Spatial and Temporal Signatures of Memorability in the Brain

Wilma A. Bainbridge (1), Seyed-Mahdi Khaligh-Razavi (2), Dimitrios Pantazis (3), Aude Oliva (2)

1 – Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA

2 – Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA

3 – McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA

Introduction: An enduring question in neuroscience is how perceived stimuli are selected for memory encoding. However, despite their interdependence, perception and memory are historically studied in isolation. Recent research (Isola et al., 2011; Bainbridge et al., 2013) has found that images have an intrinsic *memorability* – regardless of the observer, some are likely to be remembered and others forgotten. Here we investigate the neural underpinnings of memorability processing, and find evidence for a distinct neural stage between human perception and memory.

Methods: Neural signatures of memorability were measured in two different neuroimaging experiments, using 3T event-related fMRI in the spatial domain (N=16), and MEG in the temporal domain (N=14). Participants performed a perceptual task with face and scene images that were predetermined to be highly memorable or forgettable. The images were equalized on several low-level (e.g., spatial frequency, color) and high-level (e.g., aesthetics, emotion) attributes. All images were novel and presented a single time. After the scan, participants did an unexpected subsequent memory task (Wagner et al., 1998), to allow comparisons between memorability and memory encoding. FMRI data were preprocessed and univariate contrasts were defined using BrainVoyager. Regions of interest were defined using anatomically-based manual segmentation (Insausti et al., 1998). MEG data were preprocessed with spatiotemporal filters and analyzed using Brainstorm. Representational similarity analyses (RSA; Kriegeskorte et al., 2013) were conducted to compare neural pattern similarity based on memorability and subsequent memory.

Results: For the fMRI study, a significant effect of memorability was found in the medial temporal lobe, but not in early visual areas (Fig. 1). In a two-way ANOVA (memorability × subsequent memory), the perirhinal cortex (PRC) showed a significant statistical interaction between the two factors, with a main effect of memorability but no effect of subsequent memory. Post-hoc tests revealed a consistent memorability effect in the PRC regardless of subsequent memory, as well as a consistent subsequent memory effect in the parahippocampal cortex (PHC) regardless of memorability (Fig. 2). These results implicate memorability processing and memory encoding as two neurally distinct processes. Our results also suggest that PRC has a representational geometry where activation patterns within memorable images are more similar than the activation patterns within forgettable images.

For the MEG study, a significant temporal signal able to decode memorability emerged at 217ms for faces and 365ms for scenes. In contrast, subsequent memory signals are shown in previous work to occur at 400-800ms (Curran et al., 2003), while perceptual signals occur at 100ms-200ms (Liu et al., 2003; Sato et al., 1999). Furthermore, we show that memorability and subsequent memory have significantly different temporal dynamics (Fig. 3).

Conclusions: Taken together, these results provide evidence that memorability is a dissociable process from memory encoding and early perception. The fMRI work found that memorability and subsequent memory effects occur in different loci with different patterns. The MEG work found that memorability occurs after perception but before memory encoding. This *memorability stage* could be conceptualized as a stage where perceived images are tagged for their memorability – or statistical distinctiveness – to then cause important stimuli to be encoded, and unimportant stimuli to be discarded. Indeed, memorability heavily influences memory encoding; in both studies, we found participants significantly remembered memorable stimuli over forgettable stimuli. However, altogether, these results support memorability as a built-in neural stage of processing, with both distinct spatial and temporal neural signatures, guiding encoding success from perceptual inputs.

<u>3961 characters / 4000 characters</u>

References

Bainbridge, W.A., et al. (2013), 'The intrinsic memorability of face photographs', Journal of Experimental Psychology: General, vol. 142, no. 4, pp. 1323-1334.

Curran, T., et al. (2003), 'Using ERPs to dissociate recollection from familiarity in picture recognition', Cognitive Brain Research, vol. 15, no. 2, pp. 191-205.

Insausti, R., et al. (1998), 'MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices', American Journal of Neuroradiology, vol. 19, pp. 659-671.

Isola, P., et al. (2011), 'What makes an image memorable?', Proceedings of the 24th IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 145-152.

Kriegeskorte, N., et al. (2013), 'Representational geometry: integrating cognition, computation, and the brain', Trends in Cognitive Sciences, vol. 17, no. 8, pp. 401–412.

Sato, N., et al. (1999), 'Different time course between scene processing and face processing: a MEG study', Neuroreport, vol. 10, no. 17, pp. 3633-3637.

Wagner, A.D., et al. (1998), Building memories: Remembering and forgetting of verbal experiences as predicted by brain activity. Science, vol. 281, no. 5380, pp. 1188-1191.

