

Predicting visual memory across images and within individuals

Cheyenne D. Wakeland-Hart^{a,b}, Steven A. Cao^a, Megan T. deBettencourt^{a,1}, Wilma A. Bainbridge^{a,c,1}, Monica D. Rosenberg^{a,c,*1}

^a Department of Psychology, University of Chicago, Chicago, IL, USA

^b Department of Psychology, Columbia University, New York, NY, USA

^c Neuroscience Institute, University of Chicago, Chicago, IL, USA

ARTICLE INFO

Keywords:
Memorability
Sustained attention
Recognition memory

ABSTRACT

We only remember a fraction of what we see—including images that are highly memorable and those that we encounter during highly attentive states. However, most models of human memory disregard both an image's memorability and an individual's fluctuating attentional states. Here, we build the first model of memory synthesizing these two disparate factors to predict subsequent image recognition. We combine memorability scores of 1100 images (Experiment 1, $n = 706$) and attentional state indexed by response time on a continuous performance task (Experiments 2 and 3, $n = 57$ total). Image memorability and sustained attentional state explained significant variance in image memory, and a joint model of memory including both factors outperformed models including either factor alone. Furthermore, models including both factors successfully predicted memory in an out-of-sample group. Thus, building models based on individual- and image-specific factors allows for directed forecasting of our memories.

Significance statement: Although memory is a fundamental cognitive process, much of the time memory failures cannot be predicted until it is too late. However, in this study, we show that much of memory is surprisingly *predetermined* ahead of time, by factors shared across the population and highly specific to each individual. Specifically, we build a new multidimensional model that predicts memory based just on the images a person sees and when they see them. This research synthesizes findings from disparate domains ranging from computer vision, attention, and memory into a predictive model. These findings have resounding implications for domains such as education, business, and marketing, where it is a top priority to predict (and even manipulate) what information people will remember.

1. Introduction

We remember some items that we encounter in our day-to-day lives with ease. For example, we may recall a painting that we saw in a museum long after our visit. At the same time, we often fail to remember other, similar items, such as the wall art in a doctor's office. Despite the fact that what items go on to be remembered or forgotten can seem arbitrary, our memory is not completely unpredictable.

Intuitively, it seems that what we go on to remember should be largely determined by processing that occurs during and after encoding. However, recent work has discovered two factors that determine our memories ahead of time. First, individuals collectively show strong agreement in which images they will remember. That is, certain images are intrinsically memorable or forgettable (Bainbridge, Isola, & Oliva,

2013; Isola, Xiao, Torralba, & Oliva, 2011). At the same time, each individual exhibits idiosyncratic moment-to-moment sustained attention dynamics, and the attentional state leading up to the moment of encoding impacts subsequent memory (deBettencourt, Norman, & Turk-Browne, 2018). This suggests, for example, that we're unlikely to remember a forgettable art piece or recall an exhibit encountered after our sustained attention had faded during a long museum visit. Neither memorability nor sustained attention, however, feature in most models of visual memory.

By testing many individuals on diverse stimulus sets, studies of visual memory have revealed striking consistency in the pictures that are remembered or forgotten (Bainbridge et al., 2013; Isola, Parikh, Torralba, & Oliva, 2011; Isola, Xiao, et al., 2011; Isola, Xiao, Parikh, Torralba, & Oliva, 2014). In other words, ahead of time, we can make

* Corresponding author at: Department of Psychology, University of Chicago, Chicago, IL, USA.

E-mail address: mdrosenberg@uchicago.edu (M.D. Rosenberg).

¹ Authors contributed equally

predictions about which specific images will be remembered. This widespread consistency implies that memorability is a property inherent to an image itself. This memorability can be quantified through continuous recognition tasks, in which participants detect specific stimulus repeats in a stream of images. Certain images, i.e., high memorability images, are much more likely to be correctly detected in these tasks. Importantly, a memorability score measured in one experiment has been shown to successfully translate across other tasks, participants, image contexts, and delays (e.g., Bainbridge, 2020; Broers, Potter, & Nieuwenstein, 2018; Goetschalckx, Moors, & Wagemans, 2018). Furthermore, memorability is not simply a product of an item's low-level visual features (color, texture, shape, orientation, or spatial frequency) alone, nor is it a product of attractiveness or visual interest. Instead, it is distinct from other visual factors (Bainbridge, 2019; Bainbridge et al., 2013) and unaffected by reward and cognitive control (Bainbridge, 2020). In sum, research on image memorability emphasizes that individuals' memories are enhanced for specific stimuli over others.

By tracking individuals' behavior over time, on the other hand, studies of sustained attention have revealed how fluctuating attentional states impact what is later remembered (Barel & Tzischinsky, 2020; deBettencourt et al., 2018; Madore et al., 2020; Song, Finn, & Rosenberg, 2021). That is, when attention is lapsing, we can predict that the forthcoming image will be disadvantaged. Changes in sustained attention from one moment to the next can be measured via continuous performance tasks (CPTs), in which participants repeatedly make the same response to the vast majority of stimuli, but then must make a different response to a rarely presented stimulus. Sustained attentional states can be operationalized via behavioral performance on such a task, with lapsing attentional states indexed by measures such as misses (incorrect responses to infrequent target trials), false alarms (incorrect responses to frequent nontarget trials), and faster and more prepotent responses (deBettencourt, Cohen, Lee, Norman, & Turk-Browne, 2015; deBettencourt, Keene, Awh, & Vogel, 2019; Robertson, Manly, Andrade, Baddeley, & Jenny, 1997; Rosenberg, Noonan, DeGutis, & Esterman, 2013). In sum, research on sustained attention emphasizes that an individual's memories are enhanced for images that appear in engaged attentional states.

Although image memorability and sustained attentional state are each important for what we remember, no work to date has combined these two factors. Do external image features common across the population (like memorability) and internal mental states idiosyncratic to

individuals (like sustained attention) explain unique variance in what we remember? Can we make honed predictions of what people will remember based on the memorability of a given image and their sustained attentional state during that time? To ask these questions, we built a model of visual long-term memory that leverages the influence of both image memorability and individual sustained attentional state. In Experiments 1 and 2 we collected new data to measure image memorability and participant sustained attention and memory, respectively. In Experiment 3 we reanalyzed existing data measuring participant sustained attention and memory to test the replicability of the results found in Experiments 1 and 2.

Leveraging these memorability scores and behavioral measures of attentional states at each moment, we combined data across experiments to build a model of subsequent memory. Specifically, the model predicted whether images are remembered or forgotten based on image memorability scores and response time (RT) signatures of trial-to-trial fluctuations in attention. Results revealed that image memorability and sustained attentional state uniquely predicted memory, and together explained more variance in what people remembered than either factor alone. Thus, armed only with the memorability of an item and measures of someone's attentional state, we can successfully predict what individuals will go on to remember using these two factors previously described in completely separate literatures.

2. Methods

We used data from three experiments to characterize the distinct contributions of image memorability and sustained attentional state to subsequent memory (Fig. 1). In Experiment 1, we ran a large-scale crowd-sourced online experiment to derive the intrinsic memorability of 1100 scene images. In Experiment 2, we collected data as participants performed a CPT and subsequent recognition memory test with these images. In Experiment 3, we re-analyzed data from a study in which different participants performed a CPT and subsequent recognition memory test with the same images. Data from Experiments 2 and 3 were collected in different research labs at different universities. We then asked whether the intrinsic memorability of images themselves (Experiment 1)—and participants' attentional state during the CPT (Experiments 2 and 3)—uniquely predicted subsequent image memory (Fig. 2).

