AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General

ISSN: 0096-3445

2024, Vol. 153, No. 6, 1415–1424
https://doi.org/10.1037/xge0001594

Highly Memorable Images Are More Readily Perceived

Will Deng[1], Kara D. Federmeier[1, 2, 3], and Diane M. Beck[1, 2, 3]
[1] Department of Psychology, University of Illinois Urbana-Champaign
[2] Program in Neuroscience, University of Illinois Urbana-Champaign
[3] Beckman Institute for Advanced Science and Technology, University of Illinois Urbana-Champaign

Image memorability, the likelihood that a person will remember a particular image, has been shown to be an intrinsic property of the image that is distinct from many other visual and cognitive features. Research thus far has not identified particular visual features that can sufficiently explain this intrinsic memorability, but one possibility is that more and less memorable images differ in their statistical regularity (i.e., how prototypical or distinctive they are). Statistical regularity is known to affect detection time for images, such that stimuli with higher statistical regularity can be detected with shorter presentation directions. Therefore, in the present study, we probed whether memorability affects how quickly an image can be detected. High- and low-memorability images were presented in an intact/scrambled task wherein participants were asked to indicate whether they saw an intact image or noise, and we estimated the presentation duration necessary for participants to reach 70.7% accuracy. Across two experiments using different stimulus materials, we observed and then replicated that more memorable images are associated with shorter detection thresholds than those for less memorable images. The results support the idea that memorable stimuli may better match stored templates used for image perception and/or recognition.

Public Significance Statement

The likelihood that a given image will be remembered has been found to be consistent across people for a variety of image types such as natural scenes, faces, and logos. The present study tested whether image memorability affects the time it takes for people to simply detect the presence of an intact image as opposed to noise. The results show that memorability is associated with enhanced detection, suggesting that more memorable images may also be more readily perceived, consistent with them better matching our stored knowledge of visual scenes and objects.

Keywords: image memorability, detection, statistical regularity, templates

When you walk through a city street, you may be bombarded with all types of advertisements, with some being more memorable than others. Perhaps a poster on a billboard is so memorable that, after seeing it once, you immediately recognize it when you see it a second time. Interestingly, many images that are more likely to be remembered by you are also likely to be remembered by others—a property known as image memorability, which has been more formally defined by Bainbridge (2019, p. 3) as "an intrinsic, measurable property of a stimulus, measuring the likelihood for any given person to remember that image later." The likelihood of remembering an image has commonly been estimated using a continuous recognition task wherein participants are presented with a sequence of new and old images with delayed repetition and asked to identify the repeats (Bainbridge et al., 2013; Bylinskii et al., 2015; Isola et al., 2014; Khosla et al., 2015). Each image's memorability can then be derived as the hit rate for its repeats.

Across different types of stimuli, research has found a surprisingly high degree of cross-person consistency in memorability. For example, in one study that investigated the memorability of natural scenes, researchers randomly split 665 participants into two halves and

This article was published Online First May 2, 2024.
Timothy Vickery served as action editor.
Will Deng https://orcid.org/0009-0009-1120-9487
The authors wish to thank Paige Elaine Hamler for her assistance in data collection. The authors would also like to thank Pei-Ling Yang, Emily Cunningham, and Jacob Ryder for offering suggestions for programing the experiments. The authors have no known conflicts of interest to disclose. None of the ideas or data appearing in this manuscript were previously disseminated at conferences or on websites. The data, analysis code, and materials of this study are available on Open Science Framework and can be accessed at https://osf.io/548gx/?view_only=741a953ee6e44df7a9a64980ad39679e for Experiment 1 and https://osf.io/gjhfq/?view_only=440102d3f924453e9e521f91f086b37a for Experiment 2.

Will Deng served as lead for data curation, formal analysis, investigation, project administration, software, visualization, and writing–original draft and served in a supporting role for conceptualization. Kara D. Federmeier served in a supporting role for conceptualization. Diane M. Beck served as lead for conceptualization and resources. Will Deng and Diane M. Beck contributed equally to methodology. Kara D. Federmeier and Diane M. Beck contributed equally to writing–review and editing and supervision.

Correspondence concerning this article should be addressed to Will Deng, Department of Psychology, University of Illinois Urbana-Champaign, 603 East Daniel Street, Champaign, IL 61820, United States. Email: wuyizhe2@illinois.edu

calculated how well the memorability scores of 2,222 images in the first group matched the memorability scores of the same images in the second group (Isola et al., 2011). Twenty-five such random splits were made, which resulted in an average Spearman's rank correlation of .75 between the two groups. In another study, consistency across participants was found when the analysis was conducted separately for images in the same category (e.g., tower, house, and amusement park) with rank correlations ranging from .69 to .86 (Bylinskii et al., 2015). Bainbridge et al. (2013) conducted a similar analysis using 2,222 images of faces and found an average rank correlation of .68. High consistency in memorability has also been found for abstract visualization such as graphs and charts (rank correlation = .83; Borkin et al., 2013), as well as for words (Xie et al., 2020).

This consistency in stimuli that are more and less memorable is striking given that many factors that are known to affect memory are idiosyncratic: Each person has different levels, types, and timecourses of experience with any given stimulus or stimulus type and has likely formed different patterns of associations with it. Thus, it might have been expected that there would be little consistency across people in which stimuli they find easiest to remember (Bainbridge et al., 2013; Bowen et al., 2018). Yet, memorability does seem to be an intrinsic property of items and one that can be reliably estimated by convolutional neural network models trained only on images and image labels (Khosla et al., 2015; Needell & Bainbridge, 2022). A critical question, therefore, is what properties of visual stimuli contribute to this kind of consistent memorability effect.

Researchers have endeavored to pinpoint the factors that might contribute to image memorability, from color and other basic image features to semantic properties, and have further probed whether memorability could be a proxy for some other type of subjective judgment (Bainbridge et al., 2017; Broers & Busch, 2021; Bylinskii et al., 2015; Isola et al., 2011, 2014; Wakeland-Hart et al., 2022). However, there has been limited success in identifying any particular visual property that can sufficiently explain memorability. In one study that examined the memorability of natural scenes across more than 800 participants, high-memorability images had a higher hit rate (97.8%) compared to low-memorability images (69.0%) even when factors such as color, spatial frequency, size, and type (e.g., natural vs. manmade) of depicted objects and their backgrounds (e.g., indoor/outdoor) were controlled between the two image groups (Isola et al., 2011). Other work showed that memorability was not strongly correlated with participant-rated aesthetics or interestingness, even though the two were strongly associated with participants' subjective estimates of how well they thought they would remember a given image (Isola et al., 2014). Furthermore, the memorability effect has been shown to be distinct from the depth of encoding, priming, and top-down cognitive control (Bainbridge, 2020).

