IMAGE MEMORABILITY IS LINKED TO FACILITATED PERCEPTUAL AND SEMANTIC PROCESSING

BY

WILL DENG

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Psychology
in the Graduate College of the
University of Illinois Urbana-Champaign, 2024

Urbana, Illinois

Advisers:

Professor Kara D. Federmeier
Professor Diane M. Beck

ABSTRACT

Strikingly, some images are consistently more likely to be remembered compared to others – a stable, intrinsic image property that has been termed image memorability. However, the properties that afford this memory advantage have remained elusive. In prior work, we showed that more memorable images are easier to perceive, and modeling work further suggests that semantic properties contribute to variance in memorability. Thus, we hypothesize that image memorability effects arise at the interface between perception and semantic memory. To test this hypothesis, we used event-related potentials (ERPs) to measure perceptual template matching (N300) and semantic access (N400) processes in a continuous recognition memory task using high and low memorability images, each repeated once. On initial presentation, both N300 and N400 amplitudes were less negative for high memorability images, showing that memorability is linked to both facilitated high-level perceptual processing and more efficient semantic activation. High memorability images also elicited a larger N300 repetition effect compared to low memorability images, revealing that their perceptual processing benefits more from the prior exposure. The results support the idea that images that better match visual templates and elicit more targeted semantic activations are easier to identify when encountered again, and further point to a potential interplay between semantic activation and perceptual matching in supporting image memorability.

ACKNOWLEDGMENTS

I would like to thank my advisors Kara Federmeier and Diane Beck for their guidance in my research.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

Image memorability is an item-level property that describes the likelihood that someone will recognize that image when seeing it a second time. It has often been estimated using a continuous recognition task wherein participants are presented with a sequence of new and repeated 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); memorability is then measured as the average hit rate (or d-prime Bainbridge, 2017; Bainbridge & Rissman, 2018) for each item’s repetition. Importantly, memorability scores for a given image have been found to be consistent across participants, such that iterated split-half analyses show high correlations – e.g., a Spearman’s rank correlation of 0.75 for a diverse set of images (Isola et al., 2011) and 0.68 for a more constrained set of unfamiliar face images (Bainbridge et al., 2013). Additionally, convolutional neural network models trained on images and image labels have been able to produce reliable memorability estimates, which supports memorability as an intrinsic item property (Khosla et al., 2015; Needell & Bainbridge, 2022).

Despite the consistency in image memorability and the fact that models can accurately predict which images are more memorable, it has proven difficult to identify specific image properties that can explain the memorability effect. Both behavioral and fMRI experiments have shown that basic image features, such as color and spatial frequency, do not reliably predict memorability (Bainbridge et al., 2017; Bylinskii et al., 2015; Isola et al., 2014). Variation in memorability is also not strongly correlated with subjective judgments of aesthetics or interest (Isola et al., 2014) and has been shown to be distinct from top-down cognitive control such as depth of encoding and priming (Bainbridge, 2020; Wakeland-Hart et al., 2022). Neural network models that incorporate both perceptual (e.g., regions within the images) and semantic (e.g., image category) features can produce more accurate memorability estimates than models that only integrate perceptual features (Khosla et al., 2015; Needell & Bainbridge, 2022), suggesting that memorability effects may lie at the interface between high-level perceptual and semantic processing.

In our previous study (Deng et al., 2024), we investigated whether memorability may be related to statistical regularity – i.e., the degree to which information in an image matches learned visual patterns. Many models of visual recognition propose that the visual system engages in perceptual hypothesis testing (Bullier, 2001; Hochstein & Ahissar, 2002; Rao & Ballard, 1999; Bar, 2003), in which incoming visual input is compared to an internal, memory-based prediction as to what that input might be. One such model proposes that these internal predictions reflect real-world statistical regularities, which are generalized representations that, on average, allow the visual system to quickly connect the visual input to semantics (Beck et al., in revision; Center et al., 2022; Center et al., 2024; Kumar et al., 2021; Shao & Beck, 2024; Yang & Beck, 2023). On this view, statistically regular images should be easier to process, as they will more closely match the internal prediction and thus yield less prediction error. Indeed, previous work using an intact/scrambled task, wherein participants indicate whether a briefly presented stimulus is a regular image or noise, has shown that statistically regular stimuli are more readily perceived, yielding higher detection accuracy or requiring shorter presentation times to reach criterial levels of detection accuracy (Center et al., 2022; Caddigan et al., 2017; Greene et al., 2015; Smith & Loschky, 2019; Yang & Beck, 2023). For example, scene images that have been judged to be better examples of scene categories (e.g., forest, city street, beach) are more readily perceived than less representative scenes (Caddigan et al., 2017), and this has been linked to facilitated late-stage perceptual processing (Kumar et al., 2021). Given that, like statistically regular images (Shao & Beck, 2024; Torralbo et al., 2013), high memorability images have been found to elicit more similar neural activation patterns compared to low memorability images (Bainbridge et al., 2017; Bainbridge & Rissman, 2018), it is possible that memorability might be linked to statistical regularity and thus affects ease of perceptual processing. Using a staircasing procedure to control presentation duration in an intact/scrambled task with high and low memorability images, we showed and then replicated with a second stimulus set, that high memorability images require shorter presentation time to be detected (Deng et al., 2024). Thus, memorability is associated with facilitated perceptual processing, possibly because more memorable images better match memory-based templates invoked during perception. A similar ease of processing argument has been put forward by others (Broers et al., 2018; Goetschalckx et al., 2019; Han et al., 2023)

Memorability, then, may arise from processing differences at the interface wherein high-level perceptual information comes to be linked into long-term memory. Indeed, as mentioned previously, neural networks models that include conceptual as well as perceptual features can predict image memorability more accurately compared to those that only include perceptual features (Khosla et al., 2015; Needell & Bainbridge, 2022), and recent work has suggested that semantic properties may actually contribute more than visual properties to memorability for objects (Kramer et al., 2023). Thus, the nature or organization of the semantic information evoked by images may play a critical role in allowing these images to later be recognized more accurately. Further understanding of image memorability effects, therefore, would benefit from the use of event-related potentials (ERPs), which can separately measure perceptual and semantic aspects of processing evoked by images. Here, we assess image memorability effects using ERPs, focusing in particular on two ERP components, the N300 and N400, that have been associated, respectively, with high-level visual processing and long-term semantic memory access.

Because high memorability images were more readily perceived in our previous work (Deng et al., 2024), we hypothesize that memorability may modulate the amplitude of the N300 component. The N300 is a negative-going ERP response with a broadly frontal distribution that peaks around 300 ms post-stimulus-onset. It is elicited by images such as line drawings and photographs of objects (Barrett & Rugg, 1990; McPherson & Holcomb, 1999), scenes (Kumar et al., 2021), and faces (Jemel et al., 1999). Its amplitude is reduced (less negative) for images that are easier to recognize and/or categorize. For example, Schendan and Kutas (2002) presented participants with fragmented line drawings of objects and gradually decreased the levels of fragmentation; they found that N300 amplitudes were smaller for images that were closer to being identified.

