Distribution Agreement

Examining Slowed Forgetting in Recognition Memory for Emotional and Neutral Pictures

By

Susie Shepardson
Master of Arts

____

Stephan Hamann, Ph.D.
Advisor

____

Joe Manns, Ph.D.
Committee Member

____

Michael Treadway, Ph.D.
Committee Member

Accepted:

____

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

____

Date

Examining Slowed Forgetting in Recognition Memory for Emotional and Neutral Pictures

By

Susie Shepardson
B.S., University of Richmond, 2019

Advisor: Stephan Hamann, Ph.D.

An abstract of
A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies at Emory University
in partial fulfillment of the requirements for the degree of
Master of Arts
in Psychology
2021

Abstract

Examining Slowed Forgetting in Recognition Memory for Emotional and Neutral Pictures
By Susie Shepardson

Memory for emotional events (episodic emotional memory) is often stronger and more persistent compared to memory for neutral events. Although many studies have found slower forgetting for emotional vs. neutral stimuli, some key questions remain unanswered, in part because of limited relevant experimental data, methodological issues with prior studies, and a strong focus on negative emotional stimuli. Because few studies have examined forgetting for positive stimuli, it remains unclear whether forgetting is slower for positive stimuli relative to neutral stimuli. Another key question is the extent to which differential forgetting rates for negative and positive stimuli are reflected in two primary components of episodic memory, recollection (memory for contextual information) and familiarity (memory strength, independent of recollection). We examined these questions in an online experiment that assessed memory for negative, positive, and neutral pictures at three delay intervals, 10 minutes, 24 hours, and 1 week. We assessed forgetting for overall recognition performance and for estimates of recollection and familiarity across these three retention intervals for negative, positive, and neutral pictures. Forgetting for negative pictures (vs. neutral pictures) was slower for overall recognition performance and forgetting for negative pictures was also slower for both recollection and familiarity. In contrast, forgetting was not slower for positive pictures and the forgetting rates for positive and neutral pictures were broadly similar. In summary, the current study contributed to the understanding of forgetting for emotional episodic memory, finding that slowed forgetting was valence-dependent, and that slowed forgetting for negative pictures is reflected in both the recollection and familiarity components of recognition memory. These findings suggest that current theories of forgetting in emotional episodic memory, which have focused primarily on negative emotional stimuli and on recollection-based forgetting effects, need to take into account the role of negative and positive valence and effects on familiarity processes.

Examining Slowed Forgetting in Recognition Memory for Emotional and Neutral Pictures

By

Susie Shepardson
B.S., University of Richmond, 2019

Advisor: Stephan Hamann, Ph.D.

A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies at Emory University
in partial fulfillment of the requirements for the degree of
Master of Arts
in Psychology
2021

Table of Contents

List of Figures

A wealth of experimental evidence confirms that episodic memory (conscious recollection of prior events along with their context (Tulving, 1983)) is frequently enhanced for stimuli and events that are emotionally arousing (Bowen et al., 2018; Hamann, 2001; LaBar & Cabeza, 2006), relative to similar, neutral ones. Indeed, negative (unpleasant) and positive (pleasant) emotional events figure prominently in the landscape of autobiographical memory, comprising some of the most memorable and long-lasting types of episodic memories (Cabeza & St Jacques, 2007; Conway & Pleydell-Pearce, 2000; Rubin, 2005). Reviews of the emotional memory literature have emphasized the persistent, long-lived nature of emotional memories relative to similar, neutral memories as a key distinctive attribute of emotional memories, along with other key attributes, including greater memory strength, perceptual and affective vividness, and an enhanced sense of recollection and memory confidence during memory retrieval (Bowen et al., 2018; Hamann, 2001; LaBar & Cabeza, 2006; McGaugh, 2004).

What psychological and neurobiological mechanisms might account for slower forgetting for emotional episodic memories? One possibility is that emotional memories are simply stronger than neutral ones (i.e., better encoded), and thus memory performance for emotional memories is higher after a given retention interval because memory performance was higher at the time of initial encoding. However, this possibility is inconsistent with the findings of studies that have matched initial memory performance between emotional and neutral stimuli and have also found slower forgetting for emotional memories (LaBar & Cabeza, 2006; Sharot & Yonelinas, 2008). Moreover, several studies have found that the emotional memory effect (enhanced memory for emotional vs. neutral stimuli; EEM) often grows with time, for example, finding a small or nonsignificant EEM in recognition memory at short delays and a significantly larger EEM at a longer delay (typically days or weeks) (LaBar & Phelps, 1998; Sharot & Phelps, 2004; Sharot & Yonelinas, 2008). Therefore, the slowed forgetting for emotional stimuli cannot be explained in general by better initial memory for emotional stimuli.

Slower forgetting for emotional stimuli (particularly negative stimuli) has become a generally accepted empirical finding and theoretical explanations have been proposed to account for this phenomenon (Hamann, 2001; Kensinger, 2004; LaBar & Cabeza, 2006; Yonelinas & Ritchey, 2015). The most widely accepted theoretical account for slowed forgetting for emotional episodic memories is the *amygdala-mediated consolidation model* (Cahill & McGaugh, 1998; Hamann, 2001, 2009; McGaugh, 2004). In this model, when the amygdala is activated by emotionally arousing stimuli, it can modulate and enhance activity in the medial temporal lobe memory system that supports episodic memory (McGaugh, 2004; Squire & Zola-Morgan, 1991). This amygdala-mediated modulation of episodic memory is thought to occur primarily during the consolidation of the memory representation, after the initial encoding of the stimulus has completed (immediate amygdala-dependent effects of emotion on attention and memory have also been found, but these have not been generally been invoked as causal factors involved in slowed emotional forgetting (Schümann & Sommer, 2018; Talmi, 2013; Talmi et al., 2008)). Because consolidation is a process that unfolds over time, gradually stabilizing memories, this time-dependent aspect of consolidation has been used to explain why the effects of emotion on episodic memory are also time-dependent, increasing over time as the consolidation process unfolds and potentially occurring to a greater extend during sleep and other periods of increased consolidation (Payne & Kensinger, 2010, 2018; Walker, 2010).

The other major theoretical view proposed to explain slowed forgetting for emotional memory is the *emotional binding account* (Yonelinas & Ritchey, 2015). Yonelinas and Ritchey (2015) critiqued the amygdala-mediated consolidation model, pointing out a number of findings that this model either does not explain or which are not compatible with this model. They proposed an alternative model in which the amygdala creates emotion-item bindings (i.e., memory associations) for emotional stimuli but not for neutral stimuli. Specifically, they proposed that when emotional stimuli are experienced, the emotional response (i.e., affective feelings) elicited by the stimulus item are bound to the item's representation in the perirhinal cortex (a medial temporal lobe region associated with item processing and memory for items), forming an emotional binding that supports subsequent memory for that stimulus, in particular, a form of memory known as *recollection*. It is currently well accepted that episodic memory is not a unitary process but instead reflects the contributions of at least two primary underlying cognitive processes: *recollection* and *familiarity* (Diana et al., 2007; Yonelinas, 2002) (but see also (Wixted & Squire, 2011) for an alternative view). Recollection involves retrieval of qualitative contextual information about an item's prior occurrence, whereas familiarity involves a graded signal of memory strength that supports judgments of whether an item has been encountered before, even in the absence of recollection.

Recently, the generality and strength of empirical support for phenomenon of slowed forgetting has come under new scrutiny. In one such critique, Baraly et al. (2017) point out that much of the empirical support for slowed forgetting of emotional stimuli is derived from older studies (i.e., from several decades ago), and some of these studies have been shown to have important methodological flaws that undermine their conclusions in some cases. For example, a highly influential and widely cited early study by Kleinsmith & Kaplan (1963) found enhanced cued recall for neutral numbers paired with negative emotional words (relative to neutral words), but only after longer delay intervals and not at an immediate test. Although this study has been widely cited as evidence for slowed forgetting of emotional memories, Mather (2007) reviewed this literature and noted a key methodological confound: arousal level (as indexed by skin-conductance response) was confounded with serial position in Kleinsmith & Kaplan (1963)'s study. A later study that avoided this confound by including buffer items at the beginning and end of encoding lists to avoid primacy and recency effects failed to replicate the original findings (Mather, 2007; Schürer-Necker, 1990) and several other studies that used group emotion rating norms to define arousal categories also failed to replicate these effects (Mather, 2007).

Baraly et al. (2017) identified additional gaps in the current literature on forgettting and emotional memory. Relatively few contemporary studies have examined forgetting for emotional vs. neutral stimuli across more than two retention intervals. Studies that examine forgetting over multiple retention intervals can be particularly informative as they provide a more fine-grained assessment of the forgetting curve that allows additional issues to be examined, such as whether forgetting rates change across different delay intervals. Another limitation of the current literature is the wide variety of different paradigms used and sometimes conflicting results across studies (Baraly et al., 2017). To help clarify this literature, Park (2005) conducted a qualitative meta-analysis that reviewed the emotional memory forgetting literature. However, as noted by Baraly et al. (2017), this review had several important methodological limitations and examined forgetting in the context of the effects of arousal, broadly defined (e.g., effects of stimulant drugs and time-of-day effects) rather than emotional memory per se. Also, since this meta-analysis was published many years ago, it did not consider many more recent studies. Given these gaps in the current literature, additional empirical studies of forgetting for emotional stimuli over multiple retention intervals can potentially help clarify these key questions regarding emotional memory and forgetting.

An important additional question concerns whether slowed forgetting occurs to a similar extent for both negative and positive emotionally arousing stimuli. The emotional memory effect is found for both positive and negative emotionally arousing stimuli, and theoretical accounts of emotional episodic memory have highlighted similarities between the enhancing effects of emotionally arousing stimuli on memory, independent of emotional valence (Cahill & McGaugh, 1998; Hamann, 2009; Hamann et al., 1999; LaBar & Cabeza, 2006). However, most studies of emotional episodic memory and forgetting have failed to examine forgetting for positive emotional stimuli, typically focusing only on negative and neutral stimuli. Moreover, the limited existing evidence regarding slowed forgetting for positive emotional stimuli is equivocal, in part due to methodological problems with the studies that have examined this issue (Wang, 2014, 2018). Therefore, it remains unclear to what extent forgetting is slowed for positive emotional stimuli, relative to neutral stimuli. Because the amygdala-mediated consolidation model emphasizes the role of emotional arousal in the emotional enhancement of memory, rather than differences between negative and positive valence, finding broadly similar effects of negative and positive emotion on forgetting would be compatible with this theory. In contrast, the emotional binding account focuses solely on forgetting effects for negative emotional stimuli and does not make any predictions regarding forgetting for positive stimuli.

As noted above, most contemporary memory theories have focused on understanding the relative contributions of two primary memory processes mediating episodic memory: recollection and familiarity. With regard to emotional memory, interest has focused on two key questions: to what extent does emotion influence recollection and familiarity differentially, and do the effects of emotion on these two processes change over the course of forgetting. Studies comparing the effects of emotion on recollection and familiarity have consistently found that emotion enhances recollection but has little or no effect on familiarity (Yonelinas & Ritchey, 2015). Much less is known, however, about whether such effects of emotion on recollection and familiarity change over increasing memory retention intervals.

One influential study by Sharot and Yonelinas (Sharot & Yonelinas, 2008) examined recognition memory for negative and neutral pictures after a short (5 minute) and a long (24-h) delay, using a remember-know recognition paradigm to measure recollection and familiarity processes in a within-subjects design. The main finding was that recollection was enhanced for negative pictures relative to neutral pictures after a 24-h delay but not immediately after encoding at the 5-minute delay, with the net effect of slower forgetting for recollection for negative pictures relative to neutral ones. Yonelinas & Ritchey (Yonelinas & Ritchey, 2015) subsequently used these findings as key evidence supporting their emotional binding model.

