

Reduced associative memory for negative information: impact of confidence and interactive imagery during study

Jeremy B. Caplan^{a,b,1}, Tobias Sommer^{b,1}, Christopher R. Madan^{a,b,c}, Esther Fujiwara^{a,b,*}

^a University of Alberta, Edmonton, Canada

^b University Medical Center Hamburg-Eppendorf, Hamburg, Germany

^c University of Nottingham, Nottingham, United Kingdom

¹: authors contributed equally

* Correspondence

Esther Fujiwara

University of Alberta

Department of Psychiatry

1E1 WMC Health Sciences Centre

Edmonton, AB

CANADA T6G 2V2

efujiwara@ualberta.ca

Phone: +1 780 492 4104

Fax: +1 780 492 6841

Running head: Emotion reduces associative memory

Word count: 4656

24 references

Keywords: Memory, emotion, paired associate learning, imagery, metacognition

Abstract

Although item-memory for emotional information is enhanced, memory for associations between items is often impaired for negative, emotionally arousing compared to neutral information. We tested two possible mechanisms underlying this impairment, using picture pairs: 1) higher confidence in one's own ability to memorise negative information may cause participants to under-study negative pairs; 2) better interactive imagery for neutral pairs could facilitate associative memory for neutral pairs more than for negative pairs. Tested with associative recognition, we replicated the impairment of associative memory for negative pairs. We also replicated the result that confidence in future memory (judgments of learning) was higher for negative than neutral pairs. Inflated confidence could not explain the impairment of associative recognition memory: Judgements of learning were positively correlated with associative memory success for both, negative and neutral pairs. However, neutral pairs were rated higher in their conduciveness to interactive imagery than negative pairs, and this difference in interactive imagery showed a robust relationship to the associative memory difference. Thus, associative memory reductions for negative information are not due to differences in encoding effort. Instead, interactive imagery may be less effective for encoding of negative than neutral pairs.

Introduction

Emotional arousal enhances item-memory and the underlying cognitive and neural processes have been well characterized (Talmi, 2013). Conversely, emotional arousal can increase or decrease associative memory. For example, emotionally arousing stimulus features may narrow attention and increase memory for such features, at the expense of memory for their context, which can cause a net-decrease in associative memory (Kensinger, 2009). However, arousal can also increase associative memory, e.g., if the to-be-associated features act as a single, unitized item (within-object or intrinsic associations), compared to arousal-based reductions for between-item/extrinsic associations (Kensinger, 2009; Mather, 2007). Arousal-Biased Competition (ABC) Theory (Mather & Sutherland, 2011) suggests that emotional arousal will generally enhance memory for prioritized parts of an association at the expense of non-prioritized parts, irrespective if prioritization is driven by the stimulus layout or task demands. Using a verbal paired-associates task in which memorizing associations was explicitly instructed, but the stimulus layout was between-item, Madan, Caplan, Lau, and Fujiwara (2012) demonstrated that item-memory for arousing words was enhanced (relative to neutral words), but their associative binding was impaired. These verbal memory findings also extend to pictures (Bisby & Burgess, 2014).

Probing the neural mechanisms underlying this effect, Madan, Fujiwara, Caplan, & Sommer (2017) observed brain activity in extrahippocampal medial temporal lobe regions (e.g., entorhinal cortex) linked to successful neutral, but not negative-pair encoding. This was interpreted as spontaneous unitization: merging between-item associations to operate like a single item and therefore rendering successful encoding hippocampus-independent (Mayes, Montaldi, & Migo, 2007). Localization to entorhinal cortex suggested a potential role of mental imagery in this hypothetical unitization process. Interactive imagery—forming a mental image that combines elements of an association—is known to trigger unitization, and can occur

spontaneously (Ahmad & Hockley, 2014). Reduced associative memory for negative pairs could result from less effective spontaneous interactive imagery. Findings from *instructed* interactive imagery studies show that self-reported imagery success is greater for neutral than negative words pairs - although interactive imagery may require more time for neutral than negative pairs (Murray & Kensinger, 2012). That is, interactive images may emerge more quickly but not more successfully for negative pairs. Complementing these findings, reduced associative memory for negative word pairs can be prevented by instructed interactive imagery (Han, Mao, Kartvelishvili, Li, & Guo, 2018). Perhaps, spontaneous and successful interactive imagery could underlie the associative memory advantage for neutral pairs.

Results of Zimmerman and Kelley (2010) were suggestive of another, meta-cognitive mechanism underlying the associative memory reduction. Their participants encoded emotional or neutral word pairs and made judgements of learning (JOL), estimating, for each pair, how likely they would remember it in an association-memory test (cued-recall). Associative memory was better for neutral than for negative pairs, but JOLs were higher for negative than for neutral pairs and less valid predictors of actual memory success. JOLs can diverge with different materials, but their validity for memory is a separable observation. Rhodes and Castel (2008) reported greater JOLs for semantically related than unrelated word-pairs and these were linked to better cued-recall accuracy. However, greater JOLs for large-font than small-font word-pairs were unrelated to later cued-recall accuracy. Thus, JOLs can correctly or erroneously be influenced by item-properties. Subjectively harder to learn (low-JOL) materials can trigger additional study effort (as seen in self-pacing times; Koriat & Bjork, 2006; Miele, Finn, & Molden, 2011), which can then equalize accuracy differences between pair types. By this logic, Zimmerman and Kelley's findings (2010) could imply that the lower-JOL neutral pairs drew more study effort and the higher-JOL negative pairs drew less study effort, resulting in a net- memory reduction for negative pairs. Thus, high JOLs for negative

pairs may cause people to understudy them. We aimed to test whether either or both of these two hypothesized mechanisms may underlie associative memory reductions for negative pictures.

