW database, $M^{Negative\ CL} = 2.37$ vs. $M^{distractor} = 2.57$; $M^{Neutral\ CL} = 5.90$ vs. $M^{distractor} = 5.58$]). All 72 items were randomly presented.

Procedure

All participants were tested individually. Participants were informed that 120 words or word fragments (in lists of 10) would be presented on the screen. It was made clear that some words would have a missing letter and they needed to complete the fragment. They were presented with two practice examples before beginning the study phase. The phrase "next list" was shown prior to the start of each new word list. Participants were asked to either type the word (i.e. for intact words) or the correct solution (i.e. for word-fragments), and they were told that they had 15 s per word to do so. If the participants failed to correctly type the word or the solution within the 15 s time limit, then the intact version of the word was displayed on the screen for 2 s before the programme moved on to the next word. The failure rate for the correct solution was low (approx. 2%).

On completion of the encoding phase, participants were given a 5-min filler sudoku task. Following this, participants who were assigned to the delayed testing condition were thanked for their time and would return one week later to complete the recognition test. Participants assigned to the immediate testing condition completed the memory task immediately. Instructions informed participants that a series of words would appear on the screen individually (and in random order) and by pressing labelled keys on their keyboard they had to indicate whether each word had appeared in the encoding task (i.e. *old* words) or not (i.e. *new* words). We also decided to include a recollective experience judgement following each old decision. A response of *remember*, *know*, or *guess* was made by pressing the keys r, k, and g keys. Instructions for remember, know, and guess responses were taken from Dewhurst and Anderson (1999). Essentially, participants were instructed to make a remember response if they consciously recollected some aspect of a word's study presentation, a know response if a word was familiar but no contextual information was recalled, or a guess response if they were unsure whether a word had appeared or not.

Results

For direct comparison to Soraci et al. (2003), we analysed hit rates for studied list items and false alarm rates for critical lures. Hits to list items and false alarms to critical lures were analysed separately using a 2(Format: generate vs. read) × 2(Valence: negative vs. neutral) × 2(Retention Interval: immediate vs. one-week) mixed factor analysis of variance (ANOVA) with repeated measures on the first two factors. To correct for participant's response bias, signal detection parameters d' and C were computed using list item or critical lure recognition as hits and weak-related distractors as false alarms (see Arndt & Hirshman, 1998). To avoid an infinite z value in computing the d's, all hit and false-alarm rates were corrected by adding 0.5 to the frequency of hits or false alarms and dividing this adjusted frequency by $N + 1$ where N was the number of old or new trials (Snodgrass & Corwin, 1988). The analysis of d' allows us to examine the ability to discriminate hits from false alarms. For critical lures, high discrimination scores demonstrate susceptibly to false memory formation with more old responses to critical lures and new responses to distractor items. Criterion value C represents the decision criterion. The higher the value the more conservative the bias (criterion favours no responses). The lower the value, the more liberal the bias towards a "yes" response. Recollective experience measures were also analysed; however, findings are largely in line (particularly *remember* responses) with old recognition responses. For completeness, they are reported in the footnote.[1]

Recognition data

For hits to list items, there were main effects of valence, $F(1, 85) = 24.32$, $p < .001$, $\eta_p^2 = .22$, format, $F(1, 85) = 85.62$, $p < .001$, $\eta_p^2 = .50$, and retention interval, $F(1, 85) = 5.28$ $p = .02$, $\eta_p^2 = .06$. The three-way interaction was not significant ($p = .09$); however, there was a significant Format × Retention Interval interaction, $F(1, 85) = 4.84$, $p = .03$, $\eta_p^2 = .05$. Although the generation effect for list items exists at both retention intervals (both $ps < .001$), pairwise comparisons did show that hits for generate items declined ($M^{immediate} = .91$, $SE = .02$ vs. $M^{delay} = .78$, $SE = .03$, $p < .001$), but this was not the case for the read condition, ($M^{immediate} = .68$, $SE = .03$ vs. $M^{delay} = .64$, $SE = .04$, $p = .41$). Generate hits were still higher in the delay condition but dropped over time from a