However, this study by Sharot and Yonelinas (2008) also found that forgetting was slowed for *familiarity*, inconsistent with the predictions of the emotional binding model. It is unclear why this inconsistent result was not discussed by Yonelinas & Ritchey (2015). Because the enhancement of familiarity by emotion they observed was inconsistent with the findings of several previous studies that had found that emotional arousal affected recollection but not familiarity, Sharot and Yonelinas (2008) conducted a follow-up analysis of the familiarity results. The results of this follow-up analysis led them to conclude that the enhancement of familiarity for negative pictures and the finding of slowed forgetting for familiarity for negative pictures were due to an indirect effect of a source memory judgment task (asking which of two different encoding tasks had been performed with that item) that each participant completed following making a remember vs. know recognition memory decision for each item on the retrieval test. They argued that when memory for source was not available for an item, participants' confidence may have been undermined, potentially making them less inclined to respond "remember" and consequently they were more inclined to respond "know" (i.e., familiarity). After controlling for source judgment accuracy, the enhancement of familiarity for negative vs. neutral pictures was no longer statistically significant when the results from the 24 hour delay were examined. There are two potential concerns regarding this conclusion however. First, because the source memory judgments for each items were made *following* each remember-know recognition judgment, it is unclear why the latter judgments would be expected have a retroactive effect on the earlier remember-know decisions. In addition, the conclusion that remember-know recognition judgments were contaminated by the subsequent source memory test casts uncertainty on the overall validity of the remember-know results in that study as a basis for accurate estimation of recollection and familiarity. Given these potential concerns, additional data from new studies are needed before more definitive conclusions can be drawn regarding these issues.

Sharot and Yonelinas (2008) did not examine forgetting for positive emotional pictures, leaving the question of whether forgetting is slower for positive vs. neutral pictures and whether any slower forgetting is differentially manifested in the recollection and familiarity components supporting recognition memory. Regarding each theoretical model, the emotional binding view's prediction of slowed forgetting for recollection is specific to negative emotion and this view makes no specific predictions for positive emotion. The amygdala-mediated consolidation view in general predicts similar effects of positive and negative emotion on episodic memory processes, so in line with this principle, slowed forgetting would be expected to be observed for emotionally arousing stimuli regardless of positive vs. negative valence. Regarding differences between recollection and familiarity, because considerable evidence implicates the hippocampus in mediating recollection and not familiarity (but see also Wixted & Squire, 2011) and the amygdala-mediated consolidation view focuses on enhancement of hippocampal-dependent consolidation processes, this suggests that slowed forgetting would be primarily reflected in effects on recollection.

In the current study, we examined three main questions about differences in rates of forgetting for emotional vs. neutral pictures: 1) To what extent is forgetting slower for negative emotional pictures relative to neutral stimuli across multiple memory delay intervals? Based on the results of prior studies that reported slowed forgetting for negative stimuli compared to neutral stimuli (Sharot & Yonelinas, 2008; Baraly et al., 2017), we predicted that negative emotional pictures will be forgotten more slowly compared to neutral pictures, consistent with the predictions of both major theories, 2) To what extent is forgetting slower for positive emotional pictures relative to neutral stimuli across multiple memory delay intervals? Our predictions for forgetting of positive pictures were more tentative, because few prior studies have examined this issue. Given that many of the experimental effects in the emotional episodic memory literature are arousal-based and are broadly similar for both negative and positive stimuli, we tentatively predicted that forgetting for emotionally arousing positive pictures would be slowed relative to neutral pictures, 3) To what extent is slower forgetting for negative or positive pictures (relative to neutral pictures) manifested differentially for the recollection and familiarity components that support recognition memory performance? We predicted that recollection for negative pictures would be slowed relative to neutral ones, consistent with the prior literature and both theories. However, given the relatively small number of relevant prior studies, we did not have specific predictions regarding forgetting for recollection of positive pictures. Turning to familiarity and forgetting, because of the limited relevant literature we also did not have specific predictions regarding forgetting rates, for either negative or positive pictures. Notably, slowed forgetting for negative and positive familiarity would be difficult to reconcile with the predictions of the emotional binding theory.

We examined these three primary questions in an online experiment that assessed memory for negative, positive, and neutral pictures at three delay intervals, 10 minutes, 24 hours, and 1 week, in a within-subjects design. In an initial session, participants were presented with all the pictures that would serve later as targets in the recognition tests. On each encoding trial, participants briefly viewed each picture individually for 0.75 s and made an incidental indoor/outdoor decision about the picture. After a distractor task, the first recognition memory task (10 min) was given, using a remember/familiar/new recognition task to assess overall recognition performance as well as estimates of recollection and familiarity. Participants completed the second session 24 hours later and the third session 1 week after the first session. During each recognition test, one-third of the encoded pictures were presented as targets, intermixed with distractor pictures of each valence. Our primary interest was in assessing forgetting (for overall recognition performance and for estimates of recollection and familiarity) across the three retention intervals for negative, positive, and neutral stimuli.

2. Methods

2.1 *Participants*

Forty-seven undergraduate students at Emory University participated in this online study. All participants gave informed consent and received course credit for their participation. Two participants were excluded because they did not understand the retrieval instructions. Six participants were excluded because they omitted over 10% of indoor/outdoor encoding task responses during the encoding phase. The data for the 39 (10 males) remaining participants with a mean age of 18.85 (0.87) were used for the analyses in this study.

2.2 *Stimuli*

2.2.1 *Stimuli norming*

In an initial norming phase, 1,206 pictures (402 neutral, 402 negative, and 402 positive) were collected from various emotional stimuli databases and online digital photo sharing platforms including IAPS (Lang et al., 1997), OASIS (Kurdi et al., 2017), NAPS (Marchewka et al., 2014), and two online photo sharing sites, unsplash.com, and pixabay.com. For purposes of online ratings, these pictures were grouped into six rating surveys consisting of 201 pictures each, 67 of each valence. Participants rated each picture on emotional valence and emotional arousal measured using the Self-Assessment Manikin (SAM) scale. The SAM scale instructs participants to rate valence and arousal on a visually presented scale (1-9) from very negative to very positive for valence and from weak or no emotion to strong emotion for arousal (Morris, 1995). Stimulus order was counterbalanced across participants.

Each ratings survey was administered using the Qualtrics platform. 10 Emory University undergraduates (ages 18-22) took each survey. All participants gave informed consent and received course credit for their participation. Each participant's responses were examined to verify they used the entire 1-9 scale for both valence (1 – very negative: 9- very positive) and arousal (1 – weak or no emotion: 9 – strong emotion).

2.2.2 *Stimulus selection*

Using the ratings obtained from the stimulus norming phase, pictures were sorted into positive (valence rating greater than 6.00), neutral (valence ratings between 4.25 and 5.75), and negative (valence rating less than 4.00). Next, the pictures were inspected to ensure they matched their intended valence category and were assessed for visual complexity. Picture sets were selected so as to approximately match the average visual complexity of pictures across the three valence categories. Finally, pictures were selected in order to match negative and positive pictures on arousal ratings. After all selection procedures were completed, a final set of 621 selected pictures (207 per valence) were used in the current study. The mean valence ratings for the neutral, negative, and positive pictures were 5.15 (SD=0.36), 2.86 (SD=0.40), 7.18, (SD=0.47). For arousal ratings, both negative (M=5.29, SD=0.73) and positive pictures (M=5.28, SD=0.79) were rated as more arousing than neutral pictures (M=2.79, SD=0.58); $t(206) = 41.68$, $p<.001$. There was no significant difference between the arousal ratings for negative and positive, $t(206) = -0.09$, p=0.93, confirming that these sets of pictures had been matched.

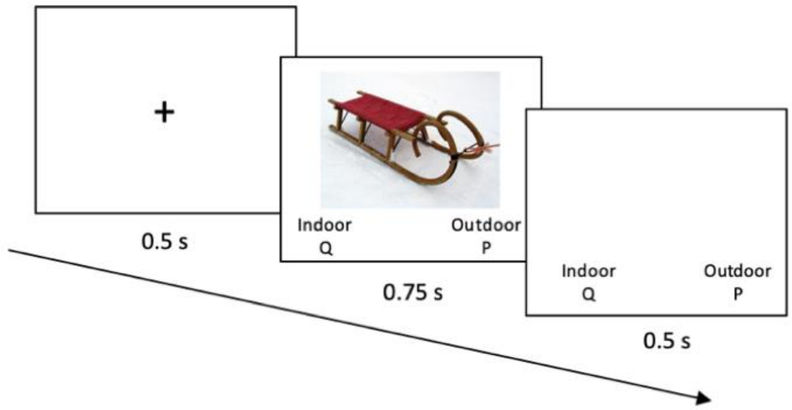
2.3 *Experimental tasks and study design*

The experiment was conducted remotely using the Zoom and Pavlovia online platforms. After participants were given instructions and had completed a practice session with a researcher in a Zoom meeting, they proceeded to complete the rest of the session alone using the Pavlovia experiment site. Given the limited control over each participant's experimental environment, an extensive pre-test and post-test questionnaire was used to gather information about each participant's testing environment. Before beginning the experiment participants provided information regarding where they were taking the experiment, the type of laptop computer and size of laptop screen they were using, and the length and width of the pictures in the experiment as they appeared on the screen, using items of known size such as a credit card. Participants were also instructed to be in a distraction-free area, to use the same laptop and location for all three sessions, and to sit with their laptop screen an elbow's length away from their body. After completing each retrieval test, participants took a survey on Qualtrics where they provided information regarding their time zone, the amount of effort they felt they expended (0-100%), whether or not they followed the instructions, and their understanding of the remember/familiar instructions.

2.3.1 *Encoding phase*

In the first session of the experiment, participants completed the encoding phase, during which they encoded 414 pictures (138 per valence). Each encoding trial began with the presentation of the picture together with the prompt for the indoor vs. outdoor incidental encoding task. The prompt for the indoor task was presented on the left side of the screen and the prompt for the outdoor decision was presented on the right side of the screen (see Figure 1). The total duration of each encoding trial was 1.75 s. Each trial started with picture presentation for 0.75 s simultaneously with the prompt for the indoor/outdoor task. Next, the picture disappeared from the screen and participants were allowed an additional 0.5 s to complete the indoor/outdoor task. Finally, a fixation cross appeared for 0.5 s before the next trial began. The trials were separated into 4 blocks of 100 trials, and there was a 15 second break between each block to reduce fatigue. Five neutral buffer pictures were presented at the beginning and at the end of the encoding list, to reduce serial position effects on memory.

Figure 1.
Encoding Phase Procedure



Note. Each encoding trial began with a 0.5 second fixation cross, followed by presentation of a picture and the indoor/outdoor prompt or 0.75 seconds, after which the picture disappeared and participants were allowed an additional 0.5 seconds to complete the indoor/outdoor task before the next trial began.

For the indoor/outdoor task, participants were instructed to press the ‘Q’ key if they thought the picture showed an indoor scene or the ‘P’ key if they thought the picture showed an outdoor scene. Participants were instructed to make a response for every picture even if they were unsure if the picture was indoor or outdoor. The purpose of the indoor/outdoor encoding task was primarily to ensure that participants viewed and attended to each stimulus and made a response. Since a substantial proportion of the pictures (approximately one-third) were difficult to classify unambiguously as being completely outdoor or indoor, to assess the quality of each subject’s responses, we focused on whether a response was made on each trial rather than on the accuracy of the indoor/outdoor response. We excluded any participant from further analysis who failed to make a response on 10% or more trials (6 participants were removed on the basis of this criterion).

2.3.2 *Delay phase*

After encoding, participants completed a trivia filler task for 10 min. Each trial of the trivia task involved viewing a trivia question for 7 seconds (answering the question silently) then viewing the answer for 3 seconds and indicating whether or not they answered the question correctly by pressing “Q” if they answered correctly and “P” if they did not.