To try to uncover the locus of memorability effects, researchers have also turned to functional magnetic resonance imaging. Participants were presented with a sequence of high- and low-memorability face images and natural scenes with predetermined memorability scores (Bainbridge et al., 2017). Memorability was correlated with activation in the ventral visual stream and in medial temporal lobe regions, and there was no interaction between memorability and image type (faces vs. natural scenes), indicating that there might be similar factors underlying memorability for different types of stimuli (Devlin & Price, 2007; Olarte-Sánchez et al., 2015). No memorability effect was found in early visual cortical areas, indicating that the effect could not be attributed to low-level visual features, supporting the findings by Isola et al. (2011). A similar image memorability effect in ventral and medial temporal regions was found in a study using 200 face images, and this activation was topographically distinct from that associated with individual memory performance when participants were asked to retrieve previously studied images (Bainbridge & Rissman, 2018). Indeed, both functional magnetic resonance imaging and magnetoencephalography data have revealed neural correlates for image memorability even when the images were not later remembered due to variation in individual performance or to extrinsic factors such as masking (Bainbridge et al., 2017; Mohsenzadeh et al., 2019). These studies suggest that the factors that affect item-level memorability are distinct from the factors that affect episodic memory in the moment and that memorability effects are related to patterns of activity in higher order perceptual areas.

Although a particular property has yet to explain memorability, the commonality across participants suggests that memorability is sensitive to higher level visual representations that are shared across people. One candidate factor for driving memorability, therefore, is real-world statistical regularity—that is, properties extracted from the visual world that aid in more efficient perception and recognition. Patterns of statistical regularity may constitute the templates used by the perceptual system to aid in perceiving and/or recognizing inputs, allowing the visual system to more quickly make contact with knowledge stored in semantic memory. Examples of highly statistically regular images would include those that are more commonly encountered and are prototypical or highly representative of their category. At the other end of the continuum are unexpected or distinctive images. Both ends of this continuum could conceivably be related to memorability.

Real-world statistical regularity has been consistently found to affect image detection: Images that are more statistically regular are more readily perceived (i.e., perceived/recognized faster and more accurately) compared to those that are more irregular (Caddigan et al., 2017; Center et al., 2022; Greene et al., 2015; Smith & Loschky, 2019; Yang & Beck, 2023). For example, natural scenes that are rated to be more representative of a particular category, such as forest, city street, or beach, are not only detected more quickly and accurately than less representative scenes (Caddigan et al., 2017), but they were also found to be more similar to each other in their basic image properties (i.e., their spatial envelopes; Oliva & Torralba, 2001). Ease of perception can be estimated with an intact/scrambled task (Caddigan et al., 2017; Smith & Loschky, 2019; Yang & Beck, 2023) in which a mixture of intact images and their phase-scrambled counterparts are briefly presented, and participants are asked to press a button after each image to indicate whether they think it was intact or scrambled. More statistically regular images of natural scenes produce a higher d' than less regular ones, indicating participants were better at discriminating intact from scrambled images when the images were more regular (Caddigan et al., 2017; Center et al., 2022; Greene et al., 2015; Smith & Loschky, 2019; Yang & Beck, 2023). The advantage of statistically regular images in the intact/scrambled task extends beyond scenes to household objects presented in typical orientations (e.g., upright, frequently encountered orientations) versus atypical orientations (Center et al., 2022; Stojanoski & Cusack, 2014) and can also be measured using an adaptive staircase algorithm to estimate the how long the images needed to be presented for participants to reach an accuracy of 70.7%. Estimated durations are longer for less statistically regular stimuli (Yang & Beck, 2023).

If more memorable images are less statistically regular, then we would expect them to be harder to detect in an intact/scrambled paradigm. Distinctiveness has long been associated with increased memorability. For example, faces that deviate more from the average face tend to be remembered better by participants (Bartlett et al., 1984; Bruce et al., 1994; Busey, 2001). Distinctiveness is hypothesized to aid memory because features of an image that are more distinct help to separate that image from others, thus reducing competition when someone is trying to remember it (Valentine, 1991; Valentine et al., 2016). Thus, one might hypothesize that a source of memorability differences could lie in how distinctive images are; that is, how much they deviate from one another and/or how much they deviate from learned patterns.

However, some studies have found that distinctiveness alone cannot sufficiently explain the intrinsic memorability effect (Bainbridge et al., 2013). Moreover, other studies have shown that, like statistically regular images (Shao & Beck, in press; Torralbo et al., 2013), high-memorability images have more similar neural patterns to each other than do low-memorability images (Bainbridge & Rissman, 2018; Bainbridge et al., 2017). Thus, another possibility is that memorability effects arise because high-memorability images are more statistically regular (or "prototypical") and, therefore, easier to perceive and/or link to memory. In either case, if memorability is related to differences in statistical regularity, then we should expect that image memorability would yield differences in an intact/scrambled paradigm. Therefore, in two experiments using two different stimulus sets, we set out to test for memorability-based differences in perception using this task, reasoning that higher memorability images may be more distinctive, requiring longer exposure times, or may be more statistically regular, requiring shorter exposure times.

Experiment 1

To examine the effect of memorability on image detection, we compared participants' performance (the duration the images needed to be presented in order for participant to reach an accuracy threshold of 70.7%) with high- and low-memorability images on the intact/scrambled task, where participants are asked to indicate whether an image is intact or scrambled under rapid presentation (Caddigan et al., 2017; Center et al., 2022; Greene et al., 2015; Smith & Loschky, 2019; Yang & Beck, 2023). The images were selected from the LaMem data set (Khosla et al., 2015). In order to verify the memorability difference among the selected images, we fitted each image to ResMem, a machine-learning model designed to predict intrinsic memorability (Needell & Bainbridge, 2022). Image size, color contrast, and categories between the image sets were controlled. If memorability is driven by distinctiveness, we would expect high-memorability images to require a longer average exposure duration than low-memorability images, whereas if memorability is driven by prototypicality, we would expect high-memorability images to require a shorter duration. If memorability has little effect on detection durations, then either statistical regularity is not related to memorability or both distinctiveness and expectedness/prototypicality play a role in memorability, resulting in a combination of fast and slow detections that average out to similar durations overall.

Method

Transparency and Openness

We report how we determined the sample size, all data exclusion, manipulations, and measures in the experiment. The data, analysis code, and materials are available on Open Science Framework and can be accessed at https://osf.io/548gx/?view_only=741a953ee6e44df7a9a64980ad39679e. The experiment was not preregistered.