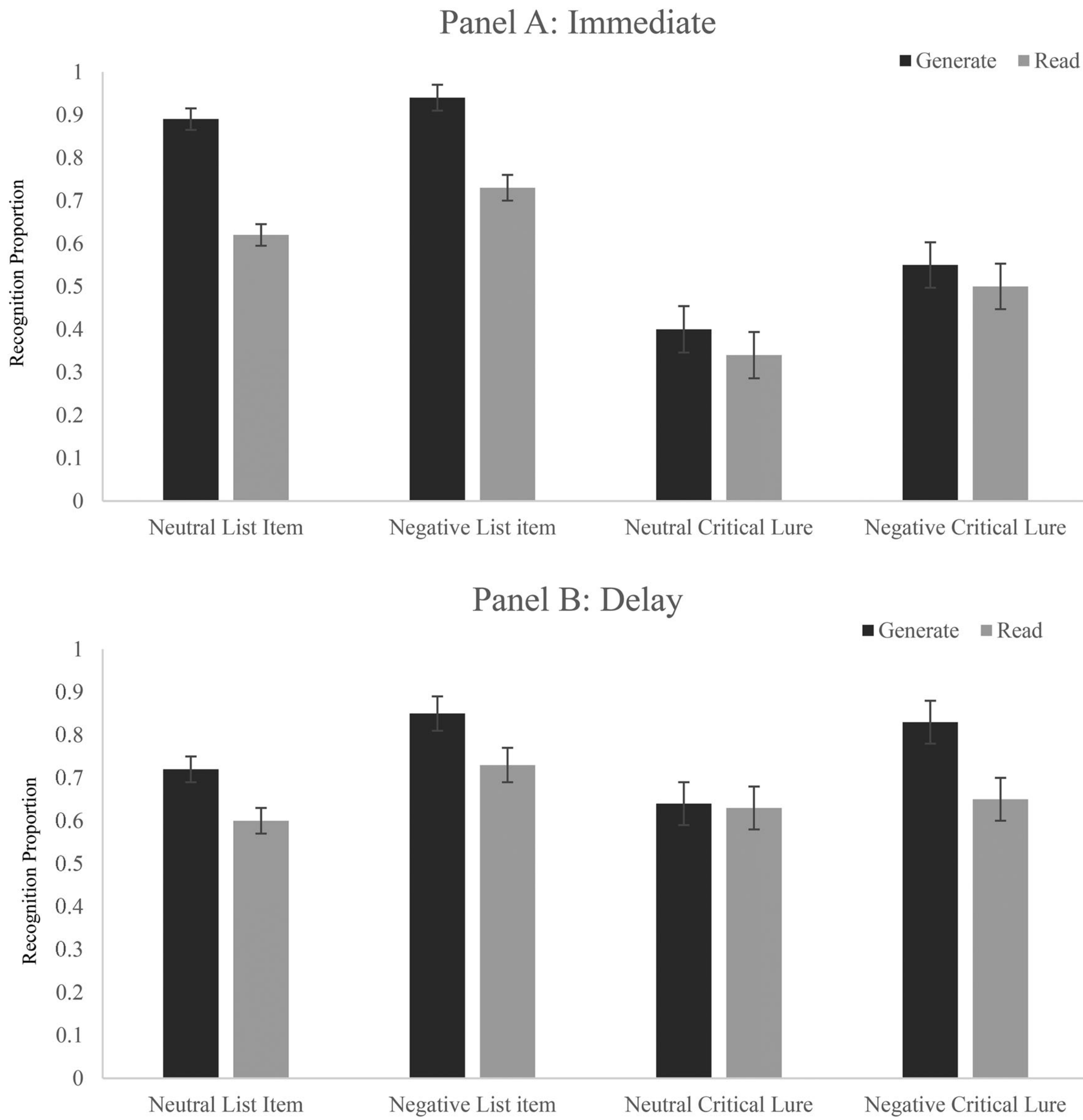

Figure 1. Mean proportions of old responses for the false recognition of critical lures and correct recognition of list items as a function of Valence and Format. Panel A presents data from the Immediate retention period and panel B presents data from the 1-week delay retention period. Error bars represent standard error.

very high ceiling. The generate effect was still evident at both time points (see Figure 1).