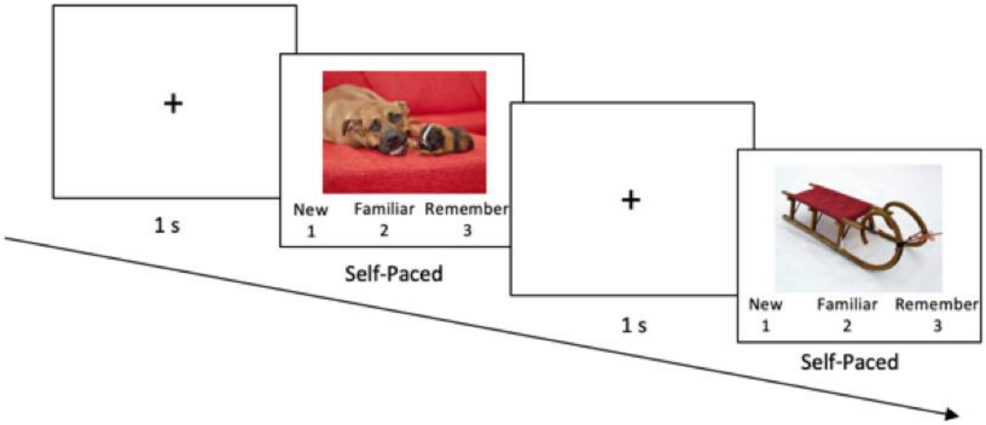
2.3.3 *Retrieval phase*

There were three separate retrieval test phases. The first occurred immediately after the delay task, at 10 min after the end of the encoding phase. The second retrieval phase occurred 24 hours after the first session, and the third retrieval phase occurred 1 week after the first session (i.e., if a participant had their first session at 9 am on Monday, their session 2 was 9 am on Tuesday and session 3 was 9 am the following Monday).

Each retrieval test included 206 pictures, 138 targets and 69 lures (46 targets and 23 lures per valence). The test was self-paced with a new/familiar/remember task (see Figure 2). Retrieval items were presented in two separate blocks of 103 pictures each, with a 15 s rest break between the blocks. Each block started with a multiple choice catch trial question intended to test the participant’s understanding of the new/familiar/remember retrieval task. To provide participants with additional practice making the keyboard responses, each retrieval test started with 4 neutral buffer pictures (that did not appear elsewhere in the experiment).

Figure 2

Retrieval Phase Procedure



Note. Each retrieval trial began with 1 second fixation followed by presentation of a picture and a new/familiar/remember task which was self-paced. Once the response of 1, 2, or 3 was pressed, the next trial began.

Each retrieval test began with a detailed instruction and practice phase where the new/familiar/remember task was explained. The instructions for this task were adapted from similar ones described in Migo et al., (2012). Participants were tested on whether they understood the recognition task instructions (specifically, whether they understood the difference between the correct meaning of a "remember" and a "familiar" response) during the instruction phase, the practice phase, during two catch trials (one at the beginning and one at the middle of the retrieval phase), and finally, in a post-test questionnaire after each retrieval phase. Participants who could not correctly explain their understanding of the recognition task instructions were excluded from further analysis (2 participants).

2.3.4 *Data analysis*

Overall picture recognition was assessed by examining the hit rates, false alarm rates, corrected recognition accuracy (hit rate minus false alarm rate), and the d' sensitivity statistic. We based our conclusions regarding memory performance and forgetting primarily on the basis of the d' and corrected recognition accuracy scores. Results for hit rates and false alarm rates were also examined to clarify the results obtained for d' and corrected recognition.

The d' measure allows memory accuracy (sensitivity) to be assessed separately from response criterion (*c*, the bias for an individual to judge an item on a recognition test as being old or new) and d' also has the advantage of being a ratio-scale measure, which is a desirable property of a memory measure in the context of assessing forgetting (MacMillan & Creelman, 2004). Participants' d' scores for each condition and delay interval were calculated using a log-linear transformation of hit and false alarm rates (Stanislaw & Todorov, 1999) to address the issue that this statistic is undefined when hit rates are 1.0 and false alarm rates are 0. Although the log-linear transformation is commonly used for this purpose, it should be noted that when hit rates are 1.0 or false alarm rates are 0, the resulting d' scores are biased and are less accurate estimates of memory sensitivity than when these boundary conditions are avoided (MacMillan & Creelman, 2004).

Mean reaction times (RTs) were also calculated for each subject for each combination of conditions to check for outliers and characterize overall performance across conditions and delay intervals. Partial Eta Squared (η^2_p) was calculated to estimate the effect size for each analysis. Recollection and familiarity processes were estimated using the independent K (IRK) procedure (Yonelinas and Jacoby 1995; Yonelinas and Levy 2002), in which "remember" responses are assumed to estimate recollection whereas familiarity is estimated as the proportion of "familiar" responses divided by the proportion of non-remember responses. The recollection and familiarity estimates were corrected for their respective false alarm rates by subtracting the proportion of "remember" responses for new (distractor) items from the proportion of "remember" responses for old (target) items, and for familiarity, using the formula familiarity = ((hit rate for "familiar" items)/(1 – hit rate for "remember" items)) – (false alarm rate for "familiar" items)/ (1- false alarm rate for "remember" items) (Yonelinas and Levy, 2002).

For each memory measure, separate repeated-measures ANOVAs were conducted with Delay (10 min, 24 h, 1 week) and Valence (negative, positive, and neutral) as within-subjects factors. Differences in forgetting rates were assessed by examining the interaction between conditions across delay intervals for each memory measure. This method is by far the one which is most often used for this purpose. However, another method of measuring forgetting has been recommended when memory performance differs at the initial, shortest memory delay interval (Loftus, 1985; Wixted, 1990). Loftus and others have noted that when memory performance at the initial time point differs significantly between conditions or groups, comparisons of forgetting can be distorted by scaling problems. An alternative method, which assesses the proportion of forgetting at each retention interval, relative to memory at the initial time point, is not affected by these scaling problems (Loftus, 1985). For the analysis of forgetting that examined recollection estimates, the proportional forgetting analysis method was because there were significant differences in memory performance between conditions at the initial, 10-min time point.

For each memory measure, additional repeated-measures ANOVAs were conducted separately comparing the negative and neutral conditions and the positive and neutral conditions. These were planned comparisons motivated by our a priori theoretical interest in comparing forgetting patterns for each emotion condition relative to the neutral condition. Simple comparisons (pairwise t tests) are reported comparing the valence conditions if there was a main effect of Valence, and these comparisons were all Bonferroni-corrected.

3. Results

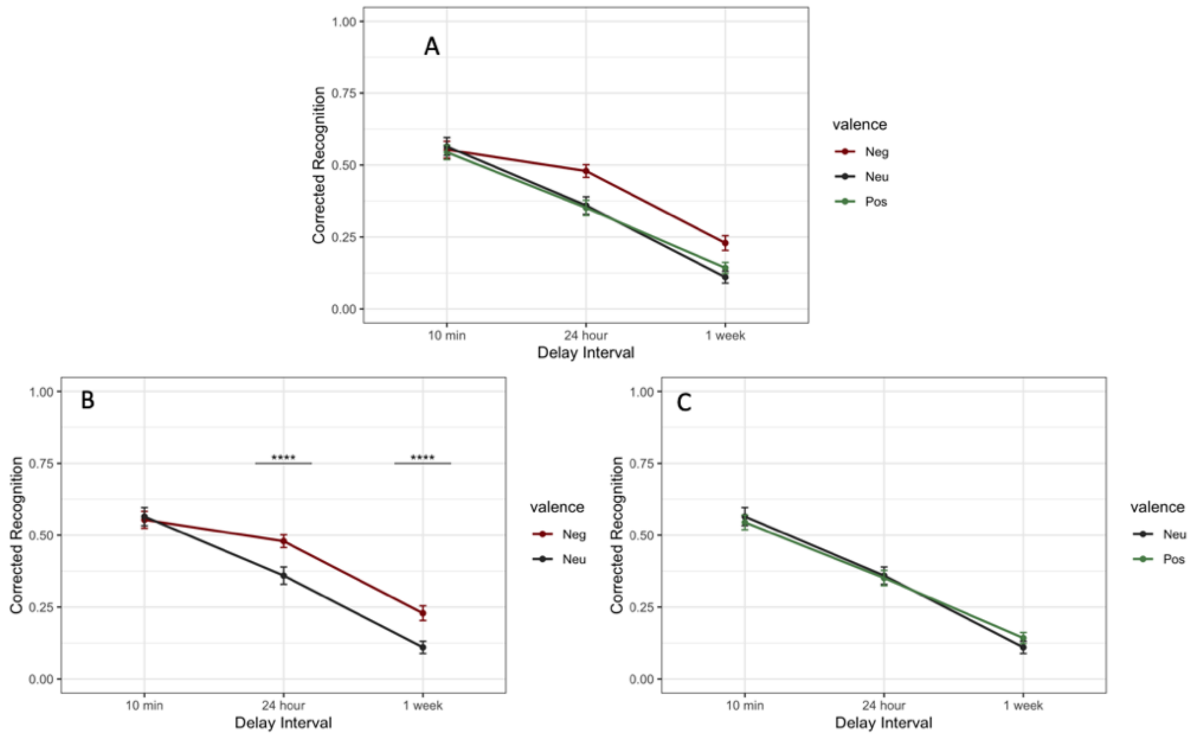
3.1 Overall picture recognition

For corrected recognition, there was a significant interaction between Delay and Valence on corrected recognition scores, $F(4, 152)=5.56$, $p<.001$, $\eta^2_p=.13$ (see Figure 3A), indicating differences in forgetting across conditions. There also were main effects of Valence, $F(2, 76)=16.30$, $p<.001$, $\eta^2_p=.30$, and Delay, $F(1.52, 57.66)=156.62$, $p<.001$, $\eta^2_p=.81$.

To assess the predictions regarding differential forgetting for negative vs. neutral pictures and positive vs. neutral pictures, planned follow-up within-subjects ANOVAs were conducted for each pair of conditions separately. The overall pattern of findings for corrected recognition and d' was quite similar. For the ANOVA with corrected recognition for negative and neutral pictures, as predicted, there was a significant interaction between Delay and Valence, $F(2, 76)=9.72$, $p<.001$, $\eta^2_p=.20$. Inspection of this interaction showed that forgetting was slower for negative pictures compared to neutral pictures (see Figure 3B). In addition, there were significant main effects of Valence, $F(1, 38)=21.42$, $p<.001$, $\eta^2_p=.36$, and Delay, $F(1.54, 58.39)=114.35$, $p<.001$, $\eta^2_p=.75$. Simple comparisons between the negative and neutral conditions on corrected recognition performance found no significant difference between these two conditions at the 10 minute delay interval, $t(38)=-0.41$, $p=.68$, whereas the negative condition had significantly higher memory performance at both 24 hours and 1 week relative to neutral pictures (ps<.001).

For corrected recognition for positive and neutral pictures, there was no interaction between Delay and Valence, $F(2, 76)=1.07$, $p=.35$, $\eta^2_p=.03$, no main effect of Valence, $F(1, 38)=0.01$, $p=.92$, $\eta^2_p<.01$, and a main effect of Delay, $F(2, 76)=159.40$, $p<.001$, $\eta^2_p=0.81$ (see Figure 3C).

Figure 3.
Forgetting Results for Corrected Recognition



Note. A) Corrected recognition performance for the valence conditions across the delay intervals. B) Corrected recognition performance for the negative and neutral conditions across the delay intervals. C) Corrected recognition performance for the positive and neutral conditions across the delay intervals. Asterisks denote significant difference between valence conditions at that delay interval.
* p<.05; ** p<.01; *** p<.001; **** p<.0001

In addition to corrected recognition, the same statistical analysis was also conducted using the signal detection theory measure d'. Because the proportions of 0 false alarm rates were relatively high at the 24 hour and 1 week delay intervals, particularly for the neutral picture condition, this had the net effect of biasing (inflating) the d' scores for the neutral condition at the 24 hour and 1 week delay intervals. The corresponding corrected recognition scores at these delay intervals for the neutral condition were not affected by this issue as it is a problem specific to calculation of d' scores.