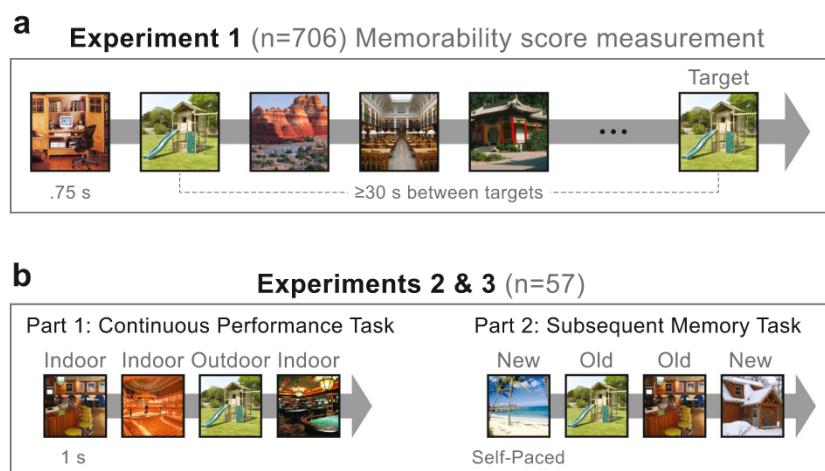


Fig. 1. a. The goal of Experiment 1 was to obtain memorability scores for a large set of real-world scene images. Participants completed a continuous recognition task to respond whenever an image repeated. Targets represented a second presentation of the same image separated by at least 30 s. b. The goal of Experiments 2 and 3 was to determine how stimulus-specific (image memorability) and individual-specific (attentional state) factors influenced long-term memory. These experiments consisted of two parts: a continuous performance task and a subsequent memory task. In the first part, participants categorized these same scene images by pressing one button to frequent (e.g., indoor; 90%) and another button to infrequent (e.g., outdoor; 10%) images. The RTs to the prepotent response (e.g., indoor) indexed their moment-to-moment attentional state. In the second part, participants performed a surprise subsequent memory task in which they reported recognition memory on a 4-point confidence rating scale for new and old indoor and outdoor images.

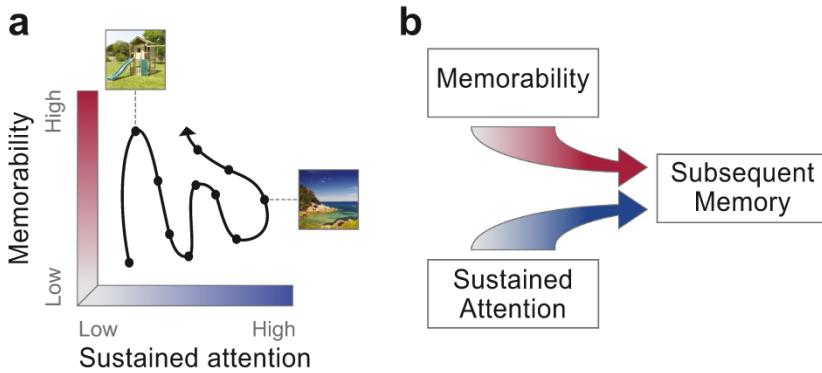


Fig. 2. Modeling subsequent memory with memorability and sustained attention. a. A depiction of memorability and sustained attention during the experiment. Memorability is plotted along the y-axis, ranging from low (gray) to high (red). Attentional state is plotted along the x-axis, ranging from low (gray) to high (blue). Each trial is indicated as a black dot, and successive trials are connected by the black line. Each trial could thus be characterized by these two dimensions, the memorability of the image or its attentional state. b. Schematic of our model in which subsequent memory is determined by two independent factors: memorability and sustained attention.

2.1. Stimuli

All experiments used 1100 color scene images from the Scene Understanding database (SUN; Xiao, Hays, Ehinger, Oliva, & Torralba, 2010). These images depicted a wide variety of representative real-world scenes, including 550 indoor scenes and 550 outdoors scenes. All images were cropped to be square and resized to 256 × 256 pixels.

2.2. Experiment 1

We measured the intrinsic memorability of 1100 color scene images in an online crowd-sourced experiment.

2.2.1. Participants

We recruited 808 adults from the United States via the online experimental platform Amazon Mechanical Turk (AMT). Participants (360 female; mean age = 39.3, range = 20–76 years) were compensated \$0.70 for participation for the approximately seven-minute experiment. Participants were required to have at least a 98% AMT approval rate, 500+ completed tasks, and an IP address in the United States. The experiment was preceded by three basic English screening questions and followed by a task instruction question to ensure quality of the data. 706 participants successfully completed the task, while the other 102 participants were excluded from the final sample because they did not complete the task, failed a periodic vigilance check, or answered at least one of the four screening questions incorrectly. Participants were recruited until we obtained 50 memory responses per stimulus, determined by prior studies as a sufficient number of participants to estimate a stable image memorability score (Isola et al., 2014). Participants provided consent in accordance with procedures approved by the University of Chicago Institutional Review Board (IRB).

2.2.2. Procedure

Participants completed a continuous recognition task (Fig. 1a; Experiment 1), in which 136 images appeared for 750 ms each with an 800 ms interstimulus interval (Bainbridge et al., 2013; Khosla, Raju, Torralba, & Oliva, 2015). Images were presented centrally against a white background. Participants were instructed to press the 'r' key when they encountered an image that was previously shown. 44 images were selected to be "target images" and repeated once during the task, at least 30 s (14 trials) apart. Non-target "filler images" were shown in between target images and a subset of them repeated in quick succession (1–5 images apart). These filler repeats were easily detectable by participants and served to maintain task vigilance. Participants who failed over 70% of the vigilance task repeats were excluded from the data analyses. Target and filler images were randomly sampled from the larger 1100 stimulus set, so that every image in the set served as a target image for at least 50 participants.

During the continuous recognition task, responses to each image were recorded. Correct identification of a repeated target image was

considered as a "hit" and failure to identify a repeated target image was considered as a "miss". Misidentification of the first presentation of a target as a repeat was classified as a "false alarm". We only analyzed responses to target images for each participant.

2.2.3. Image memorability measure

For each image, we calculated hit rate (the proportion of target repeats successfully identified) and false alarm rate (the proportion of first image presentations falsely identified) across individuals. Memorability scores for each image were calculated as the corrected recognition rate (CR) by subtracting the mean false alarm rate from the mean hit rate across participants.

To test whether image memorability was reliable across individuals, we conducted a consistency analysis in which CR was correlated between random split-halves of participants (Isola, Xiao, et al., 2011). A Spearman rank correlation was conducted across 1000 random participant split-halves, and then the correlations were averaged across iterations, resulting in an average across-participant consistency score. The across-participant consistency score was compared to a permuted chance level in which Spearman rank correlations were conducted between randomly shuffled split-halves. This analysis tests whether different groups of participants tend to consistently remember and forget the same images.

2.3. Experiment 2

The goal of this study was to ask whether image memorability and sustained attentional state predict subsequent memory. To this end, we next collected measures of attentional state as individuals performed a continuous performance task (CPT) using the same 1100 scene images characterized in Experiment 1. We also collected subsequent recognition memory performance for these images.

2.3.1. Participants

Thirty-seven participants aged 18–35 years (mean age = 24.5) were recruited via the University of Chicago Sona participant recruitment system and compensated \$10 for their participation. Three participants were excluded before data analysis because they were not right-handed or did not have normal or corrected-to-normal color vision. One additional participant was excluded due to CPT performance >3 SDs below the mean, resulting in a final sample size of 33 participants. This number of participants was based on the first sample in a previously published study on sustained attention and memory (deBettencourt et al., 2018). Participants provided written informed consent in accordance with procedures approved by the University of Chicago IRB.

2.3.2. Apparatus

Participants were seated approximately 48 cm from a CRT monitor with a 60 Hz refresh rate. Stimuli were presented using MATLAB (MathWorks, Natick, MA, USA) and the Psychophysics Toolbox v3.0.16

(Brainard, 1997; Pelli, 1997). Image stimuli subtended approximately 9.3° of visual angle on the screen. A centrally presented fixation dot subtended approximately 0.4°.

2.3.3. Procedure

Participants completed a CPT to assess sustained attention to the images from Experiment 1 (Fig. 1b; Experiments 2 & 3). During the 8:20-min task, each participant viewed 500 trial-unique scene images from the larger set of 1100. Each image was displayed for 1000 ms with no interstimulus interval. Images were presented in the center of a gray background. A black fixation dot was overlaid on the center of the image and turned white after each response.

Of the images shown, 90% (450 images) were from a frequent stimulus category (e.g., outdoor scenes) and 10% (50 images) were from an infrequent category (e.g., indoor scenes). Frequent and infrequent categories were counterbalanced across participants and images were randomly selected within each category. Participants were asked to categorize each image as an indoor scene or an outdoor scene. They were instructed to press “h” on the keyboard with their right index finger when an image belonged to the frequent category and “j” with their right middle finger when an image belonged to the infrequent category. Participants completed a short practice block (10 trials) prior to beginning the full task and repeated this practice block as necessary until they achieved at least 90% accuracy.

The CPT provided a measure of each individual’s attentional state on every trial, indexed by their RT. Previous work using this task demonstrated that fast RTs index worse attentional states and predict lapses, whereas slow RTs index better attentional states (deBettencourt et al., 2018).