Methods

Participants

We recruited 82 participants through advertisements in a local online job database. One participant providing only four memory responses was excluded. The final group contained 81 participants (17 males; age mean±standard deviation: 24.73±4.33 years, range: 18-39 years). Participants gave written informed consent and received 10 €/h for their participation. The study was approved by the local ethics committee, Board of Physicians, Hamburg, Germany.

Materials

For each participant, we used a set of 208 randomly selected pictures from Madan et al. (2017), 104 negative, arousing pictures and 104 neutral, non-arousing pictures. An independent group of 43 raters (20 males) judged arousal-levels of each picture on 9-point Self-Assessment-Manikin scales, with '9' indicating low arousal. As intended, negative pictures were rated more arousing (4.72±0.82) than neutral pictures (7.29±0.32; $t(207.87, \text{Levene's correction})=37.09$, $p<.001$, Cohen's $d=4.13$)¹. Pictures were split into List 1 and List 2, each containing 52 pairs (26 negative and 26 neutral pairs).

¹ Valence ratings, with '9' indicating positive valence, were available from 23 female raters only (Madan et al., 2018), covering all neutral pictures and 156 of the 160 possible negative pictures; valence ratings also differed as intended (negative: 3.11±0.60; neutral: 5.18±0.40; $t(266.49)=35.94$, $p<.001$, $d=3.79$). These ratings are not further detailed here.

Task

The task was kept as similar as possible to our previous fMRI experiment (see Exp. 3, Madan, et al., 2017). Each part (encoding, retrieval, imagery ratings) was preceded by five practice trials (not analysed).

Encoding and Judgments of Learning (JOLs): Both lists started with an encoding phase in which participants were explicitly asked to study the pairings and informed that their memory for each pair would be tested later. Participants provided JOLs after each pair in List 2, selecting the likelihood of later remembering the pair using a visual analogue scale (VAS) with labeled anchors ranging from 0% to 100% (Fig. 1A). JOLs were only collected in List 2 to ensure participants had experience with the associative memory task, which is necessary for memory-informed JOLs (Zimmerman & Kelley, 2010).

Figure 1

Retrieval: The two retrieval phases (Fig. 1B) each consisted of 52 trials and first tested each pair with a judgement of memory (JoM) task, to emulate cued recall, and then with a 5-AFC associative-recognition task (chance-level: 20%). During JoM, either the left or the right picture of the pair was presented and participants indicated (yes/no) whether they recalled the previous associate. For the 5-AFC associative-recognition task, the same probe picture was presented in the center of the screen, surrounded by five pictures (one target, four lures). Participants were to choose the target. Lures were always had the same valence as the target and were from the just preceding study phase. Each picture was repeated three times: Once as target and twice as lure. Lures were pseudorandomly selected such that all five recognition alternatives always had a ratio of 2:3 or 3:2 negative to neutral pictures.

Baseline and two-back tasks: Two motor baseline trials were presented after each trial and a pictorial two-back task was used to disrupt rehearsal prior to each encoding and retrieval phase (details in Madan et al., 2017). Performance in these tasks was irrelevant to the current study and not analysed.

Interactive imagery ratings: At the end of the experiment, using a VAS, participants indicated how easily they could imagine the two pictures of each pair interacting with each other (Fig. 1C), providing also the following explicit example (in German): “If one picture shows a balloon and the other picture shows a hand, an interactive image could be the hand holding the string of the balloon”. The VAS had labeled anchors ranging from “very difficult”, “medium”, to “very easy”. For analysis, VAS ratings were scaled between 0 and 100 (left or right anchor, respectively). While 39 participants judged all 104 original pairs, 41 participants judged 52 original pairs and 52 recombined pairs, with pair recombination within valence and list but not between lists. Intact or recombined pairs were presented in random order. Since interactive imagery had to be rated after retrieval to avoid influencing participants’ study strategy, ratings for recombined pairs (i.e., pairs that were never learned), allowed us to infer interactive imagery underlying (rather than following) differences in associative memory for negative and neutral pairs.