Based on findings like these, Schendan and colleagues have theorized that N300 amplitudes are linked to how readily the visual features of a stimulus can be matched to the long-term memory representation of known objects (Schendan & Lucia, 2010; Schendan, 2019). More recently, similar patterns of N300 effects have also been observed for scenes: Scenes that are good exemplars of their scene category – which, as described earlier, are detected more readily (Caddigan et al., 2017) – elicit smaller (less negative) N300s than do bad exemplars (Kumar et al., 2021). Thus, Kumar et al. (2021) suggested that the N300 might be a more general index of the ease and/or success with which incoming visual information can be matched to learned statistical regularities during perception. Given our finding that high memorability images behave like statistically regular images in a detection task, we expect that they might also elicit reduced N300s compared to low memorability images. If so, this would support the idea that more memorable images better match perceptual templates, making it easier to connect those images with long-term memory representations.

Using ERPs, we can also probe the relationship between memorability and semantic processing, as indexed by the N400, another negative-going component that peaks after the N300 (around 400 ms) and that has been linked to the access of information from long-term, multimodal (i.e., semantic) memory (Kutas & Federmeier, 2011). N400s are observed to all types of meaningful stimuli, including visual, auditory, and signed words (Kutas et al., 1987), environmental sounds (Van Petten & Rheinfelder, 1995), and line drawings, photographs, and movies (Ganis et al., 1996; Holcomb & McPherson, 1994; McPherson & Holcomb, 1999; Sitnikova et al., 2008). N400 amplitudes are reduced (become less negative) by a variety of factors that impact processing at the level of meaning. For example, it has been well-established that N400 amplitudes show a graded sensitivity to the fit between an incoming stimulus and its context, with reduced (less negative) N400s to stimuli that are related to a prior “prime” stimulus or that are more expected in a sentence, discourse, picture series, or movie (e.g., Kutas & Hillyard, 1980; Hamm et al., 2002; see review in Kutas & Federmeier, 2011). This effect is thought to arise because N400 amplitudes reflect the amount of new semantic information brought on-line by the incoming stimulus, which is correspondingly reduced when some of that information has already been activated during the processing of the context (Federmeier, 2021).

This account also explains systematic differences in N400 amplitudes to stimuli out of context (e.g., in lists). For example, for words, N400 amplitudes are graded by orthographic neighborhood size, a measure of how similar an incoming string is to other known lexical items. Both meaningful (e.g., words and acronyms) and non-meaningful (e.g., pseudowords and illegal letter strings) strings of letters elicit larger (more negative) N400 amplitudes when they have higher numbers of orthographic neighbors (Laszlo & Federmeier, 2011). One explanation for this pattern is that, when an incoming letter string is being processed, it activates semantics across the network of similar (confusable) forms, such that stimuli from a denser neighborhood activate more net semantic information, reflected in more negative N400 amplitudes (Holcomb et al., 2002). The same idea can also apply to semantic activation during image processing, wherein unrecognizable objects (Supp et al., 2005) and improbable images (Proverbio & Riva, 2009) also elicit more negative N400 amplitudes. When an image cannot be clearly comprehended, it may not be efficiently mapped onto specific semantic information and thus will elicit a higher, and possibly more variable, set of semantic activations. Extending these ideas to memorability, it is notable that low memorability images elicit neural activation patterns that are less similar to one another than those of high memorability images, which may also indicate a more variable set of semantic activations for the less memorable images (Bainbridge et al., 2017; Bainbridge & Rissman, 2018). If high memorability images better map to the activation of a specific set of semantic features, which also makes them more recognizable when viewed again, then they may also elicit smaller N400 amplitudes than do low memorability images.

We thus hypothesized that memorability might affect the amplitude of both the N300 and N400 components, with more positive amplitudes on both components for high versus low memorability images, even on initial presentation, reflecting differences in how these images are processed at the junction between high-level perception and long-term memory access. However, the fact that memorability is generally assessed based on recognition rate when images are presented for a second time makes it important to also examine ERP responses to image repetitions. We therefore adopted a continuous recognition design, presenting images of varying memorability twice across the stimulus stream and asking participants to indicate for each image whether it is new or old.

Both the N300 (Eddy et al., 2006) and N400 (Kutas & Federmeier, 2011) components have been shown to yield repetition effects in the form of less negative amplitudes when stimuli are presented for a second time after a brief delay. The N300 repetition effect has been postulated to be related to activation of learned object representations (Schendan & Maher, 2009). Canonical views of objects, which can be identified faster and more accurately because they better match stored object representations, elicit larger repetition effects (Schendan & Kutas, 2003); the same objects in non-canonical views have been found to have reduced or even absent repetition effects (Eddy & Holcomb, 2011). In turn, N400 repetition effects have been linked to the implicit maintenance of semantic/conceptual information from prior exposure (Olichney et al., 2000; Voss & Paller, 2007). Thus, we expect to see basic repetition effects on both components for all images. Of interest, then, is whether memorability might interact with the size of the repetition effect. Given that recognition rates are higher for high memorability images, we might expect to also see enhanced repetition effects for these items, suggesting that the boost to recognition at the behavioral level may be associated with processing fluency in high level perceptual processing (N300), semantic processing (N400), or both.

CHAPTER 2: EXPERIMENT

2.1 PARTICIPANTS

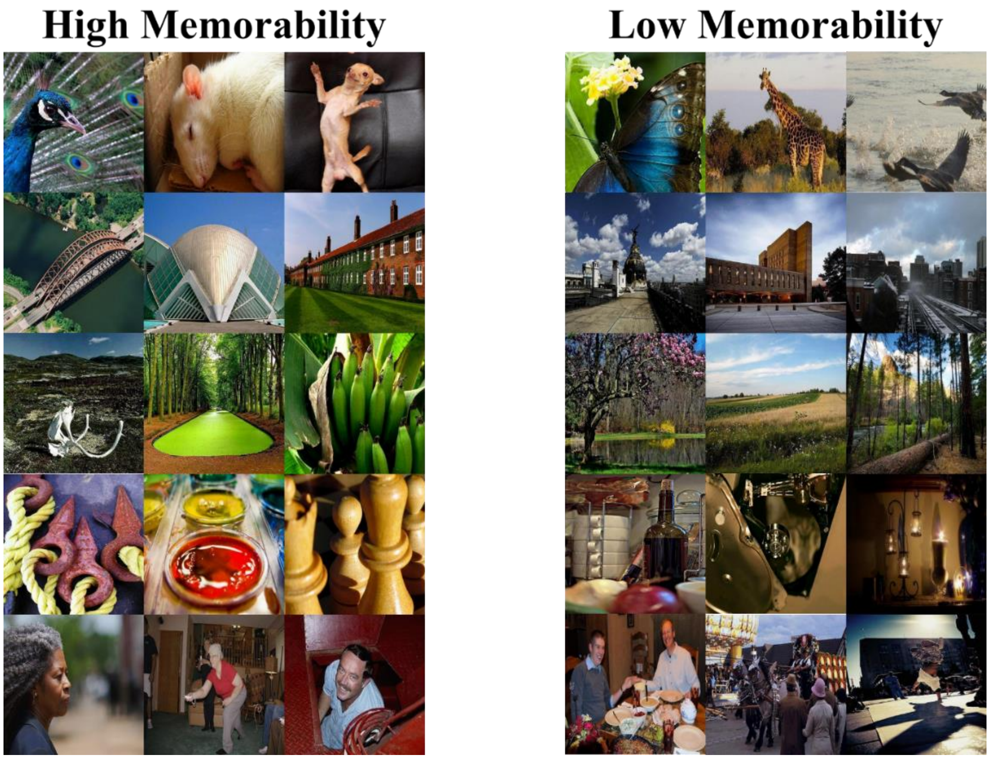
We determined the current study’s sample size based on the effect size estimate from a previous study that found a statistical-regularity-based N300 effect for visual scenes (Kumar et al., 2021); since N300 effects tend to be smaller than N400 effects, powering the study based on the N300 should also ensure sufficient power for N400 effects. An *apriori* power analysis using G*Power 3.1.9.7 for a paired sample t-test indicated that we needed 24 participants to detect an effect size of $d = 0.6$ with 80% power and a significance criterion of $\alpha = .05$. To obtain the target sample size of 24, we recruited 25 participants from the University of Illinois, who were paid for their time. Written informed consent was obtained in accordance with procedures and protocols approved by the University of Illinois Institutional Review Board. Participants all self-reported to be right-handed, have normal or corrected-to-normal vision, and no history of head trauma, seizures, neurological, or reading disorders. Data from one participant were excluded due to poor data quality, leaving 24 participants (16 self-identifying as female and 8 as male, $M_{age} = 26$) for analysis.

