



Medium strength visual long-term memories are the most fragile

Emma Megla¹ · Geoffrey F. Woodman¹

Accepted: 5 April 2021 / Published online: 13 May 2021
© The Psychonomic Society, Inc. 2021

Abstract

What memories do humans forget? One theory proposes that memories stored with moderate activation levels are weakened when faced with competitive stress so that they are particularly prone to be forgotten. However, research suggests that visual long-term memories are stronger than memories of other modalities, and therefore may never fall into this moderate activation zone. Here we tested these competing predictions by showing to-be-remembered pictures while we recorded event-related potentials (ERPs) indexing memory activation during encoding. We found that visual memories with medium levels of activation when first encoded were more prone to forgetting than memories with high or low encoding activation levels, but this only occurred if a memory was faced with competition. This study shows that we forget moderately activated memories when they are subjected to competition, regardless of the modality of experience.

Keywords Visual memory · Event-related potentials · Forgetting, human memory

Introduction

Why we forget has perplexed psychologists and neuroscientists for centuries (Ebbinghaus, 1885/1913; Freud, 1957; James, 1890). Scores of studies have shown that memories compete for encoding and retrieval in long-term memory (J. R. Anderson, 1981; M. C. Anderson, 2003; Lewis-Peacock & Norman, 2014). One theoretical view states that the memories we forget are those that have medium levels of activation, a theoretical region called the *zone of destruction*, due to stronger competitors weakening these potentially interfering

representations (Detre et al., 2013; Lewis-Peacock & Norman, 2014; Norman et al., 2007; Ritvo et al., 2019). An illustration of this idea is shown in Fig. 1.

However, it is not clear that this destructive competition plays out among visual representations in long-term memory. This is because both classic and recent research suggests that visual long-term memory representations may be too strong to be forgotten (Brady et al., 2008, 2013; Standing, 1973; Standing et al., 1970). For example, it has been proposed that memory for visual memoranda may be virtually perfect (Standing et al., 1970). Therefore, these visual representations may never be weak enough to fall into this destructive zone. Our goal in the present study was to test predictions of these competing theoretical perspectives by recording subjects' electroencephalogram (EEG) and measuring their event-related potentials (ERPs) elicited by the to-be-remembered pictures of objects. By measuring brain activity during the encoding of each picture, we can determine if a given picture is encoded with a weak, medium, or strong activation strength, and then test memory for these items after competitive stress is applied experimentally, as we describe next.

In the present study, we used a visual-long term memory task in which healthy young adults viewed pictures of to-be-remembered objects. We then had subjects restudy certain object exemplars because previous research has shown that repeating a specific exemplar of a category (e.g., a picture of a specific vase), can induce the forgetting of other categorically related exemplars that were not restudied (i.e., other vases)

Statement of Relevance

Here we show that we can measure the strength of memory encoding using electrodes placed on the surface of the head and predict which memories will be forgotten. Specifically, medium-strength visual memories, as classified based on the amplitude of a memory-related potential measured during encoding, were more prone to forgetting than visual memories encoded with lower or higher representational strength. Competition was required for these representations to be forgotten, as objects that were not faced with category competitors did not show this pattern. Our findings contribute to the view that what we forget is a combination of the competition faced by a given representation, as well as how well that representation was encoded initially.

✉ Geoffrey F. Woodman
geoffrey.f.woodman@vanderbilt.edu

¹ Department of Psychology, Vanderbilt Vision Research Center, Vanderbilt University, Wilson Hall, 111 21st Ave S, Nashville, TN 37212, USA

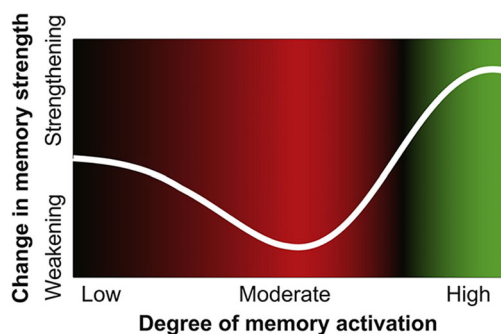


Fig. 1 Predicted effect of competition on our visual long-term memories for objects that vary in how strongly they are encoded into memory. We are supposed to see that memories with the medium activation levels at encoding are most easily forgotten after they are subjected to competition with other memories. Adapted from Detre et al. (2013)

compared to baseline objects that are also only seen once with no same-category exemplars restudied (Maxcey, Janakiefski, et al., 2019).¹ As a result, we were able to present subjects with a stream of to-be-remembered visual stimuli and induce the forgetting of certain representations by simply interleaving a repeating exemplar of that same category (see Fig. 2). Our task and stimuli distinguish the present study from previous studies that have used visual memoranda, but presented subjects with many exemplars from a couple of categories (Detre et al., 2013; Lewis-Peacock & Norman, 2014) or one exemplar from many categories (Brady et al., 2008; Standing, 1973), which could explain why memory researchers have reached such different conclusions about how well we can remember visual information.