Immediately after completing the CPT, participants performed a surprise recognition memory task (self-paced; approximately 20 min; Fig. 1b; Experiments 2 & 3) for the images. During the task, participants viewed 200 images: 100 from the frequent category and 100 from the infrequent category. Half of these images were from the CPT (“old” images) and half had not previously appeared (“new” images). In other words, memory was tested for 50/450 frequent category images and 50/50 infrequent category images seen in the CPT. Image assignment to these old vs. new conditions was random across participants and images were presented in a randomized order. Participants were instructed to indicate their memory and confidence that each image had appeared in the CPT on a scale of 1–4. A response of “1” indicated high confidence that the image had not appeared in the CPT and a response of “4” indicated high confidence that the image had appeared in the CPT. Each image remained on the screen until the participant responded, with no maximum presentation time. After each response, the image and the confidence rating were displayed for 500 ms before the next image appeared.

2.3.4. Sustained attention and subsequent memory measures

A participant’s sustained attentional state on each infrequent image category trial i was operationalized as pre-trial RT. To calculate pre-trial RT, we first detrended by calculating a linear fit over all RTs over the entire block and subtracting that from the raw RTs to control for time-dependent effects such as practice or fatigue. By subtracting a linear trend, this also effectively re-centers our RTs relative to each participant’s mean RT. Then, we calculated a trailing window of the RT, by averaging RT over the three preceding trials (trials $i-3$ to $i-1$). Analyses focused on infrequent category trials, which provide an interrogation of sustained attentional state because participants are required to switch from a habitual response to respond correctly.

To measure subsequent memory, correct high-confidence “old” responses (a response of “4”) were treated as remembered whereas all other responses were treated as forgotten (e.g., Kim, Lewis-Peacock, Norman, & Turk-Browne, 2014; Wagner et al., 1998). Only high-confidence old responses were treated as remembered because previous work has shown that high confidence hit rates are greater than high

confidence false alarm rates whereas the difference between low confidence hit rates and low confidence false alarm rates only approaches significance, suggesting that a significant proportion of low confidence responses are guesses (Turk-Browne, Yi, & Chun, 2006).

2.4. Experiment 3

Experiment 3 was a reanalysis of data previously reported in the second experiment of (deBettencourt et al., 2018). The goal of that study was to solely investigate the relationship between attention and memory.

2.4.1. Participants

Twenty-five undergraduates (15 female; mean age = 19.2 years) were recruited from Princeton University and compensated with college course credit. One participant was excluded from analysis due to sustained attention task performance >3 SDs below the mean. The final sample size, $n = 24$, included all participants who were included in previous analyses of these data. All participants reported normal or corrected-to-normal color vision and provided written informed consent in accordance with guidelines approved by the Princeton University IRB.

2.4.2. Apparatus

Participants were seated approximately 70 cm from a CRT monitor with a 100-Hz refresh rate. Stimuli were presented using MATLAB (MathWorks, Natick, MA, USA) and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Images subtended approximately 7° and the fixation dot subtended approximately 0.1° on the screen.

2.4.3. Procedure

Experiment 3 procedures were identical to those described in Experiment 2 save one key difference. In Experiment 3, infrequent images did not always appear randomly during the CPT. Rather, the order of trial types (frequent vs. infrequent) was determined in real-time based on participants’ RTs. Infrequent category images were inserted in the task if participants were responding one standard deviation faster or slower than their cumulative mean RT up to that point in the task. The intention was that this “triggering” would allow for exploration of sustained attention in extreme attentional states, when participants were very attentive or very inattentive. It was expected that there would be an even greater difference in subsequent memory for these triggered trials, as the response times preceding infrequent category images were more polarized than they were in Experiment 2. Up to 40 infrequent trials could be triggered (20 based on fast RTs, 20 based on slow RTs) during trials 51–450 of the CPT, and the other 10 trials were randomly inserted during trials 1–50 and 450–500. Fast and slow RT thresholds were computed using a “growing window” to compute cumulative mean RT after removing linear drift. There was a minimum of three frequent trials between infrequent trials to avoid contaminating the moving-window RT measure. The average number of infrequent trials per participant during the real-time period was 18.35 (SD = 5.64) for slow-triggered trials and 13.25 (SD = 2.08) for fast-triggered trials. Participants were not informed of this real-time triggering procedure.

2.4.4. Sustained attention and subsequent memory measures

Sustained attentional state (average pre-trial RT from the three preceding trials) and subsequent memory (high confidence hits) were measured in the same way as in Experiment 2.

2.5. Data analysis

2.5.1. Mixed-effects models

Generalized logistic regressions were used to predict subsequent memory (correct vs. incorrect) from two independent variables: (1) the CR memorability score of each image (from Experiment 1) and (2) pre-trial RT of each participant before they encountered each image during

the CPT (from Experiments 2 and 3; Figs. 1 & 2). In other words, the fixed effects of the model were image memorability and sustained attentional state. We compared the results of models that did and did not include an interaction term between image memorability and sustained attentional state. Intercepts for experiment and participants nested within experiment were included as random effects. Image memorability and sustained attention factors were *z*-scored within-participant to allow for comparison of beta values for each predictor across-and within-individuals. The same models without standardization show no difference in significance of each predictor but prevent the direct comparison of each predictor (Supplementary Material). The mixed-effects model formula was as follows:

$$\text{subsequent memory} \sim \text{image memorability} + \text{attentional state} + (1 | \text{experiment/participant})$$

subsequent memory corresponds to the binary memory performance metric during the surprise memory task in Experiments 2 and 3, *image memorability* corresponds to the CR for each image calculated from Experiment 1, and *attentional state* corresponds to pre-trial RT in the CPT in Experiments 2 and 3.

Mixed effects models were implemented using R's *lme4* package (Bates, Kliegl, Vasishth, & Baayen, 2015), and model optimization was performed with the limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm (Byrd, Lu, Nocedal, & Zhu, 1995) with the *optimx* package (Nash & Varadhan, 2011). Root-mean-square error (RMSE) was calculated as a measure of absolute error using the *sjstats* package in R (<https://CRAN.R-project.org/package=sjstats>).

2.5.2. Within-participant logistic regression models

Is long-term memory more dependent on intrinsic stimulus memorability for some participants, but attentional state at encoding for others? To ask this question, we ran within-participant analyses to determine the predictive power of image memorability and sustained attention for each individual. Within-participant analyses used generalized logistic regressions to predict the binary subsequent memory variable for each participant in Experiments 2 and 3. As with the mixed-effects model analyses, memory predictions were made using the memorability of each image and the preceding RT index of sustained attention at encoding for each infrequent category image. The interaction between sustained attention and memorability was not included as a predictor, as it was found to be non-significant in the across-participants mixed-effects model. The effects of experiment and participant nested within experiment were not included in these regressions because each model was fit to a single participant's data. Image memorability and sustained attention factors were *z*-scored within-participant. Thus, the within-participant model formula was as follows:

$$\text{subsequent memory} \sim \text{image memorability} + \text{attentional state}$$

The mathematical formula for this logistic regression model was:

$$P(Y_i = 1) = \frac{\exp(\beta_0 + \beta_M M_i + \beta_A A_i)}{1 + \exp(\beta_0 + \beta_M M_i + \beta_A A_i)}$$

M_i and A_i represent memorability and attentional state, respectively, for a trial i . We solve for β_M and β_A , the coefficients for memorability and attentional state, respectively, as well as the intercept β_0 . Within-participant logistic regression models were implemented using R's *glm* function, which is based on Fisher scoring (iteratively reweighted least squares).

2.5.3. Across-dataset predictions

To characterize the robustness and replicability of models predicting subsequent memory from image memorability and attentional state, we built models to predict memory using data from Experiment 2 and Experiment 3 separately. We then applied the model built in Experiment 2 to data from Experiment 3 to generate a prediction for each trial and

vice versa. We assessed models' predictive power by computing the within-participant point-biserial correlation between predicted memory accuracy and true memory accuracy (where true memory accuracy could be 1 [remembered] or 0 [forgotten]) for every individual. For each experiment, we submitted Fisher *z*-transformed within-participant correlation coefficients to a one-sample *t*-test to assess group-level significance.