2.2 MATERIALS

One hundred and fifty high memorability and 150 low memorability images were selected from the LaMem dataset using the rankings provided by the researchers (Khosla et al., 2015; see Figure 1 for image examples); these were identical to the images used in Experiment 1 of Deng, Federmeier, and Beck (2024). The same number of images for each of 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). Images were resized to 512 x 512 pixels, and the contrast was balanced between high and low memorability conditions. The memorability difference between the post-processed high and low memorability images was verified using the ResMem model with an estimated hit rate of .89 (range: .64-.98, *SD* = .06) for the high memorability images and .53 (range: .41-.75, *SD* = .09) for the low memorability images (Needell & Bainbridge, 2022). Thirty filler images (needed to create the lag structure) were selected from the same dataset with the same number of images in each of the five categories and resized to 512 x 512 pixels. The images were arranged so that every target image (high and low memorability images) repeated once with a lag between six and fifteen images, and the filler images never repeated.

Figure 1

Examples of high and low memorability images from the LaMem dataset



2.3 PROCEDURES

At the beginning of the experiment, participants were informed that they would see a series of images one by one, some of which would repeat. They were then instructed to press a button with their right index finger for each new image and to press a different button with their right middle finger for each repeated image on a Cedrus RB-830 Response Pad (Cedrus, USA). The stimuli were presented on a 60 Hz CRT monitor of resolution 1280 x 1024 using the PsychoPy 2022.2.4 package (Peirce et al., 2019) and Python (Python Software Foundation. Python Language Reference, version 3.6.6). Participants completed the experiment sitting in a sound-attenuated recording booth, 118 cm away from the monitor. They were asked to remain still and to minimize eye movements and blinks during stimulus presentation. In each trial, a blue fixation cross at the center of the screen was displayed for 1000 ms against a black background, followed by an image for 1000 ms, and then another white fixation cross for 2000 ms. Participants completed a 20-trial practice before the main experiment, which consisted of 630 trials divided into three blocks with the block order counterbalanced across participants.

2.4 EEG RECORDING PARAMETERS

We obtained continuous EEG recordings using 26 silver/silver-chloride electrodes spaced evenly over the head and mounted in an elastic cap, amplified through a BrainAmpDC amplifier (Brain Products, USA). The 26 electrodes were Midline Prefrontal (MiPf), Left and Right Medial Prefrontal (LMPf and RMPf), Left and Right Lateral Prefrontal (LLPf and RLPf), Left and Right Medial Frontal (LMFr and RMFr), Left and Right Mediolateral Frontal (LDFr and RDFr), Left and Right Lateral Frontal (LLFr and RLFr), Midline Central (MiCe), Left and Right Medial Central (LMCe and RMCe), Left and Right Mediolateral Central (LDCe and RDCe), Midline

Parietal (MiPa), Left and Right Mediolateral Parietal (LDPa and RDPa), Left and Right Lateral Temporal (LLTe and RLTe), Midline Occipital (MiOc), Left and Right Medial Occipital (LMOc and RMOc), and Left and Right Lateral Occipital (LLOc and RLOc); see Figure 2. Additional electrodes placed on the outer canthus and infraorbital ridge of each eye were used to record blinks and saccades. Recordings were referenced on-line to the left mastoid and re-referenced off-line to the average of the left and right mastoids. Electrode impedances were reduced to below 5 kΩ. EEG was amplified and digitized with a 0.02-250 Hz band pass and a sampling rate of 1000 Hz.