The advantage of this paradigm is that subjects were never told to forget any of the items. Just the opposite, people were told that they needed to remember all of the items, but that some of the items would repeat as they studied the pictures in the stream of visual stimuli. Another advantage of studying forgetting in this context is that the paradigm allows us to experimentally induce forgetting of certain pictures of objects, while ensuring that subjects have seen the critical objects the same number of times and following delays of the same length. That is, both the exemplars that are experimentally forgotten (which we will call the *related items*) and the baseline items are shown just once during the initial study phase. In addition, both of these critical types of objects are remembered across the same temporal delay, with just the restudied exemplars and new exemplars appearing in the restudy phase, as shown in Fig. 2.

We combined this visual long-term memory paradigm with recordings of subjects' ERPs because the literature has shown that the amplitude of subjects' frontal ERPs measured during

memory encoding can provide a measure of memory strength (Paller et al., 2007; Rugg & Curran, 2007; Servant et al., 2018). Specifically, the frontal positivity, also known as the FN400 or N400, tracks the fidelity of long-term memory storage, with its sensitivity sufficient to measure the strength of memory encoding on a single trial (Fukuda & Woodman, 2015; Kutas & Federmeier, 2011). In addition, although it is unknown whether this ERP component has sufficient sensitivity to pick up on the non-monotonic changes in memory encoding strength in the present paradigm, previous work has shown that its amplitude tracks non-monotonic learning curves (Servant et al., 2018). For the present purposes, we did not focus on the canonical debate about whether this component measures the strength of familiarity or implicit memory (Paller et al., 2007; Rugg & Curran, 2007), but simply used this waveform as a way of covertly measuring memory storage strength from the brain.

If subjects forget the medium strength representations of objects in visual long-term memory, then when we measure the frontal positivity elicited by category exemplars during study, we should find that it is those exemplars that elicit medium amplitude positivities during encoding that are subsequently forgotten by the time subjects are tested at the end of the experiment. In contrast, if visual long-term memories are uniformly strong, then we should instead find that the amplitude of the frontal positivity elicited by items will be uniformly high and its amplitude will be unrelated to which items are forgotten.

Methods

Participants

We ran twenty-two participants (14 females, $M_{\text{age}} = 24.5$, $SD_{\text{age}} = 4.7$), guided by an a priori power analysis using effect sizes derived from previous work measuring similar brain potentials in other tasks (Fukuda & Woodman, 2015; Reinhart & Woodman, 2014). This estimate showed that a sample of 20 subjects would be sufficient to detect effects of the same magnitude 80% of the time at an alpha level of 0.05. All provided informed consent prior to experimental procedures as approved by the Vanderbilt University Institutional Review Board. Participants received a compensation of \$15/hour. All reported normal or corrected-to-normal visual acuity, normal color vision, and no history of neurological problems.

Stimuli

The stimuli were images of everyday objects centered on the monitor, and subtending $\sim 7^\circ \times 7^\circ$. The image set consisted of 32 categories, with 21 images in each category, for a total of

¹ Similar consequences of accessing a memory representation has been shown in the laboratory when subjects retrieve words from memory (Anderson et al., 1994). However, research with words suggests that the mere exposure of a subset of the exemplars is not sufficient to trigger this form of forgetting (Maxcey, Janakiefski, et al., 2019b), whereas it is with visual stimuli.