We conducted permutation tests to further inspect the robustness of the results. In these permutation tests, we compared cross-experiment model performance to a null distribution of performance values. We estimated this distribution by randomly shuffling subsequent memory accuracy within-subject in the test set (i.e., Experiment 2 data for the model trained on Experiment 3 and tested on Experiment 2; Experiment 3 data for the model trained on Experiment 2 and tested on Experiment 3). We then correlated predicted memory scores with these shuffled observed memory scores within-subject, Fisher *z*-transformed the within-subject correlation coefficients, averaged these values across participants, and converted the average value back to Pearson *r*. We repeated this process 10,000 times to generate 10,000 null mean *r* values, separately for the model trained on Experiment 2 and tested on Experiment 3 and the model trained on Experiment 3 and tested on Experiment 2. We calculated a permutation (i.e., non-parametric *p*-value) as $p = (1 + \text{number null mean } r \text{ values} \geq \text{true mean } r \text{ value})/10,001$ (Supplementary Material).

2.5.4. Statistics

To compare the predictive power of image memorability and sustained attention for subsequent memory and effects of interest between groups, we calculated *t*-statistics, effect sizes, and Bayes factors where appropriate. All *t*-tests are two-tailed and assume equal variance between groups. Cohen's *d* statistics were calculated with the *effsize* package in R as a measure of effect size. Bayes factors (BF) were computed with the *BayesFactor* package in R and are reported as measures of relative evidence in favor of the alternative or null hypothesis. Bayes factors >3 indicate substantial evidence in favor of the alternative hypothesis, whereas Bayes factors $<1/3$ indicate substantial evidence in favor of the null (Wetzels et al., 2011). All data and code are available in a repository on the Open Science Framework (<https://osf.io/6uc48/>).

3. Results

3.1. Validating measures of image memorability and attentional state

3.1.1. Image memorability is reliable across individuals

We first asked whether there were specific scene images in our stimulus set that were more memorable or forgettable across individuals. To evaluate the consistency of corrected recognition (CR) performance for images in the continuous recognition task (Fig. 1a), we correlated the values for specific scene images obtained from different split-halves of the Experiment 1 Amazon Mechanical Turk sample ($n = 706$). Consistent with prior work, image memorability was highly reliable (mean Spearman's rank correlation across 1000 split halves: $\rho = 0.35$, 95% CI [0.32, 0.38], non-parametric $p < 0.001$; Fig. 3a). In other words, in our online sample, individuals tended to remember and forget the same images.

3.1.2. Pre-trial RT predicts attention lapses

We next asked whether RTs in the CPT reliably indexed attentional state in Experiment 2 ($n = 33$) and Experiment 3 ($n = 24$). To do so, we calculated a measure of pre-trial attentional state operationalized as the detrended RT before each infrequent trial, by first subtracting the linear trend across the entire block and then averaging the three preceding trials. We compared whether the pre-trial RT differed prior to attention lapses (incorrect responses to infrequent category trials) than non-lapses (correct responses to infrequent category trials). Consistent with prior work, participants' mean pre-trial RT was slower preceding correct vs.

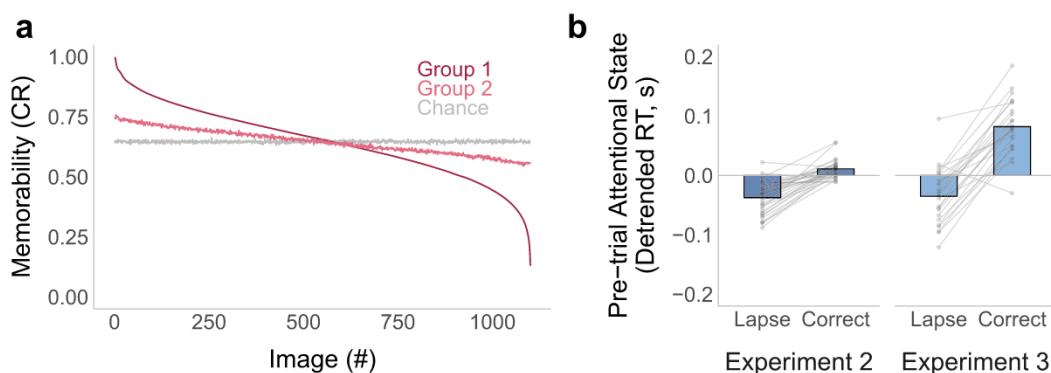


Fig. 3. Validating measures of memorability and attentional state. a. High consistency in the memorability of images. Random participant split-halves (Group 1 and Group 2) tend to remember and forget the same images (as measured by CR score), in comparison to a permuted chance level (gray). Lines in these plots show average memory performance (y-axis) across 1000 iterations, for each image ranked from highest to lowest CR score (x-axis). b. Pre-trial reaction time predicted lapses within both Experiment 2 (dark blue) and Experiment 3 (light blue). Faster response time is indicative of an individual being in a worse attentional state, while slower response time is indicative of a participant being in a better attentional state. Each gray dot represents an individual participant, with lines connecting data from the same participant. The height of the bar indicates the mean.

incorrect infrequent images in Experiment 2 (mean pre-trial RT difference = 0.048 s [0.037, 0.059], $t_{32} = 9.05$, $p < 0.001$, Cohen's $d = 2.19$) and Experiment 3 (mean difference = 0.117 s [0.094, 0.140], $t_{23} = 10.33$, $p < 0.001$, Cohen's $d = 2.49$). Thus, in these two independent datasets, pre-trial RT is a reliable measure of attentional state in that it predicts lapses (Fig. 3b).

3.2. Predicting subsequent memory from image memorability and attentional state

Each image presented in Experiments 2 and 3 can be operationalized via its memorability (obtained via online data collection from the separate pool of participants in Experiment 1) and sustained attentional state (obtained at each moment from each participant in Experiments 2 and 3). Can we use these image-specific measures of memorability and individual-specific measures of attentional state to predict what individuals will go on to remember (Fig. 2)?

3.2.1. Image memorability and attentional state uniquely predict subsequent memory

To understand the consequences of memorability and sustained attention for subsequent memory, we created separate mixed-effects models for each factor after collapsing across data from both Experiments 2 and 3. First, a model based on image memorability alone significantly predicted subsequent memory (Table 1). In addition, a model based on attentional state alone also significantly predicted subsequent memory (Table 1). That is, independent models using each factor reliably predict memory.

Although we see this relationship collapsing across the studies, we

were also interested in demonstrating replicability within each independent dataset. Therefore, we next examined the influence of image memorability and attentional state on subsequent memory by building separate models using data from Experiments 2 and 3 (Fig. 4). First, image memorability (M) remained a significant predictor of subsequent memory for both Experiment 2 ($\beta_M = 0.32$, $SE_M = 0.06$, $z_M = 5.75$, $p_M < 0.001$) and Experiment 3 ($\beta_M = 0.22$, $SE_M = 0.07$, $z_M = 2.95$, $p_M = 0.003$). Attentional state (A) also remained a significant predictor of subsequent memory for both Experiment 2 ($\beta_A = 0.12$, $SE_A = 0.05$, $z_A = 2.23$, $p_A = 0.026$) and Experiment 3 ($\beta_A = 0.18$, $SE_A = 0.07$, $z_A = 2.40$, $p_A = 0.016$). That is, the influence of memorability and attention on memory behavior was reliable across the two datasets.

3.2.2. Memorability and attentional state maintain unique contributions in a combined model of memory

Do image memorability and attentional state explain unique or overlapping variance in subsequent memory? To ask this question, we built a combined additive model including both image memorability and attentional state as predictors of subsequent memory. This model included data from both Experiment 2 and 3 to maximize power. Results revealed that both memorability and attentional state still predicted subsequent memory in this combined model (Table 1; Fig. 4). This pattern of results remained consistent when the interaction between image memorability and attentional state was included as a predictor in the model, with no significant effect of the interaction itself (Table 1). Results were also consistent in additive models built using data from Experiment 2 ($\beta_M = 0.33$, $SE_M = 0.06$, $z_M = 5.83$, $p_M < 0.001$; $\beta_A = 0.13$, $SE_A = 0.06$, $z_A = 2.43$, $p_A = 0.015$) and Experiment 3 ($\beta_M = 0.22$, $SE_M = 0.07$, $z_M = 2.93$, $p_M = 0.003$; $\beta_A = 0.18$, $SE_A = 0.07$, $z_A = 2.38$, $p_A = 0.016$).

Table 1

Results of mixed effects logistic regression models using image memorability, sustained attentional state, or both to predict subsequent image memory. Models included data from both Experiment 2 and 3 to maximize power. Intercepts for experiment and participants nested within experiment were included as random effects. Image memorability and sustained attention factors were z-scored within-participant to allow for comparison of beta values for each predictor across and within-individuals.