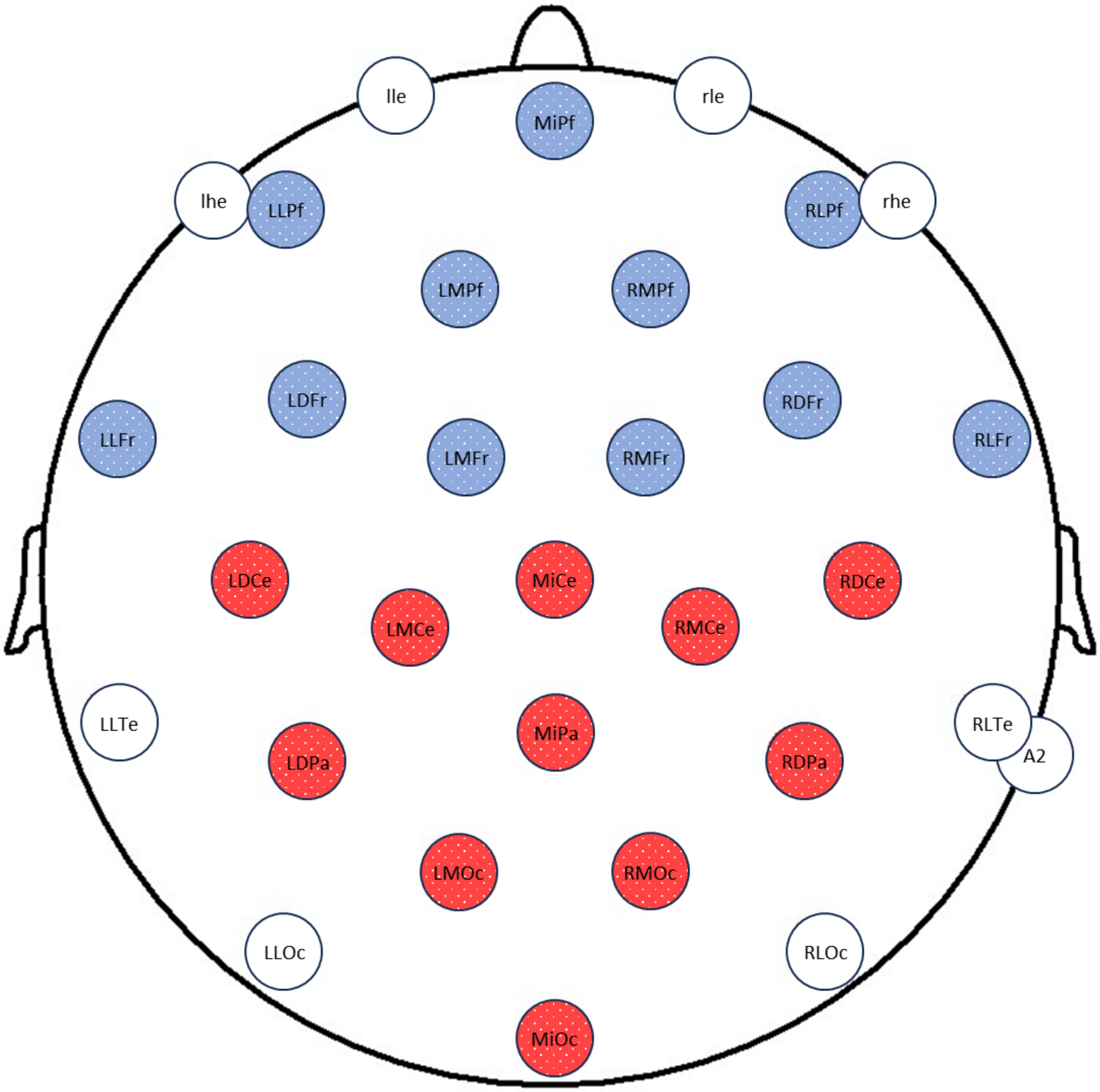
2.5 ERP DATA PROCESSING AND ANALYSIS

Trial-level EEG data were processed with EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014) toolboxes in Matlab. Each trial was a 1100 ms epoch from -200 to 900 ms time-locked to stimulus onset and baseline corrected with the mean amplitude of the 200 ms window prior to stimulus onset. We applied a 30 Hz low-pass filter and did artifact rejection for blinks, saccades, drift, and excessive muscle activity. Blinks and saccades were identified using thresholds calibrated for each participant in a condition-blind manner using visual inspection. Independent component analysis using AMICA was applied to participant data if there were blinks in more than 25% of the epochs ($N = 6$; Delorme et al., 2012); in this case, components were removed if they correlated with the eye channels at above .5. Epochs with artifacts that were not corrected were removed from all participant EEG data. Participants with greater than 25% of epochs removed due to artifacts even after applying ICA were excluded from analysis ($N = 1$ out of 25). On average, 85% of trials were retained with a range of 75-98% across participants (new high memorability: 87%, old high memorability: 88%, new low memorability: 81%, old low memorability: 84%).

To examine N300 effects, which are known to peak around 300 ms and to be largest over the front of the head (Hamm et al., 2002; Kumar et al., 2021; Schendan & Lucia, 2010), we computed mean amplitudes for each trial in an *apriori* 250-350 ms time window over all 11 frontal channels (MiPF, LLPf, RLPf, LMPf, RMPf, LLFr, LDFr, LMFr, RMFr, RDFr, RLFr; see Figure 2). N400 effects to color pictures also show a frontally skewed distribution (Federmeier & Kutas, 2002; Ganis et al., 1996), so we used this same frontal region of interest (ROI) to characterize the N400, measuring mean amplitudes in an *apriori* 350-500 ms window. Given that the distribution of the N400 tends to be broad, for completeness, we also measured mean amplitudes between 350-500 ms over a central-posterior ROI (LDCe, LMCe, MiCe, RMCe, RDCe, LDPa, MiPa, RDPa, LMOc, MiOc, RMOc; see Figure 2), with the expectation that this would show the same pattern as in the frontal ROI. We then used the same 11 channel posterior ROI to characterize the LPC, in a time window of 500-800 ms. The single trial ERP signals were fitted to linear mixed-effect models with dummy-coded categorical fixed effects of old/new (repeated vs. new images) and memorability (high vs. low memorability), random intercepts for participants and images, and random slopes of old/new and memorability for participants (Bates et al., 2014). Significance testing of fixed effects was estimated with t-tests using the Satterthwaite method in the lmerTest package in R (Kuznetsova et al., 2017).

Figure 2

Electrode Layout with Frontal and Central-Posterior ROIs



Note. Blue circles indicate channels in the frontal ROI, and red circles indicate channels in the central-posterior ROI.

CHAPTER 3: RESULTS

3.1 BEHAVIORAL RESULTS

To check for behavioral effects of image memorability, we compared the hit rates between high and low memorability images with a non-parametric Wilcoxon signed-rank test (functionally similar to a Mann-Whitney U test; Field et al., 2012) due to a violation of normality. The average hit rate of high memorability images ($M = .94$) was significantly higher than low memorability images ($M = .89$) according to a paired one-sided Wilcoxon signed-rank test ($M_{diff} = .05$, 95% CI = [.03, .07]), $V = 283$, $p < .001$, $r = .78$. As predicted, participants were more likely to recognize high memorability than low memorability images, even with the short repetition interval used in this study.

3.2 ERP RESULTS

Grand average ERPs grouped by memorability and old/new status at all 26 scalp electrode sites are shown in Figure 3.

3.2.1 N300

A mixed-effect model for the N300 time window with memorability and old/new as fixed effects revealed main effects of both memorability ($t = 5.75$, $p < .001$) and old/new status ($t = 12.33$, $p < .001$) and a significant interaction between the two ($t = 3.53$, $p < .001$). A simple comparison showed a reliable effect of memorability on the N300 even for first presentation of the pictures, such that new low memorability images ($M = -8.25$ µV) elicited more negative N300 amplitudes than new high memorability images ($M = -6.11$ µV), $M_{diff} = 2.14$ µV, $t = 5.75$, *p* < .001. Repetition reduced N300 amplitudes overall, but the decrease in N300 amplitude from new to repeated items was 3.51 µV, *t* = 12.33, *p* < .001 for high memorability images and 2.54 µV, *t* = 8.79, *p* < .001 for low memorability images. The larger decrease in amplitudes from repetition for high memorability images (see Figure 4 & 5) resulted in an overall larger memorability effect on second presentation (second presentation N300 amplitudes of -5.71 µV for low memorability versus -2.60 µV for high memorability).