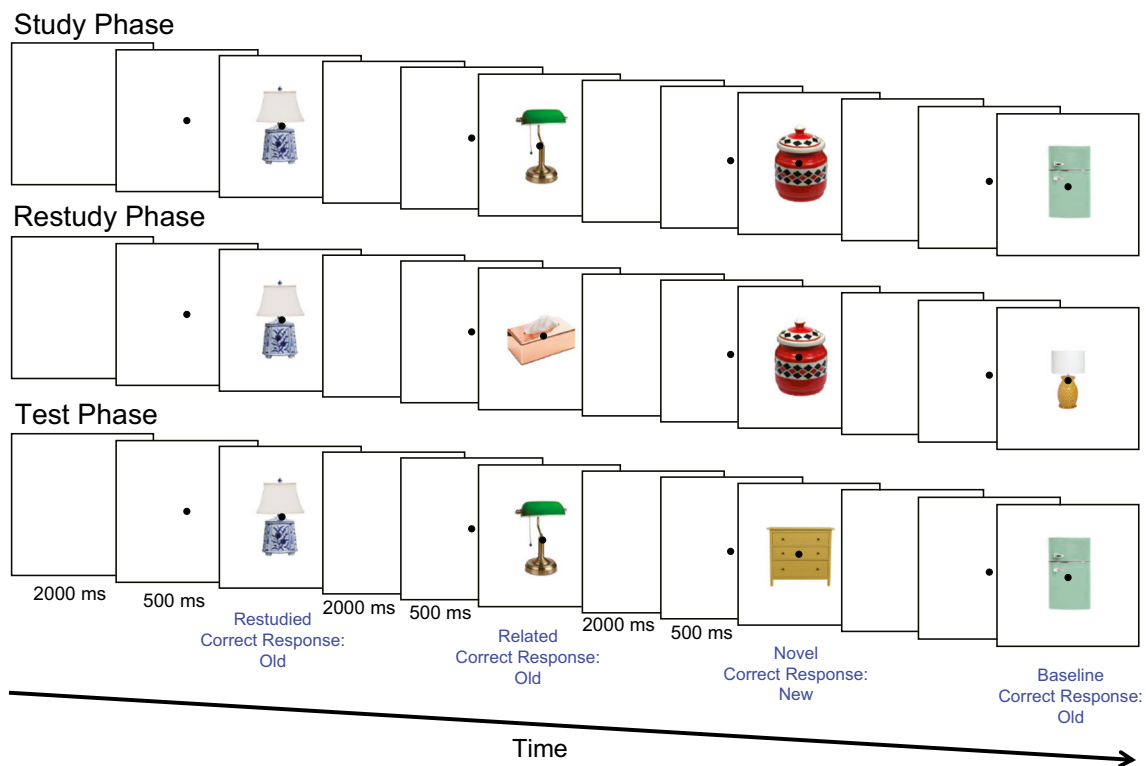


Fig. 2 An example of the stimuli and overview of the different phases of the experiment. In the study phase, a central fixation dot was presented for 500 ms followed by stimulus presentation in which a single object was presented on the screen for 2000 ms. Participants were instructed to maintain fixation for the duration of the trial while attempting to memorize all images presented for a later memory test. There was a 2000 ms intertrial interval in which participants could blink or move

their eyes between trials. In the restudy phase, half of the objects from half of the categories in the study phase were presented on the screen a total of three times throughout the phase. An equal number of novel objects from each category were also presented during this restudy phase. In the test phase, participants performed an old-new recognition task with an equal number of restudied, related, novel, and baseline objects presented

672 images. The full set of stimuli can be found here: <https://osf.io/yhaqn/>.

Procedure

An overview of the experiment structure is shown in Fig. 2. We instructed participants to memorize each object for a later memory test. Each trial started with a black fixation dot (6.89 cd/m^2) presented on a gray screen (30.5 cd/m^2 ; white in Fig. 2) for 500 ms followed by the stimulus presentation for 2000 ms. The fixation dot remained on the screen during this time to encourage participants to refrain from blinking or making eye movements. Following stimulus presentation, there was a 2000 ms inter-trial interval when participants could blink and move their eyes. Participants received a break every 64 images. This continued until all 384 images from 32 object categories (12 exemplars each) had been shown.

In the restudy phase, half of the studied images from half of the original categories were presented again. These categories and images were randomly chosen for each participant as well as presented in a random order. The images were shown a total of three times during the

restudy phase. An equal number of novel images in each of the categories were also presented. Before starting the restudy phase, we told participants that both old and new images would be presented and to restudy the old images while memorizing the new ones.

The restudy phase produced three different stimulus types: (1) restudied objects, or object exemplars that were restudied in the restudy phase (e.g., the blue lamp in Fig. 2), (2) related objects, or objects whose *category* was restudied, but the object exemplars themselves were not (e.g., the green desk lamp in Fig. 2), and (3) baseline objects, or objects whose entire category was absent from the restudy phase.

In the test phase, an equal number of all object types (i.e., restudied, related, and baseline) were randomly presented on the screen in addition to novel objects, which were divided equally between all categories (i.e., three novel objects for each of the 32 categories). Participants performed an old-new recognition task using a keyboard with either the *f* or the *j* key corresponding to old, with the keys counterbalanced between participants. Trials terminated with the keyboard response and were followed by the 2000 inter-trial interval. Before the test phase began, we told participants that 75% of the images would be old.

EEG Acquisition

We recorded the EEG during all phases of the experiment from a 20-channel cap with channels located according to the International 10-20 system (F3, F4, C3, C4, P3, P4, PO3, PO4, O1, O2, PO7, PO8, T3, T4, T5, T6, Fz, Cz, Pz). During recording, we kept impedance values below 4 k Ω . Data were referenced online to the right mastoid and re-referenced offline to the average of the left and right mastoid electrodes. We placed an electrode approximately 1 cm lateral to the outer canthi of each eye in addition to an electrode underneath the right eye to monitor eye movements and blinks. All channels were band-pass filtered from 0.01–100 Hz and digitized at 250 Hz.

