Predictive processing of upcoming scene views in immersive environments: evidence from continuous flash suppression

Anna Mynick (Anna.R.Mynick.gr@dartmouth.edu)
Psychological & Brain Sciences, Dartmouth College, 3 Maynard Street
Hanover, NH 03755 USA

Michael A. Cohen (mcohen@amherst.edu)
Psychology, Amherst College, 25 East Drive
Amherst, MA 01002 USA

Kala Goyal (Kala.Goyal@dartmouth.edu)
Psychological & Brain Sciences, Dartmouth College, 3 Maynard Street
Hanover, NH 03755 USA

Adithi Jayaraman (Adithi.Jayaraman.24@dartmouth.edu)
Psychological & Brain Sciences, Dartmouth College, 3 Maynard Street
Hanover, NH 03755 USA

Caroline E. Robertson (Caroline.E.Robertson@dartmouth.edu)
Psychological & Brain Sciences, Dartmouth College, 3 Maynard Street
Hanover, NH 03755 USA

Abstract:

As we move our eyes, heads, and bodies through the world, we operate with large sections of our environment out of view. How do we compensate for our limited field of view to interact effectively in 360° environments? Here, we test the idea that memory for an environment supplies predictions of upcoming scene views across head turns, modulating the speed at which scene views enter perceptual awareness. Participants (N=64) first learned a set of immersive, real-world scenes in head mounted virtual reality (VR). They then completed a perceptual test in which they head-turned toward target scene views presented in either spatially congruent or incongruent locations. Using continuous flash suppression (CFS) to initially mask the target views, we found that incongruent scene views enter perceptual awareness faster than congruent ones, suggesting that predictions of upcoming scene views are generated across head turns, and that prediction violations are prioritized in perceptual awareness.

Keywords: predictive processing; virtual reality; scene perception; memory

Introduction

Our visual environment extends beyond our field of view, perpetually leaving large sections of our surroundings out of sight. Yet, we operate effectively in 360° environments. How do we overcome our limited field of view to interact fluidly in immersive, real-world places?

Predictions represent one potential mechanism for facilitating fluid interaction in immersive environments. Past work indicates that the visual system generates predictions of upcoming visual information across small-scale visual actions – saccades (Duhamel et al., 1992; Herwig & Schneider, 2014; Osterbrink & Herwig, 2021). But, it remains unclear whether predictions are also generated for large-scale visual actions, such as head turns and body movements. Critically, in contrast to saccades, head turns and body movements draw entirely new visual information into view. As a result, predictions of upcoming information across these actions would likely require information from memory to generate.

In screen-based contexts, memory-based predictions of upcoming visual information have been shown to affect the speed at which stimuli enter perceptual awareness (Jiang et al., 2007; Pinto et al., 2015). Here, we test the idea that memory for a familiar immersive environment may be used to generate predictions of upcoming scene views across head turns (Mynick et al., 2023), modulating how quickly upcoming scene views enter perceptual awareness. In the present work, we measure the speed of perceptual breakthrough (Stein et al., 2011) for scene views drawn from familiar 360° scenes, which we present either in learned, spatially-congruent locations, or in novel, spatially incongruent locations. In brief, we find evidence that unexpected