3.2.2 N400

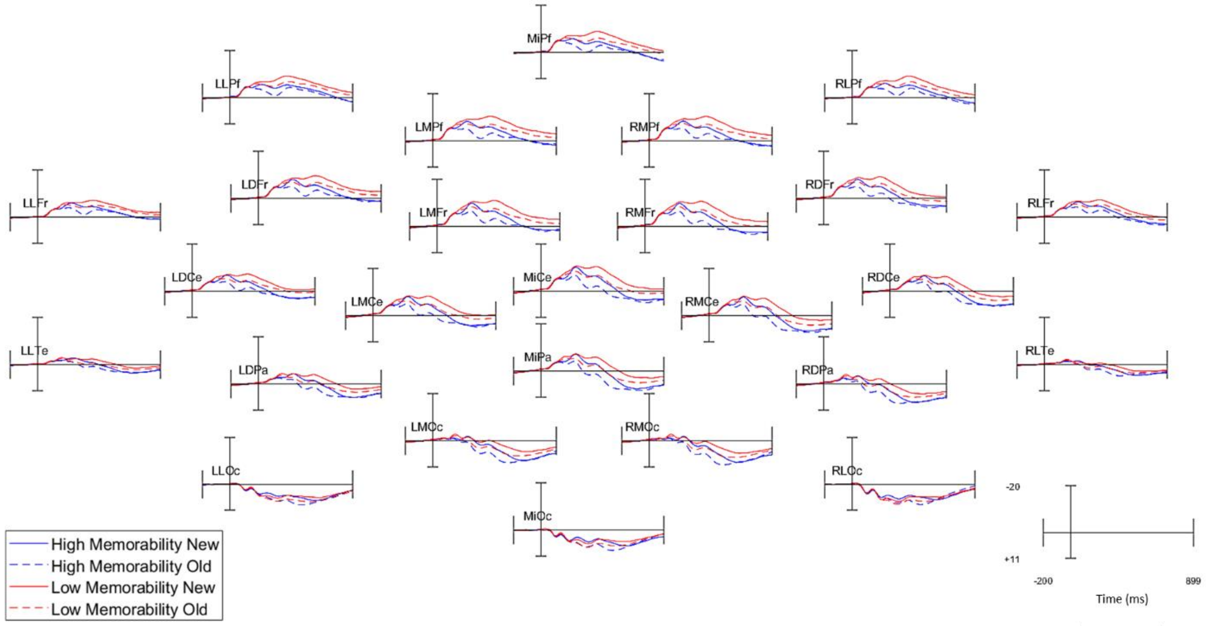
The same mixed-effect model for the N400 time window at the frontal ROI revealed main effects of memorability (*t* = 11.96, *p* < .001) and old/new status (*t* = 8.33, *p* < .001) but no significant interaction (*t* = 1.61, *p* = .109). The memorability effect reflected the fact that low memorability images (*M* = -8.82 µV) elicited more negative N400 amplitudes than did high memorability images (*M* = -4.81 µV) on first presentation. N400 amplitudes were then decreased with repetition, to a similar degree for high and low memorability images (M_{diff} = 2.63 µV). As predicted, the same N400 patterns were found in the posterior ROI—main effects of memorability (*t* = 9.57, *p* < .001) and old/new status (*t* = 8.77, *p* < .001), with no significant interaction (*t* = 1.71, *p* = .087).

3.2.3 LPC

The same mixed-effect model conducted in the LPC time window revealed main effects of memorability (*t* = 12.31, *p* < .001) and old/new status (*t* = 2.97, *p* = .005) and a significant interaction between the two (*t* = 3.76, *p* < .001). For new items, there was a memorability effect, with more positive responses to high (*M* = 4.84 µV) than to low (*M* = 0.91 µV) memorability images, M_{diff} = 3.99 μV, t = 12.31, p < .001. Repetition led to more positive LPC amplitudes overall, but the decrease in LPC amplitudes from new to repeated items was 1.08 μV, t = 2.97, p < .005 for high memorability images and 2.17 μV, t = 5.92, p < .001 for low memorability images. Thus, repetition effects on the LPC were enhanced for low compared to high memorability images.

Figure 3

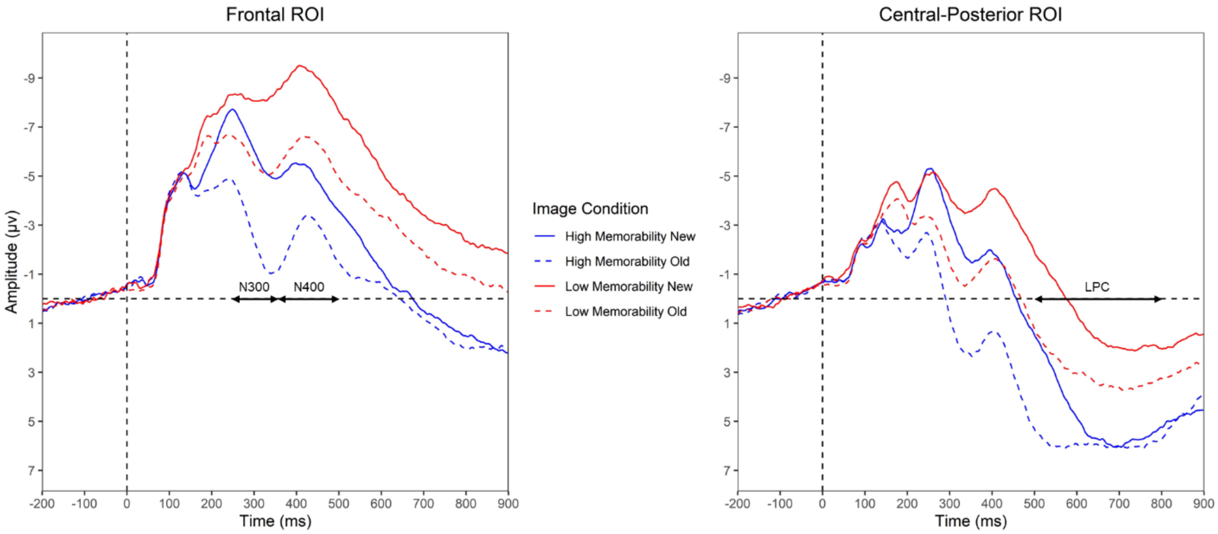
Grand Average ERPs at all 26 Electrode Sites.



Note. Plotted are responses to high memorability images (blue) and low memorability images (red) on initial presentation (New; solid lines) and when repeated (Old; dashed lines).

Figure 4

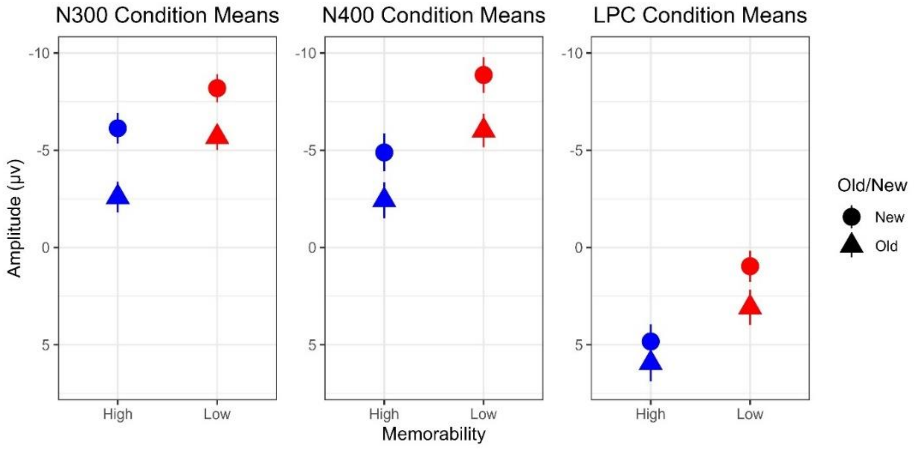
Grand average ERPs at the Frontal and Central-Posterior ROIs.