For critical lure false-alarm rates, there was a main effect of format, $F(1, 85) = 10.30$, $p = .002$, $\eta_p^2 = .11$, a main effect of valence, $F(1, 85) = 14.06$, $p < .001$, $\eta_p^2 = .14$, and the main effect of retention interval, $F(1, 85) = 21.85$, $p < .001$, $\eta_p^2 = .20$. These findings were qualified by a significant Valence × Format × Retention Interval interaction, $F(1, 85) = 3.97$, $p = .049$, $\eta_p^2 = .05$. Two separate Valence × Format analyses were conducted for immediate and one-week retention. Bonferroni pairwise comparisons suggested no generation effect for negative ($p = .17$) or neutral ($p = .13$) critical lure false alarms at the immediate test. However, after the 1-week retention interval, whilst there was no generation effect for neutral critical lures ($p = .88$), there were more false alarms to negative critical lures in the generate vs. read condition ($p < .001$) (see Figure 1).

The false-alarm rates for weak related and unrelated distractors cannot be analysed by format, so

Table 1. Proportionate means (and Standard Deviations) for old responses to unrelated and weak related distractors and recollective experience judgements to list items, critical lures, and distractors as a function of valence, format, and retention interval.

	Neutral stimuli				Negative stimuli			
	Immediate		Delay (1-week)		Immediate		Delay (1-week)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
				Generate				
List Items								
Old responses	.89	.16	.72	.29	.94	.10	.85	.17
Remember responses	.61	.32	.41	.31	.71	.28	.53	.26
Know responses	.19	.23	.22	.25	.18	.23	.23	.21
Guess responses	.09	.15	.10	.13	.05	.11	.10	.13
Critical Lures								
Old responses	.41	.33	.64	.38	.55	.35	.84	.23
Remember responses	.11	.20	.26	.34	.20	.31	.43	.31
Know responses	.17	.25	.26	.28	.21	.25	.26	.29
Guess responses	.14	.19	.12	.10	.14	.21	.14	.20
				Read				
List Items								
Old responses	.63	.25	.60	.28	.73	.24	.68	.26
Remember responses	.30	.25	.27	.29	.42	.32	.34	.24
Know responses	.18	.18	.19	.17	.21	.23	.21	.20
Guess responses	.15	.15	.15	.17	.11	.16	.14	.14
Critical Lures								
Old responses	.33	.35	.63	.35	.49	.37	.65	.30
Remember responses	.11	.17	.29	.32	.13	.24	.21	.27
Know responses	.09	.15	.18	.15	.23	.26	.34	.29
Guess responses	.13	.21	.16	.21	.13	.21	.10	.17
Weak related distractors								
Old responses	.12	.14	.40	.23	.25	.22	.57	.21
Remember responses	.02	.05	.11	.18	.08	.13	.16	.16
Know responses	.04	.07	.15	.15	.08	.11	.22	.17
Guess responses	.06	.13	.14	.17	.09	.13	.19	.16
Unrelated distractors								
Old responses	.06	.09	.29	.18	.19	.21	.49	.20
Remember responses	.006	.02	.05	.07	.02	.05	.13	.12
Know responses	.004	.02	.05	.08	.05	.09	.20	.13
Guess responses	.05	.08	.13	.13	.12	.14	.17	.15

instead separate 2(Valence: negative vs. neutral) × 2 (Retention Interval: immediate vs. one-week) ANOVAs were conducted. For related distractors, there were significant main effects of valence, $F(1, 85) = 38.87$, $p < .001$, $\eta_p^2 = .31$, and retention interval, $F(1, 85) = 66.45$, $p < .001$, $\eta_p^2 = .44$, but no interaction, $F(1, 85) = .60$, $p = .44$, $\eta_p^2 = .01$. For unrelated distractors, there were significant main effects of valence, $F(1, 85) = 83.92$, $p < .001$, $\eta_p^2 = .50$, and retention interval, $F(1, 85) = 59.10$, $p < .001$, $\eta_p^2 = .41$, and also interaction, $F(1, 85) = 10.71$, $p = .002$, $\eta_p^2 = .11$, but analysis of simple main effects showed a similar pattern of results to the main effect of valence and retention interval, with higher false alarms to negative items, and more false alarms after a one-week retention interval (all *ps* < .001; see Table 1).