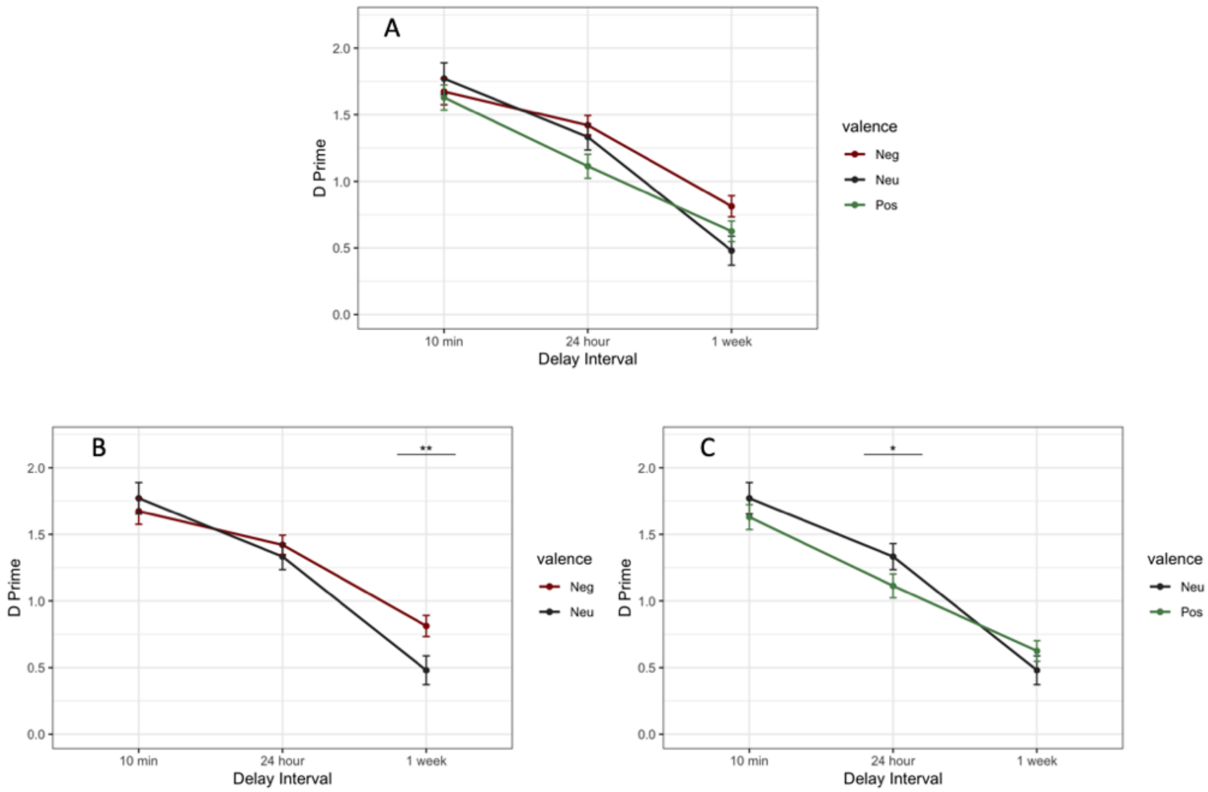
As shown in Figure 4A, there was a significant interaction between Delay and Valence for d' scores, $F(4, 152)=3.28$, $p=.013$, $\eta^2_p=.08$. In addition, there were main effects for both Valence, $F(2, 76)=5.41$, $p=.006$, $\eta^2_p=.13$, and Delay, $F(1.64, 62.39)=98.11$, $p<.001$, $\eta^2_p=.72$.

To assess the predictions regarding differential forgetting for negative vs. neutral pictures and positive vs. neutral pictures, planned follow-up within-subjects ANOVAs were conducted for each pair of conditions separately with d' scores as the dependent variable. For the ANOVA for negative and neutral pictures, as predicted, there was a significant interaction between Delay and Valence, $F(2, 76)=4.84$, $p=.01$, $\eta^2_p=.11$, indicating that forgetting was slower for negative pictures compared to neutral ones (see Figure 4B). In addition, there was a main effect of Delay, $F(2, 76)=70.51$, $p<.001$, $\eta^2_p=.65$, but no main effect of Valence, $F(1, 38)=3.69$, $p=.06$, $\eta^2_p=.09$. Simple comparisons between the negative and neutral conditions on d' scores showed that memory was higher for negative pictures at the 1 week delay interval, $t(38)=3.23$, $p=.003$, but not at 10 minutes ($p=.29$) or 24 hours ($p=.38$).

The corresponding ANOVA was conducted comparing positive and neutral pictures on d' scores. There was no significant interaction between Delay and Valence, $F(2, 76)=2.83$, $p=.065$, $\eta^2_p=.07$, no main effect of Valence, $F(1, 38)=1.80$, $p=.19$, $\eta^2_p=.05$, and a main effect of Delay, $F(2, 76)=83.42$, $p<.001$, $\eta^2_p=.69$ (see Figure 4C).

Figure 4.

Forgetting Results for d' Scores



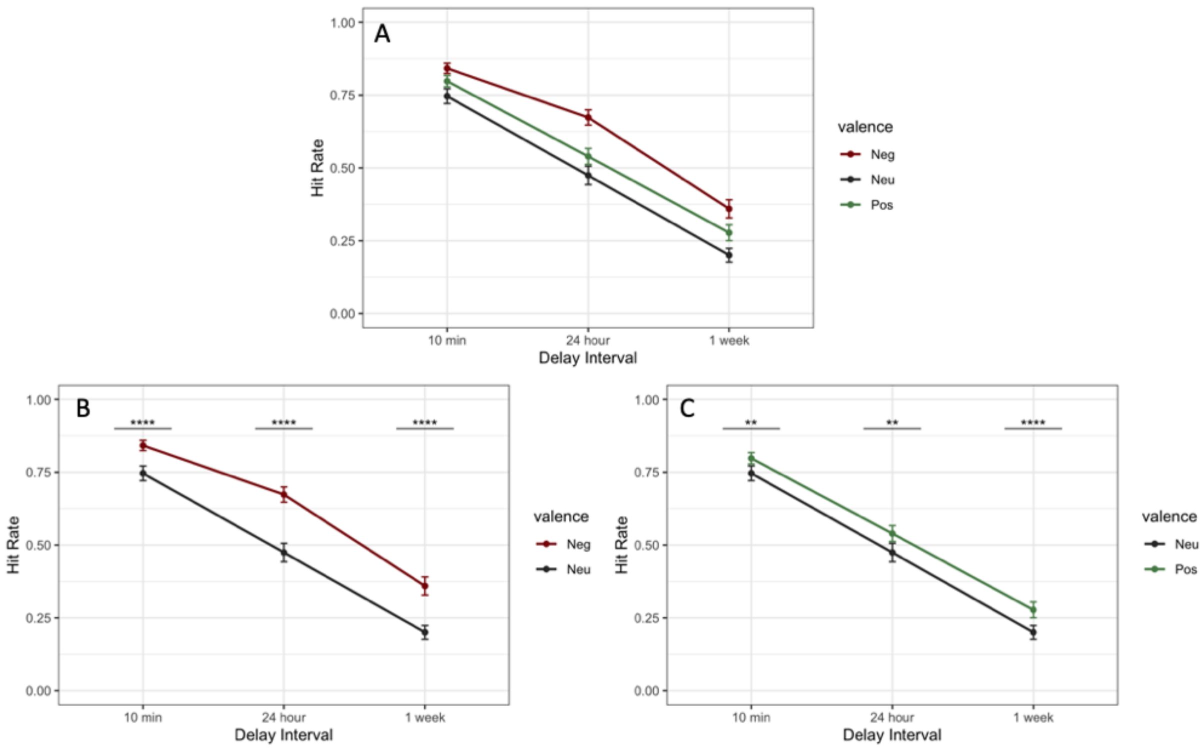
Note. A) D' performance for the valence conditions across the delay intervals. B) D' performance for the negative and neutral conditions across the delay intervals. C) D' performance for the positive and neutral conditions across the delay intervals. Asterisks denote significant difference between valence conditions at that delay interval.
* p<.05; ** p<.01; *** p<.001; **** p<.0001

Turning to the analysis of hit rates, as shown in Figure 5A, there was a significant interaction between Delay and Valence, $F(4, 152)=5.67$, $p<.001$, $\eta^2_p=.13$. There were main effects for both Valence, $F(2, 76)=70.26$, $p<.001$, $\eta^2_p=.65$, and Delay, $F(2, 76)=283.75$, $p<.001$, $\eta^2_p=.88$.

To assess the predictions regarding differential forgetting for negative vs. neutral pictures and positive vs. neutral pictures, planned follow-up within-subjects ANOVAs were conducted for each pair of conditions separately with hit rates as the dependent variable. For the ANOVA with the negative and neutral conditions, as predicted, there was a significant interaction between Delay and Valence for hit rates, $F(2, 76)=9.12$, $p<.001$, $\eta^2_p=.19$, indicating that forgetting was slower for negative compared to neutral (see Figure 5B). In addition, there were main effects of Valence, $F(1, 38)=113.45$, $p<.001$, $\eta^2_p=.75$, and Delay, $F(2, 76)=261.75$, $p<.001$, $\eta^2_p=.87$. Simple comparisons between the negative and neutral conditions on hit rate performance found negative hit rates were significantly higher than neutral, $t(38)=5.99$, $p<.001$, at the 10 minute delay, and this effect was found at both later delay intervals as well (*p*s<0.001).

For the ANOVA with the positive and neutral conditions, there was no significant interaction between Delay and Valence, $F(2, 76)=0.59$, $p=.56$, $\eta^2_p=.02$. However, there were main effects of Valence, $F(1, 38)=28.62$, $p<.001$, $\eta^2_p=.43$ and Delay, $F(2, 76)=256.73$, $p<.001$, $\eta^2_p=.87$ (see Figure 5C). Simple comparisons between the positive and neutral conditions on hit rate performance found positive hit rates were significantly higher than neutral, $t(38)=-2.76$, $p=.009$, at the 10 minute delay, and this effect was found at both later delay intervals as well (*p*s<.01).