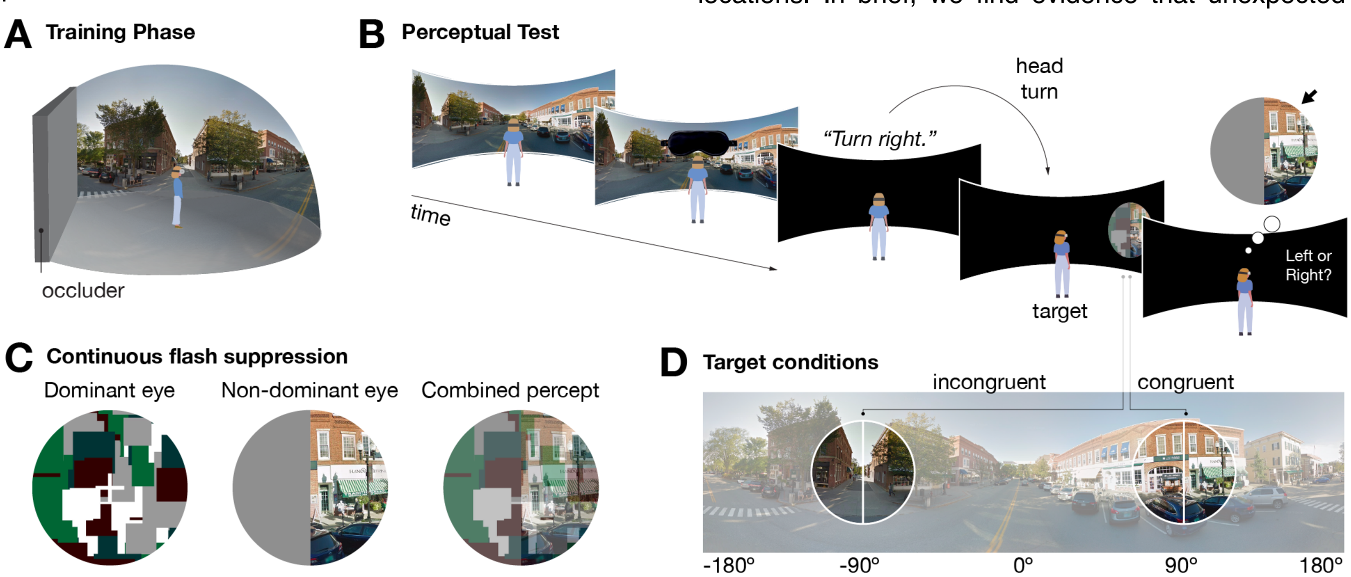

Figure 1: **A.** In the Training Phase, participants studied immersive scenes in head-mounted VR. Their instructions were simply, "look around like you normally would". **B.** On each trial of the Perceptual Test, participants were primed with a studied scene, which was then occluded. Participants next turned (left/right) toward a target scene view and indicated when they detected the target. They then indicated which side of the circle the target had occupied. **C.** Targets, presented to the non-dominant eye, were initially masked with a dynamic visual mask presented to the dominant eye. **D.** Targets were either presented in spatially congruent locations (e.g. left scene view shown on the participants' left) or spatially incongruent locations (e.g. left scene view shown on the participants' right).

scene views enter perceptual awareness faster than expected views, in keeping with a predictive processing account of visual processing ((Walsh et al., 2020), whereby expected input is suppressed to promote unexpected input for further processing.

Methods

The main experiment contained a Training Phase and a Perceptual Test. Both took place in head mounted virtual reality (VR). In the Training Phase, participants (N=64) learned a set of immersive, real-world, scenes (Fig. 1a) depicting locations around the local college campus. Participants were instructed to "look around like you normally would". In each trial of the subsequent Perceptual Test, a studied scene (prime) was presented, then fully occluded (Fig. 1b). Participants then turned 90º (left or right) toward a target image. The target image was presented to the non-dominant eye, and was initially masked by a dynamic Mondrian presented to the dominant eye (continuous flash suppression; CFS) (Fig. 1c). Participants' task was to detect the target, which either contained a spatially congruent view of the primed scene (e.g. the left view following a left head turn) or a spatially incongruent view (e.g. the right view following a left head turn) (Fig. 1d). To ensure true target detection, only half of the target was displayed (a semi-circle) and participants indicated which side of the circle the target was on (left/right). Crucially, the trial structure was balanced such that each target was presented both in a congruent location and in an incongruent location. To confirm that participants had strong recall for the studied scenes, participants also completed an Explicit Memory Test in VR at the end of the testing session. On each trial, a 45º section of a studied scene was presented directly in front of the participant, and participants indicated whether the image section had appeared on their left, right, or center during the Training Phase.

Results

Participants demonstrated high accuracy for target detection in the Perceptual Test (% accuracy: M=96.42, SD=6.81), indicating adherence to task instructions. We also observed high accuracy in the in the Explicit Memory Test (% accuracy: M=83.52, SD=16.03) (Fig. 2a), indicating strong recall of the spatial layout of the studied scenes from the Training Phase.

To determine whether memory for a studied scene modulates the speed of perceptual awareness for upcoming scene views, we compared response times (RTs) for detection of spatially congruent versus incongruent scene views during CFS. Participants detected incongruent scene views faster on average than congruent ones (p<.01) (Fig. 2b), suggesting that memory-based predictions of upcoming scene views affect the speed at which scene views enter perceptual awareness.