EEG Analysis

We detected trials containing blinks, amplifier saturation, or excessive noise in the EEG by first running each subject's data through the EEGLAB Toolbox function `eegthresh.m` (Delorme & Makeig, 2004). This rejected any trials with voltages greater than +100 μ V or less than -100 μ V. Next, we used a split-half sliding window approach (window size = 200 ms, step size = 10 ms, threshold = 10 μ V), as used in Adam et al. (2018), on the remaining trials to further reject any trials with eye movements. This approach placed a 200 ms window every 10 ms from the beginning to the end of a trial in the difference HEOG signal (left HEM – right HEM). If the HEOG difference from the first half to the second half of the window was greater than 10 μ V, then the trial was rejected. An average of 9.33% of study trials and 9.13% of restudy trials were rejected for each participant.

ERP Analysis

We measured the amplitude of subjects' ERPs across the midline electrode Fz following our previous studies of the frontal positivity (Reinhart & Woodman, 2014; Servant et al., 2018) and used the measurement window of 200 – 1000 ms following stimulus onset as used previously to calculate mean amplitude (Fukuda & Woodman, 2015). Due to subjects responding prior to 1000 ms in the test phase, the frontal positivity measurement window was truncated to 200 – 800 ms for this phase. Analyses were performed on baseline corrected, but unfiltered data so that our measurements were not contaminated by filtering (JASP Team, 2020). For visualization purposes only, trials were low-pass filtered using the EEGLAB Toolbox function `eegfilt.m` with a half-amplitude low-pass cutoff at 30 Hz (Delorme & Makeig, 2004).

Our first step was to measure the amplitude of the frontal positivity elicited by each stimulus during the study phase. This approach was based on a recent study showing that the amplitude of the frontal positivity provides a trial-by-trial

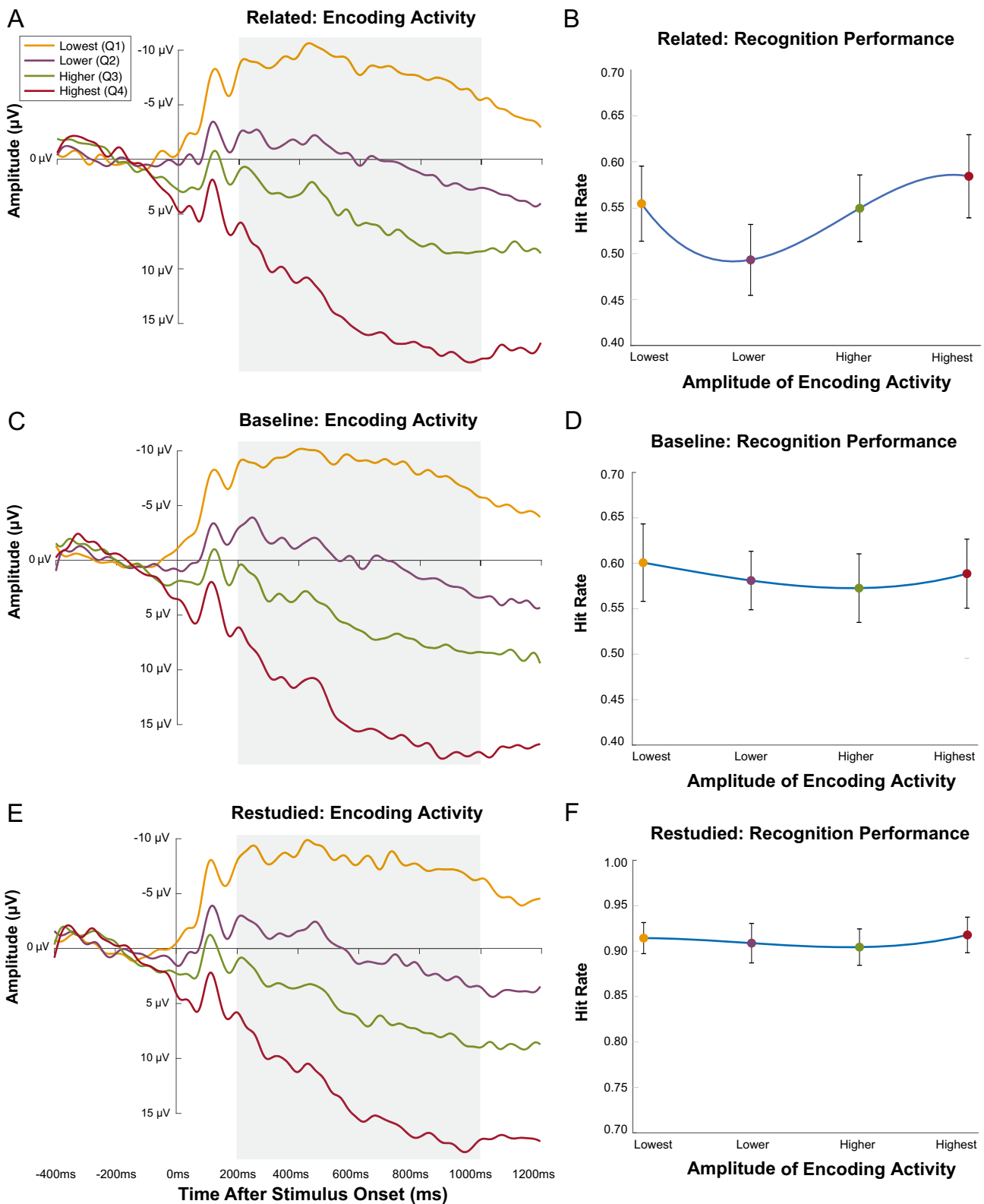
measure of the encoding quality, with more positive potentials resulting in better recognition performance in a subsequent memory test, and this amplitude being unrelated to the memorability or physical characteristics of the individual stimulus (Fukuda & Woodman, 2015). Measuring the frontal positivity for each stimulus allowed us to determine which objects were encoded with medium levels of activation as measured by the amplitude of their frontal positivity. To determine this, we divided objects into quartiles based on the amplitude of the frontal positivity elicited by that stimulus with this EEG data first baseline corrected from -400 to 0 ms to ensure trials were not sorted based on pre-stimulus noise. Objects that elicited the lowest quarter of frontal positivities comprised Quartile 1 (Q1). In contrast, objects encoded with the highest quarter of frontal positivity comprised Quartile 4 (Q4). The middle two quartiles, Quartile 2 (Q2) and Quartile 3 (Q3), contained the objects that were encoded with moderate levels of activation, or the objects thought to lie in the zone of destruction. We performed this sorting process separately for each subject and object type (i.e., restudied, related, and baseline).