Materials

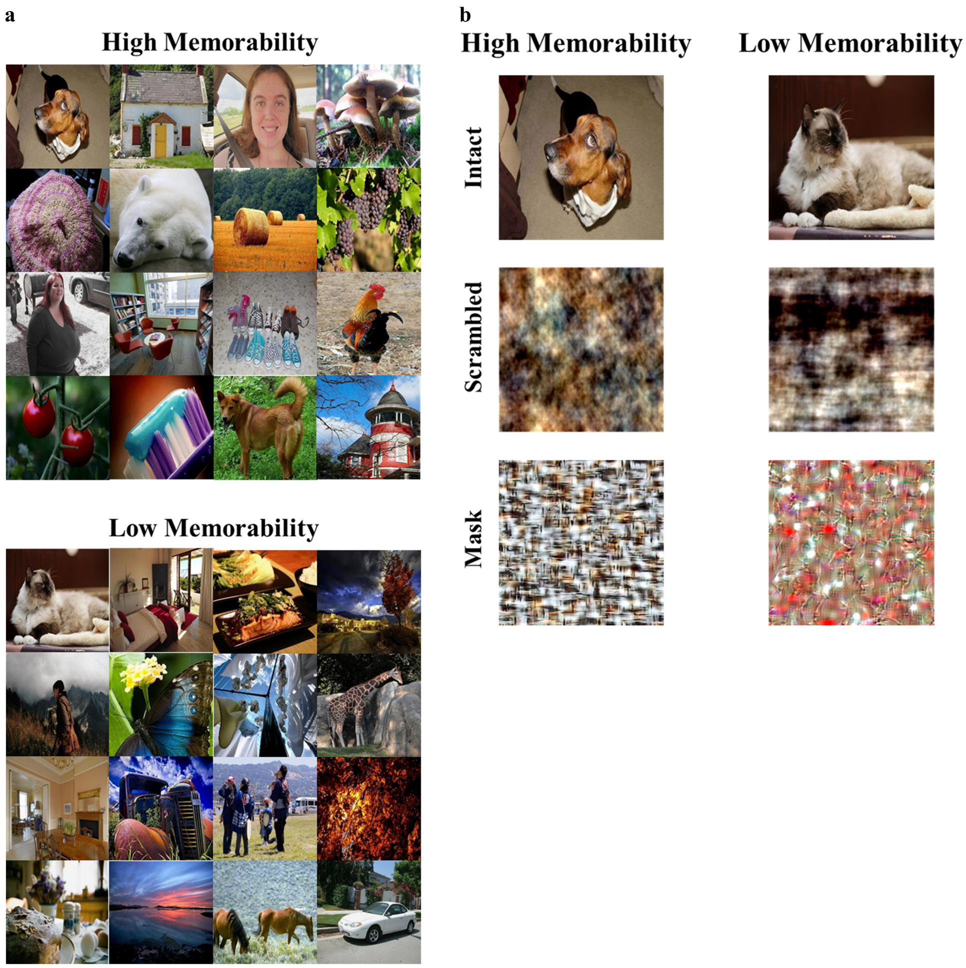
One hundred fifty high-memorability and 150 low-memorability natural scenes were selected from the LaMem data set using the rankings provided by the researchers (Khosla et al., 2015). The same number of images for each of the five experimenter-determined categories (animal, architecture, nature, object, and people) were selected for the two conditions to broadly control for any category effect (Bylinskii et al., 2015). In order to adapt the images to the intact/scrambled task, they were resized to 512 × 512 pixels, and the contrast was balanced between high- and low-memorability images. The resizing caused changes in aspect ratio for some images, but these changes, characterized as the absolute value of the difference in aspect ratio between the original and resized images, were similar across the memorability conditions (high, $M = 0.317$, $SD = 0.163$; low, $M = 0.427$, $SD = 0.173$). Three hundred scrambled images were created from each of the 300 intact images by randomizing the phase of their power spectrum while leaving its amplitude unchanged (Sadr & Sinha, 2004). A perceptual mask, comprised of colorful synthetic texture (Portilla & Simoncelli, 2000) shown to be an effective mask of scenes (Loschky et al., 2010), was randomly assigned to each intact and scrambled image. Examples of image sets from high- and low-memorability conditions are shown in Figure 1a. Intact and scrambled high- and low-memorability images with masks are shown in Figure 1b. The memorability difference between the post-processed high- and low-memorability images was verified using the ResMem model with an average estimated hit rate of 0.89 (range = 0.64–0.98, $SD = 0.06$) for the high-memorability images and 0.53 (range = 0.41–0.75, $SD = 0.09$) for the low-memorability images while categories, color contrast, and image size were controlled between the two conditions (Needell & Bainbridge, 2022).

Procedure

In the intact/scrambled task, each trial consisted of a fixation cross, a target image (intact or scrambled), and then a perceptual mask (see Figure 2). Participants were asked to indicate whether the target image is intact or scrambled by pressing the left versus the right Control keys on a keyboard. The assignment of the keys for intact/scrambled was counterbalanced across participants. Participants were told to respond quickly without sacrificing accuracy. Images were presented on an 85 Hz CRT monitor of resolution 1,280 × 960 using the PsychoPy 2021.2.3 package (Peirce et al., 2019) and Python (Python Software Foundation. Python Language Reference, Version 3.6.6). Participants viewed the images with their chin on a chinrest 59 cm away from the monitor.

Practice. Participants first completed 60 practice trials. In the practice, the target images were always presented for 118 ms and the masks were presented for 494 ms. Feedback was given for each trial; the word "incorrect" appeared in red in the middle of the screen along with a beep sound (100 Hz tone) for incorrect responses, and a black "correct" appeared for correct responses.

Figure 1
Examples of Stimuli



Note. (a) Examples of high- and low-memorability images in Experiment 1. (b) Examples of intact, scrambled, and mask images in Experiment 1. The first row shows examples of intact images in Experiment 1. The second row shows the full-phase scrambled version of the intact images. The third row shows examples of the perceptual masks that were randomly assigned to an intact or scrambled image. Adapted from the LaMem dataset; "Understanding and Predicting Image Memorability at a Large Scale," by A. Khosla, A. S. Raju, A. Torralba, and A. Olivia, 2015, International Conference on Computer Vision (https://ieeexplore.ieee.org/document/7410632). See the online article for the color version of this figure.

The images used in the practice trials were from the same general categories as those used in the main experiment but did not overlap with images in the main experiment.