	Single factor model				Joint model				
	Memorability		Sustained attention		Without interaction term		With interaction term		
	Mem.	Attn.	Mem.	Attn.	Mem.	Attn.	Mem.	Attn.	Mem. * Attn. Interaction
β	0.29	–	–	0.14	0.29	0.15	0.29	0.15	0.063
SE	0.04	–	–	0.04	0.04	0.04	0.05	0.04	0.05
z	6.37	–	–	3.23	6.43	3.35	6.36	3.33	1.39
p -value	< 0.001	–	–	0.0013	< 0.001	< 0.001	< 0.001	< 0.001	0.17
RMSE	0.449		0.451		0.447		0.447		
Akaike Information Criterion (AIC)	3182.3		3213.4		3173.1		3173.2		

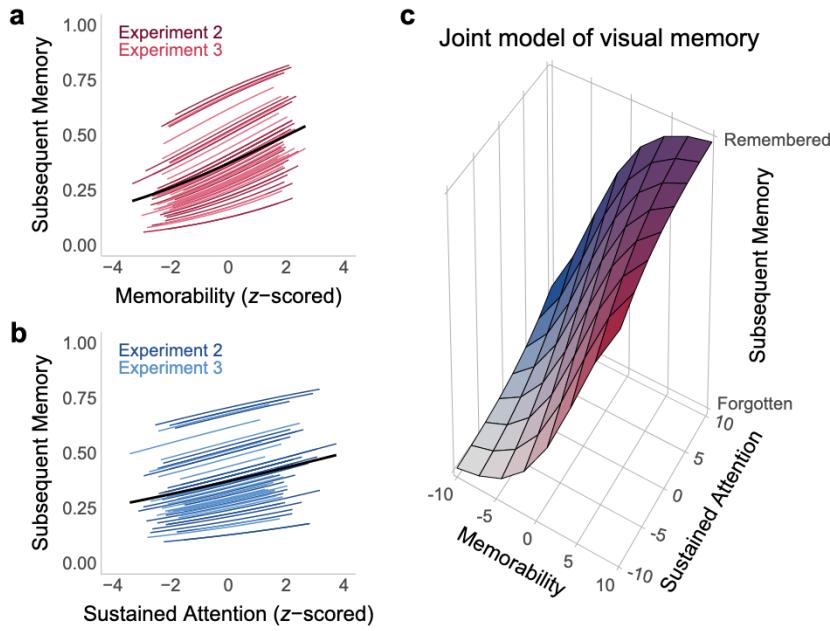


Fig. 4. Modeling visual memory. a. Modeling the influence of memorability on subsequent memory. We related image memorability (operationalized as corrected recognition score from Experiment 1, z-scored across images and within individuals) to subsequent recognition memory with mixed effects logistic regression. Regression estimates for individual participants are depicted for both Experiment 2 (dark red) and Experiment 3 (light red). The black line represents the average regression estimate from the model that combines across all participants from both experiments. b. Modeling the influence of attentional state on subsequent memory. We related attentional state (operationalized as pre-trial RT from the preceding three images, z-scored within individuals) to subsequent recognition memory with mixed effects logistic regression. Regression estimates for individual participants are depicted for both Experiment 2 (dark blue) and Experiment 3 (light blue). The black line represents the average regression estimate from the model that combines across all participants from both experiments. c. Joint model of memorability and attention for subsequent memory. We depict the multidimensional topography of subsequent recognition memory determined jointly from memorability and attentional state. Red colors indicate regions where image memorability supports memory, blue colors indicate regions where attention supports memory, and purple indicates regions where both factors support memory. The slope of the curve corresponds to the average regression estimates, combining across all participants from both experiments.

0.017) separately. They were additionally consistent in interactive models built using data from Experiment 2 ($\beta_M = 0.33$, $SE_M = 0.06$, $z_M = 5.80$, $p_M < 0.001$; $\beta_A = 0.13$, $SE_A = 0.06$, $z_A = 2.41$, $p_A = 0.016$; $\beta_{\text{interact}} = 0.04$, $SE_{\text{interact}} = 0.06$, $z_{\text{interact}} = 0.76$, $p_{\text{interact}} = 0.45$) and Experiment 3 ($\beta_M = 0.21$, $SE_M = 0.07$, $z_M = 2.81$, $p_M = 0.005$; $\beta_A = 0.18$, $SE_A = 0.07$, $z_A = 2.38$, $p_A = 0.017$; $\beta_{\text{interact}} = 0.10$, $SE_{\text{interact}} = 0.07$, $z_{\text{interact}} = 1.35$, $p_{\text{interact}} = 0.18$) separately.

We examined model fits using AIC, which assesses model fit while penalizing model complexity. Based on AIC, the best performing model in the sample collapsed across experiments was the additive model combining image memorability and sustained attentional state, followed by the interactive model including image memorability, sustained attentional state, and their interaction. Because overall model fit was numerically worse for the interactive than the additive joint model (AIC: 3173.2 vs. 3173.1), we focus on the additive model in the remainder of the text. The next-best performing model was the model utilizing memorability alone and then the model utilizing attention alone (Table 1). A likelihood ratio test comparing the joint and single-factor models demonstrated that the additive model significantly outperformed the memorability-only model ($\chi^2 = 11.236$, $p < 0.001$) and the attention-only model ($\chi^2 = 42.294$, $p < 0.001$). That is, these data reveal that while behavior was influenced by both memorability and attentional state, memorability appeared to exert a stronger influence in this task.

3.2.3. Low trial-by-trial influences of memorability on current attentional state

Although memorability and attentional state predicted unique variance in subsequent memory performance, it is possible that these two factors could also interact. To examine these potential interactions, we first tested whether image memorability on a given trial would immediately impact attentional state on that trial. We correlated the memorability score for each image with the linearly-detrended RT to that

image and repeated this analysis separately for each participant. To eliminate trials contaminated by errors and response switching, we restricted this analysis to correct trials from the frequent category in the CPT, although results fully replicated when examining all trials (Supplementary Material). We observed numerically small Spearman rank correlations between the memorability of an image and the measured attentional state to that image (Exp 2: mean within-participant $\rho = 0.031$, range: $[-0.16, 0.11]$; Exp 3: mean within-participant $\rho = 0.0079$, range: $[-0.090, 0.092]$). Group-level *t*-tests comparing Fisher *z*-transformed Spearman rho values to zero revealed a significant relationship between RT and memorability in Experiment 2 ($t_{32} = 3.06$, $p = 0.004$, Cohen's $d = 0.53$, $BF = 8.68$) but not in Experiment 3 ($t_{23} = 0.75$, $p = 0.46$, Cohen's $d = 0.15$, $BF = 0.28$). In other words, participants on average responded more slowly (i.e., were more attentive) to more memorable images in Experiment 2 but not in Experiment 3. Future work can explore this intriguing (albeit inconsistent) result and investigate whether more memorable images may subtly modulate sustained attention fluctuations. In either case, the low correlation between trial-by-trial memorability and attention, in combination with the significance of both memorability and attention as predictors in combined mixed effects models, demonstrates that memorability and attention are not redundant predictors of subsequent memory.

3.3. Predicting memory at the level of the individual

Although a model combining all participants showed significant contributions of both image memorability and sustained attentional state to subsequent memory, we also wanted to examine these factors at the level of individual participants. Within-participant logistic regression models revealed that memory in Experiment 2 was significantly predicted by both image memorability (mean within-participant $\beta = 0.38$ [0.24, 0.52], $t_{32} = 5.39$, $p < 0.001$, Cohen's $d = 0.94$) and attention (mean within-participant $\beta = 0.12$ [0.016, 0.22], $t_{32} = 2.36$, $p = 0.025$,

Cohen's $d = 0.41$). When conducting logistic regression models within participants for Experiment 3, we also observed a reliable relationship for both memorability ($\beta = 0.25 [0.053, 0.44]$, $t_{23} = 2.64$, $p = 0.015$, Cohen's $d = 0.54$) and attention ($\beta = 0.26 [0.063, 0.47]$, $t_{23} = 2.72$, $p = 0.012$, Cohen's $d = 0.55$). That is, both memorability and attention remained reliable at the individual participant level in both studies.