Results

Associative Memory

Subjective retrieval success (JoM) was statistically similar for negative and neutral pairs in both lists (not shown), unrelated to objective associative memory performance (5-AFC), and hence not further analysed. To test the expected associative memory reduction for negative

pairs (Madan et al., 2017), a repeated-measures ANOVA on 5-AFC associative-recognition accuracy (Fig. 2A) and factors List (1, 2) and Emotion (negative pairs, neutral pairs) showed main effects of Emotion, $F(1,80)=19.84$, $p<.001$, $\eta_p^2=.20$, and List, $F(1,80)=20.99$, $p<.001$, $\eta_p^2=.21$. Associative recognition was lower for negative than neutral pairs and for List 1 than List 2. Main effects were qualified by a List x Emotion interaction, $F(1,80)=5.08$, $p=.027$, $\eta_p^2=.060$. Simple effects showed that the associative memory reduction for negative (minus neutral) pairs was more pronounced in List 1 than in List 2, $t(80)=2.26$, $p=.027$, $d=0.25$ ($M_{List1}=-0.08\pm0.16$; $M_{List2}=-0.03\pm0.13$). However, in both lists, association memory was significantly lower for negative compared to neutral pictures (List 1: $M_{negative}=0.47\pm0.18$, $M_{neutral}=0.55\pm0.20$; $t(80)=4.43$, $p<.001$, $d=0.49$; List 2: $M_{negative}=0.57\pm0.19$, $M_{neutral}=0.61\pm0.23$; $t(80)=2.43$, $p=.017$, $d=0.27$; Fig. 2A). Differences in negative/neutral memory did not change over the course of the task (see Supplemental).

Figure 2

Confidence: Judgments of Learning (JOLs) and Associative Memory

Consistent with Zimmerman and Kelley (2010), JOLs were greater for negative than neutral pairs ($t(80)=3.46$, $p<.001$, $d=0.38$), despite the associative memory reduction. To test the validity of the JOLs for associative memory success within-subjects, a repeated-measures ANOVA was conducted on JOLs (Fig. 2B), with factors Emotion (negative, neutral) and Memory (5-AFC: correct, incorrect). Outcomes revealed main effects of Emotion, $F(1,78)=14.1$, $p<.001$, $\eta_p^2=.15$ and Memory, $F(1,78)=121.2$, $p<.001$, $\eta_p^2=.61$, but no significant interaction, $F(1,78)=3.38$, $p=.070$, $\eta_p^2=.04$. Participants were more confident in their memory for negative compared to neutral pairs. This difference in JOLs did not change over the course of the task (see Supplemental). Participants were also more confident for later correctly retrieved pairs compared to incorrect pairs, suggesting validity of the JOLs, and the lack of a significant interaction suggested comparable validity for both pair types.

Between-subjects, JOLs correlated with associative memory accuracy for neutral pairs, $r(79)=.35$, $p=.001$, but not for negative pairs, $r(79)=.17$, $p=.12$, suggesting reduced coupling between memory-confidence and memory-accuracy for negative pairs. Lower validity of JOLs for negative pairs could therefore imply that participants are understudying negative pairs, reducing later associative memory. The difference in accuracy (5-AFC: neutral minus negative) should then correlate negatively with the difference in JOLs (neutral minus negative). However, critically, the correlation was *positive*, $r(79)=.44$, $p<.001$ (List 2) (see Supplemental for a similar correlation to List 1 accuracy differences). The difference in JOLs was also unrelated to the *change* in associative memory difference from List 1 to List 2, $r(79)=-.017$, $p=.88$. Therefore, participants' differences in confidence for learning negative versus neutral pairs accurately reflected differences in their later associative memory for these pairs.

To test whether differences in JOLs were driven by differences in arousal rather than valence, we used arousal ratings from an independent sample (see Methods) and correlated them with the JOLs. After Fisher z-transformation of the correlations, t-tests against zero were conducted. With a higher score indicating lower arousal, the correlation was negative for negative pairs, $t(80)=3.23$, $p=.002$, $d=0.36$, indicating that participants guessed they were more likely to remember higher-arousing negative pairs. For neutral pairs, the correlation was positive, $t(79)=3.65$, $p<.001$, $d=0.41$, suggesting higher arousal decreased confidence in future memory for neutral pairs.

Interactive Imagery and Associative Memory

To test whether higher interactive imagery explained better memory for neutral pairs, a three-way ANOVA was conducted on imagery ratings (see Fig. 2C), with factors List, Emotion, and Memory. A main effect of Memory, $F(1,72)=200$, $p<.001$, $\eta_p^2=.74$, showed substantially higher imagery ratings for correct than incorrect pairs ($M_{\text{correct}}=57.59\pm 13.56$; $M_{\text{incorrect}}=35.84\pm 14.69$).

While the main effect of Emotion was not significant, $F(1,72)=1.43$, $p=.24$, $\eta_p^2=.019$, pertinent to our hypothesis, an Emotion by Memory interaction was observed, $F(1,72)=8.06$, $p=.01$, $\eta_p^2=.10$. For correctly remembered pairs, imagery ratings were higher for neutral ($M_{\text{neutral}}=60.02\pm 15.72$) than negative pairs ($M_{\text{negative}}=55.14\pm 14.13$), $t(72)=2.79$, $p=.01$, $d=0.33$. No difference in ratings emerged for incorrect pairs, $t(72)=0.70$, $p=.49$, $d=0.08$. Thus, imagery ratings were substantially higher for remembered than for forgotten pairs overall, and this effect was stronger for neutral than negative pairs. Other significant effects were a main effect of List, $F(1,72)=27.6$, $p<.001$, $\eta_p^2=.28$, and a List by Memory interaction, $F(1,72)=23.1$, $p<.001$, $\eta_p^2=.24$. For correct pairs only, imagery ratings were higher on List 2 ($M_{\text{correct}}=62.70\pm 14.72$) than List 1 ($M_{\text{correct}}=52.47\pm 14.70$), $t(72)=6.28$, $p<.001$, $d=0.73$.