Note. Plotted are responses to high memorability images (blue) and low memorability images (red) on initial presentation (New; solid lines) and when repeated (Old; dashed lines). Time windows used to assess the N300, N400, and LPC are marked.

Figure 5

Average Amplitudes of ROIs Grouped by Memorability and Old/New Status for the N300, N400, and *LPC Time Windows*



CHAPTER 4: GENERAL DISCUSSION

We used event-related potentials (ERPs) to investigate how high and low memorability images are processed at perceptual and semantic levels, both on first apprehension and when encountered for a second time. Previous findings suggest that high memorability images may also be more statistically regular, as, similar to patterns observed for other images characterized by higher statistical regularity – e.g., objects seen from typical viewpoints and representative examples of natural scenes (Center et al., 2022; Caddigan et al., 2017; Greene et al., 2015; Smith & Loschky, 2019; Yang & Beck, 2023) – high memorability images are more readily perceived, requiring less time to be detected in an intact/scrambled task (Deng et al., 2024). Because statistical regularity is based on matching visual input to internal predictions from learned visual patterns (Beck et al., in revision; Center et al., 2022; Center et al., 2024; Kumar et al., 2021; Rao & Ballard, 1999; Shao & Beck, 2024; Yang & Beck, 2023), we hypothesized that the facilitated perceptual processing in memorable images could be related to the activation of semantic information in long-term memory. Thus, our main analyses in the present work focused on the N300 and N400 ERP components, which have been associated with high-level perceptual processing (e.g., template matching; Kumar et al., 2021) and semantic activation (Kutas & Federmeier, 2011), respectively. Because memorability measures the likelihood of recognizing an image that has been seen before, we used a continuous recognition task wherein a mixture of high and low memorability images was presented twice with a lag of 6-15 images, and participants were asked to indicate whether each image was new or repeated. Based on the ERP results for both the first and second presentations, we show that image memorability may arise due to an interplay between perceptual matching and semantic activation.

We predicted that the enhanced perceptual processing associated with memorability would be reflected in N300 amplitudes. Consistent with our prediction, at the first presentation, there was a small but significant N300 difference between the high and low memorability images. High memorability images elicited smaller (less negative) N300 amplitudes than did low memorability images, consistent with enhanced template matching that can aid in identification/recognition (Schendan & Kutas, 2002; Kumar et al., 2021). The N300 difference not only coincides with our previous behavioral findings that high memorability images are more readily perceived (Deng et al., 2024), but also provides further support for the idea that high memorability images are more statistically regular. Previous work has shown that scenes that are more statistically regular are both more readily perceived (Caddigan et al., 2017) and elicit smaller N300s (Kumar et al., 2021), identical to the pattern seen here for high (versus low) memorability images. The commonality with statistical regularity is an interesting one. We have argued that real-world statistical regularities, that is, those regularities built up over a lifetime, should produce generalized benefit such that, on average, the visual system is able to more quickly connect the visual input to knowledge (Beck et al., in revision; Center et al., 2024; Center et al., 2022; Kumar et al., 2021; Shao & Beck, 2024; Yang & Beck, 2023). By extension, this experience-based perceptual fluency may also confer a memory advantage.

Moreover, the perceptual fluency associated with memorable images seems to be accompanied by differences in semantic processing. We were interested in possible memorability-based differences in semantic activation since semantic features of images have been found to improve model prediction of memorability scores (Khosla et al., 2015; Needell & Bainbridge, 2022) and may even contribute more to memorability than visual features (Kramer et al., 2023). Based on fMRI data showing more homogeneous neural activation patterns for high memorability images (Bainbridge et al., 2017; Bainbridge & Rissman, 2018) as well as smaller N400s for statistically regular images (Kumar et al., 2021), we predicted that high memorability images may be more easily recognized because they are more efficiently mapped to specific semantic features. The more efficient and/or less variable mapping onto semantic information should thus result in an overall lower level of semantic activation, which would affect N400 amplitudes (Kutas & Federmeier, 2011). Consistent with our prediction, we observed that high memorability images elicit significantly smaller (less negative) N400s at first presentation. This finding is important because, to our knowledge, this is the first direct evidence that image memorability is related to semantic activation patterns during initial processing.

Because memorability is based on population-level recognition performance from tasks wherein people need to identify an image when seeing it a second time, we also examined repetition effects on the N300 and N400. Repetition effects were measured as changes in amplitude from the first to the second presentation, which allows us to probe for changes during perceptual processing and semantic activation as a result of the previous exposure. Replicating past work, both the N300 and N400 components showed a basic repetition effect, with less negative amplitudes on second presentation, showing that both aspects of processing are facilitated by the prior exposure. The size of the N400 repetition effects did not vary based on memorability. However, there was a larger N300 repetition effect for high memorability images compared to low memorability images. Thus, the perceptual processing of high memorability images benefitted more from repetition, consistent with earlier work showing that canonical views of objects elicit larger repetition effects (Schendan & Kutas, 2003). These results suggest that more easily recognized and more statistically regular images both benefit more from repetition, and that this benefit seemed to be localized at the level of late-stage perceptual processing indexed by the N300 (cf. Schendan & Maher, 2009).

High memorability images, therefore, are characterized by a reduced level of semantic activation at both the first and second presentation and by having particularly enhanced perceptual processing after initial exposure. One possibility is that the patterns of semantic and perceptual processing are linked, in that more targeted semantic activation may help to refine/restrict the perceptual templates to which the visual input is matched, which, in turn, facilitates perceptual processing at the second presentation. In contrast, less focused and more extensive semantic activation may be less informative feedback to perceptual templates. This is consistent with studies showing that, although knowledge of learned image categories can lead to faster visual perception (Maier & Rahman, 2019), additional semantic information, such as the category's functional significance, can actually hamper instead of facilitate perception, leading to slower processing times (Maier et al., 2014). Thus, a large amount of semantic information activated by an image may in fact create a perceptual disadvantage. In turn, consistency at the perceptual-semantic interface can affect recognition. For example, two-tone face images designed so that they appear like faces to all participants (semantically unambiguous to the face category) are more likely to be recognized at the second presentation than images that may not be seen as faces by all participants (Brady et al., 2019). It is possible, then, that semantic features make important contributions to memorability (Kumar et al., 2021) because semantic information serves the critical function of training the representations that determine how readily images are perceived/recognized later on.