Results

The first step in assessing the hypothesis that activation strength determines the fragility of human visual long-term memories is to measure the variability of activation strength at encoding. Figure 3 shows that when we sorted the stimuli into quartiles based on the amplitudes of their frontal positivities elicited by presenting each object in the study phase, we observed a large spread in the mean amplitudes across bins (i.e., approximately 30 μ V of range, with similar ranges observed across object types as shown in Fig. 3), resulting in a significant effect of bin using a one-way ANOVA across the four bins ($F(3,63) = 452.7$, $p < 0.001$, $\eta_p^2 = 0.956$) as well as pre-planned pairwise comparisons (Q1 vs. Q2 ($t(21) = -17.48$, $p < .001$, $d = -3.726$) Q2 vs. Q3 ($t(21) = -19.16$, $p < .001$, $d = -4.086$) Q3 vs. Q4 ($t(21) = -20.22$, $p < .001$, $d = -4.312$)).

This initial observation demonstrates that our measure of memory strength exhibits sufficient variability at encoding so as to result in distinguishable bins that differ in amplitude of the frontal positivity. Moreover, this is a general observation, with the baseline and restudied objects also exhibiting substantial variability in the amplitude of the frontal positivities elicited during the initial encoding events (see Fig. 3C & E). Next we ask whether the objects that have elicited these different measures of encoding activity exhibited different degrees of fragility when faced with competition.

The mean hit rates across subjects for the related objects across the four amplitude bins are shown in Fig. 3B. As you can see, following the restudy phase in which the related items



faced competition from subjects restudying certain exemplars from certain categories of objects, subjects’ memory for the related items was fairly accurate if that particular object had

elicited a particularly low or high amplitude frontal positivity. However, the objects that elicited medium amplitude frontal positivities were more easily forgotten.

◀ **Fig. 3** Study phase frontal positivity and behavioral performance for all object types. **A** Study phase grand-average ERP waveforms from electrode Fz (over the frontal lobe, along the midline) for related objects with quartiles separated based on memory activation (i.e., binned by amplitude of frontal positivity). The measurement window used to calculate the frontal positivity amplitude is shown in gray (i.e., 200–1000ms). **B** Hit rates for related objects during the test phase for each memory activation quartile as determined from encoding amplitude. Error bars show standard errors of the mean. **C** Study phase grand-average ERP waveforms for baseline objects separated by memory activation. **D** Hit rates for baseline objects during the test phase for each memory activation quartile. **E** Study phase grand-average ERP waveforms for restudied objects separated by memory activation. **F** Hit rates for baseline objects during the test phase for each memory activation quartile. Note that the y-axis range differs for restudied objects