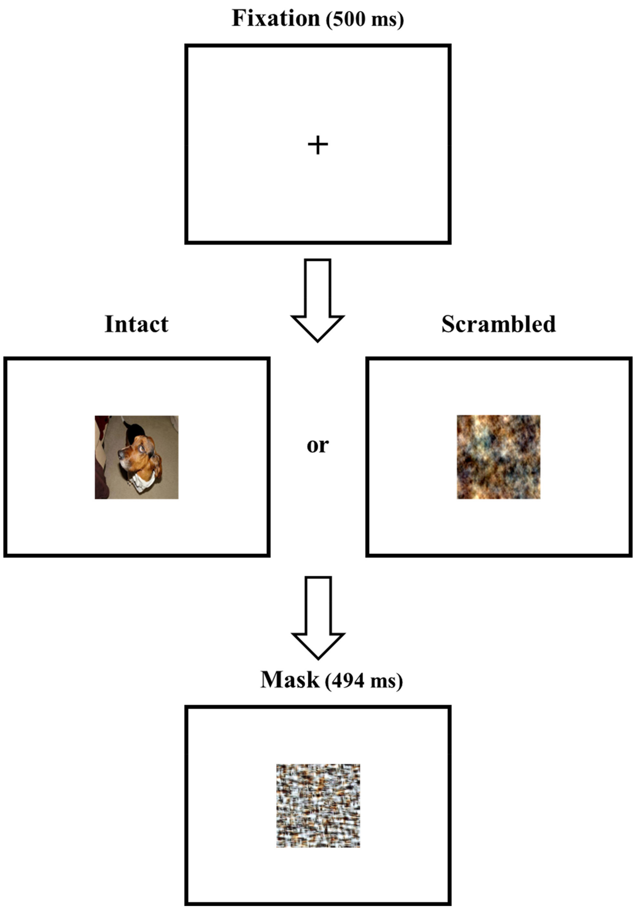
Main Experiment. For the main experiment, the duration of the target images for a given trial was determined by a one-up two-down staircase that is designed to converge on an accuracy threshold of 70.7% (Wetherill & Levitt, 1965). Each staircase consisted of 300 trials (150 intact and 150 scrambled in randomized order), an initial duration of 118 ms, a step size of 12 ms, and a maximum duration capped at 247 ms. The duration only started changing after the first six trials, which were excluded from the analysis. The masks were presented for 494 ms, and no feedback was given. Participants completed a staircase of high-memorability images and a separate staircase of low-memorability images with the order of the two counterbalanced across participants.

Participants

To determine our sample size, we recruited 13 participants for a pilot study. All participants in the pilot and the main experiment were recruited from the University of Illinois Urbana-Champaign participant pool and compensated with course credits. Written informed consent was obtained in accordance with procedures and protocols approved by the University of Illinois Institutional Review Board. All had self-reported normal or corrected-to-normal vision.

We determined a priori that participant data would be excluded from analysis if their staircase duration ever reached 200 ms in the last 50 trials, suggesting a significant deviation from the average performance seen in previous research (Caddigan et al., 2017; Yang & Beck, 2023). Participants would also be excluded if their overall accuracy in a staircase was lower than 60% (the chance

Figure 2
The Procedure of the Intact/Scrambled Task



Note. In the main experiment, the duration of the intact or scrambled image was determined by a one-up two-down staircase to achieve 70.7% accuracy. After the image was displayed, the trial would only proceed after a response was detected. Adapted from the LaMem dataset; "Understanding and Predicting Image Memorability at a Large Scale," by A. Khosla, A. S. Raju, A. Torralba, and A. Olivia, 2015, International Conference on Computer Vision (https://ieeexplore.ieee.org/document/7410632). See the online article for the color version of this figure.

level is 50%), suggesting that the staircase was not converging. Finally, they would be excluded if their staircase data failed to be fitted onto a two-parameter Weibull function, which was used to estimate the duration at 70.7% accuracy threshold (Wichmann & Hill, 2001). Of the 13 pilot participants, three were excluded for failing to be fitted onto a Weibull function (optimal parameters were not found). Using data from the remaining participants, we computed an effect size of $d = 0.86$ and determined the sample size using G*Power 3.1.9.7 for a paired sample t test (Faul et al., 2007). The power analysis indicated that we needed 13 participants to detect an effect size of $d = 0.86$ with 80% power and a significance criterion of $\alpha = .05$.

Given exclusion patterns in the pilot, we expected that we might need to exclude up to 30% of participants, so we recruited 25 participants for the main experiment to ensure we would meet or exceed our target sample size of 13. In the experimental sample, no participant was excluded because of high duration or low accuracy. Five participants were excluded because their data failed to be fitted by the Weibull function. Thus, data are reported from a final sample of 20 participants (11 self-identified as women and nine as men, $M_{\text{age}} = 20.34$). Participants provided their gender and age information when they signed up to create an account in the University of Illinois Urbana-Champaign participant pool. The available options for gender were man, woman, nonbinary, agender, genderqueer, questioning, other, and decline to answer.

Results

The estimated duration at 70.7% accuracy threshold for the high- and low-memorability images was derived by fitting the staircase data with a two-parameter Weibull function (Manning et al., 2018; Wichmann & Hill, 2001). The average sum of squared errors of the fits was 0.20, range = 0.01–0.44, $SD = 0.12$. Because the duration of the images was controlled based on the number of frames displayed by the CRT monitor in the experiment (one frame correlates to approximately 11.8 ms), the estimated duration was also in the unit of frames. We then compared the estimated duration between the high-memorability and low-memorability images using a nonparametric Wilcoxon signed-rank test (functionally similar to a Mann–Whitney U test; Field et al., 2012) due to violations of normality. Contrary to the prediction that memorability might be related to distinctiveness (statistical irregularity), the average duration of the low-memorability images ($M = 4.20$ frames or 49.41 ms) was significantly higher than the average duration of the high-memorability images ($M = 3.23$ frames or 38.00 ms) according to a two-sided Wilcoxon signed-rank test ($M_{\text{difference}} = 0.97$ frames or 11.41 ms, 95% confidence interval (CI) = [0.50, 1.39]), $V = 194$, $p < .001$, $r = .74$ (see Figure 3). This means that participants were better at detecting high-memorability images than low-memorability images. Thus, the results supported the hypothesis that more memorable images are easier to perceive, raising the possibility that they are more statistically regular than the low-memorability images. However, before drawing such a conclusion, we asked whether this result would generalize to a new set of high- and low-memorability images (taken from the FIGRM data set; Bylinskii et al., 2015)

Experiment 2

Experiment 2 followed the same intact/scrambled task procedures as Experiment 1 but with a separate set of natural scene images. We aimed to replicate the effect that high-memorability images would require a lower estimated duration to achieve 70.7% accuracy than low-memorability images, indicating that high-memorability images are easier to perceive.