Since imagery ratings were acquired after the memory test, better memory for neutral pairs could have been the cause rather than consequence of higher interactive imagery. To address this ambiguity, we analysed imagery ratings for recombined pairs, pictures that were never encoded together and therefore uncontaminated by participants' own memory (acquired in $N=41$). Neutral recombined pairs were rated higher in interactive imagery than negative recombined pairs, $t(40)=2.14$, $p=.038$, $d=0.33$, supporting our interpretation that interactive imagery may have promoted better associative memory for neutral pairs.

Imagery ratings were then correlated with associative memory accuracy, showing positive and similarly sized correlations for both pair types and lists (List 1: $r_{\text{negative}(79)}=.44$, $p<.001$, $r_{\text{neutral}(79)}=.58$, $p<.001$; $Z_{\text{diff}}=1.19$, $p=.23$. List 2: $r_{\text{negative}(79)}=.43$, $p<.001$, $r_{\text{neutral}(79)}=.56$, $p<.001$; $Z_{\text{diff}}=1.08$, $p=.28$). The difference in interactive imagery ratings between negative and neutral pairs was related to the difference in associative memory, $r(79)=.46$ and $r(79)=.36$ for Lists 1 and 2, respectively, both $p<.001$. Thus, participants who judged interactive imagery for neutral pairs higher than for negative pairs also had better memory for neutral than negative pairs.

Arousal ratings were positively correlated with interactive imagery ratings ($t_{\text{negative}(79)}=4.45$, $p<.001$, $d=0.50$; $t_{\text{neutral}(79)}=3.44$, $p=.001$, $d=0.38$), indicating that participants judged interactive images easier to form for less arousing pairs, regardless of valence.

Hierarchical regression

The relative contribution of JOLs and interactive imagery ratings to explain the associative memory advantage for neutral over negative pairs (5-AFC accuracy: neutral minus negative, over both lists) was then assessed by hierarchical regressions. JOLs were predictors in the first block and interactive imagery ratings (averaged across both lists) were used in the second block. Both models were significant (model 1: $R^2=.31$; $F(2,78)=17.47$, $p<.001$; model 2: $R^2=.39$; $F(4,76)=12.31$, $p<.001$), as was the increase in variance explanation by adding imagery ratings in block 2 ($\Delta R^2=.08$, $p=.007$). In the final model 2, better associative memory for neutral than negative pairs was reflected in higher JOLs ($\beta=.51$, $p=.001$) and higher interactive imagery ratings ($\beta=.44$, $p=.004$) for neutral pairs, but lower JOLs ($\beta=-.29$, $p=.04$) and lower interactive imagery ratings for negative pairs ($\beta=-.40$, $p=.004$). Reversing the order of the two blocks showed similar outcomes (not shown).

Discussion

Associative memory for negative pairs was worse than for neutral pairs, replicating previous results (Bisby & Burgess, 2014; Bisby, Horner, Horlyck, & Burgess, 2016; Madan et al., 2012; Madan, et al., 2017; Zimmerman & Kelley, 2010). Based on previous findings of relationships between JOLs, allocation of study resources, and later memory (e.g., Koriat & Bjork, 2006), we tested whether participants' high confidence to remember negative pairs may trigger less study effort, resulting in lower associative memory for negative pairs. Clearly, our results refute this interpretation. Even though participants' confidence in memorizing negative pairs was

inflated, JOLs and interactive imagery ratings accurately reflected performance, regardless of valence: Better associative memory for neutral than negative pairs correlated positively with JOLs (and imagery ratings) for neutral pairs, but negatively for negative pairs.

What could be the origin of higher JOLs for emotional materials? Previous item-memory studies using positive and neutral pictures matched in arousal (Tauber, Dunlosky, Urry, & Opitz, 2017) or semantic relatedness (Exp. 2 of Hourihan & Bursey, 2017) also showed higher JOLs for positive than neutral pictures. Hourihan, Fraundorf, and Benjamin (2017) showed similar results for arousal-matched words: negative words, regardless of arousal, evoked higher JOLs than neutral words. Thus, valence-based, semantic features rather than arousal-related properties increased JOLs for emotional materials in these studies. These findings were interpreted as reflecting participants' explicit, conscious belief of having better memory for materials belonging to an 'emotional' semantic category, leading to inflated JOLs. The current study was not designed to disambiguate influences of valence from those of arousal on JOLs and associative memory. However, we observed an interesting interaction between arousal ratings and picture valence on JOLs. For negative pairs, arousal ratings (acquired in a different cohort) were *positively* correlated to the negative pairs' JOLs, but *negatively* correlated to the neutral pairs' JOLs. To speculate, if valence were the primary origin of differences in JOLs for emotional versus neutral materials (Hourihan & Bursey, 2017; Hourihan et al., 2017; Tauber et al., 2017), higher arousal in negative pictures may corroborate their perception as belonging to a 'negative' category and relate positively to their JOLs as we observed. By this logic, higher arousal in neutral pictures would be incongruent with their perception as 'neutral' and therefore relate negatively to their JOLs.