Figures

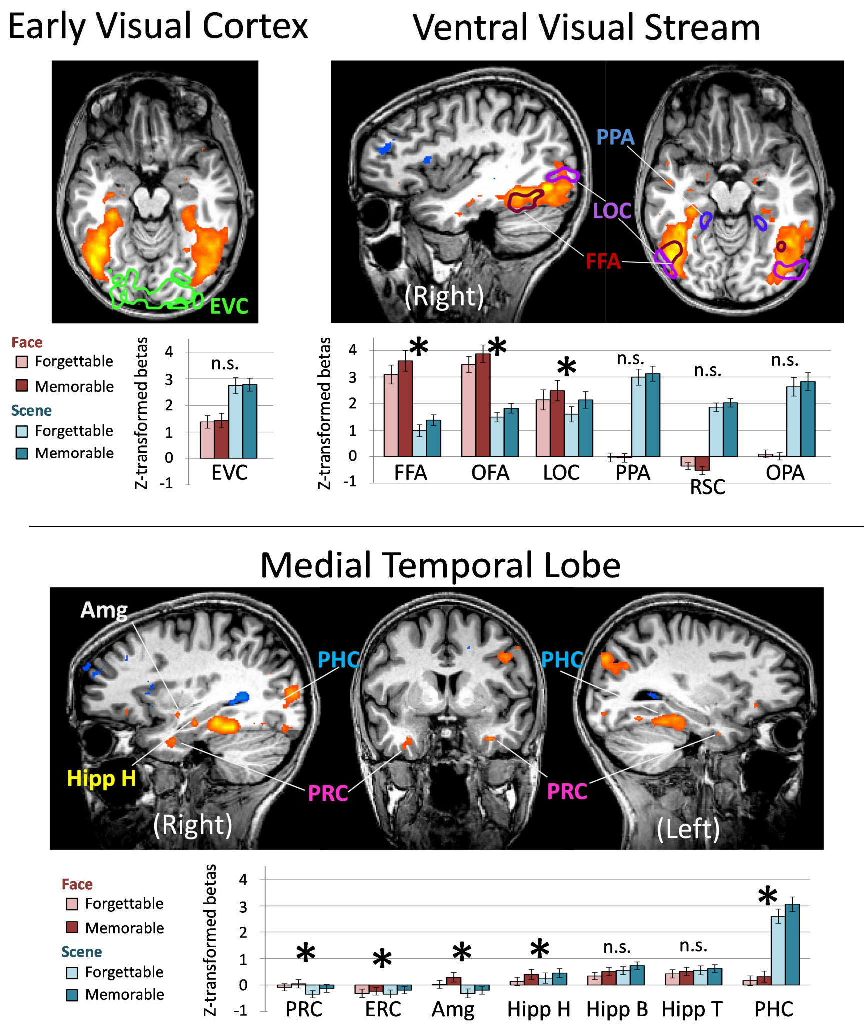

Fig. 1. fMRI results of memorable versus forgettable images (N = 16, random effects, FDR < 0.05). The outline of representative anatomical ROIs is also overlaid on the activation maps. A significant effect is found in the ventral visual stream, but not the early visual cortex. For the medial temporal lobe, there is a significant effect in the perirhinal cortex, parahippocampal cortex, amygdala, and hippocampus head. Bar graphs show ROI analyses of z-transformed betas for each condition (faces in red, scenes in blue, and the more memorable conditions in darker

colors). Asterisks indicate significance in a 3-way ANOVA (memorability, subsequent memory, stimulus type) of a memorability effect.