Figure 5.
Forgetting Results for Hit Rates



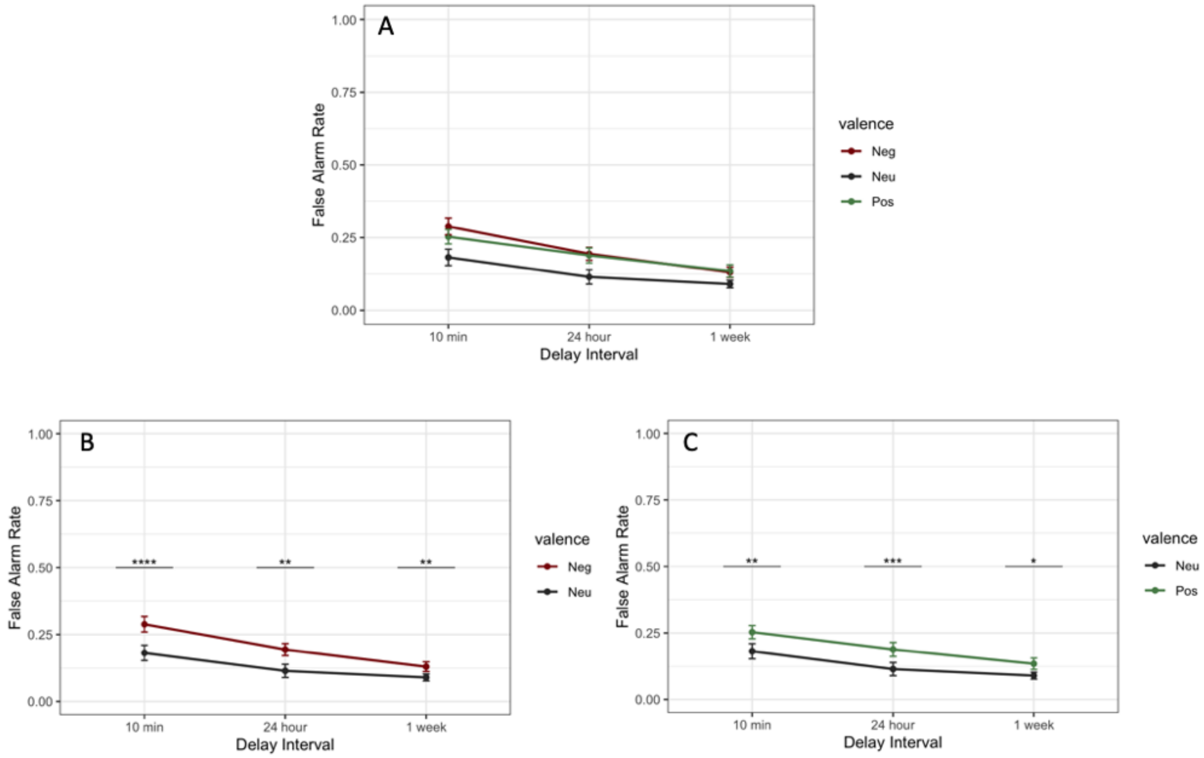
Note. A) Hit rate performance for the valence conditions across the delay intervals. B) Hit rate performance for the negative and neutral conditions across the delay intervals. C) Hit rate performance for the positive and neutral conditions across the delay intervals. Asterisks denote significant difference between valence conditions at that delay interval.
* p<.05; ** p<.01; *** p<.001; **** p<.0001

For false alarm rates, there was no interaction between Delay and Valence on false alarm rates, $F(4, 152)=1.76$, $p=.14$, $\eta^2_p=.04$ (see Figure 6A) indicating no significant differences in patterns of false alarm rates across conditions. There were main effects of Valence, $F(1.71, 64.98)=20.14$, $p<.001$, $\eta^2_p=.35$, and Delay, $F(1.72, 65.53)=20.51$, $p<.001$, $\eta^2_p=.35$.

To assess the predictions regarding differential forgetting for negative vs. neutral pictures and positive vs. neutral pictures, planned follow-up within-subjects ANOVAs were conducted for each pair of conditions separately. For the ANOVA with false alarm rates for negative and neutral pictures, as predicted, there was a significant interaction between Delay and Valence, $F(2, 76)=4.00$, $p=.02$, $\eta^2_p=.09$. Inspection of this interaction showed that false alarm rates declined slower for negative pictures compared to neutral pictures (see Figure 6B). In addition, there were also main effects of Valence, $F(1, 38)=27.21$, $p<.001$, $\eta^2_p=.42$, and Delay, $F(2, 76)=17.97$, $p<.001$, $\eta^2_p=.32$. Simple comparisons between the negative and neutral conditions on false alarm rate performance found that negative false alarm rates were significantly higher than neutral, $t(38)=5.42$, $p<.001$ at the 10 minute delay, and this effect was found at both later delay intervals as well (*p*s<.01).

For false alarm rates for positive and neutral pictures, there was no interaction between Delay and Valence, $F(2, 76)=0.68$, $p=.51$, $\eta^2_p=.02$, but there were main effects of Valence, $F(1, 38)=40.99$, $p<.001$, $\eta^2_p=.52$, and Delay, $F(2, 76)=14.28$, $p<.001$, $\eta^2_p=.27$ (see Figure 6C). Simple comparisons between the positive and neutral conditions on false alarm rate performance found positive false alarm rates were significantly higher than neutral, $t(38)=-3.53$, $p=.001$ at the 10 minute delay, and this effect was found at both later delay intervals as well (*p*s<.05).

Figure 6.
Forgetting Results for False Alarms



Note. A) False alarm rate performance for the valence conditions across the delay intervals. B) False alarm rate performance for the negative and neutral conditions across the delay intervals. C) False alarm rate performance for the positive and neutral conditions across the delay intervals. Asterisks denote significant difference between valence conditions at that delay interval.
* p<.05; ** p<.01; *** p<.001; **** p<.0001

3.2 Analysis of Recollection and Familiarity

To address our third main question, regarding potential differences in forgetting rates for recollection and familiarity, we examined differences in forgetting for estimates of recollection and familiarity, derived from the remember/familiar recognition task, for negative, positive, and neutral pictures. As noted in the Methods section, when memory performance differs significantly between experimental conditions at an initial time point, this raises the concern that comparisons of forgetting may be distorted by scaling effects. A recommended solution is to transform memory scores for later delay intervals into proportional forgetting scores, based on the corresponding initial level of memory performance, because this proportional measure is unaffected by possible scaling issues.

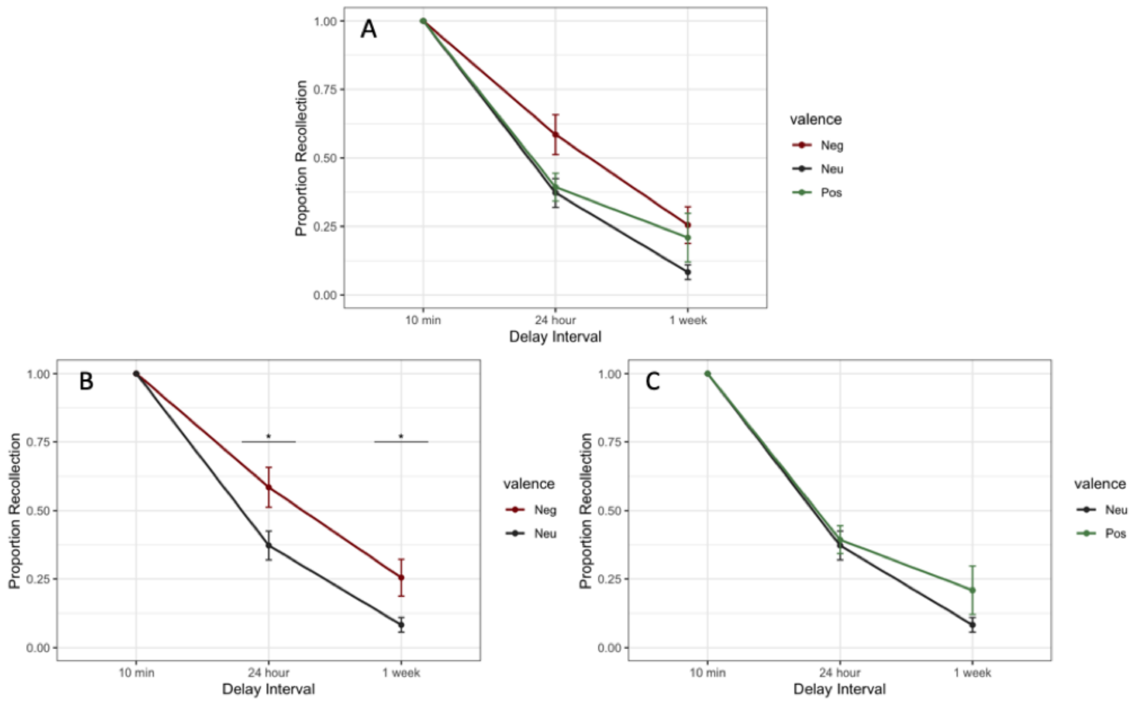
3.2.1 Recollection

We first examined whether the average recollection estimates for the negative, positive, and neutral conditions differed significantly at the first delay interval of 10 min. As an initial step, we excluded from further analysis the results from 4 participants who had initial recollection estimates near floor (less than .05) at the initial, 10 min time point, since initial near-floor memory scores precluded assessment of forgetting across the subsequent two delay intervals. As shown in Figure 8B and 8C, in line with previous findings (Ochsner, 2000), recollection estimates were higher for negative pictures than for neutral or positive pictures at the initial 10 minute delay interval, and were in turn higher for positive pictures than neutral pictures. Because of these significant differences at the initial 10 min delay interval, we converted the recollection scores at each delay interval into proportional forgetting scores by dividing the recollection scores at each delay by the corresponding initial recollection score, separately for each valence condition.

For proportional recollection, there was a significant interaction between Delay and Valence on corrected recognition scores, $F(2.01, 68.25)=2.56$, $p=0.08$, $\eta^2_p=.07$ (see Figure 7A), indicating differences in forgetting across conditions. There also were main effects of Valence, $F(1.70, 57.73)=5.87$, $p<.001$, $\eta^2_p=.15$ and Delay, $F(2, 68)=141.31$, $p<.001$, $\eta^2_p=.81$. To assess the predictions regarding differential forgetting for negative vs. neutral pictures and positive vs. neutral pictures, planned follow-up within-subjects ANOVAs were conducted for each pair of conditions separately. For the ANOVA with corrected recognition for negative and neutral pictures, as predicted, there was a significant interaction between Delay and Valence, F(2, 68)=4.316, p=.017, η^2_p=.11. Inspection of this interaction showed that forgetting was slower for negative pictures compared to neutral pictures (see Figure 7B). In addition, there were also main effects of Valence, F(1, 34)=10.04, p=.003, η^2_p=.23, and Delay, F(1.55, 52.54)=133.93, p<.001, η^2_p=.80. Simple comparisons between the negative and neutral conditions on proportional recollection scores found significant differences between these two conditions at the 24 hour delay, t(34)=2.60, p=.014 and at the 1 week delay, t(34)=2.68, p=.011.

For proportional recollection for positive and neutral pictures, there was no interaction between Delay and Valence, F(1.59, 54.15)=1.51, p=.23, η^2_p=.04, no main effect of Valence, F(1, 34)=1.34, p=.25, η^2_p=.04, and a main effect of Delay, F(1.65, 55.99)=146.60, p<.001, η^2_p=.81(see Figure 7C).