Signal detection measures

Analysing discrimination values first, for list items, there were main effects of valence, $F(1, 85) = 21.21$, $p < .001$, $\eta_p^2 = .20$, format, $F(1, 85) = 96.45$, $p < .001$, $\eta_p^2 = .53$, and retention interval, $F(1, 85) = 84.89$, $p < .001$, $\eta_p^2 = .50$. Similar to recognition data, there was only one significant interaction between Format and Retention Interval, $F(1, 85) = 6.46$, $p = .01$, $\eta_p^2 = .07$, but pairwise comparisons showed that discrimination decreased for read vs. generate at both test intervals, and discrimination declined for both format types, in line with the main effects patterns (all *ps* < .001). For critical lures, there were main effects of valence, $F(1, 85) = 14.54$, $p < .001$, $\eta_p^2 = .15$, format, $F(1, 85) = 10.74$, $p = .002$, $\eta_p^2 = .11$, and retention interval, $F(1, 85) = 4.89$, $p = .03$, $\eta_p^2 = .05$. There was a significant 2-way Valence × Retention Interval, $F(1, 85) = 5.05$, $p = .03$, $\eta_p^2 = .06$, with a 3-way interaction, $F(1, 85) = 3.69$, $p = .06$, $\eta_p^2 = .04$ which suggested a pattern of results mirrored the recognition data and Bonferroni pairwise comparisons indicated no difference in susceptibility to negative ($p = .15$) or neutral ($p = .16$) critical lures across format at immediate test. After a delay, although,

Table 2. Signal detection measures of Discriminability (d′) and Criterion Bias (C) for list items, critical lures as a function of valence and retention interval.

	Neutral								Negative							
	d′				C				d′				C			
Retention interval	Immediate		1 week		Immediate		1 week		Immediate		1 week		Immediate		1 week	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Generate																
Critical lure	1.22	.90	1.09	.87	.81	.40	.24	.60	1.07	.86	.74	.61	.43	.60	-.39	.46
List item	2.50	.61	1.37	.76	.18	.35	.10	.58	2.21	.77	.93	.62	-.14	.40	-.49	.45
Read																
Critical lure	1.06	.81	1.07	.81	.90	.49	.25	.56	.91	.81	.30	.69	.51	.65	-.17	-.43
List item	1.75	.68	1.06	.86	.55	.44	.25	.52	1.59	.90	.44	.64	.17	.51	-.24	.56

again no difference for neutral critical lures across format ($p = .87$), susceptibility to negative critical lures was higher in the generate vs. read condition ($p < .001$).

For criterion *C*, responses were more liberal for negative compared to neutral list items, $F(1, 85) = 93.86$, $p < .001$, $\eta_p^2 = .53$, for the generate compared to read condition, $F(1, 85) = 96.45$, $p < .001$, $\eta_p^2 = .53$, and after a one-week compared to immediate retention, $F(1, 85) = 10.93$, $p = .001$, $\eta_p^2 = .11$. There was a significant interaction between Format × Retention Interval, $F(1, 85) = 6.46$, $p = .01$, $\eta_p^2 = .07$ and Valence × Retention Interval, $F(1, 85) = 4.25$, $p = .04$, $\eta_p^2 = .05$, the pattern of results mirrors the main effects for format and retention. For critical lures, responses bias was more liberal for negative compared to neutral valence, $F(1, 85) = 61.02$, $p < .001$, $\eta_p^2 = .42$, for the generate compared to read, $F(1, 85) = 10.74$, $p = .002$, $\eta_p^2 = .11$, and after a one-week retention, $F(1, 85) = 54.77$, $p < .001$, $\eta_p^2 = .39$. Similar to discrimination analysis, there was a 3-way interaction, $F(1, 85) = 3.69$, $p = .06$, $\eta_p^2 = .04$, which implied that there was no response bias differences at immediate test ($p^{negative} = .15$ and $p^{neutral} = .16$), but after one-week, there was more liberal responding to negative valence critical lures in the generate vs. read condition ($p < .001$) but no difference for neutral critical lures, ($p = .87$), see Table 2.

The signal detection analysis indicates that discrimination between target (or critical lure) and distractor was better for immediate compared to delayed testing, for negative compared to neutral stimuli and for generate vs. read. Once broken down, the three-way interaction for critical lures, mirrored a similar finding in the recognition rates. There was an expected increase in liberal responding for negative compared to neutral valanced items and an increase in liberal responding after a period of delay. These findings are further reviewed in the discussion.

Discussion

The present experiment explored the effect of generative processing on false memories. The purpose of this study was not only to attempt a replication of the original findings from Soraci et al. (2003), but also to examine whether there are conditions where a generation effect could increase false memories for critical lures in the DRM paradigm. For correct recognition, we saw a robust generation advantage at both retention intervals and for neutral and negative list items. We also replicated Soraci et al.'s findings in that generative processing did not significantly increase false recognition for the immediate test condition and this was the case for both valence types. These findings differ from the *more is less* pattern found by Toglia et al. (1999) and Thapar and McDermott (2001) in which manipulations that facilitated correct recollection (i.e. levels of processing and blocked thematic presentation) also increased false recollection.