Besides the N300 and N400 components, which were the main focus of this study, we also examined the late positive component (LPC), which follows the N400 in time, because it has been associated with explicit recollection (Curran, 2000; Finnigan et al., 2002) and recognition confidence (Addante et al., 2012) in memory tasks, even for stimuli without learned semantics (Bermúdez-Margaretto et al., 2015; Voss & Paller, 2007). Responses to low memorability images showed a typical “old/new effect”, with increased LPC amplitudes on second presentation compared to first presentation, suggesting that prior exposure enhanced brain activity associated with explicit recollection. High memorability images elicited larger (more positive) LPC signals even on the first presentation. Prior work has found that LPC responses are graded not only by recent exposure, but also by cumulative lifetime exposure (Yang et al., 2019). Thus, this pattern provides further support for the idea that higher memorability images may tend to be more statistically regular. Different from the pattern for low memorability images, high memorability images did not show a repetition effect on the LPC. This memorability-based difference in repetition effect pattern on the LPC is analogous to that documented for high versus low frequency words, where only low frequency words elicit enhanced LPCs upon repetition (Rugg, 1990). Overall, LPC responses thus provide additional evidence supporting enhanced recollection for high memorability images, possibly because they are more similar to past experience.

In conclusion, we not only confirmed that high memorability is associated with facilitated perceptual processing but extended prior work on memorability by showing memorability-based differences in the amount of semantic activation. Critically, we found that high memorability images, which are associated with higher recognition rate, actually trigger less semantic activation. Coupled with the facilitated perceptual processing, a picture emerges whereby the perceptual features of more memorable images not only map more efficiently onto object representations but the associated concepts themselves are more well-defined and easier to access. This allows us to propose a potential interplay between semantic activation and perceptual processing, where more efficient semantic activation facilitates the formation and refinement of effective templates for perception. Such efficiency and precision results in more memorable items.

REFERENCES

Addante, R. J., Ranganath, C., & Yonelinas, A. P. (2012). Examining ERP correlates of recognition memory: Evidence of accurate source recognition without recollection. *NeuroImage*, *62*(1), 439-450. https://doi.org/10.1016/j.neuroimage.2012.04.031

Bainbridge, W. A. (2017). The memorability of people: Intrinsic memorability across transformations of a person's face. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(5), 706–716. https://doi.org/10.1037/xlm0000339

Bainbridge, W. A. (2020). The resiliency of image memorability: A predictor of memory separate from attention and priming. *Neuropsychologia*, *141*, 1-12. 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*, 1-11. 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

Barrett, S. E., & Rugg, M. D. (1990). Event-related potentials and the semantic matching of pictures. *Brain and Cognition*, *14*(2), 201-212. https://doi.org/10.1016/0278-2626(90)90029-N

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting Linear Mixed-Effects Models using lme4. arXiv:1406.5823. http://arxiv.org/abs/1406.5823

Beck, D. M., Center, E. G., & Shao, Z. (2024). The role of real-world statistical regularities in visual perception. Manuscript submitted for publication.

Bermúdez-Margaretto, B., Beltrán, D., Domínguez, A., & Cuetos, F. (2015). Repeated exposure to "meaningless" pseudowords modulates LPC, but not N(FN)400. *Brain Topography*, *28*, 838-851. https://doi.org/10.1007/s10548-014-0403-5

Brady, T. F., Alvarez, G. A., & Störmer, V. S. (2019). The role of meaning in visual memory: Face-selective brain activity predicts memory for ambiguous face stimuli. *The Journal of Neuroscience*, *39*(6), 1100–1108. https://doi.org/10.1523/JNEUROSCI.1693-18.2018

Broers, N., Potter, M. C., & Nieuwenstein, M. R. (2018). Enhanced recognition of memorable pictures in ultra-fast RSVP. *Psychonomic Bulletin & Review*, *25*(3), 1080–1086. https://doi.org/10.3758/s13423-017-1295-7

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

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

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), 1–11. https://doi.org/10.1167/17.1.21

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), 1–14. https://doi.org/10.1167/jov.22.12.1

Center E. G., Federmeier, K. D., & Beck, D. M. (2024). The brain's sensitivity to real-world statistical regularity does not require full attention. *Journal of Cognitive Neuroscience*, 1-26. https://doi.org/10.1162/jocn_a_02181

Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition*, *28*, 923- 938. https://doi.org/10.3758/BF03209340

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9-21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Delorme, A., Palmer, J., Onton, J., Oostenveld, R., & Makeig, S. (2012). Independent EEG sources are dipolar. *PLOS ONE*, *7*(2): e30135. https://doi.org/10.1371/journal.pone.0030135

Deng, W., Federmeier, K. D., & Beck, D. M. (2024). Highly memorable images are more readily perceived. *Journal of Experimental Psychology: General.* Advance online publication. https://doi.org/10.1037/xge0001594

Eddy, M. D., & Holcomb, P. J. (2011). Invariance to rotation in depth measured by masked repetition priming is dependent on prime duration. *Brain Research*, *1424*, 38–52. https://doi.org/10.1016/j.brainres.2011.09.036

Eddy, M. D., Schmid, A., & Holcomb, P. J. (2006). Masked repetition priming and event-related brain potentials: A new approach for tracking the time-course of object perception. *Psychophysiology*, *43*, 564-568. https://doi.org/10.1111/j.1469-8986.2006.00455.x

Federmeier, K. D. (2021). Connecting and considering: Electrophysiology provides insights into comprehension. *Psychophysiology*, *59*(1):e13940. 10.1111/psyp.13940

Federmeier, K. D., & Kutas, M. (2002). Picture the difference: Electrophysiological investigations of picture processing in the two cerebral hemispheres. *Neuropsychologia*, *40*(7), 730-747. https://doi.org/10.1016/S0028-3932(01)00193-2

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

Finnigan, S., Humphreys, M. S., Dennis, S., & Geffen, G. (2002). ERP 'old/new' effects: Memory strength and decisional factor(s). *Neuropsychologia*, *40*(13), 2288-2304. https://doi.org/10.1016/S0028-3932(02)00113-6

Ganis, G., Kutas, M., & Sereno, M. I. (1996). The Search for "Common Sense": An Electrophysiological Study of the Comprehension of Words and Pictures in Reading. *Journal of Cognitive Neuroscience*, *8*(2), 89–106. https://doi.org/10.1162/jocn.1996.8.2.89

Goetschalckx, L., Moors, P., Vanmarcke, S., & Wagemans, J. (2019). Get the picture? Goodness of image organization contributes to image memorability. *Journal of Cognition*, *2*(1), 1-27. https://doi.org/10.5334/joc.80

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*, 1239–1251. https://doi.org/10.3758/s13414-015-0859-8

Hamm, J. P., Johnson, B. W., & Kirk, I. J. (2002). Comparison of the N300 and N400 ERPs to picture stimuli in congruent and incongruent contexts. *Clinical Neurophysiology*, *113*(8), 1339–1350. https://doi.org/10.1016/S1388-2457(02)00161-X

Han, S., Rezanejad, M., & Walther, D. B. (2023). Memorability of line drawings of scenes: The role of contour properties. *Memory & Cognition*. https://doi.org/10.3758/s13421-023-01478-4

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

Holcomb, P. J., Grainger, J., & O'Rourke, T. (2002). An Electrophysiological Study of the Effects of Orthographic Neighborhood Size on Printed Word Perception. *Journal of Cognitive Neuroscience*, *14*(6), 938–950. https://doi.org/10.1162/089892902760191153

Holcomb, P. J., & McPherson, W. B. (1994). Event-related brain potentials reflect semantic priming in an object decision task. *Brain & Cognition*, *24*(2), 259-276. https://doi.org/10.1006/brcg.1994.1014

Isola, P., Xiao, J., Parikh, D., Torralba, A., Oliva, A. (2011), What makes an image memorable? *CVPR 2011*, 145-152. 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), 469-1482. 10.1109/TPAMI.2013.200.