Figure 7.
Forgetting Results for Proportional Recollection Scores



Note. 4 participants removed due to low initial recollection values. A) Proportional recollection performance for the valence conditions across the delay intervals. B) Proportional recollection performance for the negative and neutral conditions across the delay intervals. C) Proportional recollection performance for the positive and neutral conditions across the delay intervals. Asterisks denote significant difference between valence conditions at that delay interval.
* p<.05; ** p<.01; *** p<.001; **** p<.0001

Although the initial differences in recollection at the 10 min delay interval required an analysis of proportional forgetting scores for the analysis of recollection, for comparison with other studies that have not analyzed proportional forgetting scores, we also conducted the same analysis on the original recollection estimates. As shown in Figure 8A, this 3 (Valence) x 3 (Delay) ANOVA with untransformed recollection estimates as the dependent variable found a significant interaction between Valence and Delay, $F(3.27, 111.23)=4.22$, $p=.006$, $\eta^2_p=.11$. However, inspection of this interaction showed that it was likely an artifact of floor effects in the recollection estimates for the neutral condition. There were significant main effects for both

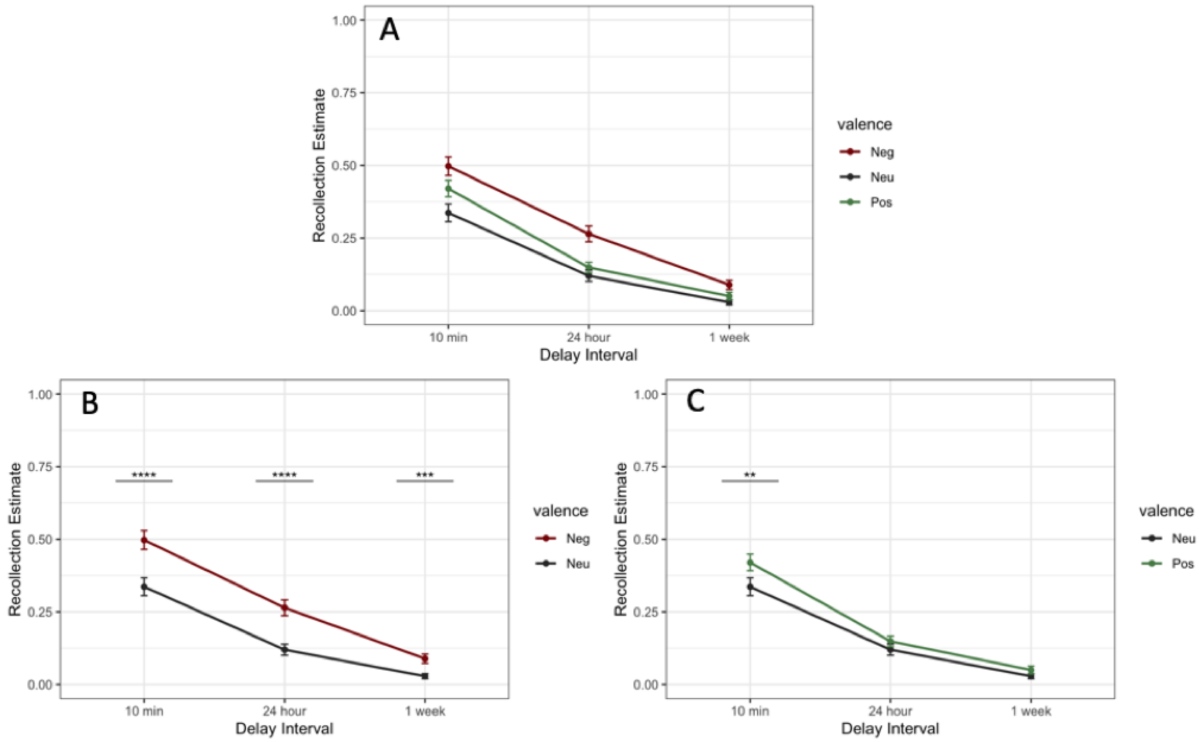
Valence, $F(1.69, 57.44)=41.30$, $p<.001$, $\eta^2_p=.55$ and Delay, $F(1.65, 56.26)=133.72$, $p<.001$, $\eta^2_p=.80$.

To assess the predictions regarding differential forgetting for negative vs. neutral pictures and positive vs. neutral pictures, planned follow-up within-subjects ANOVAs were conducted for each pair of conditions separately. For the ANOVA with original recollection estimates for negative and neutral pictures, as predicted, there was a significant interaction between Delay and Valence, $F(2, 68)=6.35$, $p=.003$, $\eta^2_p=.16$. Inspection of this interaction showed that forgetting was slower for negative pictures compared to neutral pictures (see Figure 8B). In addition, there were also main effects of Valence, $F(1, 34)=63.80$, $p<.001$, $\eta^2_p=.65$, and Delay, $F(1.63, 55.25)=108.75$, $p<.001$, $\eta^2_p=.76$. Simple comparisons between the negative and neutral conditions on original recollection estimates found significant differences between these two conditions at the 10 minute delay, $t(34)=5.93$, $p<.001$) as well as at the two later delays (ps<.001).

For original recollection estimates for positive and neutral pictures, there was no interaction between Delay and Valence, $F(1.39, 47.15)=2.84$, $p=0.08$, $\eta^2_p=0.08$, but there were main effects of Valence, $F(1, 34)=18.82$, $p<.001$, $\eta^2_p=.36$ and Delay, $F(1.58, 53.84)=132.84$, $p<.001$, $\eta^2_p=.80$ (see Figure 8C). Simple comparisons between the positive and neutral conditions on original recollection estimates found significant differences between these two conditions at the 10 minute delay, $t(34)=-3.05$, $p=.004$ but not at the two later delays (ps>0.05).

Figure 8.

Forgetting Results for Untransformed Recollection Estimates



Note. 4 participants removed due to low initial recollection values. A) Original recollection estimates for the valence conditions across the delay intervals. B) Original recollection estimates for the negative and neutral conditions across the delay intervals. C) Original recollection estimates for the positive and neutral conditions across the delay intervals. Asterisks denote significant difference between valence conditions at that delay interval.
* p<.05; ** p<.01; *** p<.001; **** p<.0001

3.2.2 Familiarity

We first excluded from further analysis the results from 6 participants who had initial familiarity estimates near floor (less than .05) at the initial, 10 min time point, since initial near-floor memory scores precluded assessment of forgetting across the subsequent two delay intervals. As shown in Figure 9B and 9C, familiarity estimates did not differ for negative, neutral, or positive pictures at the initial 10 minute delay interval, therefore, the transformation to proportional scores was not warranted.

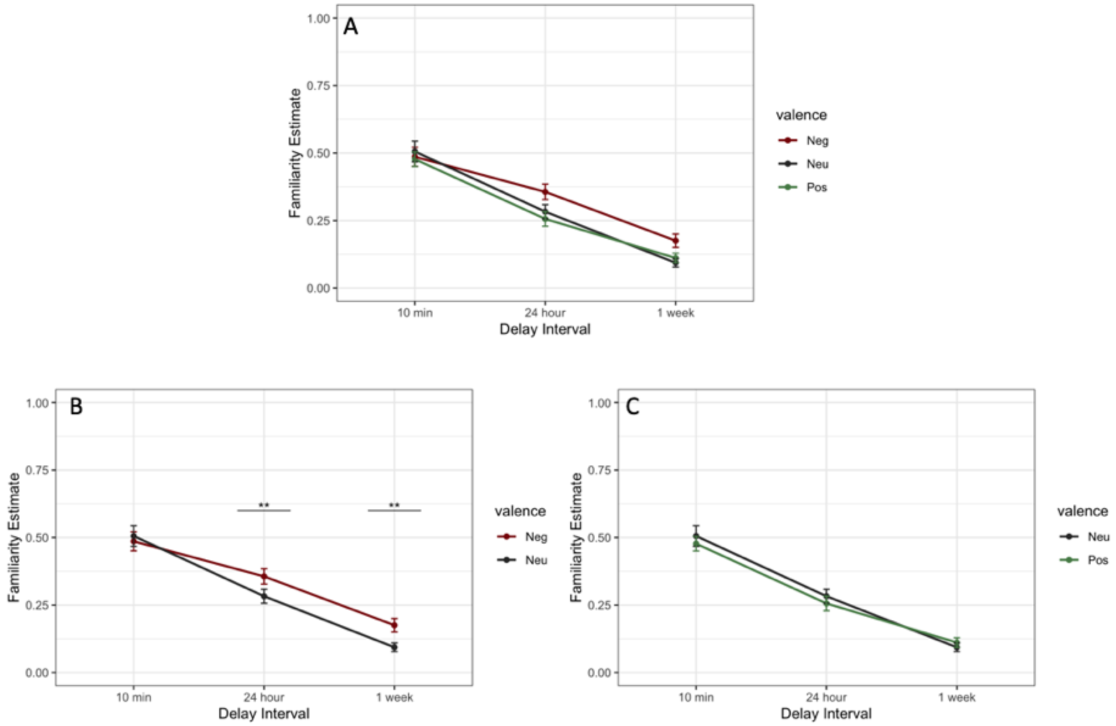
As shown in Figure 9A, this 3 (Valence) x 3 (Delay) ANOVA with familiarity estimates as the dependent variable found a significant interaction between Valence and Delay, *F*(4, 128)=2.91, *p*=.024, η^2_p=.08. There were significant main effects for both Valence, *F*(2, 64)=8.66, *p*<.001, η^2_p=.21 and Delay, *F*(1.56, 49.87)=70.86, *p*<.001, η^2_p=.69.

To assess the predictions regarding differential forgetting for negative vs. neutral pictures and positive vs. neutral pictures, planned follow-up within-subjects ANOVAs were conducted for each pair of conditions separately. For the ANOVA with familiarity estimates for negative and neutral pictures, as predicted, there was a significant interaction between Delay and Valence, *F*(2, 64)=4.73, *p*=.001, η^2_p=.13. Inspection of this interaction showed that forgetting was slower for negative pictures compared to neutral pictures (see Figure 9B). In addition, there were also main effects of Valence, *F*(1, 32)=8.06, *p*=.008, η^2_p=.20, and Delay, *F*(1.52, 48.59)=54.34, *p*<.001, η^2_p=.63. Simple comparisons between the negative and neutral conditions on original recollection estimates found no significant difference between these two conditions at the 10 minute delay, *t*(32)=-0.69, *p*=.49, but there were significant differences between these two conditions at the two later delays (*p*s<0.01).

For familiarity estimates for positive and neutral pictures, there was no interaction between Delay and Valence, *F*(2, 64)=1.42, *p*=.25, η^2_p =.04, no main effect of Valence, *F*(1, 32)=0.81, *p*=.38, η^2_p =.03, and a significant main effect of Delay, *F*(2, 64)=84.11, *p*<.001, η^2_p =.72 (see Figure 9C).

Figure 9.

Forgetting Results for Familiarity Estimates



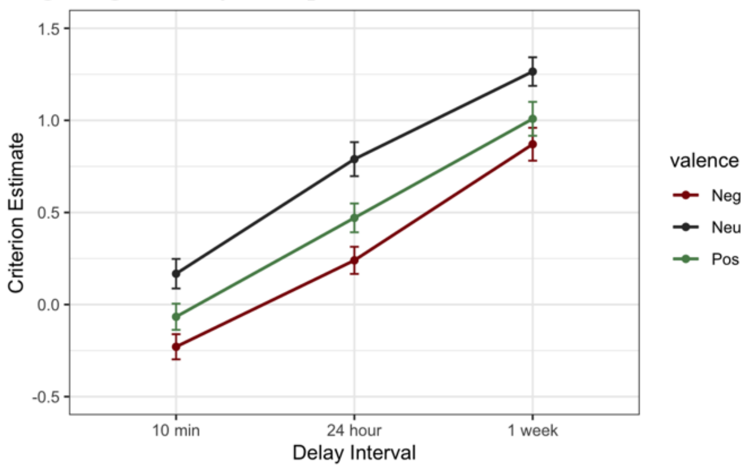
Note. 6 participants removed due to low initial familiarity values. A) Familiarity estimates for the valence conditions across the delay intervals. B) Familiarity estimates for the negative and neutral conditions across the delay intervals. C) Familiarity estimates for the positive and neutral conditions across the delay intervals. Asterisks denote significant difference between valence conditions at that delay interval.
* $p<.05$; ** $p<.01$; *** $p<.001$; **** $p<.0001$

3.3 Recognition response criteria

As briefly explained above, in signal detection theory the response criterion *c* represents the bias for an individual to judge an item on a recognition test as being old or new (MacMillan & Creelman, 2004). Typically, as a delay increases, participants' response criteria decrease (become more lenient) meaning that they tend to respond "yes I remember" more frequently (Singer & Wixted, 2006). However, in the current study, as shown in Figure 10, as delay increased, participants' response criteria became more strict (i.e., the criterion values *c* increased) and therefore participants had a greater tendency to respond "no".

As shown in Figure 10, this 3 (Valence) x 3 (Delay) ANOVA with response criterion as the dependent variable found no interaction between Valence and Delay, $F(4, 152)=1.262$, *p*=.29, η^2_p =.03. There were significant main effects for both Valence, *F*(1.73, 65.65)=61.812, *p*<.001, η^2_p =.62 and Delay, *F*(1.63, 62.05)=137.38, *p*<.001, η^2_p =.78. Simple comparisons between the delay intervals found negative response criteria significantly increased from the 10 minute delay to the 24 hour delay (*t*(38)=-9.18, *p*<.001) and from the 24 hour delay to the 1 week delay (*t*(38)=-7.29, *p*<.001). The same effects were found for neutral and positive (*p*s<.001). Simple comparisons between the valence conditions found the response c for each valence were significantly different from each other at each delay interval (*p*s<.01) except for positive and negative at the 1 week delay (*t*(38)=-1.82, *p*=.23).