3.3.1. Examining the relative contributions of memorability and attentional state

While both memorability and attention reliably predicted memory within participants, we were also interested in quantifying the relative influence of each factor on memory. We observed that the coefficient for memorability was significantly stronger than the coefficient for sustained attention in Experiment 2 (mean difference in $\beta = 0.26 [0.010, 0.43]$, $t_{32} = 3.28$, $p = 0.0025$, Cohen's $d = 0.75$, BF = 14.39; Fig. 5a). However, we did not observe this difference in Experiment 3 (mean difference in $\beta = -0.019 [-0.28, 0.24]$, $t_{23} = -0.15$, $p = 0.88$, Cohen's $d = -0.04$, BF = 0.22). In fact, this may reveal an interesting difference between the two studies. The trial structure of Experiment 3 was specifically designed to boost the influence of attention on memory, by presenting certain trials during particularly high or low attentional states (see *Methods* for detail). This suggests the exciting possibility that through careful experiment design we may be able to manipulate the relative influences of attention or memorability.

We were further interested in examining memorability and attention interactions from person to person. One possibility is that the same individuals tend to demonstrate a strong influence of both memorability and attention. If so, then the coefficients for memorability and attention will be correlated across participants. Conversely, individuals whose memories are driven by attentional fluctuations may be less influenced by the memorability of the images (or vice versa), predicting a negative correlation between these factors across participants. Finally, memorability and attention may represent independent dimensions for later memory, and then we would not necessarily predict a relationship across participants. To test these hypotheses, we correlated the coefficients for memorability and attentional state across participants. Interestingly, neither experiment revealed a reliable relationship between the coefficients (Exp 2: $r = 0.078$, $p = 0.67$; Exp 3: $r = 0.22$, $p = 0.29$; Fig. 5b). This provides further evidence that memorability and attention have dissociable influences on memory.

3.3.2. These factors remain consistent regardless of individual differences in memory

Memorability and attentional state predicted memory across the population. Was this effect driven by individuals with overall better memory? That is, do image memorability and/or attentional state better predict memory for individuals who better remember task stimuli overall? To ask this question, we separately examined individuals with better and worse recognition memory performance (defined via median split within experiment). In data collapsed across Experiments 2 and 3,

mixed-effects models including both memorability and attention predicted subsequent memory in both the better-memory group ($\beta_M = 0.31$, $SE_M = 0.06$, $z_M = 5.18$, $p_M < 0.001$; $\beta_A = 0.18$, $SE_A = 0.06$, $z_A = 2.60$, $p_A = 0.0094$) and the worse-memory group ($\beta_M = 0.26$, $SE_M = 0.07$, $z_M = 3.85$, $p_M < 0.001$; $\beta_A = 0.14$, $SE_A = 0.07$, $z_A = 2.14$, $p_A = 0.032$). Within-participant coefficients did not differ between groups (Exp 2: memorability: $t_{31} = -0.03$, $p = 0.97$, BF = 0.33; attention: $t_{31} = 0.73$, $p = 0.47$, BF = 0.41; Exp 3: memorability: $t_{22} = 1.26$, $p = 0.22$, BF = 0.66; attention: $t_{22} = -0.33$, $p = 0.74$, BF = 0.39), with Bayes factors providing anecdotal evidence in favor of the null hypothesis. This suggests that individuals did not likely achieve better memory performance by relying more heavily on image memorability and/or attentional state.

3.4. Across-sample prediction with this model

Up to this point, we have discussed how both memorability and attentional state are reliable, robust, and independent predictors of subsequent memory. A model of human memory relying on these factors would be most useful if it could generalize across datasets and be used to make predictions for unseen, out-of-sample data. To assess the generalizability of models, we developed a model of subsequent memory from one dataset (e.g., Experiment 2) and applied it to the other dataset (e.g., Experiment 3) and vice versa. If these models of subsequent memory successfully generalize, then we would expect successful prediction between these datasets. Alternatively, if models do not successfully generalize, this suggests that the models are relying upon information specific to each experiment (e.g., differences in individuals or tasks). We trained a logistic regression model to predict subsequent memory on data from Experiment 2 and applied it to data from Experiment 3. Then, we correlated the model-predicted and true memory performance for each participant and examined whether these correlations were significant at the group level. Indeed, a model derived from data from Experiment 2 successfully predicted memory in Experiment 3 (mean within-participant $r = 0.13 [0.05, 0.20]$, $t_{23} = 3.44$, parametric $p = 0.0022$, permutation $p = 0.0005$, Cohen's $d = 0.70$, BF = 17.50). Furthermore, we replicated these findings by training on data from Experiment 3 and applying it to data from Experiment 2 (mean within-participant $r = 0.15 [0.09, 0.20]$, $t_{32} = 5.36$, parametric $p < 0.001$, permutation $p < 0.0001$, Cohen's $d = 0.93$, BF = 2922.41). Thus, models trained on memorability and attentional state generalize across independent datasets—collected at different institutions with different participant samples and different experimental procedures—to predict memory.

4. Discussion

What affects our long-term memories? We posited that certain factors are highly specific to the information to be encoded, but shared across individuals, whereas other factors may be highly specific to individuals, regardless of what information is being encoded. To

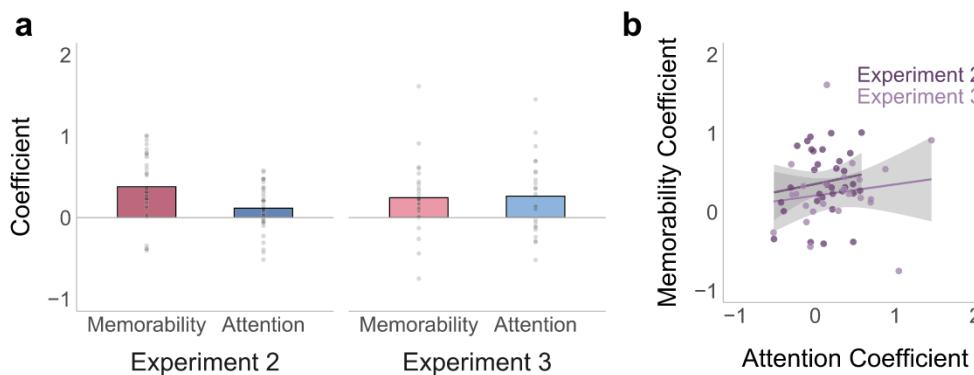


Fig. 5. Examination of memorability and attention within participants. a. Coefficients for memorability (blue) and attention (red) both independently reliably predict subsequent memory in Experiments 2 and 3. Each gray dot represents an individual participant's beta coefficients on a given factor resulting from the within-participant logistic regression. The height of the bar indicates the population mean. b. Individual differences in memorability and attention. Correlation between the degree to which sustained attention and memorability predict subsequent memory, separately for Experiment 2 (dark purple) and 3 (light purple). Gray regions indicate the 95% confidence interval.

characterize both image- and individual-specific factors that impact what we remember, we built a model that predicted visual long-term memory from an image's memorability and an individual's moment-to-moment attentional state. We computed memorability scores for a large set of real-world images using a large crowd-sourced online experiment. We then combined these image memorability scores with a behavioral index of sustained attention measured in a continuous performance task. Results revealed that image memorability and sustained attention uniquely predicted subsequent memory in two independent datasets. That is, images that are more memorable and/or that appeared during a better attentional state were more likely to be remembered. Furthermore, models including both memorability and attention better predicted recognition memory than models including either factor alone. Thus, memory is best considered a function of what information is presented when to whom, and models that only consider one or the other factor (or neither) will fail to fully explain what we remember.

4.1. Dissociable influences of memorability and attentional state

In building a model of memory, we determined that memorability and attentional state each independently predict what information is later remembered. Interestingly, we did not find clear evidence of a trade-off of the influence of memorability and attention on subsequent memory. That is, it was not the case that people for whom memory was more heavily influenced by attention state were less heavily influenced by image memorability or vice versa. Furthermore, the weight of each factor did not differ between participants with overall better or worse memory; in other words, it was not the case that individuals who used more attention-based or memorability-based strategies showed higher performance. Finally, we observed no consistent relationship between the two factors within individuals: People responded more slowly to more memorable images in one dataset but not the other. This result partially aligns with prior work reporting that memorability effects are not influenced by bottom-up or top-down attention (Bainbridge, 2020). Importantly, however, our experiments were not designed to specifically investigate individual differences in memory or the factors predicting it. Future work can more powerfully explore how individuals rely on different external and internal factors, and how these factors predict memory for different stimuli in a variety of tasks. Overall, these results highlight the critical importance of considering both attention and memorability when modeling memory.