To provide statistical support for this observation we performed a two-way repeated measures ANOVA of accuracy across object type (restudied, related, and baseline) and quartile bins (Q1, Q2, Q3, versus Q4) defined by the amplitude of the frontal positivity at encoding. This ANOVA yielded a significant main effect of object type ($F(2,42) = 78.225$, $p < .001$, $\eta_p^2 = 0.788$), although not for quartile bins ($F(3,63) = 2.270$, $p = 0.089$, $\eta_p^2 = 0.098$), nor object type \times quartile bin interaction ($F(6,126) = 1.283$, $p = 0.270$, $\eta_p^2 = 0.058$). Pre-planned pairwise comparisons showed that the related objects that elicited medium-low amplitudes (i.e., bin Q2; mean = 49.3% correct) were more prone to forgetting than related objects that elicited the smallest frontal positivities at encoding (i.e., bin Q1; mean = 55.4% correct, $t(21) = -3.182$, $p = 0.004$, $d = -0.678$) or the largest frontal positivities (i.e., bin Q4; mean = 58.4% correct, $t(21) = -3.371$, $p = 0.003$, $d = -0.719$). However, related objects in the medium-high bin (i.e., bin Q3; mean = 54.9% correct) were not forgotten more frequently than objects in the high or low bins (Q3 vs. Q1 ($t(21) = -0.188$, $p = 0.853$, $d = -0.049$) Q3 vs. Q4 ($t(21) = -1.299$, $p = 0.208$, $d = -0.277$)). Interestingly, previous experiments using words as memoranda have previously observed that items encoded with medium-low strength may be particularly fragile (Newman & Norman, 2010), and is the pattern that we see here with pictures of objects as well. Thus, the pattern of effects is consistent with the predictions of the zone of destruction account of forgetting in human memory.

Although the pattern of results obtained from our analyses of the related objects are consistent with the prediction that medium strength memories should be more easily forgotten when faced with competition, our analyses suggested that this did not occur for the related items, which were seen the same number of times, but did not have competition from restudied exemplars in their category. To assess this quantitatively, we analyzed the frontal positivities and the subsequent behavior elicited by the restudied and baseline objects. Here we found no evidence

that stimuli that elicited medium strength frontal positivities were easier to forget. Specifically, one-way repeated measures ANOVAs did not show a significant effect of bin (Q1, Q2, Q3, versus Q4) for either baseline ($F(3,63) = 0.255$, $p = 0.858$, $\eta_p^2 = 0.012$) or restudied objects ($F(3,63) = 0.272$, $p = 0.845$, $\eta_p^2 = 0.013$), although both of these object types differ from the pattern we observed across related objects. Thus, unlike the related objects that were subjected to competitive stress during the middle restudy phase of the experiment, the restudied and baseline objects exhibited a generally linear trend relating amplitude of the frontal positivity to how easy it was to remember a given object.

To ensure that these behavioral results were not simply due to response bias, we also calculated sensitivity indexes (P_r , d' , and A') for each memory activation quartile within each object type, as well as a measure of bias (C). These means are shown in Table S1. We found the same pattern of performance using these metrics as we observed with hit rate, showing that subjects' hit rates were not simply an artifact of response bias.

Recall that we also measured subjects' ERPs during the restudy phase of the experiment. We first baseline corrected this EEG data using the interval from -200 to 0 ms relative to stimulus onset. This baseline was 200 ms shorter than we used to baseline correct the encoding ERPs, allowing us to lose fewer trials to blink artifacts occurring between memory test events. Subjects' grand-average ERP waveforms during this restudy phase are shown in Fig. 4. The waveforms recorded during this phase, in which objects were repeatedly reshowed, allowed us to verify that the amplitude of the frontal positivity systematically changes as a memory is strengthened. As shown, there was a parametric increase in amplitude of the positivity with each exposure of the to-be-remembered object. A one-way repeated measures ANOVA showed a significant effect of object repetition (new, restudy repetition 1, 2, versus 3) ($F(3,63) = 3.761$, $p = 0.015$, $\eta_p^2 = 0.152$). Follow-up analyses showed that the pairwise comparisons of neighboring conditions were not significant for the first couple of stimulus presentations (novel vs. restudy repetition 1 ($t(21) = -0.895$, $p = 0.381$, $d = 0.191$) restudy repetition 1 vs. restudy repetition 2 ($t(21) = -1.904$, $p = 0.071$, $d = -0.406$) restudy repetition 2 vs. restudy repetition 3 ($t(21) = 0.276$, $p = -0.785$, $d = 0.059$)). However, there was a significant difference between the amplitude of the frontal positivity elicited by novel objects and restudied objects on their later repetitions (novel vs. restudy repetition 2 ($t(21) = -3.346$, $p = 0.003$, $d = -0.713$) novel vs. restudy repetition 3 ($t(21) = -2.958$, $p = 0.008$, $d = -0.631$)). Thus, we saw that each learning event systematically increased the amplitude of the positivity that we measured, confirming our assumption based on previous research that the amplitude of the frontal positivity tracks the strength of the memory representation.