Importantly, this outcome changed when the retention interval was delayed. After one week, there was still no difference in false recognition between generate and read for neutral stimuli, but false recognition was higher in the generate compared to read condition for negative stimuli. Therefore, for negative valence, after delay, we see the generation advantage for false recognition that is typically seen for correct recognition. Our findings appear to suggest that the distinctiveness effect caused by generative processing may drive down our susceptibility to false recognition at immediate test, but over time that distinctiveness decays. We argue that negative valanced lists have benefited from enhanced relational processing compared to neutral lists and that this could explain the increase in false recognition that we see.

To expand, we hypothesised that negative false memories may benefit from more over a delay from the generation effect because of the nature of negative stimuli and the way they are encoded, stored, and retrieved. We can turn to memory-based accounts of false memory production to explain these findings. We typically see an enhanced negative-emotional false memory effect which has been attributed to a greater reliance on relational information or meaning-based representations for valenced stimuli (Howe et al., 2010) and fewer overlapping theme nodes (Otgaar et al., 2016) leading to easier associative activation or gist formation (Bookbinder & Brainerd, 2016). Over time, the richer negative stimuli foment which leads to an increase in false memories. In addition, we argue that such rich stimuli likely allow for better relational processing during generation. Research has shown that there are conditions under which generative processing can enhance inter-target processing. For example, if the study list is structured such that information from earlier targets is useful in the generation of subsequent targets, that is they are semantically related, then inter-target relational processing will be enhanced. This occurs because when target items are members of a common category, then processing their common characteristics assists generation (e.g. McDaniel et al., 1988). If negative-emotional lists are more semantically dense, we hypothesise that generative processing will aid the activation of semantic associates and process common category or associative information. Although item-specific detail decays for both neutral and negative stimuli, increasing false recognition overall after a delay, enhanced associative/relational processing during generation for negative lists may explain the presence of the generation effect for negative but not neutral false memories at the one-week retention interval.

We should note that false alarms to list distractors were high for negative valence matched items compared to neutral. This drives down memory sensitivity measures for negative stimuli, although still showing the same generation pattern in the data similar to raw recognition responses. This finding appears to be emerging in the emotional false memory literature that measures recognition (e.g. Budson et al., 2006; Hellenthal et al., 2019; Howe et al., 2010). Participants appear to be less willing to reject negative-emotional stimuli at the test. It is likely that the semantically cohesive inter- and intra-list items make it difficult to discriminate information that has a high familiarity level. This can be seen in more liberal responding to negative stimuli, especially in the generation condition, where relational processing appears to be more enhanced. Indeed, the previous research has suggested that our response bias does shift for negative stimuli, where items are more densely organised and thus harder to differentiate at retrieval (see Hellenthal et al., 2019; Howe et al., 2010). Understanding threshold settings and their impact on false recognition for emotional stimuli is an important avenue for future research, especially if we are more willing to accept information that has a high level of emotional salience attached.

To summarise, Soraci et al. (2003) suggest that generative processing may be qualitatively different from other encoding manipulations known to enhance false memory formation. Instead, generation has a robust effect on recognition because it preserves the distinctiveness of list items during acquisition through enhanced item-specific processing. This, in turn, supports the distinctiveness of presented over non-presented items at retrieval. After long retention intervals, item-specific processing fades but relational detail is consolidated. Because we know negative stimuli are richly represented, having many densely integrated semantic associates, such stimuli aid generation through inter-target processing. After delay, it is this processing that wins out, increasing false memories for negative-emotional stimuli. The results of this study provide insights into the importance of relational or meaning-based processing for negative-emotional stimuli especially over long-term retention.