Method

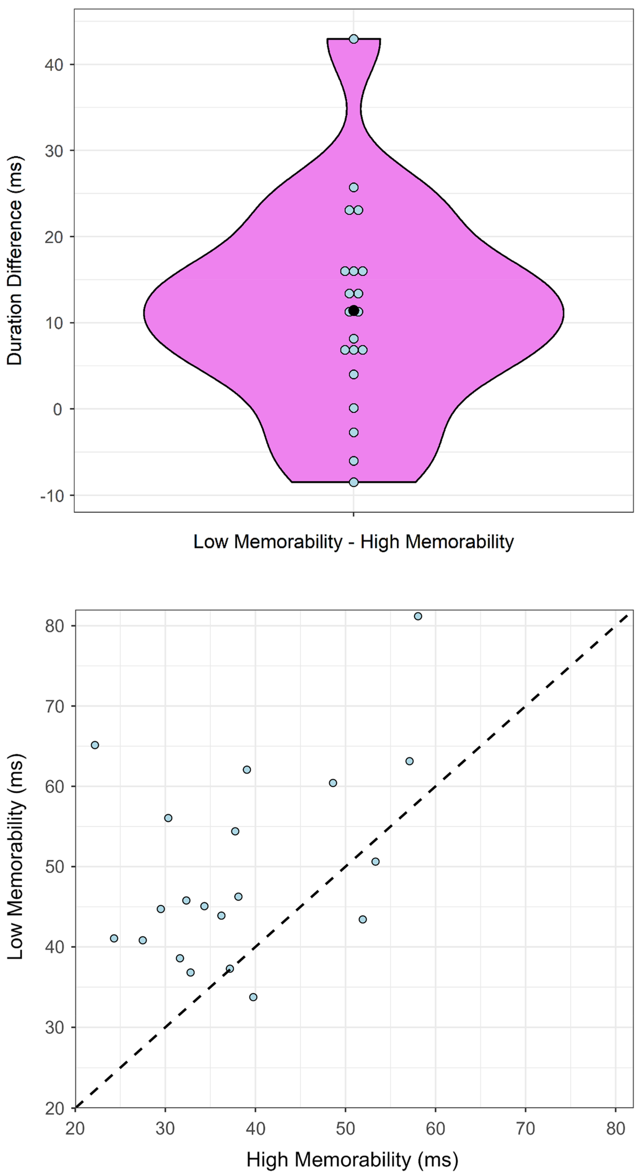
Transparency and Openness

Same as Experiment 1; for data, analysis code, and materials, see https://osf.io/gjhfq/?view_only=440102d3f924453e9e521f91f086b37a.

Materials

Eighty high-memorability and 80 low-memorability natural scenes were selected from the FIGRIM data set according to the

Figure 3
Estimated Durations for Low- Versus High-Memorability Images in Experiment 1



Note. The top panel shows the within-subject differences between the low- and high-memorability conditions. The larger black dot represents the overall mean difference in duration between low- and high-memorability images, and the smaller blue (light gray) dots represent the difference in duration for individual participants. The bottom panel shows the subject-level estimated durations for the low- and high-memorability conditions. See the online article for the color version of this figure.

memorability scores provided by the researchers (Bylinskii et al., 2015). Similar to Experiment 1, the same number of images were selected for each of the four categories (castle, golf course, house, and skyscraper) for the two memorability conditions. The procedures used to adapt the images to the intact/scrambled task were identical to Experiment 1. The memorability difference between the postprocessed high- and low-memorability images was verified using the ResMem model with an average estimated hit rate of 0.64 (range = 0.44–0.88, $SD = 0.10$) for the high-memorability images and 0.58 (range = 0.41–0.77, $SD = 0.09$) for the low-memorability images while categories, color contrast, and image size were controlled between the two conditions (Needell & Bainbridge, 2022). The difference in the memorability manipulation was thus lower than in Experiment 1.

Procedure

Experiment 2 followed the same procedure as Experiment 1, except the number of trials for each staircase was 160 in the main experiment (80 intact and 80 scrambled in randomized order).

Participants

We again used a pilot study ($N = 36$) to determine our sample size. As in Experiment 1, all participants in both the pilot and main study were recruited from the University of Illinois Urbana-Champaign participant pool and compensated with course credits. Written informed consent was obtained in accordance with procedures and protocols approved by the University of Illinois Institutional Review Board. All reported normal or corrected-to-normal vision.

We expected the required sample size to be larger than in Experiment 1 because of the smaller difference in memorability between the high- and low-memorability images (see the Materials section). Of the 36 pilot participants recruited, one was eliminated for having low accuracy, six were eliminated for high duration, and four were eliminated because their data failed to be fitted to a Weibull function. The remaining participants produced a memorability effect in the same direction as Experiment 1 with effect size of $d = 0.46$. A power analysis, using G*Power 3.1.9.7 for a paired sample t test and an effect size of $d = 0.46$ (Faul et al., 2007), indicated that we needed 40 participants to detect a similar-sized effect with 80% power with a two-sided test and 31 participants with a one-sided test and a significance criterion of $\alpha = .05$. We decided to obtain a sample size powered for a two-sided test, even though we planned to use a one-sided test in the main analysis to test whether we replicated the main finding of Experiment 1 that higher memorability images required reduced exposure durations (see the Results section).

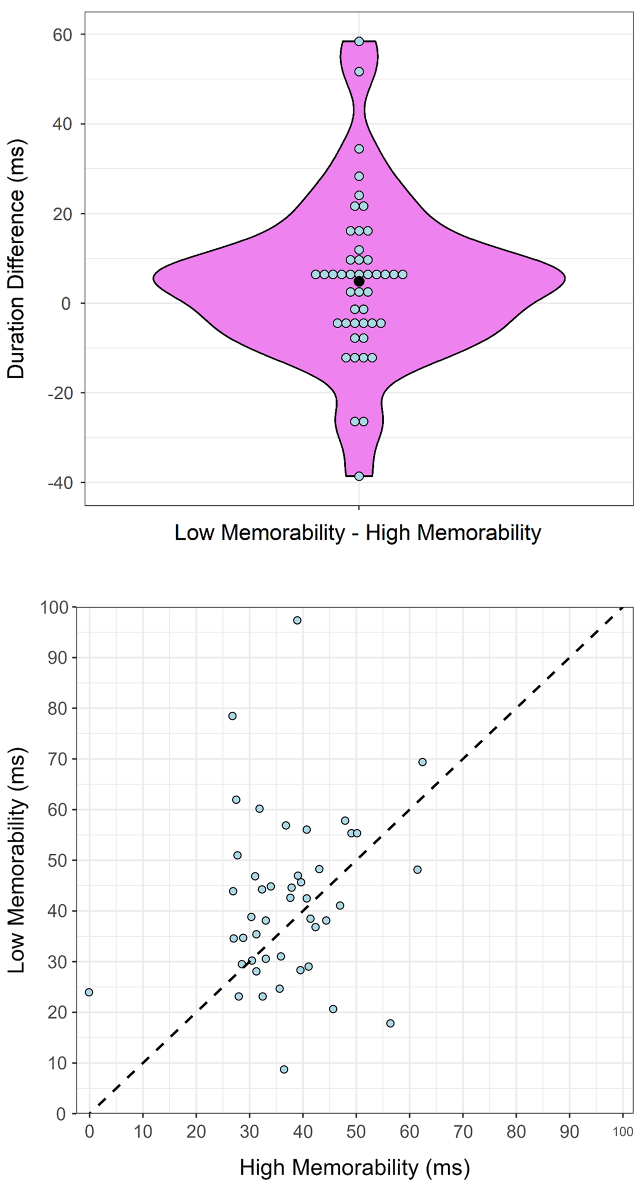
Assuming a possible exclusion rate (using identical criteria as for Experiment 1) of around 30%, we recruited 65 participants to ensure we would meet or exceed our target sample size of 40. One participant was eliminated for low accuracy, six participants were eliminated for high duration, and 13 participants were eliminated for failing to be fitted onto the Weibull function. Thus, we report data from a final sample of 45 participants (34 self-identified as women and 12 as men, $M_{age} = 20.20$).