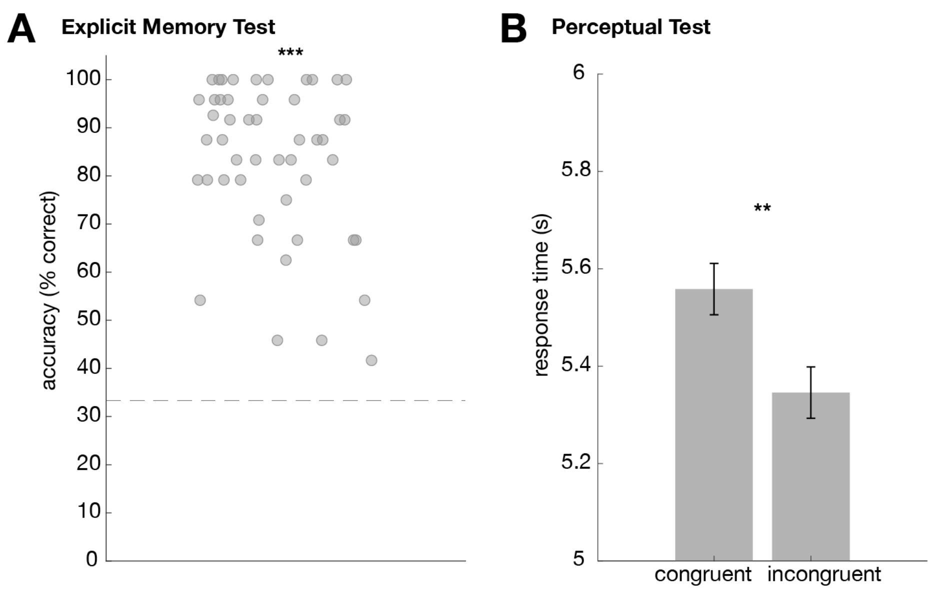

Figure 2: **A.** Explicit Memory Test accuracy. On each trial of the Explicit Memory Test, participants judged whether a 45º scene view had appeared on their left, right, or center during the Training Phase. Each point represents one participant. Dashed line represents chance accuracy. **B.** Mean response times for congruent and incongruent scene views. Bars represent mean across participants for each condition. Error bars represent within-subject standard error. ** p <=.01, *** p <= .001

Acknowledgments

This work was supported by a grant from the National Institutes of Health (R01MH130529) to CER.

References

Duhamel, J.-R., Colby, C. L., & Goldberg, M. E. (1992). The Updating of the Representation of Visual Space in Parietal Cortex by Intended Eye Movements. *Science*, *255*(5040), 90–92. https://doi.org/10.1126/science.1553535

Herwig, A., & Schneider, W. X. (2014). Predicting object features across saccades: Evidence from object recognition and visual search. *Journal of Experimental Psychology: General*, *143*(5), 1903–1922. https://doi.org/10.1037/a0036781

Jiang, Y., Costello, P., & He, S. (2007). Processing of Invisible Stimuli: Advantage of Upright Faces and Recognizable Words in Overcoming Interocular Suppression. *Psychological Science*, *18*(4), 349–355. https://doi.org/10.1111/j.1467-9280.2007.01902.x

Mynick, A., Steel, A., Jayaraman, A., Botch, T. L., Burrows, A., & Robertson, C. E. (2023). *Memory-based predictions prime perceptual judgments across head turns in immersive, real-world scenes*. https://doi.org/10.31234/osf.io/nkdjv

Osterbrink, C., & Herwig, A. (2021). Prediction of complex stimuli across saccades. *Journal of Vision*, *21*(2), 10. https://doi.org/10.1167/jov.21.2.10

Pinto, Y., Van Gaal, S., De Lange, F. P., Lamme, V. A. F., & Seth, A. K. (2015). Expectations accelerate entry of visual stimuli into awareness. *Journal of Vision*, *15*(8), 13. https://doi.org/10.1167/15.8.13

Stein, T., Hebart, M. N., & Sterzer, P. (2011). Breaking Continuous Flash Suppression: A New Measure of Unconscious Processing during Interocular Suppression? *Frontiers in Human Neuroscience*, *5*. https://doi.org/10.3389/fnhum.2011.00167

Walsh, K. S., McGovern, D. P., Clark, A., & O’Connell, R. G. (2020). Evaluating the neurophysiological evidence for predictive processing as a model of perception. *Annals of the New York Academy of Sciences*, *1464*(1), 242–268. https://doi.org/10.1111/nyas.14321