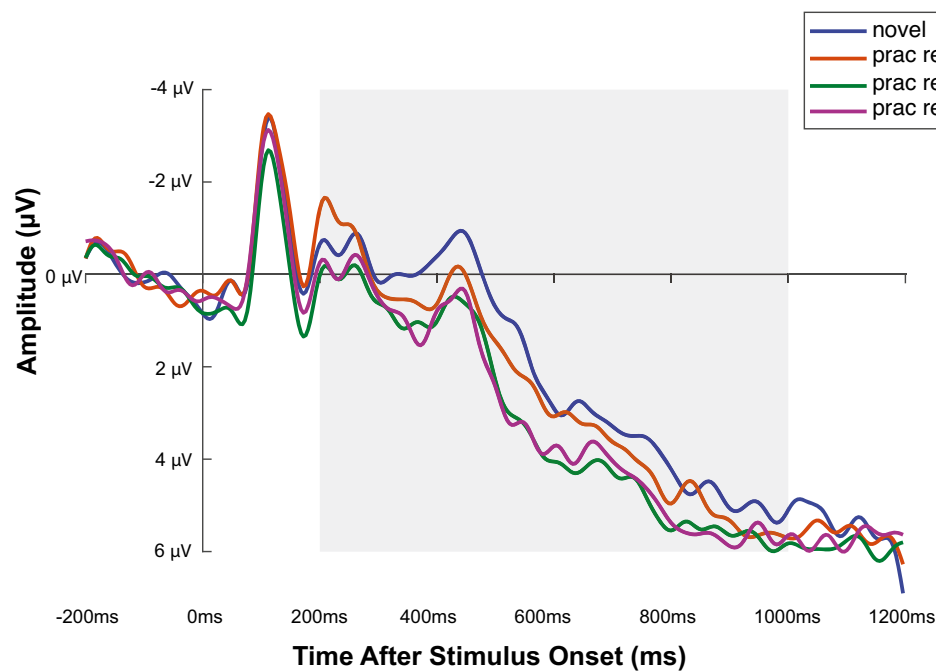


Fig. 4 Restudied phase frontal positivity. These grand average ERP waveforms were measured at electrode Fz and were elicited by the object when they were new (novel) and each restudy presentation (i.e.,

restudy rep 1 = the first time a picture is restudied). The measurement window used to calculate the frontal positivity amplitude is shown in gray (i.e., 200 – 1000ms)

Discussion

The goal of this study was to determine whether moderately activated visual long-term memory representations are more susceptible to forgetting (Detre et al., 2013; Lewis-Peacock & Norman, 2014). Despite existing evidence that visual representations are particularly strong, our findings show that when representations are stored with medium levels of activation, they are especially vulnerable to forgetting when faced with competition from other memories. This pattern of memorability was true only of representations that faced competition. Memoranda from categories that were not subjected experimentally to competition did not show this telltale pattern. Nor did those that were restudied enough so that they were too strong to be in the zone of destruction.

A reader may be confused as to why we saw forgetting in one medium-strength quartile and not the other. Specifically, pictures in the Q2 bin (i.e., medium-low activation) exhibited weakened visual memories following competition, whereas the other moderate quartile, the Q3 bin (i.e., medium-high), did not. Although these results seem at odds, they are consistent with previous empirical studies that have observed that fragility was particularly true of medium-low strength memories (Newman & Norman, 2010). We believe this pattern is due to inherent quality of the memories in the medium-low quartile itself, but whether this pattern is indicative of the nature of the underlying memory mechanisms at work to mitigate interference in memory will require additional work.

Specifically, it would be ideal to fit our data with a polynomial function to understand the true nature of the function (Detre et al., 2013).

Our observations indicate that the mechanisms of memory that determine the memorability of an event may operate according to similar principles regardless of the sensory modality that we gain experience through, although, further studies are needed to directly compare memory for verbal and visual materials within the same subjects using the approach we used here. Classic studies of visual long-term memory had suggested that our memory for information from this modality might be almost perfect (Brady et al., 2008; Standing, 1973). By measuring the variability of memory encoding electrophysiologically, we were able to observe differences in encoding strength, and watch as those translated into differences in recognition memory if a given memory representation was subsequently faced with competition. This pattern distinguishes the present study from previous work that presented subjects with visual memoranda to localize certain stimulus-selective regions of cortex using neuroimaging (Detre et al., 2013; Lewis-Peacock & Norman, 2014). In the previous work, subjects were presented with many exemplars from just a couple of categories (e.g., pictures of faces and pictures of scenes). Our findings indicate that under this high degree of within-category competition, forgetting will be much more extreme than when subjects are presented with pictures of objects from different categories, such as in the work suggesting that visual long-term memory for pictures

is nearly perfect (e.g., Brady et al., 2008; Standing, 1973). Thus, our findings demonstrate the generality of models of human memory that account for forgetting and appear to reconcile discrepant findings about whether visual information is forgettable.