Results

The estimated duration at 70.7% accuracy threshold for the high- and low-memorability images was derived by fitting the staircase data with a two-parameter Weibull function (Manning et al., 2018; Wichmann & Hill, 2001). The average sum of squared errors of the fits was 0.24, range = 0.01–1.12, $SD = 0.17$. Consistent with the findings in Experiment 1, the average duration of the low-memorability images ($M = 3.55$ frames or 41.76 ms) was significantly higher than the average duration of the high-memorability images ($M = 3.13$ frames or 36.82 ms) according to a one-sided Wilcoxon signed-rank test ($M_{\text{difference}} = 0.41$ frames or 4.82 ms, 95% CI = [0.04, $-\infty$]), $V = 687.5$, $p = .028$, $r = .29$ (see Figure 4). As predicted, participants were better at detecting high-memorability images than low-memorability images.

Figure 4

Estimated Durations for Low- Versus High-Memorability Images in Experiment 2



Note. The top panel shows the within-subject differences between the low- and high-memorability conditions. The larger black dot represents the overall mean difference in duration between low- and high-memorability images, and the smaller blue (light gray) dots represent the difference in duration for individual participants. The bottom panel shows the subject-level estimated durations for the low- and high-memorability conditions. See the online article for the color version of this figure.

Discussion

Past research has consistently shown a high degree of cross-person consistency in image memorability (the likelihood for a particular image to be remembered; Bainbridge et al., 2013; Borkin et al., 2013; Bylinskii et al., 2015; Isola et al., 2011). However, no visual properties have been found to sufficiently explain the memorability effect (Bainbridge et al., 2017; Broers & Busch, 2021; Bylinskii et al., 2015; Isola et al., 2011, 2014; Wakeland-Hart et al., 2022). In the present study, we show that highly memorable images are also more readily perceived under brief presentation.

To our knowledge, the present study is the first to show that the intrinsic memorability of visual stimuli affects image detection. Results from both Experiments 1 and 2 indicated that high-memorability images were easier to detect in the intact/scrambled task than low-memorability images. The effect persisted with images from different data sets (LaMem for Experiment 1 and FIGRIM for Experiment 2; Bylinskii et al., 2015; Khosla et al., 2015) and, in Experiment 2, could be observed even when the difference in memorability between the high- and low-memorability image sets was relatively small (0.64 vs. 0.58, respectively). Importantly, participants had not viewed the images prior to the intact/scrambled task, so the advantage is not due to the participants' episodic memory for the stimuli but properties of the images themselves. This detection advantage was observed even though the categories, stimulus contrast, and image size were controlled between the high- and low-memorability images. Indeed, past studies have suggested that low-level visual features (e.g., contrast and color) and scene categories are weak predictors of image memorability (Bainbridge et al., 2017; Dubey et al., 2015; Isola et al., 2014).

Instead, the present findings raise the possibility that highly memorable images are more statistically regular than low-memorability images. Statistical regularities, or learned patterns extracted from the visual environment and used to aid perception, have consistently been shown to affect how readily images are perceived in an intact/scrambled paradigm (Caddigan et al., 2017; Center et al., 2022; Greene et al., 2015; Smith & Loschky, 2019; Yang & Beck, 2023). In keeping with various kinds of statistically regular images, highly memorable images were discriminated from noise more readily. Because the intact/scrambled task is a perceptual task, the greater detectability—and, by inference, higher statistical regularity—of high-memorability images concurs with findings that memorability contains a perceptual component (Bainbridge & Rissman, 2018; Bainbridge et al., 2017). Like more statistically regular stimuli (Shao & Beck, in press; Torralbo et al., 2013), more memorable stimuli have more similar neural patterns in ventral visual stream regions (Bainbridge & Rissman, 2018; Bainbridge et al., 2017), suggesting that memorability, at least in part, relates to visual properties of an image. This is consistent with the findings that neural network models are able to predict image memorability with only visual information (Khosla et al., 2015).

We have argued in previous papers that statistical regularity can come in many forms: representativeness/prototypicality (Caddigan et al., 2017; Center et al., in press; Kumar et al., 2021; Shao & Beck, in press; Torralbo et al., 2013), canonicity (Center et al., 2022), familiarity (Smith & Loschky, 2019; Yang & Beck, 2023), and probability (Greene et al., 2015). Importantly, all such forms of statistical regularity have been shown to affect detection, such that the more statistically regular images are detected more accurately and quickly than the statistically irregular images. The key concept tying them together is that all such forms serve as better priors in a Bayesian sense. Many models of visual recognition propose that the visual system engages in perceptual hypothesis testing (Bar, 2003; Bullier, 2001; Hochstein & Ahissar, 2002; Rao & Ballard, 1999), in which incoming visual input is compared to an internal prediction (i.e., a template) as to what that input might be. Efficient processing in such systems depends, therefore, on minimizing prediction error. We argue that statistically regularities serve to, on average, minimize prediction error with respect to what the input is (i.e., recognition). The findings in the current experiments raise the possibility that images that better fit representational templates are not only easier to perceive but, at least in some cases, are also easier to remember. Such a hypothesis fits with recent data showing that object memorability positively correlates with typicality (Kramer et al., 2023), one of the factors that confer statistical regularity.

Of course, it need not be the case that the factors conferring the greater detectability of more memorable images are also what solely or directly yield a memory advantage for those items. Perceptual templates are also deeply related to semantic processing, and a number of semantic models postulate that our semantic knowledge of a visual category may include a visual template (Barsalou et al., 2003; Thompson-Schill, 2003). The view that the visual system uses statistical regularities to predict incoming stimuli thus also leaves room for a semantic component to memorability effects. For example, it could be that semantic properties of more memorable images make their visual templates easier to access, in turn facilitating detection. Alternatively, high-level perceptual properties of more memorable stimuli could bring online statistically regular visual templates, which then better connect to semantic representations, and it is the stronger connection with semantics that confers the memory advantage. Notably, neural network models that integrate both perceptual and conceptual features have produced more accurate estimates of image memorability compared to those that only integrate perceptual features (Khosla et al., 2015; Needell & Bainbridge, 2022). Moreover, memorability effects have also been documented for words and, in that case, are associated with semantic rather than perceptual features (Xie et al., 2020). Finally, a recent paper has shown that, across different image categories, semantic properties explain more variance in memorability than do visual properties (Kramer et al., 2023). Thus, the memorability effect could be an intermediate-stage property of perceptual as well as conceptual processing that encompasses both predictive coding for late-stage visual processing and semantic access, such that memorable stimuli may be both more statistically similar and easier to process semantically. Indeed, more prototypical scene images, which engender better detection in the intact/scrambled task, have been found in an event-related potential study (Kumar et al., 2021) to elicit facilitated processing both on the N300, a late visual component that has been linked to visual template matching (Schendan, 2019) and on the N400 component, which is linked to semantic access (Federmeier, 2022). As a result, high-memorability images' detection advantage in the intact/scrambled task may be attributed to a combination of facilitated predictive coding and ease of access to semantic knowledge.