Figure 10.
Forgetting Results for Response Criterion



Note. Average criterion c scores for each valence conditions across the delay intervals.

4. Discussion

The current study investigated three main questions concerning differences in rates of forgetting for emotional vs. neutral pictures. For the first two questions, which concerned the extent to which forgetting is slower for negative and positive emotional pictures relative to neutral stimuli, the results were relatively clear and readily interpretable . For negative pictures, overall forgetting assessed with recognition memory was slower relative to neutral pictures across the three delay intervals (10 minutes, 24 hours, and 1 week), consistent with several other studies. Also, whereas there was no EEM for negative pictures at the shortest delay interval, a significant EEM was observed at the latter two delay intervals, again consistent with prior findings that the EEM often is small or absent at a short delay and grows over time (Sharot & Yonelinas; Yonelinas & Ritchey, 2015).

In contrast, for positive emotional pictures, although we had tentatively predicted slower forgetting on the basis of the amygdala-modulated consolidation view, in fact there was no evidence for overall slower forgetting relative to neutral pictures, and the forgetting curves for positive and neutral pictures were very similar, particularly for corrected recognition performance (Figure 3). An EEM for positive pictures was observed for hit rates but not for the corrected recognition or d' measures. EEM effects are frequently smaller for positive emotional stimuli than for negative emotional stimuli, even when positive and negative stimuli are matched on normative arousal, as was the case in the current study (Talmi et al., 2007).

Our third primary question concerned the extent to which any potential slowed forgetting for negative or positive pictures (relative to neutral pictures) is reflected differently in the recollection and familiarity components that contribute to recognition memory performance. Although floor effects and differences in initial memory performance complicated the interpretation of the results, particularly for recollection, some relevant new findings were obtained. Consistent with our predictions, recollection for negative pictures (as assessed by proportional forgetting scores because of differences between valence conditions in initial levels of memory) was slowed relative to neutral pictures, consistent with both main theories and with the findings of most prior relevant studies. In contrast, we did not find evidence for slowed forgetting for recollection of positive pictures relative to neutral ones, and had made no specific predictions regarding forgetting for recollection for positive pictures.

Turning to familiarity, we observed slower forgetting for familiarity for negative pictures relative to neutral ones. This result is contrary to the predictions of the emotional binding view, which specifies that only recollection should be enhanced for negative emotional stimuli. However, this result it is broadly compatible with the amygdala-mediated consolidation view, which allows for modulation of familiarity processes. No evidence for slowed forgetting for familiarity for positive pictures was found, a finding that is compatible with both theories. The slower forgetting for familiarity for negative pictures matches a similar result reported by Sharot & Yonelinas (2008). Although slowed forgetting for familiarity was attributed in that study to the influence of a subsequent, separate source judgement, our current findings, which were obtained in the absence of this source judgment task, suggest that slowed forgetting for familiarity for negative pictures does occur and is not due to the influence of other memory tests.

The goal of the current study was not to determine which of the two major theories is more compatible with our findings. This was in part because gaps in the predictions of these theories made it difficult to directly compare between them. Nevertheless, it is useful and informative to discuss the current findings in the context of these two theories. To recap the major findings, forgetting was slower for negative pictures for overall recognition (i.e., for corrected recognition, d', and hit rates), consistent with the predictions of both theories, and this slower forgetting extended across all three delay intervals up to the 1 week delay. Another key finding was that slower forgetting was not observed for positive pictures, even though the positive pictures were matched to the negative pictures on normative rated arousal, and the enhancement of recognition for positive pictures was also weaker than the corresponding effect for negative pictures, being observed only for hit rates. Because both theories focus on the effects of negative emotion and do not make specific explicit predictions regarding positive emotion, these findings regarding memory for positive emotion are broadly consistent with both theories. Although these findings would not be predicted based on the amygdala-modulated consolidation view's proposal that emotional arousal is the primary factor influencing the amygdala's memory-enhancing functions, this view does acknowledge that there may be additional valence-specific effects on memory.

The finding that familiarity for negative pictures was slowed relative to neutral pictures appears difficult to reconcile with the emotional binding view, as this result is contrary to an explicit prediction of the model. The current findings replicate those of Sharot and Yonelinas (2008), who also found slower forgetting for familiarity for negative vs. neutral pictures, using a similar remember/know task, although they concluded that this finding was spurious and was due to the effects of including a source memory task. Finally, the null effects of positive emotion on forgetting for both recollection and familiarity are equally compatible with both theories.

Two studies by Wang, (Wang, 2014, 2018) addressed similar questions to those addressed in the current study, but the findings of these studies are difficult to interpret because of methodological issues and anomalous patterns of results. Wang (2014) attempted to analyze forgetting rates of positive, neutral, and negative pictures across three delay intervals. However, this first study used a between-subjects design with only 20 subjects in each group leaving the study considerably underpowered to detect differences in forgetting effects. There were also patterns of results that point to other potential issues in the study including a lack of forgetting in all three valence conditions between the 5 minute and 24 hour delays, and in the analyses of recollection, though the statistics were not reported, it appears that negative and neutral had no differences in recollection at any of the delay intervals. Finally, the study also attempted to analyze familiarity but the familiarity estimates were calculated incorrectly.

These methodological and result pattern issues were addressed by Wang in a subsequent study in 2018 (Wang, 2018). Wang (2018) used a within-subjects design that was properly powered, however other methodological issues persisted as well as the anomalous patterns of results. For example, Wang (2018) increased the number of delays from three to seven, but did not increase the stimuli set and therefore only 10 pictures of each valence were presented at each delay. This study did have evidence of forgetting in the overall recognition unlike in Wang (2014), but the problems with negative and neutral recollection persisted and at one delay interval, negative recollection was significantly lower than neutral. Familiarity was correctly calculated in Wang (2018), but the analyses found no forgetting or differences between the valence conditions. Given these anomalous results for recollection and familiarity, there is a possibility that the participants did not have a strong understanding of the instructions for retrieval. These methodological problems and issues with the results of these two studies make it difficult to interpret their findings with respect to the main questions we focused on in the current study.

There were some limitations associated with the current study. First, floor effects were present in some conditions, which complicated the interpretation of some results. False alarm rates for neutral pictures exhibited floor effects at the 24 hour and 1 week delays, thus contributing to inflated scores in d’ for the neutral condition at those delays. There were also floor effects in recollection and familiarity estimates for some conditions. Some subjects had recollection or familiarity estimates at floor at the first delay interval and therefore had to be excluded from the analyses of recollection and familiarity, and there were also floor effects at the 1 week delay which made it difficult to interpret the interactions.

Second, for recollection estimates, initial values differed significantly between valence conditions at the initial, 10 min delay and therefore we used proportional forgetting scores, which are often recommended to deal with possible scaling effect issues (Wixted, 1990; Ritchey, Dolcos, & Cabeza, 2008). Although the use of proportional scores to assess forgetting is well motivated theoretically, to assess the robustness of this finding across different measurement scales and performance levels, future studies should reexamine this question under conditions where initial recollection levels are matched and no conversion to proportional forgetting scores is required.

Finally, in the current study we used normative ratings to select the sets of negative, positive, and neutral stimuli and to match negative and positive stimuli on mean arousal, but emotion ratings for the stimuli from the subjects in this study were not collected, in part because of the large number of pictures and time constraints for the online study, which already required subjects to participate in three relatively long experimental sessions. Therefore, it is possible that stimuli did not consistently elicit the intended emotional responses in some subjects. However, normative ratings of arousal are often highly correlated with individual arousal ratings and with EEM effects, especially when obtained from the same population (Canli et al. 2000; Kensinger & Schacter 2006; Touryan et al. 2007). Future studies should acquire post-retrieval emotion ratings and psychophysiological responses for stimuli address this limitation. Another direction for further study is to determine whether similar effects are found using recall tasks. Because performance on recall tasks is widely considered to be primarily mediated by recollection (Baraly et al. 2017; Craik & Lockhart 1972), emotional forgetting effects in recollection can be assessed more directly using a recall task, obviating the need to estimate recollection from remember/familiar recognition task performance.

In conclusion, the current study showed slowed forgetting for negative vs. neutral pictures stimuli in overall recognition performance. Slowed forgetting for negative pictures was also found for both of the main component processes contributing to recognition memory, recollection and familiarity. In contrast, forgetting was not slower for positive pictures and the forgetting results for positive and neutral pictures were broadly similar. Importantly, this study highlighted that the slowed forgetting of emotional stimuli occurs in both recollection and familiarity but only for negative stimuli and not positive. In summary, the current study contributed to the understanding of forgetting for emotional episodic memory, finding that slowed forgetting was valence dependent, and that slowed forgetting for negative pictures is reflected in both the recollection and familiarity components of recognition memory. These findings suggest that current theories of forgetting for emotional episodic memory, which have focused primarily on negative emotional stimuli and on recollection-based forgetting effects, need to take into account the role of negative and positive valence and effects on familiarity processes.

References

Baraly, K. T. A., Hot, P., Davidson, P. S., & Talmi, D. (2017). *How emotional arousal enhances episodic memory*.

Bowen, H. J., Kark, S. M., & Kensinger, E. A. (2018). NEVER forget: Negative emotional valence enhances recapitulation. *Psychonomic Bulletin & Review*, *25*(3), 870–891. https://doi.org/10.3758/s13423-017-1313-9

Cabeza, R., & Stjacques, P. (2007). Functional neuroimaging of autobiographical memory. *Trends in Cognitive Sciences*, *11*, 219–227. https://doi.org/10.1016/j.tics.2007.02.005

Cahill, L., & McGaugh, J. L. (1998). Mechanisms of emotional arousal and lasting declarative memory. *Trends Neurosci*, *21*, 294–299.

Canli, T., Zhao, Z., Brewer, J., Gabrieli, J. D., & Cahill, L. (2000). Event-related activation in the human amygdala associates with later memory for individual emotional experience. Journal of Neuroscience, 20(19), RC99-RC99.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. Journal of verbal learning and verbal behavior, 11(6), 671-684.

Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. Psychological Review; Psychological Review, 107, 261.

Diana, R., Yonelinas, A., & Ranganath, C. (2007). Imaging recollection and familiarity in the medial temporal lobe: A three-component model. *Trends in Cognitive Sciences*, *11*, 379–386. https://doi.org/10.1016/j.tics.2007.08.001

Du, X., Zhan, L., Chen, G., Guo, D., Li, C., Moscovitch, M., & Yang, J. (2019). Differential activation of the medial temporal lobe during item and associative memory across time. *Neuropsychologia*, *135*, 107252. https://doi.org/10.1016/j.neuropsychologia.2019.107252

Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Sciences*, *5*(9), 394–400.

Hamann, S. (2009). The human amygdala and memory. In P. J. ; P. Whalen Elizabeth A. (Ed.), *The human amygdala* (p. 429). Guilford Press.

Hamann, S. B., Ely, T. D., Grafton, S. T., & Kilts, C. D. (1999). Amygdala activity related to enhanced memory for pleasant and aversive stimuli. *Nature Neuroscience*, *2*(3), 289–293.

Kensinger, E. A. (2004). Remembering emotional experiences: The contribution of valence and arousal. Reviews in the Neurosciences, 15, 241–252.

Kensinger, E. A., & Schacter, D. L. (2006). Processing emotional pictures and words: Effects of valence and arousal. Cognitive, Affective, & Behavioral Neuroscience, 6(2), 110-126.