Jemel, B., George, N., Olivares, E., Fiori, N., & Renault, B. (1999). Event-related potentials to structural familiar face incongruity processing. *Psychophysiology*, *36*(4), 437–452. https://doi.org/10.1111/1469-8986.3640437

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. 10.1109/ICCV.2015.275.

Kramer, M. A., Hebart, M. N., Baker, C. I., & Bainbridge, W. A. (2023). The features underlying the memorability of objects. *Science Advances*, *9*(17), 1-14. 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), 1-14. https://doi.org/10.1093/texcom/tgab030

Kutas, M., & Hillyard, S. A. (1980). Reading Senseless Sentences: Brain Potentials Reflect Semantic Incongruity. *Science*, *207*(4427), 203–205. https://doi.org/10.1126/science.7350657

Kutas, M., Neville, H. J., & Holcomb, P. J. (1987). A preliminary comparison of the N400 response to semantic anomalies during reading, listening, and signing. *Electroencephalogr. Clin. Neurophysiol*. (Suppl. 39), 325–330.

Kutas, M., & Federmeier, K. D. (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP). *Annual Review of Psychology*, *62*(1), 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13). https://doi.org/10.18637/jss.v082.i13

Laszlo, S., & Federmeier, K. D. (2011). The N400 as a snapshot of interactive processing: Evidence from regression analyses of orthographic neighbor and lexical associate effects. *Psychophysiology*, *48*(2), 176–186. https://doi.org/10.1111/j.1469-8986.2010.01058.x

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00213

Maier, M., Glage, P., Hohlfeld, A., & Rahman, R. A. (2014). Does the semantic content of verbal categories influence categorical perception? An ERP study. *Brain and Cognition*, *91*, 1–10. https://doi.org/10.1016/j.bandc.2014.07.008

Maier, M., & Rahman, R. A. (2019). No matter how: Top-down effects of verbal and semantic category knowledge on early visual perception. *Cognitive, Affective Behavioral Neuroscience*, *19*, 859–876. https://doi.org/10.3758/s13415-018-00679-8

McPherson, W. B., & Holcomb, P. J. (1999). An electrophysiological investigation of semantic priming with pictures of real objects. *Psychophysiology*, *36*(1), 53–65. 10.1017/S0048577299971196

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

Olichney, J. M., Van Petten, C., Paller, K. A., Salmon, D. P., Iragui, V. J., & Kutas, M. (2000). Word repetition in amnesia: Electrophysiological measures of impaired and spared memory. *Brain*, *123*(9), 1948-1963. https://doi.org/10.1093/brain/123.9.1948

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*, 195-203. https://doi.org/10.3758/s13428-018-01193-y

Proverbio, A. M., & Riva, F. (2009). RP and N400 ERP components reflect semantic violations in visual processing of human actions. *Neuroscience Letters*, *459*(3), 142–146. https://doi.org/10.1016/j.neulet.2009.05.012

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*, 79–87. https://doi.org/10.1038/4580

Rugg, M.D. (1990). Event-related brain potentials dissociate repetition effects of high-and low-frequency words. *Memory & Cognition*, *18*, 367–379. https://doi.org/10.3758/BF03197126

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* (Vol. 71, pp. 303–386). Academic Press.

Schendan, H., & Kutas, M. (2002). Neurophysiological evidence for two processing times for visual object identification. *Neuropsychologia*, *40*(7), 931–945. https://doi.org/10.1016/S0028-3932(01)00176-2

Schendan, H. E., & Kutas, M. (2003). Time Course of Processes and Representations Supporting Visual Object Identification and Memory. *Journal of Cognitive Neuroscience*, *15*(1), 111– 135. https://doi.org/10.1162/089892903321107864

Schendan, H. E., & Lucia, L. C. (2010). Object-sensitive activity reflects earlier perceptual and later cognitive processing of visual objects between 95 and 500ms. *Brain Research*, *1329*, 124–141. https://doi.org/10.1016/j.brainres.2010.01.062

Schendan, H. E., & Maher, S. M. (2009). Object knowledge during entry-level categorization is activated and modified by implicit memory after 200 ms. *NeuroImage*, *44*(4), 1423-1438. https://doi.org/10.1016/j.neuroimage.2008.09.061

Shao, Z., & Beck, D. M. (2024). Is attention necessary for the representational advantage of good exemplars over bad exemplars? *European Journal of Neuroscience*, *59*(9), 2353–2372. https://doi.org/10.1111/ejn.16291

Sitnikova, T., Holcomb, P. J., Kiyonaga, K. A., & Kuperberg, G. R. (2008). Two neurocognitive mechanisms of semantic integration during the comprehension of visual real-world events. *Journal of Cognitive Neuroscience*, *20*(11), 2037-2057. https://doi.org/10.1162/jocn.2008.20143

Smith, M. E., & Loschky, L. C. (2019). The influence of sequential predictions on scene-gist recognition. *Journal of Vision*, *19*(12), 1–24. https://doi.org/10.1167/19.12.14

Supp, G. G., Schlögl, A., Fiebach, C. J., Gunter, T. C., Vigliocco, G., Pfurtscheller, G., & Petsche, H. (2005). Semantic memory retrieval: Cortical couplings in object recognition in the N400 window: Cortical couplings in object recognition. *European Journal of Neuroscience*, *21*(4), 1139–1143. https://doi.org/10.1111/j.1460-9568.2005.03906.x

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), e58594. https://doi.org/10.1371/journal.pone.0058594

Van Petten, C., & Rheinfelder, H. (1995). Conceptual relationships between spoken words and environmental sounds: Event-related brain potential measures. *Neuropsychologia*, *33*(4), 485-508. https://doi.org/10.1016/0028-3932(94)00133-A

Voss, J. L., & Paller, K. A. (2007). Neural correlates of conceptual implicit memory and their contamination of putative neural correlates of explicit memory. *Learning & Memory*, *14*(4), 259–267. https://doi.org/10.1101/lm.529807

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*, 1-11. https://doi.org/10.1016/j.cognition.2022.105201

Yang, PL., & 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

Yang, H., Laforge, G., Stojanoski, B., Nichols, E. S., McRae, K., & Köhler, S. (2019). Late positive complex in event-related potentials tracks memory signals when they are decision relevant. *Scientific Reports*, *9*, 1-15. https://doi.org/10.1038/s41598-019-45880-y