As noted in the introduction, it was reasonable to hypothesize that memorability might be associated with distinctiveness, given the evidence that images that stand out from the rest are more likely to be remembered; for example, a face that looks more atypical may be better remembered among other faces (Bartlett et al., 1984; Bruce et al., 1994; Busey, 2001; Valentine, 1991; Valentine et al., 2016). However, the fact that high-memorability images were more easily perceived in the intact/scrambled task contradicts the hypothesis that memorability is driven by distinctiveness, since statistical irregularity has been associated with slower detection, not faster. It could be that the impact of distinctiveness versus statistical regularity in memorability varies with task demands or across stimulus type. For example, distinctiveness might play a more important role in memorability for stimuli that are highly similar to one another, such as faces. On the other hand, studies using both faces and scenes have found that distinctiveness cannot sufficiently explain memorability and that more memorable images may in fact be more statistically similar to one another (Bainbridge et al., 2013, 2017).

That said, it is likely that there are many factors that determine memorability, and our method may be sensitive to just some of them. Given earlier work on distinctiveness, it may be that distinctiveness is a factor that determines memorability for some images. Indeed, using the layers of a convolutional neural network as a proxy for similarity in early vision (lower layers of the network) versus later vision (later layers of the network), Koch et al. (2020) found that memorability increased for images that were more distinctive in the early vision but more similar in later vision (i.e., typical). Given this, it will be important to explore additional data sets and to understand the relative contribution of such factors. Other methods, such as event-related potentials, could also help dissociate the contributions of early perceptual and later semantics aspects of processing to these effects.

Critically, however, our results from two distinct data sets suggest typicality and/or statistical regularity may be the stronger driver of memorability. The same factors that confer a memory advantage also, on average, confer a perceptual advantage. Thus, it seems that a critical part of memorability arises from the ease and/or reliability with which incoming visual information can be connected with existing long-term (semantic) memory representations, which is facilitated by prototypicality or, more generally, statistical regularity.

Constraints on Generality

In our study, we presented photographic images to undergraduate students from the University of Illinois Urbana-Champaign participant pool in an intact-scrambled task. This population was chosen because memorability scores collected from healthy young participants in the lab have shown high consistency with memorability scores collected through online norming experiments (Bylinskii et al., 2015). Similarly, the perceptual difference based on statistical regularity in the intact-scrambled task has been replicated in many studies with healthy young participants (Caddigan et al., 2017; Center et al., 2022; Greene et al., 2015; Smith & Loschky, 2019; Yang & Beck, 2023). We thus chose the same population in our studies in order to draw comparisons across the studies. We therefore expect our results to generalize to all healthy, educated young adults and to real-world object- or scene-centric images. We have no reason to believe that the results depend on other characteristics of the participants or context, assuming that there is a memorability difference from the norming procedure described in the paper.

References

Bainbridge, W. A. (2019). Memorability: How what we see influences what we remember. In K. D. Federmeier & D. M. Beck (Eds.), *The psychology of learning and motivation: Knowledge and vision; The psychology of learning and motivation: Knowledge and vision* (pp. 1–27). Elsevier Academic Press. 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*, Article 107408. https://doi.org/10.1016/j.neuropsychologia.2020.107408

Bainbridge, W. A., Dilks, D. D., & Oliva, A. (2017). Memorability: A stimulus-driven perceptual neural signature distinctive from memory. *NeuroImage*, *149*, 141–152. 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), 1323–1334. 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), Article 8679. https://doi.org/10.1038/s41598-018-26467-5

Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience*, *15*(4), 600–609. https://doi.org/10.1162/089892903321662976

Barsalou, L. W., Simmons, W. K., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, *7*(2), 84–91. https://doi.org/10.1016/S1364-6613(02)00029-3

Bartlett, J. C., Hurry, S., & Thorley, W. (1984). Typicality and familiarity of faces. *Memory & Cognition*, *12*(3), 219–228. https://doi.org/10.3758/BF03197669

Borkin, M. A., Vo, A. A., Bylinskii, Z., Isola, P., Sunkavalli, S., Oliva, A., & Pfister, H. (2013). What makes a visualization memorable? *IEEE Transactions on Visualization and Computer Graphics*, *19*(12), 2306–2315. https://doi.org/10.1109/TVCG.2013.234

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

Broers, N., & Busch, N. A. (2021). The effect of intrinsic image memorability on recollection and familiarity. *Memory & Cognition*, *49*(5), 998–1018. https://doi.org/10.3758/s13421-020-01105-6

Bruce, V., Burton, M. A., & Dench, N. (1994). What's distinctive about a distinctive face? *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *47*(1), 119–141. https://doi.org/10.1080/14640749408401146

Bullier, J. (2001). Integrated model of visual processing. *Brain Research Reviews*, *36*(2–3), 96–107. https://doi.org/10.1016/S0165-0173(01)00085-6

Busey, T. A. (2001). 5 formal models of familiarity and memorability in face recognition. In M. J. Wenger & J. T. Townsend (Eds.), *Computational, geometric, and process perspectives on facial cognition: Contexts and challenges* (pp. 147–192). Lawrence Erlbaum Associates.

Bylinskii, Z., Isola, P., Bainbridge, C., Torralba, A., & Oliva, A. (2015). Intrinsic and extrinsic effects on image memorability. *Vision Research*, *116*(Pt. B), 165–178. https://doi.org/10.1016/j.visres.2015.03.005

Caddigan, E., Choo, H., Fei-Fei, L., & Beck, D. M. (2017). Categorization influences detection: A perceptual advantage for representative exemplars of natural scene categories. *Journal of Vision*, *17*(1), Article 21. https://doi.org/10.1167/17.1.21

Center, E. G., Federmeier, K. D., & Beck, D. M. (in press). *The brain's sensitivity to statistical regularity does not require full attention*. University of Illinois.

Center, E. G., Gephart, A. M., Yang, P.-L., & Beck, D. M. (2022). Typical viewpoints of objects are better detected than atypical ones. *Journal of Vision*, *22*(12), Article 1. https://doi.org/10.1167/jov.22.12.1

Devlin, J. T., & Price, C. J. (2007). Perirhinal contributions to human visual perception. *Current Biology*, *17*(17), 1484–1488. https://doi.org/10.1016/j.cub.2007.07.066

Dubey, R., Peterson, J., Khosla, A., Yang, M., & Ghanem, B. (2015). What makes an object memorable? *International conference on computer vision (ICCV)* (pp. 2089–2097). IEEE.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Federmeier, K. D. (2022). Connecting and considering: Electrophysiology provides insights into comprehension. *Psychophysiology*, *59*(1), Article e13940. https://doi.org/10.1111/psyp.13940

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. SAGE.