JOLs were reliably correlated to memory accuracy for neutral, but not negative pairs, implying lower validity of the JOLs for negative associative memory (Zimmerman & Kelley, 2010), resembling previous item-memory findings (Hourihan & Bursey, 2017; Hourihan et al.,

2017; Tauber et al., 2017). However, the critical question here was whether increased JOLs for negative pairs could result in less encoding and thereby cause the associative memory *reduction* for negative pairs. This was not the case: Participants who were more confident in retrieving neutral than negative pairs also showed better memory for neutral than negative pairs. Conversely, participants with the strongest memory reduction for negative pairs showed the least increase of JOLs for negative over neutral pairs. Thus, greater memory confidence for negative picture pairs could not have caused less encoding and subsequently reduced memory for these pairs.

Associative memory reductions for negative pairs were less pronounced in List 1 than List 2, which was not expected and could point to possible changes in the task due to the presence of JOL ratings in List 2. Methodologically, JOLs would not have been interpretable if made naively (cf., Zimmerman & Kelley, 2010), hence their inclusion only in List 2. As Koriat and Bjork (2006) argued, so-called foresight bias in JOLs can be reduced when participants have experience with the task. Although we cannot directly speak to the mechanism within a single experiment, the JOLs could have subtly changed the attentional focus in List 2, dampening the impact of negative pairs in List 2 due to the concurrent JOL task (although habituation was similar in both lists; see Supplemental). Thus, focusing participants on evaluating their ability to encode the pairs may have directed (some of) their attention away from potential attention-seizing features of the negative pairs which could have made associative memory-reducing effects for negative pairs smaller in List 2 than List 1 (cf. Mather & Sutherland, 2011). Interestingly, Maddox et al. (2012) observed that associative encoding of neutral items (words) requires more controlled attention than encoding of negative pairs and that dividing attention during encoding may allow automatic associative encoding of emotional pairs. Assuming that the JOLs in List 2 act like a secondary, attention-dividing task, the resulting net-reduction of negative associative memory should become weaker, as we observed.

Evidently, these interpretations are rather speculative and a third list may be a useful follow-up study to clarify the interactive effects of emotions, list, and JOL presence/absence.

Interactive imagery ratings were higher for neutral than negative remembered pairs but not forgotten pairs, suggesting that spontaneous interactive imagery may be more effective for neutral pairs and lead to superior associative memory. Never-studied (recombined) negative pairs were also rated lower in imagery than never-studied neutral pairs, addressing a potential confound in this experiment, i.e., differences in memory causing (rather than following) differences in imagery ratings. Previous studies have shown that *explicit* interactive imagery instructions can increase associative memory for neutral materials more effectively than for negative materials (word pairs: Murray & Kensinger, 2012; picture triplets: Bisby, Horner, Bush, & Burgess, 2018). Interestingly, Han et al. (2018) could eliminate the known associative memory reductions for negative (word) pairs by instructing unitization via interactive imagery and this effect was accompanied by recovery of memory-relevant electroencephalographic event-related potentials (parietal late positive potential LPP; frontal N400 wave). Thus, although speculative, these and the current findings suggest less effective *spontaneous* interactive imagery during study of negative compared to neutral pairs, which could then produce the observed net-reduction in associative memory for negative pairs. However, these studies used rather different materials and tasks, limiting their applicability to the current study. Furthermore, since interactive imagery ratings could not be acquired *during* encoding, the current study cannot finally answer whether participants engaged in such encoding strategy. Future studies may employ negative and neutral pairs matched in their conduciveness to interactive imagery. Associative memory reductions for imagery-matched negative pairs, relative to neutral pairs, may then be less ambiguously interpreted as differences in use of *spontaneous* interactive imagery. Finally, although visual differences between the picture types may have influenced the outcomes, objectively matching visual properties between pairs would

likely be insufficient, e.g., emotional pictures can be perceived as visually more complex than neutral pictures (see Madan et al., 2018).

