NeuroImage 149 (2017) 141–152

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Memorability: A stimulus-driven perceptual neural signature distinctive from memory

Wilma A. Bainbridge[a,*], Daniel D. Dilks[b], Aude Oliva[c]

[a] Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, United States
[b] Department of Psychology, Emory University, 201 Dowman Drive, Atlanta, GA 30322, United States
[c] Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, United States

ARTICLE INFO

Keywords:
Stimulus memorability
Medial temporal lobe
Memory
Perception

ABSTRACT

A long-standing question in neuroscience is how perceptual processes select stimuli for encoding and later retrieval by memory processes. Using a functional magnetic resonance imaging study with human participants, we report the discovery of a global, stimulus-driven processing stream that we call *memorability*. Memorability automatically tags the statistical distinctiveness of stimuli for later encoding, and shows separate neural signatures from both low-level perception (memorability shows no signal in early visual cortex) and classical subsequent memory based on individual memory. Memorability and individual subsequent memory show dissociable neural substrates: first, memorability effects consistently emerge in the medial temporal lobe (MTL), whereas individual subsequent memory effects emerge in the prefrontal cortex (PFC). Second, memorability effects remain consistent even in the absence of memory (i.e., for forgotten images). Third, the MTL shows higher correlations with memorability-based patterns, while the PFC shows higher correlations with individual memory voxels patterns. Taken together, these results support a reformulated framework of the interplay between perception and memory, with the MTL determining stimulus statistics and distinctiveness to support later memory encoding, and the PFC comparing stimuli to specific individual memories. As stimulus memorability is a confound present in many previous memory studies, these findings should stimulate a revisitation of the neural streams dedicated to perception and memory.

Introduction

While visual perception and memory are classically supported by functionally distinct cortical circuits (visual perception: Kanwisher and Dilks, 2013; memory: Wixted and Squire, 2011), the transitional steps from perception to memory remain under debate. Memory is thought of as a function of an individual observer, yet there are many consistencies across people in memory performance based on the stimulus (Bainbridge et al., 2013; Isola et al., 2011b). Could the brain be sensitive to a high-level perceptual stimulus property that guides memory encoding?

This question touches upon the concept of *memorability* – an intrinsic, perceptual stimulus property correlated with the likelihood of an image being later remembered or forgotten. Behavioral and computational work finds that memorability is a highly consistent statistical property of visual information, regardless of the observer (Bainbridge et al., 2013; Isola et al., 2011b), and can be computationally predicted (Khosla et al., 2015, 2013). Memorability, like many image properties (e.g., aesthetics, emotional content (Datta et al., 2008)), remains robust over different time lags (Isola et al., 2013) and viewing contexts (Bylinskii et al., 2015), and can be consistently measured in differing visual stimulus sets such as faces (Bainbridge, in press; Bainbridge et al., 2013), scenes (Isola et al., 2011b), and graphs and infographics (Borkin et al., 2013). A corpus of behavioral results has also shown that memorability can be isolated from other stimulus properties known to affect perception and memory. For instance, Isola et al. (2013, 2014) show that properties such as aesthetics, emotionality, subjective judgments of memorability, and low-level visual features are uncorrelated with scene memorability (Isola et al., 2013). Bainbridge et al. (2013) tested a comprehensive set of twenty face attributes to find that these attributes explained less than half of the variance in face memorability. Importantly, memorability is also found to be distinct from other cognitive processes known to influence memory behavior, and is dissociable from bottom-up attention, top-down attention, and priming effects (Bainbridge, in preparation). Memorability is thus particularly unique in that it serves as a

* Corresponding author. Present Address: NIMH Laboratory of Brain and Cognition, 10 Center Drive Room 3N228, Bethesda, MD 20814, United States.
E-mail addresses: wilma@mit.edu, wilma.bainbridge@gmail.com (W.A. Bainbridge), dilks@emory.edu (D.D. Dilks), oliva@mit.edu (A. Oliva).

http://dx.doi.org/10.1016/j.neuroimage.2017.01.063
Received 19 October 2016; Accepted 26 January 2017
Available online 27 January 2017