Kleinsmith, L. J., & Kaplan, S. (1963). Paired-associate learning as a function of arousal and interpolated interval. Journal of Experimental Psychology, 65(2), 190–193. https://doi.org/10.1037/h0040288

Kurdi, B., Lozano, S., & Banaji, M. R. (2017). Introducing the open affective standardized image set (OASIS). Behavior research methods, 49(2), 457-470.

LaBar, Kevin S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience*, *7*, 54–64. https://doi.org/10.1038/nrn1825

LaBar, K.S., & Phelps, E. A. (1998). Arousal-mediated memory consolidation: Role of the medial temporal lobe in humans. *Psychological Science*, *9*, 490–493.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). International affective picture system (IAPS): Technical manual and affective ratings. NIMH Center for the Study of Emotion and Attention, 1, 39-58.

Loftus, G. R. (1985). Evaluating forgetting curves. Journal of Experimental Psychology: Learning, Memory, and Cognition, 11(2), 397.

Mather, M. (2007). Emotional Arousal and Memory Binding: An Object-Based Framework. *Perspectives on Psychological Science*, *2*(1), 33–52. https://doi.org/10.1111/j.1745-6916.2007.00028.x

Macmillan, N. A., & Creelman, C. D. (2004). Detection theory: A user's guide. Psychology press.

Marchewka, A., Żurawski, Ł., Jednoróg, K., & Grabowska, A. (2014). The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database. Behavior research methods, 46(2), 596-610.

McGaugh, J. L. (2004). The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annual Review of Neuroscience*, *27*(1), 1–28. https://doi.org/10.1146/annurev.neuro.27.070203.144157

Migo, E. M., Mayes, A. R., & Montaldi, D. (2012). Measuring recollection and familiarity: Improving the remember/know procedure. Consciousness and cognition, 21(3), 1435-1455.

Morris, J. D. (1995). Observations: SAM: the Self-Assessment Manikin; an efficient cross-cultural measurement of emotional response. Journal of advertising research, 35(6), 63-68.

Ochsner, K. N. (2000). Are affective events richly recollected or simply familiar? The experience and process of recognizing feelings past. Journal of experimental psychology: general, 129(2), 242.

Park, J. (2005). Effect of Arousal and Retention Delay on Memory: A Meta-Analysis. Psychological Reports, 97(2), 339–355. https://doi.org/10.2466/pr0.97.2.339-355

Payne, J. D., & Kensinger, E. A. (2010). Sleep's Role in the Consolidation of Emotional Episodic Memories. *Current Directions in Psychological Science*, *19*(5), 290–295. https://doi.org/10.1177/0963721410383978

Payne, J. D., & Kensinger, E. A. (2018). Stress, sleep, and the selective consolidation of emotional memories. *Current Opinion in Behavioral Sciences*, *19*, 36–43. https://doi.org/10.1016/j.cobeha.2017.09.006

Ritchey, M., Dolcos, F., & Cabeza, R. (2008). Role of amygdala connectivity in the persistence of emotional memories over time: An event-related fMRI investigation. Cerebral Cortex, 18(11), 2494-2504.

Rubin, D. C. (2005). A basic-systems approach to autobiographical memory. *Current Directions in Psychological Science*, *14*(2), 79–83.

Sadeh, T., Ozubko, J. D., Winocur, G., & Moscovitch, M. (2016). Forgetting Patterns Differentiate Between Two Forms of Memory Representation. *Psychological Science*, *27*(6), 810–820. https://doi.org/10.1177/0956797616638307

Schümann, D., & Sommer, T. (2018). Dissociable contributions of the amygdala to the immediate and delayed effects of emotional arousal on memory. *Learning & Memory*, *25*(6), 283–293. https://doi.org/10.1101/lm.047282.117

Schürer-Necker, E. (1990). Arousal and paired-associate learning. *The Pavlovian Journal of Biological Science*, *25*(4), 195–200. https://doi.org/10.1007/BF02900702

Sharot, T., & Phelps, E. A. (2004). How arousal modulates memory: Disentangling the effects of attention and retention. *Cognitive, Affective, & Behavioral Neuroscience*, *4*(3), 294–306. https://doi.org/10.3758/CABN.4.3.294

Sharot, T., & Yonelinas, A. P. (2008). Differential time-dependent effects of emotion on recollective experience and memory for contextual information. *Cognition*, *106*(1), 538–547. https://doi.org/10.1016/j.cognition.2007.03.002

Squire, L., & Zola-Morgan, S. (1991). The medial temporal lobe memory system. *Science*, *253*(5026), 1380–1386. https://doi.org/10.1126/science.1896849

Singer, M., & Wixted, J. T. (2006). Effect of delay on recognition decisions: Evidence for a criterion shift. Memory & Cognition, 34(1), 125-137.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. Behavior research methods, instruments, & computers, 31(1), 137-149.

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

Talmi, D., Anderson, A. K., Riggs, L., Caplan, J. B., & Moscovitch, M. (2008). Immediate memory consequences of the effect of emotion on attention to pictures. *Learning & Memory*, *15*(3), 172–182.

Talmi, D., Schimmack, U., Paterson, T., & Moscovitch, M. (2007). The role of attention and relatedness in emotionally enhanced memory. Emotion, 7(1), 89.

Touryan, S. R., Marian, D. E., & Shimamura, A. P. (2007). Effect of negative emotional pictures on associative memory for peripheral information. Memory, 15(2), 154-166.

Tulving, E. (1983). Elements of Episodic Memory. Clarendon Press.

Walker, M. P. (2010). Sleep, memory and emotion. *Progress in Brain Research*, *185*, 49.

Wang, B. (2014). Effect of Time Delay on Recognition Memory for Pictures: The Modulatory Role of Emotion. *PLOS ONE*, *9*(6), e100238. https://doi.org/10.1371/journal.pone.0100238

Wang, B. (2018). Retention interval modulates the effect of negative arousing pictures on recognition memory. *Memory*, *26*(8), 1105–1116. https://doi.org/10.1080/09658211.2018.1427763

Wixted, J. T. (1990). Analyzing the empirical course of forgetting. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16(5), 927.

Wixted, J. T., & Squire, L. R. (2011). The medial temporal lobe and the attributes of memory. *Trends in Cognitive Sciences*, *15*(5), 210–217. https://doi.org/10.1016/j.tics.2011.03.005

Yang, J., Zhan, L., Wang, Y., Du, X., Zhou, W., Ning, X., Sun, Q., & Moscovitch, M. (2016). Effects of learning experience on forgetting rates of item and associative memories. *Learning & Memory*, *23*(7), 365–378. https://doi.org/10.1101/lm.041210.115

Yonelinas, A. P. (2002). The Nature of Recollection and Familiarity: A Review of 30 Years of Research. *Journal of Memory and Language*, *46*(3), 441–517. https://doi.org/10.1006/jmla.2002.2864

Yonelinas, A. P., & Jacoby, L. L. (1995). The relation between remembering and knowing as bases for recognition: Effects of size congruency. Journal of memory and language, 34(5), 622-643.

Yonelinas, A. P., & Levy, B. J. (2002). Dissociating familiarity from recollection in human recognition memory: different rates of forgetting over short retention intervals. Psychonomic bulletin & review, 9(3), 575-582.

Yonelinas, A. P., & Ritchey, M. (2015). The slow forgetting of emotional episodic memories: An emotional binding account. *Trends in Cognitive Sciences*, *19*(5), 259–267. https://doi.org/10.1016/j.tics.2015.02.009

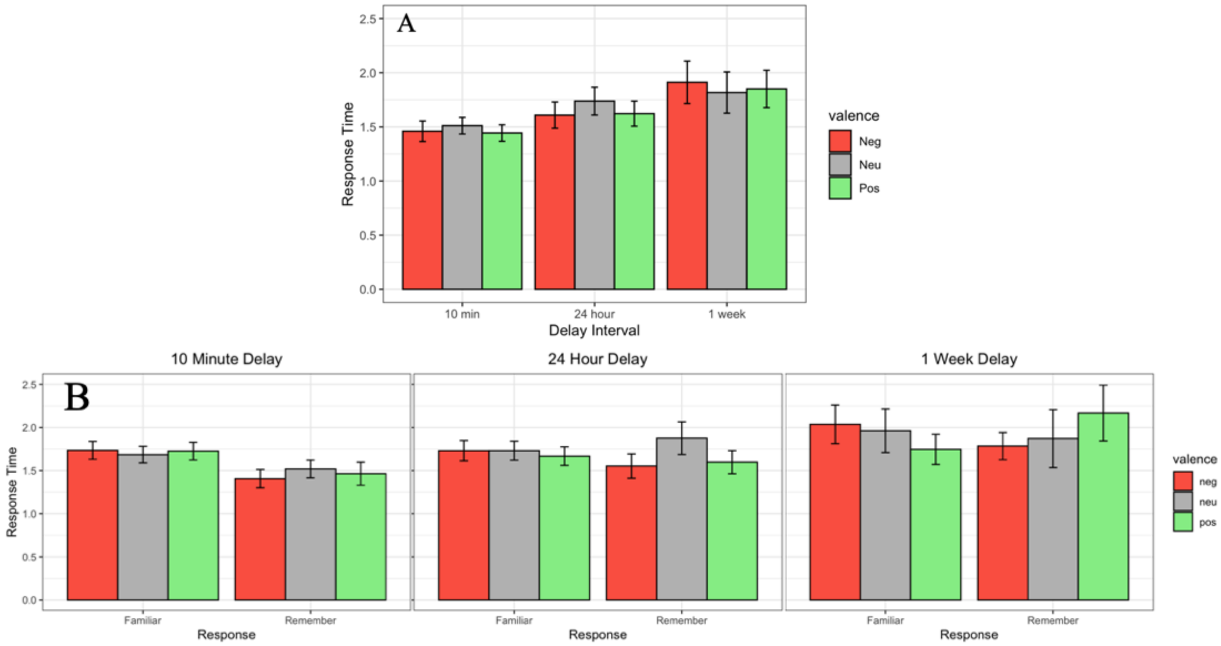
Appendix

1. Response times

As shown in Figure 1A, to characterize overall performance of response times (RTs) across conditions and delay intervals, RTs for hits were analyzed with a 3 (Valence) x 3 (Delay) repeated measures ANOVA. There was no interaction between Valence and Delay, *F*(2.76, 99.34)=1.54, *p*=.21, η^2_p =.04 or main effect of Valence, *F*(2, 72)=0.26, *p*=.77, η^2_p =.01. There was a main effect of Delay, *F*(1.33, 48.03)=8.91, *p*=.002, η^2_p =.20 characterized by the slowest RT at the 10 minute delay (*M*=1.47, *SE*=0.05) and longest at the 1 week delay (*M*=1.86, *SE*=0.11).

Finally, as shown in Figure 1B, to characterize performance of RTs across the responses of either “Familiar” or “Remember”, RTs were analyzed with a 2 (Response: Familiar, Remember) x 3 (Valence) x 3 (Delay) repeated measures ANOVA. There was no significant three-way interaction, *F*(4, 56)=1.8, *p*=.14, η^2_p =.11. There were also no significant 2-way interactions (*F*s<2.2). However, there were main effects for Response, *F*(1, 14)=8.5, *p*=.011, η^2_p =.38, with recollection (*M*=1.67, *SE*=0.059) lower than familiarity (*M*=1.78, *SE*=0.050) and Delay, *F*(1.35, 18.9)=4.68, *p*=0.034, η^2_p =.25 with the slowest RT at the 10 minute delay (*M*=1.59, *SE*=0.04) and longest at the 1 week delay (*M*=1.93, *SE*=0.10). Importantly, there were no main effects of Valence in either RT analysis, suggesting that any effects of valence on memory were unlikely to be due to a speed-accuracy tradeoff.

Figure 1.
Response time analysis



Note. A) Response times for the valence conditions across the delay intervals. B) Response times for familiarity and recollection responses at each delay.