Note

1. Recollective experience (remember/know/guess judgements) was analysed using the same 2(Format: generate vs. read) × 2(Valence: negative vs. neutral) × 2(Retention Interval: immediate vs. delay) mixed factor analysis of variance (ANOVA) used to analyse old recognition rates. Although Soraci et al. (2003) did not use a recollective experience measure, this is common in DRM studies measuring recognition (Roediger & McDermott, 1995) to demonstrate a strong subjective experience of remembering the false item. Analyses associated with remember responses revealed a similar patten to those findings presented in the old recognition analysis. *Remember responses*: For list items, *remember* responses were higher for negative vs. neutral hits, $F(1, 85) = 18.63$, $p < .001$, $\eta_p{}^2 = .18$, and for generate vs. read, $F(1, 85) = 82.38$, $p < .001$, $\eta_p{}^2 = .49$. There was a retention interval main effect, $F(1, 85) = 7.49$, $p = .008$, $\eta_p{}^2 = .08$, and one significant Format × Retention Interval interaction, $F(1, 85) = 7.26$, $p = .009$, $\eta_p{}^2 = .08$, pairwise comparisons indicated that, for remember responses, the generation effect was only present at immediate ($p < .001$), and not 1-week delay ($p = .27$), albeit in the same direction. For critical lures, *remember* responses associated with false recognition were higher for generate vs. read, $F(1, 85) = 8.33$, $p = .005$, $\eta_p{}^2 = .09$ and higher after one-week, $F(1, 85) = 14.90$, $p < .001$, $\eta_p{}^2 = .15$. There was no significant main effect of valence, $F(1, 85) = 3.18$, $p = .08$, $\eta_p{}^2 = .04$ but the three-way interaction was significant, $F(1, 85) = 4.14$, $p < .05$, $\eta_p{}^2 = .05$. When separated by retention interval, Bonferroni pairwise comparisons showed no differences between format for neutral and negative valence ($p = .86$ and $p = .11$) at immediate test, however, after one-week, there were more remember responses associated with false recognition in the generate condition compared to the read condition for negative valence only ($p < .001$). There was no format difference for neutral valence ($p = .56$). For false alarms to related, $F(1, 85) = 10.42$, $p = .002$, $\eta_p{}^2 = .11$ and unrelated, $F(1, 85) = 29.65$, $p < .001$, $\eta_p{}^2 = .26$ distractors there were more *remember* responses for negative vs. neutral valence, and more for immediate versus one-week ($F(1, 85) = 13.42$, $p < .001$, $\eta_p{}^2 = .14$ and $F(1, 85) = 30.57$, $p < .001$, $\eta_p{}^2 = .26$, respectively). There was a Valence X Retention Interval interaction for unrelated distractors, $F(1, 85) = 12.70$, $p < .001$, $\eta_p{}^2 = .13$, and pairwise comparisons indicated that the valence difference occurred at the one week interval only ($p = .009$ vs. $p = .13$). *Know responses*: For list items, there were no significant main effects or interactions (all $ps > .05$). *Know* responses associated with critical lure false recognition were higher for negative vs. neutral valence, $F(1, 85) = 13.78$, $p < .001$, $\eta_p{}^2 = .14$, and higher after one-week, $F(1, 85) = 6.06$, $p = .02$, $\eta_p{}^2 = .07$. There was one significant two-way Valance × Format interaction, $F(1, 85) = 6.03$, $p = .02$, $\eta_p{}^2 = .07$. Bonferroni pairwise comparisons revealed no difference in format for neutral valence ($p = .46$) but more know responses in the generate vs. read condition for negative valence critical lures. This did not interact with retention interval. For false alarms to related, $F(1, 85) = 8.99$, $p = .004$, $\eta_p{}^2 = .10$ and unrelated, $F(1, 85) = 62.17$, $p < .001$, $\eta_p{}^2 = .42$ distractors there were more *know* responses for negative vs. neutral valence, and more for immediate versus one-week ($F(1, 85) = 36.49$, $p < .001$, $\eta_p{}^2 = .30$ and $F(1, 85) = 46.85$, $p < .001$, $\eta_p{}^2 = .36$, respectively). There was a Valence × Retention Interval interaction for unrelated distractors, $F(1, 85) = 16.82$, $p < .001$, $\eta_p{}^2 = .17$, but pairwise comparisons revealed same patterns as the main effects (all $ps < .05$). *Guess responses*: For list items, although more guessing were made for the read compared to generate condition, $F(1, 85) = 12.09$, $p < .001$, $\eta_p{}^2 = .13$, there were no other significant main effects or interactions (all $ps > .05$). For critical lures, there were no significant main effects or interactions (all $ps > .05$). For related distractors, guessing was higher after one week delay, $F(1, 85) = 12.58$, $p < .001$, $\eta_p{}^2 = .13$, but no main effect of valence or significant interaction ($ps > .05$). For unrelated distractors, guessing was higher for negative items, $F(1, 85) = 12.59$, $p < .001$, $\eta_p{}^2 = .13$, and after one week delay, $F(1, 85) = 7.91$, $p = .006$, $\eta_p{}^2 = .09$, but no interaction, $F(1, 85) = .52$, $p = .47$, $\eta_p{}^2 = .01$ (see Table 1 for all descriptive statistics).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by a project grant from the Economic and Social Research Council, ESRC grant ref: ES/L00853X/1.