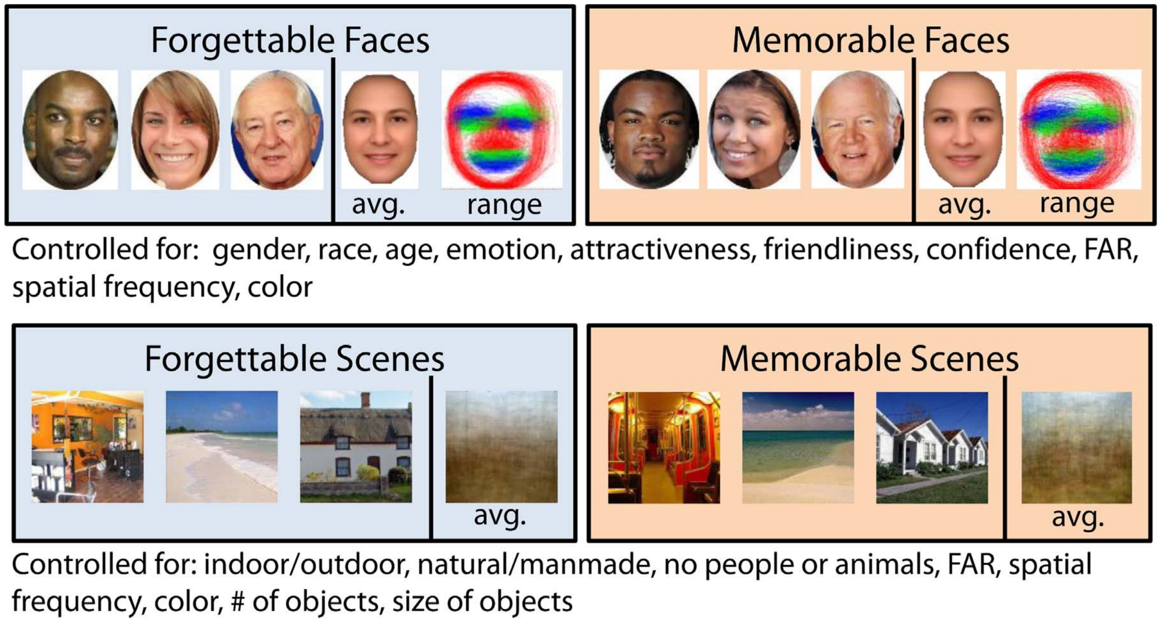

Fig. 1. Example stimuli used in the study, varying along two factors: 1) memorability (forgettable/memorable), and 2) stimulus type (face/scene). The stimuli were controlled for several low-level (e.g., color, spatial frequency) and high-level properties (e.g., emotion, attractiveness) that could be potentially confounded with memorability. Note that while it is not visually obvious which images would be memorable or forgettable, these images still result in very different memory behavior. The "avg." images show the average face and scene image within each condition. Faces were averaged using Active Appearance Model averaging (Cootes et al., 2001) while scenes were average based on average color at each pixel. The "range" for the faces shows the outlines of the facial features (e.g., eyes, nose, lips) for each stimulus. As one can see, between memorable and forgettable images, the averages are extremely similar and in fact show no significant differences. The face images shown here were not used in the actual study but are visually similar example images at opposite ends of memorability, from the public domain and available for publication. The average and range are determined from the stimuli used in the study.

consistent, measurable, and predictable signature of the encoding of a visual stimulus.

Memorability may thus help pinpoint high-level perceptual processing that precedes memory encoding. One particularly interesting candidate region for memorability is the medial temporal lobe (MTL), here defined as the set of structures encompassing the hippocampus, perirhinal cortex, entorhinal cortex, amygdala, and parahippocampal cortex. It is still debated whether the MTL is a memory-based region sensitive to novelty and familiarity (Brown and Aggleton, 1991; Daselaar et al., 2006; Desimone, 1996; Eichenbaum et al., 2007; Kim, 2011; Kumaran and Maguire, 2009; Rissman et al., 2010), or a high-level perceptual region for object discrimination and identification (Buckley and Gaffan, 1998; Cate and Köhler, 2006; Devlin and Price, 2007; Olarte-Sánchez et al., 2015). Several individual subsequent memory studies identify two main regions – the MTL and the prefrontal cortex (PFC) – that show higher activation for stimuli that are later remembered than those that are forgotten by an individual participant (Brewer et al., 1998; Wagner et al., 1998). However, the PFC shows the strongest effects of subsequent individual memory (Kim, 2011) and has been proposed to be a locus of familiarity separate from the MTL (Kafkas and Montaldi, 2014).

Given that memorable images are more likely to be remembered by individuals, most behavioral and neuroimaging studies may have confounded some effects attributed to individual subsequent memory (based on neural comparisons with previous individual experiences) with stimulus-driven memorability (based on statistical properties of a stimulus). In fact, behavioral work has identified that these are two separate factors that influence later memory behavior; stimulus memorability determines ultimate memory behavior as much as all other factors combined, including individual experience and environmental factors (Bainbridge et al., 2013). Here, we demonstrate that stimulus-driven memorability effects can be neurally disentangled from individual subsequent memory effects: the MTL appears to be involved with a rapid processing stream sensitive to a normative perceptual marker of memory (memorability), while the PFC appears to be involved in processing individual memory.

Using rapid event-related functional magnetic resonance imaging (fMRI) in a perceptual task with stimuli of predetermined memorability, we discover the human neural substrates of memorability, for both face and scene stimuli. First, we find clear sensitivity to memorability along the ventral visual stream (VVS) and the MTL for both face and scene categories, while no sensitivity exists in early visual cortex (EVC). This sensitivity to memorability is found to exist even in the absence of memory (i.e., for forgotten images). Lastly, we discover a double dissociation between stimulus memorability (in the MTL) and individual memory (in the PFC) in both univariate and multivariate measures. Taken together, these results point to a specialized neural stream that calculates stimulus-driven signatures of what people ultimately remember, at the speeds of perception.

Materials and methods

Participants

Eighteen adults (eight female, average age of 25.9 years) participated in the main event-related experiment. Seventeen separate adults (eight female, average age of 24.8 years) participated in a block-design study (see Supplementary material, S4), which serves as a replication of the results. Participants were healthy, right-handed, with normal or corrected vision, and US citizens, to reduce out-group effects (Chiroro and Valentine, 1995). They consented following guidelines approved by the MIT Institutional Review Board for fMRI studies and were compensated \$30/h. Two participants in the event-related experiment and one participant in the block-design experiment were excluded from analyses due to an inability to localize any ROIs at our chosen threshold of $p < 0.001$.

Stimuli

720 stimuli were put into four conditions of 180 each – 1) forgettable faces, 2) memorable faces, 3) forgettable scenes, and 4) memorable scenes. Stimuli were selected to be on opposite ends of