Greene, M. R., Botros, A. P., Beck, D. M., & Fei-Fei, L. (2015). What you see is what you expect: Rapid scene understanding benefits from prior experience. *Attention Perception, & Psychophysics*, *77*(4), 1239–1251. https://doi.org/10.3758/s13414-015-0859-8

Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron*, *36*(5), 791–804. https://doi.org/10.1016/S0896-6273(02)01091-7

Isola, P., Xiao, J., Parikh, D., Torralba, A., & Oliva, A. (2011), What makes an image memorable? *CVPR 2011* (pp. 145–152). https://doi.org/10.1109/CVPR.2011.5995721

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

Khosla, A., Raju, A. S., Torralba, A., & Oliva, A. (2015). Understanding and predicting image memorability at a large scale [Conference presentation]. *2015 IEEE international conference on computer vision, Santiago, Chile*. https://doi.org/10.1109/ICCV.2015.275

Koch, G. E., Akpan, E., & Coutache, M. N. (2020). Image memorability is predicted by discriminability and similarity in different stages of a convolutional neural network. *Learning & Memory*, *27*, 503–509. https://doi.org/10.1101/lm.051649.120

Kramer, M. A., Hebart, M. N., Baker, C. I., & Bainbridge, W. A. (2023). The features underlying the memorability of objects. *Science Advances*, *9*(17), Article eadd2981. https://doi.org/10.1126/sciadv.add2981

Kumar, M., Federmeier, K. D., & Beck, D. M. (2021). The N300: An index for predictive coding of complex visual objects and scenes. *Cerebral Cortex Communications*, *2*(2), Article tgab030. https://doi.org/10.1093/texcom/tgab030

Loschky, L. C., Hansen, B. C., Sethi, A., & Pydimarri, T. N. (2010). The role of higher order image statistics in masking scene gist recognition. *Attention, Perception, & Psychophysics*, *72*(2), 427–444. https://doi.org/10.3758/APP.72.2.427

Manning, C., Jones, P. R., Dekker, T. M., & Pellicano, E. (2018). Psychophysics with children: Investigating the effects of attentional lapses on threshold estimates. *Attention, Perception, & Psychophysics*, *80*(5), 1311–1324. https://doi.org/10.3758/s13414-018-1510-2

Mohsenzadeh, Y., Mullin, C., Oliva, A., & Pantazis, D. (2019). The perceptual neural trace of memorable unseen scenes. *Scientific Reports*, *9*(1), Article 6033. https://doi.org/10.1038/s41598-019-42429-x

Needell, C. D., & Bainbridge, W. A. (2022). Embracing new techniques in deep learning for estimating image memorability. *Computational Brain & Behavior*, *5*(2), 168–184. https://doi.org/10.1007/s42113-022-00126-5

Olarte-Sánchez, C. M., Amin, E., Warburton, E. C., & Aggleton, J. P. (2015). Perirhinal cortex lesions impair tests of object recognition memory but spare novelty detection. *European Journal of Neuroscience*, *42*(12), 3117–3127. https://doi.org/10.1111/ejn.13106

Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision*, *42*(3), 145–175. https://doi.org/10.1023/A:1011139631724

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. (2019). Psychopy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Portilla, J., & Simoncelli, E. P. (2000). A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer Vision*, *40*(1), 49–70. https://doi.org/10.1023/A:1026553619983

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. https://doi.org/10.1038/4580

Sadr, J., & Sinha, P. (2004). Object recognition and random image structure evolution. *Cognitive Science*, *28*(2), 259–287. https://doi.org/10.1016/j.cogsci.2003.09.003

Schendan, H. E. (2019). Memory influences visual cognition across multiple functional states of interactive cortical dynamics. In K. D. Federmeier (Ed.), *Psychology of learning and motivation* (pp. 303–386). Academic Press. https://doi.org/10.1016/bs.plm.2019.07.007

Shao, Z., & Beck, D. M. (in press). *Is attention necessary for the representational advantage of good exemplars over bad exemplars?* University of Illinois.

Smith, M. E., & Loschky, L. C. (2019). The influence of sequential predictions on scene-gist recognition. *Journal of Vision*, *19*(12), Article 14. https://doi.org/10.1167/19.12.14

Stojanoski, B., & Cusack, R. (2014). Time to wave good-bye to phase scrambling: Creating controlled scrambled images using diffeomorphic transformations. *Journal of Vision*, *14*(12), Article 6. https://doi.org/10.1167/14.12.6

Thompson-Schill, S. L. (2003). Neuroimaging studies of semantic memory: Inferring "how" from "where". *Neuropsychologia*, *41*(3), 280–292. https://doi.org/10.1016/S0028-3932(02)00161-6

Torralbo, A., Walther, D. B., Chai, B., Caddigan, E., Fei-Fei, L., & Beck, D. M. (2013). Good exemplars of natural scene categories elicit clearer patterns than bad exemplars but not greater BOLD activity. *PLOS ONE*, *8*(3), Article e58594. https://doi.org/10.1371/journal.pone.0058594

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology*, *43*(2), 161–204. https://doi.org/10.1080/14640749108400966

Valentine, T., Lewis, M. B., & Hills, P. J. (2016). Face-space: A unifying concept in face recognition research. *Quarterly Journal of Experimental Psychology*, *69*(10), 1996–2019. https://doi.org/10.1080/17470218.2014.990392

Wakeland-Hart, C. D., Cao, S. A., deBettencourt, M. T., Bainbridge, W. A., & Rosenberg, M. D. (2022). Predicting visual memory across images and within individuals. *Cognition*, *227*, Article 105201. https://doi.org/10.1016/j.cognition.2022.105201

Wetherill, G., & Levitt, H. (1965). Sequential estimation of points on a psychometric function. *British Journal of Mathematical and Statistical Psychology*, *18*(1), 1–10. https://doi.org/10.1111/j.2044-8317.1965.tb00689.x

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics*, *63*(8), 1293–1313. https://doi.org/10.3758/BF03194544

Xie, W., Bainbridge, W. A., Inati, S. K., Baker, C. I., & Zaghloul, K. A. (2020). Memorability of words in arbitrary verbal associations modulates memory retrieval in the anterior temporal lobe. *Natural Human Behavior*, *4*(9), 937–948. https://doi.org/10.1038/s41562-020-0901-2

Yang, P. L., & Beck, D. M. (2023). Familiarity influences visual detection in a task that does not require explicit recognition. *Attention, Perception, & Psychophysics*, *85*, 1127–1149. https://doi.org/10.3758/s13414-023-02703-7

Received September 29, 2023
Revision received February 27, 2024
Accepted March 12, 2024 ■