In our previous study, extra-hippocampal medial temporal lobe activity promoted successful neutral, but not negative pair encoding, which instead required hippocampal activity (Madan et al., 2017). We had tentatively interpreted this as reflecting better unitization via interactive imagery during neutral pair encoding. Supporting this interpretation, here we observed higher interactive imagery ratings for neutral pairs correlated with superior memory for neutral pairs. Why would negative pairs be harder to unitize through interactive imagery? Individual negative items trigger a range of deeper processing mechanisms compared to neutral items (Markovic, Anderson, & Todd, 2014). For example, we had observed that negative pictures attract more narrow visual attention to themselves, preventing attention to their pairing (Madan et al., 2017). A substantial proportion of variance in the associative memory difference remained unexplained. Therefore, interactive imagery is only but one plausible mechanism underlying reduced associative memory for negative information. These valence-based differences in hypothesized spontaneous interactive imagery may further interact with differences in semantic relatedness between negative and neutral materials (Talmi, 2013), the perceived plausibility of the presented associations (Bisby et al., 2014), and other factors whose possible synergistic or antagonistic combined effects remain to be tested.

Conclusions

This study replicates and extends previous observations of an associative memory reduction for negative pictures compared to neutral pictures. Two potential mechanisms underlying this reduction were tested. Although judgements of learning for negative pairs were elevated, relative to neutral pairs and relative to memory accuracy, reduced negative associative memory was not predicted by exaggerated judgements of learning. Instead, more effective

spontaneous interactive imagery may underlie superior associative memory for neutral information.

Our findings also imply that reducing valence-based associative memory biases would unlikely benefit from metamemory training; while such training may render more accurate judgements of learning, their relationships with associative memory success would be unaffected. However, targeting interactive imagery might boost associative memory for negative information.

References

- Ahmad, F. N., & Hockley, W. E. (2014). The role of familiarity in associative recognition of unitized compound word pairs. *Quarterly Journal of Experimental Psychology*, *67*(12), 2301-2324. doi: 10.1080/17470218.2014.923007
- Bisby, J. A., & Burgess, N. (2014). Negative affect impairs associative memory but not item memory. *Learning and Memory*, *21*(1), 21-27. doi: 10.1101/lm.032409.113
- Bisby, J. A., Horner, A. J., Bush, D., & Burgess, N. (2018). Negative emotional content disrupts the coherence of episodic memories. *Journal of Experimental Psychology: General*, *147*(2), 243-256. doi: 10.1037/xge0000356
- Bisby, J. A., Horner, A. J., Horlyck, L. D., & Burgess, N. (2016). Opposing effects of negative emotion on amygdalar and hippocampal memory for items and associations. *Social, Cognitive, & Affective Neuroscience*, *11*(6), 981-990. doi: 10.1093/scan/nsw028
- Han, M., Mao, X., Kartvelishvili, N., Li, W., & Guo, C. (2018). Unitization mitigates interference by intrinsic negative emotion in familiarity and recollection of associative memory: Electrophysiological evidence. *Cognitive, Affective, and Behavioral Neuroscience*, *18*(6), 1259-1268. doi: 10.3758/s13415-018-0636-y
- Hourihan, K. L., & Bursey, E. (2017). A misleading feeling of happiness: metamemory for positive emotional and neutral pictures. *Memory*, *25*(1), 35-43. doi: 10.1080/09658211.2015.1122809
- Hourihan, K. L., Fraundorf, S. H., & Benjamin, A. S. (2017). The influences of valence and arousal on judgments of learning and on recall. *Memory and Cognition*, *45*(1), 121-136. doi: 10.3758/s13421-016-0646-3

- Kensinger, E. A. (2009). Remembering the Details: Effects of Emotion. *Emotion Review*, 1(2), 99-113. doi: 10.1177/1754073908100432
- Koriat, A., & Bjork, R. A. (2006). Illusions of competence during study can be remedied by manipulations that enhance learners' sensitivity to retrieval conditions at test. *Memory and Cognition*, 34(5), 959-972. doi: 10.3758/BF03193244
- Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, 1(4), 476-490. doi: 10.3758/BF03210951
- Madan, C. R., Bayer, J., Gamer, M., Lonsdorf, T. B., & Sommer, T. (2018). Visual complexity and affect: ratings reflect more than meets the eye. *Frontiers in Psychology*, 8, 2368. doi: 10.3389/fpsyg.2017.02368
- Madan, C. R., Caplan, J. B., Lau, C. S., & Fujiwara, E. (2012). Emotional arousal does not enhance associative memory. *Journal of Memory and Language*, 66(4), 695-716. doi: 10.1016/j.jml.2012.04.001
- Madan, C. R., Fujiwara, E., Caplan, J. B., & Sommer, T. (2017). Emotional arousal impairs associative memory: Roles of amygdala and hippocampus. *Neuroimage*, 156, 14-28. doi: 10.1016/j.neuroimage.2017.04.065
- Maddox, G. B., Naveh-Benjamin, M., Old, S., & Kilb, A. (2012). The role of attention in the associative binding of emotionally arousing words. *Psychonomic Bulletin & Review*, 19(6), 1128-1134. doi: 10.3758/s13423-012-0315-x
- Markovic, J., Anderson, A. K., & Todd, R. M. (2014). Tuning to the significant: neural and genetic processes underlying affective enhancement of visual perception and memory. *Behavioral Brain Research*, 259, 229-241. doi: 10.1016/j.bbr.2013.11.018
- Mather, M. (2007). Emotional arousal and memory binding: An object-based framework. *Perspectives on Psychological Science*, 2(1), 33-52. doi: 10.1111/j.1745-6916.2007.00028.x