What causes the destruction of the medium strength memories? Theoretical perspectives propose that the medium strength memory traces are destroyed because they are particularly potent distractors in memory. But the mechanism of their demise is far less well defined. For example, it is possible that the medium strength memories are actively inhibited (Anderson, 2003; Anderson et al., 1994). Alternatively, it is possible that the restudy events in our experiment simply make the matching memory representation impossible to access because the practiced items out compete the related items for retrieval. If true, it should be possible to decode neural activity and see that related objects memory probes elicit retrieval of the practiced competitors (Lewis-Peacock & Norman, 2014). Our future research is targeted at answering these questions.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.3758/s13423-021-01929-6>.

Acknowledgements We would like to thank Ashleigh Maxcey for stimulating this work through her extensive discussions with both of the authors. David Sutterer, Sisi Wang, and Chong Zhao provided vital guidance with coding and data analysis. The present work was supported by grants from the National Institutes of Health (R01-EY019882, R01-MH110378, and P30-EY08126) to G.F.W.

References

- Adam, K. C. S., Robison, M. K., & Vogel, E. K. (2018). Contralateral Delay Activity Tracks Fluctuations in Working Memory Performance. *Journal of Cognitive Neuroscience*, 30(9), 1229–1240.
- Anderson, J. R. (1981). Interference: The relationship between response latency and response accuracy. *Journal of Experimental Psychology: Human Learning and Memory*, 7(5), 326–343.
- Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory and Language*, 49(4), 415–445.
- Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(5), 1063.
- Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences*, 105(38), 14325–14329.
- Brady, T. F., Konkle, T., Gill, J., Oliva, A., & Alvarez, G. A. (2013). Visual Long-Term Memory Has the Same Limit on Fidelity as Visual Working Memory. *Psychological Science*, 24(6), 981–990.
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9–21.
- Detre, G. J., Natarajan, A., Gershman, S. J., & Norman, K. A. (2013). Moderate levels of activation lead to forgetting in the think/no-think paradigm. *Neuropsychologia*, 51(12), 2371–2388.
- Ebbinghaus, H. (1885). *Über das Gedchtnis: New York by Teachers College, Columbia University*.
- Freud, S. (1957). Repression. In *The Standard Edition of the Complete Psychological Works of Sigmund Freud, Volume XIV (1914-1916): On the History of the Psycho-Analytic Movement, Papers on Metapsychology and Other Works* (pp. 141–158).
- Fukuda, K., & Woodman, G. F. (2015). Predicting and Improving Recognition Memory Using Multiple Electrophysiological Signals in Real Time. *Psychological Science*, 26(7), 1026–1037.
- James, W. (1890). *The principles of psychology*. New York: Holt.
- JASP Team. (2020). *JASP (Version 0.14) [Computer Software]*. <https://jasp-stats.org/>
- Kutas, M., & Federmeier, K. D. (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP). *Annual Review of Psychology*, 62(1), 621–647.
- Lewis-Peacock, J. A., & Norman, K. A. (2014). Competition between items in working memory leads to forgetting. *Nature Communications*, 5(1), 5768.
- Maxcey, A. M., Janakiewski, L., Megla, E., Smerdel, M., & Stallkamp, S. (2019). Modality-specific forgetting. *Psychonomic Bulletin & Review*, 26(2), 622–633.
- Newman, E. L., & Norman, K. A. (2010). Moderate Excitation Leads to Weakening of Perceptual Representations. *Cerebral Cortex*, 20(11), 2760–2770.
- Norman, K. A., Newman, E. L., & Detre, G. (2007). A neural network model of retrieval-induced forgetting. *Psychological Review*, 114(4), 887.
- Paller, K. A., Voss, J. L., & Boehm, S. G. (2007). Validating neural correlates of familiarity. *Trends in Cognitive Sciences*, 11(6), 243–250.
- Reinhart, R. M. G., & Woodman, G. F. (2014). High Stakes Trigger the Use of Multiple Memories to Enhance the Control of Attention. *Cerebral Cortex*, 24(8), 2022–2035.
- Ritvo, V. J. H., Turk-Browne, N. B., & Norman, K. A. (2019). Nonmonotonic Plasticity: How Memory Retrieval Drives Learning. *Trends in Cognitive Sciences*, 23(9), 726–742.
- Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, 11(6), 251–257.
- Servant, M., Cassey, P., Woodman, G. F., & Logan, G. D. (2018). Neural bases of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(3), 440–464.
- Standing, L. (1973). Learning 10000 pictures. *Quarterly Journal of Experimental Psychology*, 25(2), 207–222.
- Standing, L., Conezio, J., & Haber, R. N. (1970). Perception and memory for pictures: Single-trial learning of 2500 visual stimuli. 19, 73–74.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.