References

Arndt, J., & Hirshman, E. (1998). True and false recognition in MINERVA2: Explanation from a global matching perspective. *Journal of Memory and Language*, *39*(3), 371–391. https://doi.org/10.1006/jmla.1998.2581

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

Bradley, M. M., & Lang, P. J. (1999). *Affective Norms for English Words (ANEW): Instruction manual and affective ratings*.

Technical report C-1, The Center for Research in Psychophysiology, University of Florida.

Brainerd, C. J., 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*, 137–138.

Brainerd, C. J., & Reyna, V. F. (2001). Fuzzy-trace theory: Dual processes in memory, reasoning, and cognitive neuroscience. In H. W. Reese & R. Kail (Eds.), *Advances in child development and behavior* (pp. 41–100). Academic Press. https://doi.org/10.1016/S0065-2407(02)80062-3

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*(2), 71–78. https://doi.org/10.1097/01.wnn.0000213905.49525.d0

Chechile, R. A., & Soraci, S. A. (1999). Evidence for a multiple-process account of the generation effect. *Memory (Hove, England)*, *7*(4), 483–508. https://doi.org/10.1080/741944921

Colbert, J. M., & McBride, D. M. (2007). Comparing decay rates for accurate and false memories in the DRM paradigm. *Memory & Cognition*, *35*(7), 1600–1609. https://doi.org/10.3758/BF03193494

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

Dewhurst, S. A., & Anderson, S. J. (1999). Effects of exact and category repetition in true and false recognition memory. *Memory and Cognition*, *27*, 665–673.

Gardiner, J. M., & Hampton, J. A. (1988). Item-specific processing and the generation effect: Support for a distinctiveness account. *American Journal of Psychology*, *101*(4), 495–504. https://doi.org/10.2307/1423227

Hellenthal, M. V., Knott, L. M., 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

Hirshman, E., & Bjork, R. A. (1988). The generation effect: Support for a two-factor theory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *14*(3), 484–494. https://doi.org/10.1037/0278-7393.14.3.484

Howe, M. L. (2005). Children (but not adults) can inhibit false memories. *Psychological Science*, *16*(12), 927–931. https://doi.org/10.1111/j.1467-9280.2005.01638.x

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

McDaniel, M. A., Wadill, P. J., & Einstein, G. O. (1988). A contextual account of the generation effect: A three-factor theory. *Journal of Memory and Language*, *27*(5), 521–536. https://doi.org/10.1016/0749-596X(88)90023-X

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

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

Roediger, H. L., & McDermott, B. K. (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

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, *4*(6), 592–604. https://doi.org/10.1037/0278-7393.4.6.592

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, *117*(1), 34–50. https://doi.org/10.1037/0096-3445.117.1.34

Soraci, S. A., Carlin, M. T., Toglia, M. P., Chechile, R. A., & Neuschatz, J. S. (2003). Generative processing and false memories: When there is no cost. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(4), 511–523. https://doi.org/10.1037/0278-7393.29.4.511

Soraci, S. A., Jr., Franks, J. J., Bransford, J. D., Chechile, R. A., Belli, R. F., Carr, M., & Carlin, M. T. (1994). Incongruous item generation effects: A multiple-cue perspective. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(1), 1–12. https://doi.org/10.1037/0278-7393.20.1.67

Stadler, M. A., Roediger, H. L., & McDermott, I. I. I., & B, K. (1999). Norms for word lists that create false memories. *Memory & Cognition*, *27*(3), 494–500. https://doi.org/10.3758/BF03211543

Thapar, A., & McDermott, K. B. (2001). False recall and false recognition induced by presentation of associated words: Effects of retention interval and level of processing. *Memory & Cognition*, *29*(3), 424–432. https://doi.org/10.3758/BF03196393

Toglia, M. P., Neuschatz, J. S., & Goodwin, K. A. (1999). Recall accuracy and illusory memories: When more is less. *Memory (Hove, England)*, 7(2), 233–256. https://doi.org/10.1080/741944069