- Mather, M., & Sutherland, M. R. (2011). Arousal-biased competition in perception and memory. *Perspectives on Psychological Science*, *6*(2), 114-133. doi: 10.1177/1745691611400234
- Mayes, A., Montaldi, D., & Migo, E. (2007). Associative memory and the medial temporal lobes. *Trends in Cognitive Science*, *11*(3), 126-135. doi: 10.1016/j.tics.2006.12.003
- Miele, D. B., Finn, B., & Molden, D. C. (2011). Does easily learned mean easily remembered? It depends on your beliefs about intelligence. *Psychological Science*, *22*(3), 320-324. doi: 10.1177/0956797610397954
- Murray, B. D., & Kensinger, E. A. (2012). The effects of emotion and encoding strategy on associative memory. *Memory and Cognition*, *40*(7), 1056-1069. doi: 10.3758/s13421-012-0215-3
- Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: evidence for metacognitive illusions. *Journal of Experimental Psychology: General*, *137*(4), 615-625. doi: 10.1037/a0013684
- Talmi, D. (2013). Enhanced emotional memory: Cognitive and neural mechanisms. *Current Directions in Psychological Science*, *22*, 430-436. doi: 10.1177/0963721413498893
- Tauber, S. K., Dunlosky, J., Urry, H. L., & Opitz, P. C. (2017). The effects of emotion on younger and older adults' monitoring of learning. *Neuropsychology, development, and cognition*, *24*(5), 555-574. doi: 10.1080/13825585.2016.1227423
- Zimmerman, C. A., & Kelley, C. M. (2010). "I'll remember this!" Effects of emotionality on memory predictions versus memory performance. *Journal of Memory and Language*, *62*(3), 240-253. doi: 10.1016/j.jml.2009.11.004

Figure legends

Figure 1. Overview of the experimental procedure. (A) Encoding task in List 1 and List 2, including Judgements of Learning (JOL) in List 2. Examples show neutral-neutral pairs. (B) Recognition task with Judgements of Memory (JOM) and Five Alternative-Force-Choice Associative Recognition (5-AFC). (C) Interactive imagery task.

Figure 2. (A) Associative recognition accuracy (Assoc., in %) and Judgement of Learning (JOL) ratings (in %) as a function of negative (Neg) or neutral (Neu) pair emotionality. (B) JOL as a function of associative recognition accuracy (correct, incorrect) and pair emotionality. (C) Interactive imagery ratings for studied (correct/incorrect in the associative recognition task) and unstudied (recombined) negative and neutral pairs. Error bars are 95%-confidence intervals, corrected for interindividual variability (Loftus & Masson, 1994).