4.2. Predicting memory in multiple datasets

A strength of this work is that it leverages multiple datasets collected both online and in person at multiple institutions to ensure that findings are replicable and generalizable. Even though memorability scores were recorded from an online sample (Experiment 1), they predicted memory performance in two entirely different participant samples (Experiments 2 and 3). For Experiments 2 and 3, although the task paradigms were largely similar, trials in Experiment 3 were presented in non-random order to specifically target extremely attentive or inattentive states. Despite these differences in paradigm, participants, and institution, results were largely similar, and—demonstrating robustness and replicability—models defined in one experimental sample generalized to predict memory in the other. Additionally, because cross-dataset comparison is an extremely conservative test of model generalizability, capturing even a modest statistically significant proportion of the variance is theoretically impactful.

4.3. Influence of memorability vs. attentional state

Although we observed robust evidence of both attention and memorability effects on memory, there were hints that, at least within these datasets, memorability exerted a stronger effect. In the mixed-effects model, the model with a memorability factor better fit the data

than the model with an attention factor. We did not find that one factor significantly outperformed the other across both datasets, but the influence of memorability was numerically greater in both samples and reliably so in Experiment 2. We found similar trends in Experiment 3, but these results were less reliable likely due to task manipulations which targeted more extreme attentional states and thus may have boosted the relative contributions of sustained attention. This suggests that stimulus features have an impressively strong influence over our later memories, equal to or even stronger than our level of engagement at any given moment.

4.4. Other measures of sustained attention and memorability

In this experiment, we exclusively operationalized sustained attention via RT. While this measure of attentional state is robust and strongly predictive of lapses, research has also characterized sustained attentional states via other measures of behavioral task performance, including response time variability (Fortenbaugh et al., 2015; Rosenberg et al., 2013) and task accuracy (Decker, Finn, & Duncan, 2020). Sustained attentional states can be further characterized using a variety of physiological measures including pupil size (Keene, deBettencourt, Awh, & Vogel, 2021), EEG patterns (deBettencourt, Williams, Vogel, & Awh, 2021), and fMRI connectivity patterns (Song & Rosenberg, 2021). Similarly, while the current study operationalizes memorability as CR, prior work has shown high consistency across individuals in their hit rates and false alarm rates (Bainbridge et al., 2013), as well as eye-movement patterns (Bylinskii, Isola, Bainbridge, Torralba, & Oliva, 2015). Memorability can also be predicted by deep learning neural networks (Khosla et al., 2015; Needell & Bainbridge, 2022) and results in specific fMRI patterns in visual and mnemonic areas of the brain (Bainbridge, Dilks, & Oliva, 2017; Bainbridge & Rissman, 2018). Future work could explore how these different behavioral and physiological signatures of sustained attention and memorability interact with memory.

4.5. Other external and internal factors

Although the current work characterizes the contribution of two important and understudied factors to memory, the RMSE of the logistic regression model including image memorability and sustained attentional state is 0.45, suggesting that further improvements in memory predictions are possible. Other research has highlighted the role of other external (stimulus-specific) and internal (person-specific) factors that influence long-term memory, including image context (e.g., Bodrogi & Tarczali, 2001), emotional state (e.g., Murray, Holland, & Kensinger, 2013), and prior experiences and familiarity (e.g., Yonelinas, Kroll, Dobbins, & Soltani, 1999). The current work thus introduces a useful framework for characterizing the contribution of multiple predictors of memory, which can be applied to understand the variance in memory explained by each. Future studies could expand the model introduced here by integrating these factors to even more comprehensively model and characterize the suite of factors that act on memory with unique (or overlapping) contributions.

4.6. To what degree is memory predetermined?

One key similarity between memorability and attentional state is that they can both be quantified prior to the presentation of the stimulus. By the time you encounter a striking painting in a museum or an uninspiring piece of wall art in a doctor's office, the properties of the image are already set, and you are already in a given attentional state. Therefore, our model highlights that some variance in our memory can be attributed to pre-established influences, rather than processes that only occur during and following that image presentation. Further, these successful predictions can be made in a relatively context-free manner—we do not need to know an individual's prior experiences (even

within the same experiment!) to predict their performance on a given trial with some degree of accuracy.

That our memories may be partly predetermined suggests tantalizing possibilities for influencing and manipulating what we remember. For example, experimenters could select highly memorable and forgettable images to induce certain memories (Needell & Bainbridge, 2022), or monitor someone's attentional fluctuations over time to identify the best time to present to-be-remembered material (deBettencourt et al., 2018, 2019). Utilizing both approaches, one could make a particularly unforgettable experience, presenting memorable items when attention is high. Alternatively, in educational scenarios, it may be preferable to trade off these two factors, presenting forgettable items when attention is high and vice versa to maximize learning outcomes. Finally, there is also emerging evidence that both attentional state (deBettencourt et al., 2015; Rosenberg et al., 2016) and memorability can be manipulated (Goetschalckx, Andonian, Oliva, & Isola, 2019; Khosla, Bainbridge, Torralba, & Oliva, 2013). We could use these approaches to modify either attentional state or item memorability to create videos, visual art, and educational content that is inherently memorable, and intentionally shown at attentive moments to an individual.

In sum, visual long-term memory research has largely overlooked the importance of considering *what* we are remembering and the *state* we are in when we encounter it. However, combining the memorability of images, with the attentional state of individuals, we can now better predict what one will remember.

Data and code availability

The de-identified data for all experiments and materials for all data analyses are posted at <https://osf.io/6uc48/>. Experiments were not formally preregistered.

CRediT authorship contribution statement

Cheyenne D. Wakeland-Hart: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. **Steven A. Cao:** Investigation, Data curation, Writing – review & editing. **Megan T. deBettencourt:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization, Supervision. **Wilma A. Bainbridge:** Conceptualization, Methodology, Software, Validation, Formal analysis, Data curation, Writing – original draft, Visualization, Supervision, Funding acquisition. **Monica D. Rosenberg:** Conceptualization, Methodology, Software, Validation, Formal analysis, Data curation, Writing – original draft, Visualization, Supervision, Funding acquisition.

Declaration of Competing Interest

None.

Acknowledgements

This research was supported by the National Science Foundation BCS-2043740 (M.D.R.), National Institute of Health grant F32MH115597 (M.T.d.B.), and The University of Chicago Summer Institute in Social Research Methods and Micro-Metcalf Programs (C.D. W.-H.) and College Curriculum Innovation Fund (M.D.R.).

Appendix A. Supplementary material

Supplementary material to this article can be found online at <https://doi.org/10.1016/j.cognition.2022.105201>.