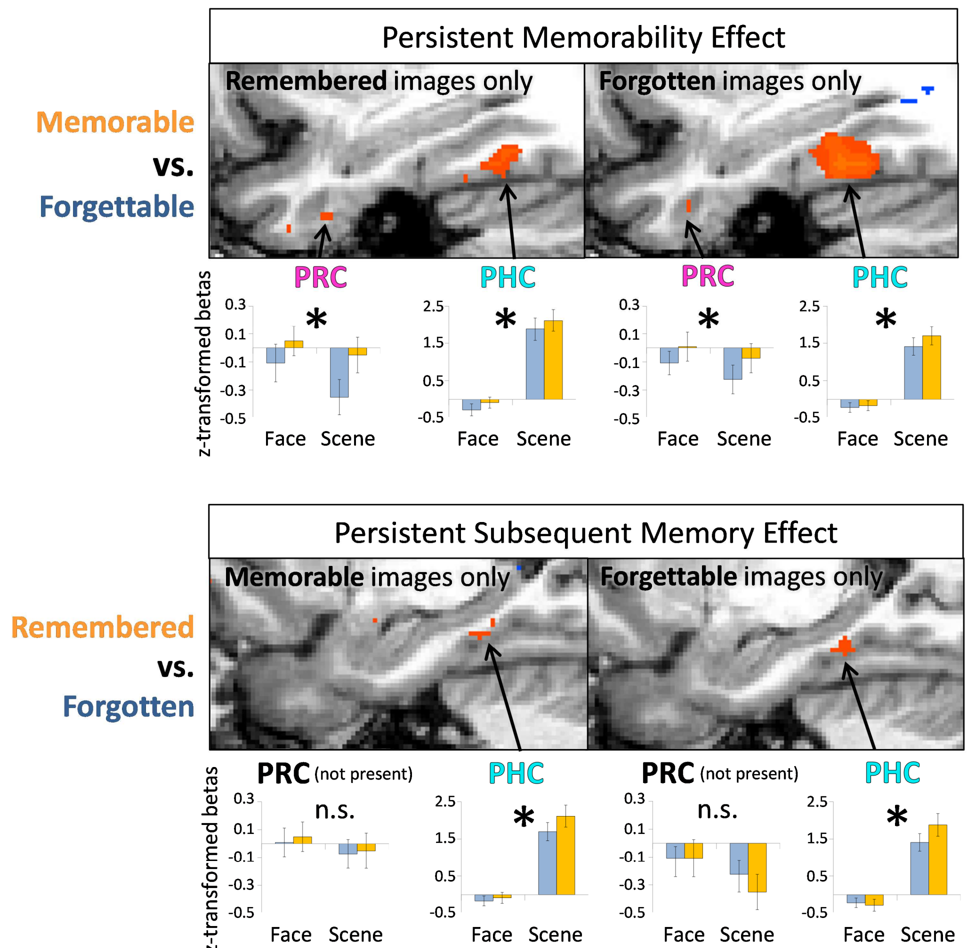

Fig. 2. The effect of one factor (memorability or memory encoding) when holding the other constant, in the fMRI study. (Top) The memorability contrast for when subsequent memory is held constant – looking at either only remembered images (left) or only forgotten images (right). Both share areas of activation, with significant memorability effects in both the perirhinal cortex (PRC) and parahippocampal cortex (PHC). Asterisks here show a significant main effect of memorability in a 2-way ANOVA (memorability × subsequent memory) in both regions. (Bottom) The subsequent memory contrast for when memorability is held constant – looking at either only memorable (left) or only forgettable images (right). Both share the PHC as the only region of significant activation for memory encoding. Asterisks show a significant main effect of memory encoding only in the PHC.

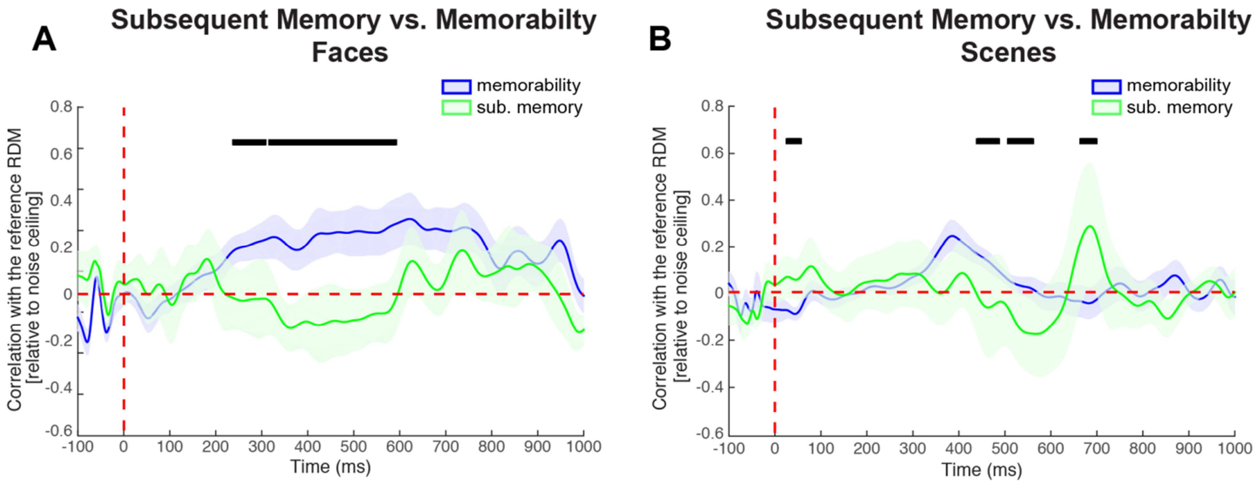

Fig. 3. MEG results of memorability and subsequent memory for faces and scenes. At each timepoint a MEG representational dissimilarity matrix (RDM) of the face (A) or scene (B) stimuli was correlated with a memorability model RDM (blue curve) and a subsequent memory model RDM (green curve). For both faces and scenes, memorability shows a significant effect, and memory encoding appears to occur later than memorability (particularly for scene stimuli). Shaded error bars are SEM across subjects. Horizontal black lines indicate time points with significant difference between the two curves (two-sided sign-rank test, FDR corrected at 0.01).