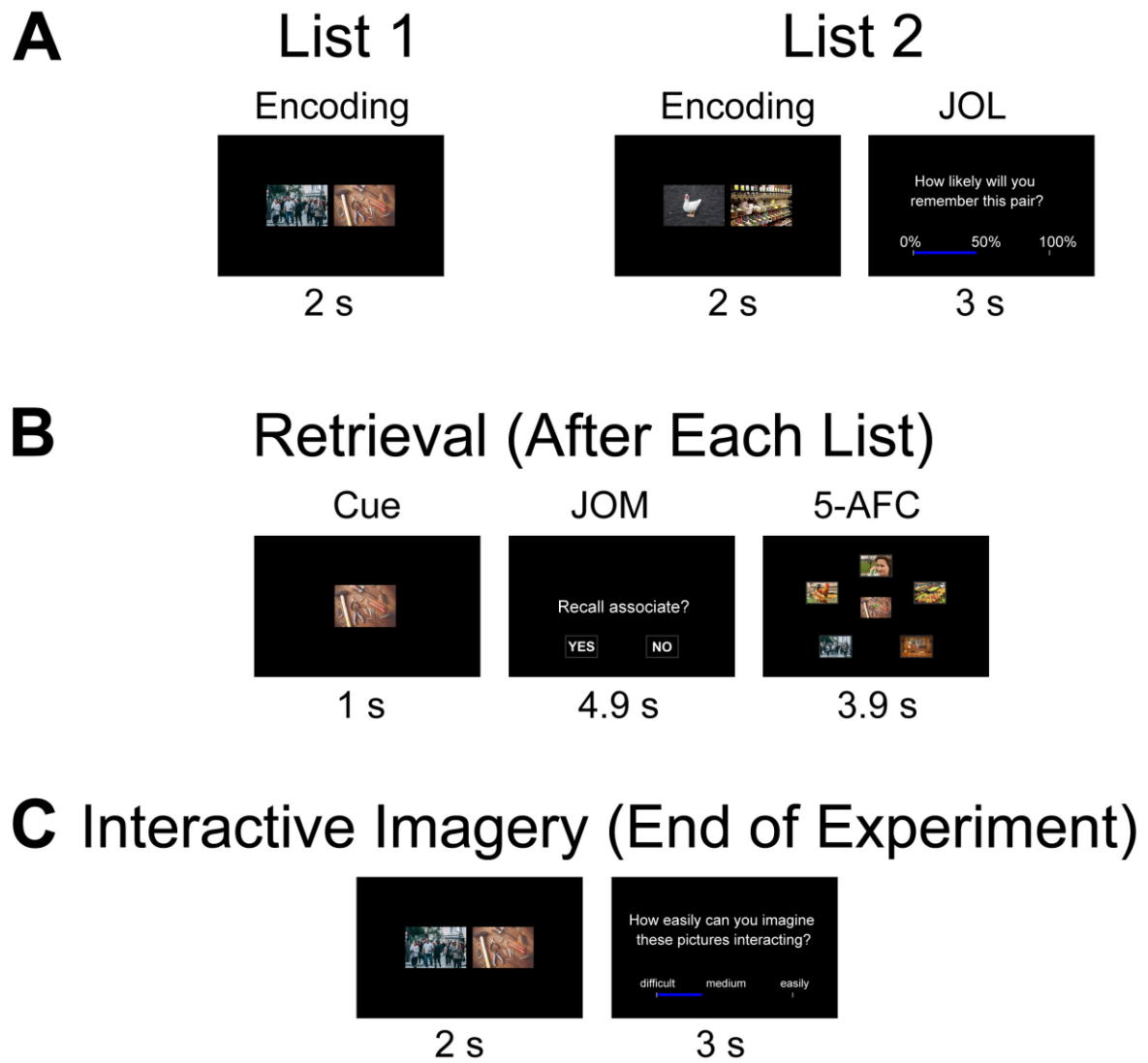


Figure 1

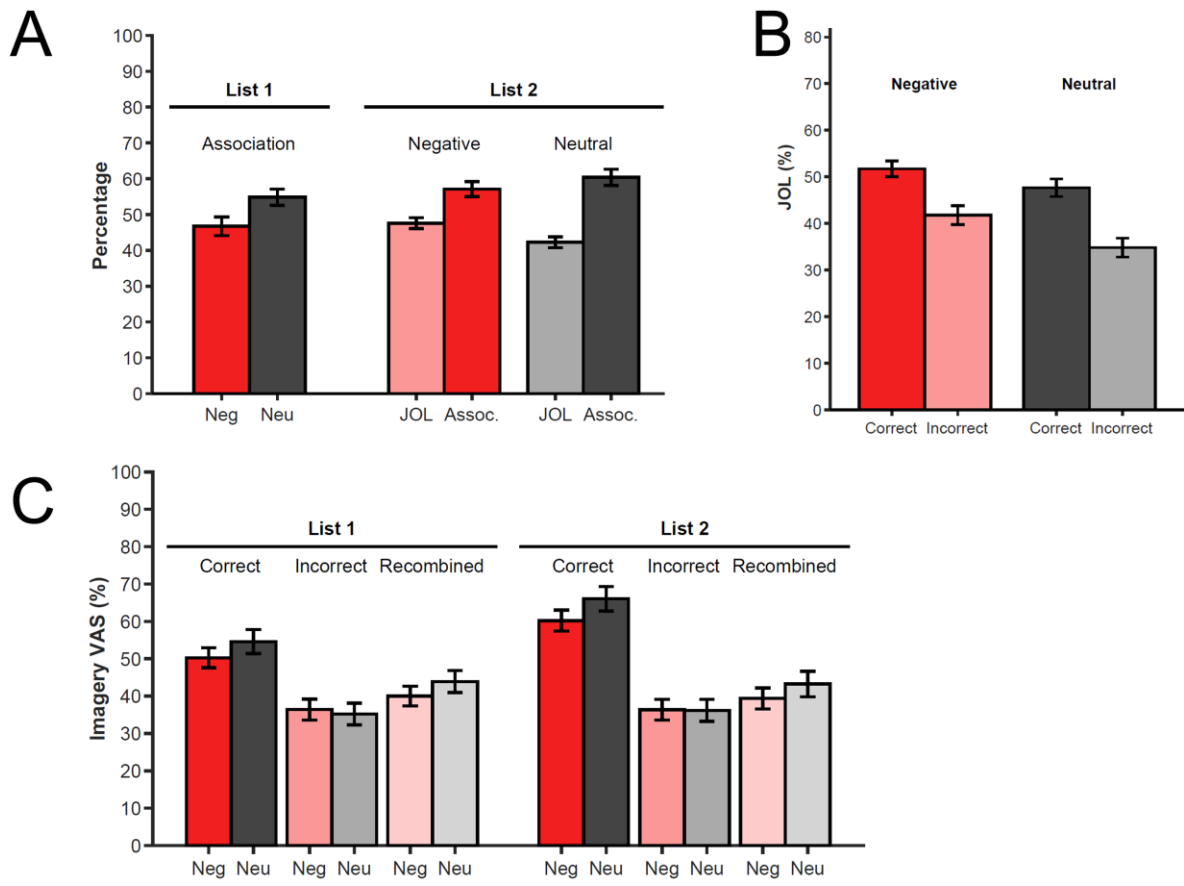


Figure 2