References

- Bainbridge, W. A. (2019). Memorability: How what we see influences what we remember. *Psychology of Learning and Motivation - Advances in Research and Theory*, 70, 1–27. <https://doi.org/10.1016/bs.plm.2019.02.001>
- Bainbridge, W. A. (2020). The resiliency of image memorability: A predictor of memory separate from attention and priming. *Neuropsychologia*, 141. <https://doi.org/10.1017/CBO9781107415324.004>
- Bainbridge, W. A., Dilks, D. D., & Oliva, A. (2017). Memorability: A stimulus-driven perceptual neural signature distinctive from memory. *NeuroImage*, 149. <https://doi.org/10.1016/j.neuroimage.2017.01.063>
- Bainbridge, W. A., Isola, P., & Oliva, A. (2013). The intrinsic memorability of face photographs. *Journal of Experimental Psychology: General*, 142(4). <https://doi.org/10.1037/a0033872>
- Bainbridge, W. A., & Rissman, J. (2018). Dissociating neural markers of stimulus memorability and subjective recognition during episodic retrieval. *Scientific Reports*, 8(1), 1–11. <https://doi.org/10.1038/s41598-018-26467-5>
- Barel, E., & Tzischinsky, O. (2020). The relation between sustained attention and incidental and intentional object-location memory. *Brain Sciences*, 10(3). <https://doi.org/10.3390/brainsci10030145>
- Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *ArXiv*, 1506, 04967. <https://doi.org/10.48550/arXiv.1506.04967>
- Bodrogi, P., & Tarczali, T. (2001). Colour memory for various sky, skin, and plant colours: Effect of the image context. *Color Research and Application*, 26(4), 278–289. <https://doi.org/10.1002/col.1034>
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4), 433–436. <https://doi.org/10.1163/156856897X00357>
- Broers, N., Potter, M. C., & Nieuwenstein, M. R. (2018). Enhanced recognition of memorable pictures in ultra-fast RSVP. *Psychonomic Bulletin and Review*, 25(3), 1080–1086. <https://doi.org/10.3758/s13423-017-1295-7>
- Bylinskii, Z., Isola, P., Bainbridge, C., Torralba, A., & Oliva, A. (2015). Intrinsic and extrinsic effects on image memorability. *Vision Research*, 116, 165–178. <https://doi.org/10.1016/j.visres.2015.03.005>
- Byrd, R., Lu, P., Nocedal, J., & Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *Journal of Scientific Computing*, 16(5), 1190–1208. <https://doi.org/10.1137/0916069>
- deBettencourt, M. T., Cohen, J. D., Lee, R. F., Norman, K. A., & Turk-Browne, N. B. (2015). Closed-loop training of attention with real-time brain imaging. *Nature Neuroscience*, 18(3), 470–478. <https://doi.org/10.1038/nn.3940>
- deBettencourt, M. T., Keene, P. A., Awh, E., & Vogel, E. K. (2019). Real-time triggering reveals concurrent lapses of attention and working memory. *Nature Human Behaviour*, 3, 808–816. <https://doi.org/10.1038/s41562-019-0606-6>
- deBettencourt, M. T., Norman, K. A., & Turk-Browne, N. B. (2018). Forgetting from lapses of sustained attention. *Psychonomic Bulletin and Review*, 25(2), 605–611. <https://doi.org/10.3758/s13423-017-1309-5>
- deBettencourt, M. T., Williams, S. D., Vogel, E. K., & Awh, E. (2021). Sustained attention and spatial attention distinctly influence long-term memory encoding. *Journal of Cognitive Neuroscience*, 33(10), 2132–2148. https://doi.org/10.1162/jocn_a_01748
- Decker, A., Finn, A., & Duncan, K. (2020). Errors lead to transient impairments in memory formation. *Cognition*, 204, Article 104338. <https://doi.org/10.1016/j.cognition.2020.104338>
- Fortenbaugh, F. C., Degutis, J., Germine, L., Wilmer, J. B., Grosso, M., Russo, K., & Esterman, M. (2015). Sustained attention across the life span in a sample of 10,000: Dissociating ability and strategy. *Psychological Science*, 26(9), 1497–1510. <https://doi.org/10.1177/0956797615594896>
- Goetschalckx, L., Andonian, A., Oliva, A., & Isola, P. (2019). GANalyze: Toward visual definitions of cognitive image properties. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 5744–5753). <https://doi.org/10.1167/jcvi.2011.297>
- Goetschalckx, L., Moors, P., & Wagenaars, J. (2018). Image memorability across longer time intervals. *Memory*, 26(5), 581–588. <https://doi.org/10.1080/09658211.2017.1383435>
- Isola, P., Parikh, D., Torralba, A., & Oliva, A. (2011). Understanding the intrinsic memorability of images. In *Advances in neural Information Processing Systems 24: 25th Annual Conference on Neural Information Processing Systems 2011, NIPS 2011* (pp. 1–9). <https://doi.org/10.1167/12.9.1082>
- Isola, P., Xiao, J., Parikh, D., Torralba, A., & Oliva, A. (2014). What makes a photograph memorable? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(7), 1469–1482. <https://doi.org/10.1109/TPAMI.2013.200>
- Isola, P., Xiao, J., Torralba, A., & Oliva, A. (2011). What makes an image memorable? *IEEE Conference on Computer Vision and Pattern Recognition*, 145–152. <https://doi.org/10.1109/CVPR.2011.5995721>
- Keene, P. A., deBettencourt, M. T., Awh, E., & Vogel, E. K. (2021). Pupilometry signatures of sustained attention and working memory. *BioRxiv*. <https://doi.org/10.1101/2021.01.18.426969>
- Khosla, A., Bainbridge, W. A., Torralba, A., & Oliva, A. (2013). Modifying the memorability of face photographs. *Proceedings of the IEEE International Conference on Computer Vision*. <https://doi.org/10.1109/ICCV.2013.397>
- Khosla, A., Raju, A. S., Torralba, A., & Oliva, A. (2015). Understanding and predicting image memorability at a large scale. In *Proceedings of the IEEE International Conference on Computer Vision, 2015 International Conference on Computer Vision, ICCV 2015* (pp. 2390–2398). <https://doi.org/10.1109/ICCV.2015.275>
- Kim, G., Lewis-Peacock, J. A., Norman, K. A., & Turk-Browne, N. B. (2014). Pruning of memories by context-based prediction error. *Proceedings of the National Academy of Sciences of the United States of America*, 111(24), 8997–9002. <https://doi.org/10.1073/pnas.1319438111>

- Madore, K. P., Khazenzon, A. M., Backes, C. W., Jiang, J., Uncapher, M. R., Norcia, A. M., & Wagner, A. D. (2020). Memory failure predicted by attention lapsing and media multitasking. *Nature*, 587(7832), 87–91. <https://doi.org/10.1038/s41586-020-2870-z>
- Murray, B. D., Holland, A. C., & Kensinger, E. A. (2013). Episodic memory and emotion. In M. D. Robinson, E. Watkins, & E. Harmon-Jones (Eds.), *Handbook of cognition and emotion* (pp. 156–175). The Guilford Press.
- Nash, J. C., & Varadhan, R. (2011). Unifying optimization: Algorithms to aid software system users: Optimx for R. *Journal of Statistical Software*, 43(9), 1–14. <https://doi.org/10.18637/jss.v043.i09>
- Needell, C. D., & Bainbridge, W. A. (2022). Embracing new techniques in deep learning for estimating image memorability. *Computational Brain & Behavior*. <https://doi.org/10.1007/s42113-022-00126-5>
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. In *Spatial Vision*, 10(4), 437–442. <https://doi.org/10.1163/156856897X00366>
- Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. B., & Jenny, Y. (1997). “Oops!”: Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia*, 35(6), 747–758. [https://doi.org/10.1016/s0028-3932\(97\)00015-8](https://doi.org/10.1016/s0028-3932(97)00015-8)
- Rosenberg, M. D., Noonan, S., DeGutis, J., & Esterman, M. (2013). Sustaining visual attention in the face of distraction: A novel gradual-onset continuous performance task. *Attention, Perception, & Psychophysics*, 75(3), 426–439. <https://doi.org/10.3758/s13414-012-0413-x>
- Rosenberg, M. D., Zhang, S., Hsu, W. T., Scheinost, D., Finn, E. S., Shen, X., ... Chun, M. M. (2016). Methylphenidate modulates functional network connectivity to enhance attention. *Journal of Neuroscience*, 36(37), 9547–9557. <https://doi.org/10.1523/JNEUROSCI.1746-16.2016>
- Song, H., Finn, E. S., & Rosenberg, M. D. (2021). Neural signatures of attentional engagement during narratives and its consequences for event memory. *Proceedings of the National Academy of Sciences of the United States of America*, 118(33), Article e2021905118. <https://doi.org/10.1073/pnas.2021905118>
- Song, H., & Rosenberg, M. D. (2021). Predicting attention across time and contexts with functional brain connectivity. *Current Opinion in Behavioral Sciences*, 40, 33–44. <https://doi.org/10.1016/j.cobeha.2020.12.007>
- Turk-Browne, N. B., Yi, D. J., & Chun, M. M. (2006). Linking implicit and explicit memory: Common encoding factors and shared representations. *Neuron*, 49(6), 917–927. <https://doi.org/10.1016/j.neuron.2006.01.030>
- Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., ... Buckner, R. L. (1998). Building memories: Remembering and forgetting of verbal experiences as predicted by brain activity. *Science*, 281(5380), 1188–1191. <https://doi.org/10.1126/science.281.5380.1188>
- Wetzel, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E. J. (2011). Statistical evidence in experimental psychology: An empirical comparison using 855 t tests. *Perspectives on Psychological Science*, 6(3), 291–298. <https://doi.org/10.1177/1745691611406923>
- Xiao, J., Hays, J., Ehinger, K. A., Oliva, A., & Torralba, A. (2010). SUN database: Large-scale scene recognition from abbey to zoo. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 3485–3492). <https://doi.org/10.1109/CVPR.2010.5539970>
- Yonelinas, A. P., Kroll, N. E. A., Dobbins, I. G., & Soltani, M. (1999). Recognition memory for faces: When familiarity supports associative recognition judgments. *Psychonomic Bulletin and Review*, 6(4), 654–661. <https://doi.org/10.3758/BF03212975>