Decoding the contents of perceived and imagined navigation through virtual reality environments

Nicco Reggente[1], Joey K.Y. Essoe[1], Ivana Jevtic[1], Jesse Rissman[1,2]

[1]Dept. of Psychology, [2]Dept. of Psychiatry & Biobehavioral Sciences, University of California, Los Angeles

Introduction

- Navigating about one's environment is a multi-faceted effort requiring a faithful representation of the visuospatial layout and one's position and orientation within a space.
- Mental imagery is known to play a key role in successful navigation; whether planning out one's own route or providing directions to others, we must mentally simulate a trajectory through space and conjure up representations of pertinent contextual details.
- In this fMRI study, subjects first learned their way around three unique virtual reality environments, and later were scanned while viewing or imagining specific navigational routes through each of these worlds.
- We used a searchlight-based multi-voxel pattern analysis approach to examine how the neural representations of one's environment and navigational trajectory change as a function of whether the navigation is perceived or imagined.

fMRI Methods

MRI Acquisition (Siemens 3T Tim Trio)

CMRR Multiband EPI parameters:
TR = 1.5s; TE = 33ms; Flip Angle = 90º; # slices = 60
Voxel Size = 2.25 x 2.25 x 2.5 mm^3
Multiband acceleration factor = 3

Preprocessing

B0 unwarping with fieldmap and mag maps (y direction)
Signal Loss Threshold = 10%
High pass filter cutoff = 100s
Motion Correction accomplished with MCFLIRT

Spatial Alignment with ANTS

The middle volume of each run (8 in total) was used to create a subject-specific BOLD template space to which all volumes were aligned.

Each subject's hi-res anatomical image (MPRAGE) was aligned to the BOLD template, and these in turn were warped to MNI space using diffeomorphic image registration (SyN).

All analyses were conducted in subject-specific BOLD template space, and the resulting stat-maps were warped to MNI space to facilitate group level analyses.

Classification

Whole-Brain Searchlight with a 5-voxel radius
nu-SVC multi-class support vector machine (libSVM)
Polynomial kernel; Penalty =1
Leave-One-Run-Out Cross Validation
Patterns for each trial were created by averaging TRs 4:7 after the trial onset.

Classification was concerned with decoding:
1) Which World a subject was viewing / imagining (Lagoon, Toon, or Ruin).
2) Which Direction (Clockwise, Counter-Clockwise) a subject was viewing / imagining.

Classification architecture took on 3 forms:
1. Training and Testing exclusively with Video Viewing or Mental Imagery
2. Training on Video Viewing and Testing on Mental Imagery
3. Training on Video Viewing and Mental Imagery and Testing on Mental Imagery

(NOTE: in all scenarios, testing set data were always from an independent run)

Resultant Searchlight maps were submitted to a voxel-wise one-sample t-test to assess if the distribution of accuracy values across subjects was unlikely to have come from a normal distribution centered around chance (33% for Which World; 50% for Which Direction). Across-subject mean accuracy maps were thresholded to display only voxels with a t-test p-value <.05 (uncorrected) and a mean accuracy > 39% for Which World and accuracy > 54% for Which Direction.

Design & Stimuli

Day 1 (Lab)

- After a VR "orientation session" subjects were teleported, in a random order, to one of three virtual environments, each populated with 8 distinctive landmarks at the cardinal locations along the periphery:

Lagoon World

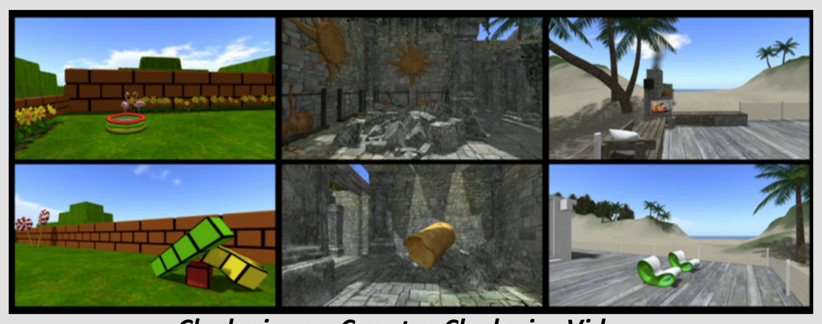
Toon World

Ruin World

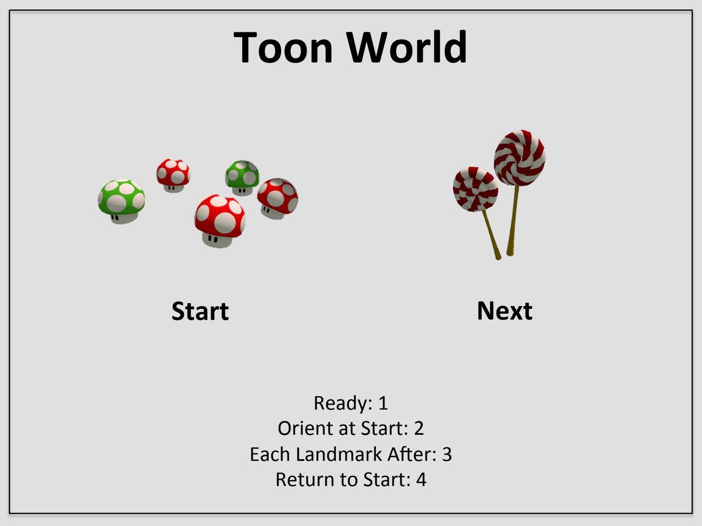
- Subjects (N=15) were given five minutes to collect 20 tokens scattered throughout the environment, using any remaining time to freely explore. Subjects were explicitly instructed to take note of each environment and relevant landmarks. Subjects were then cycled back through each environment and given three minutes to collect the same tokens as quickly as possible.
- Subjects also participated in "guided-navigation tasks" where they were instructed to move about each environment so as to pass by a designated sequence of major landmarks located at cardinal directions around the periphery.

Day 2 (Lab & Scanner)

- Subjects returned to the lab where they completed more landmark-to-landmark guided-navigation tasks.
- In the scanner, subjects viewed a series of 30s video clips. Each video started at one of the four corner landmarks in a world and followed a trajectory around the perimeter of the environment before arriving back at the starting landmark.
- For each starting landmark, the video traversed the environment in either a clockwise or counter-clockwise direction. This resulted in 8 videos per world (24 total, broken up into 4 runs of 6 trials each).
- Following the video viewing runs, subjects performed a new task in which they were to mentally imagine themselves moving about the environments.
- On each trial, subjects were instructed to mentally orient themselves at a specified landmark in a specified world. They were also told which landmark they should pass next as they circumnavigated the environment back to the starting landmark (half of the trajectories involved clockwise movement and the other half counter-clockwise). This resulted in 8 mental imagery trials per world (24 total, broken up into 4 runs of 6 trials each).
- An active baseline task was used between trials. Subjects had to respond whether the multiple of two numbers on the screen was odd or even.



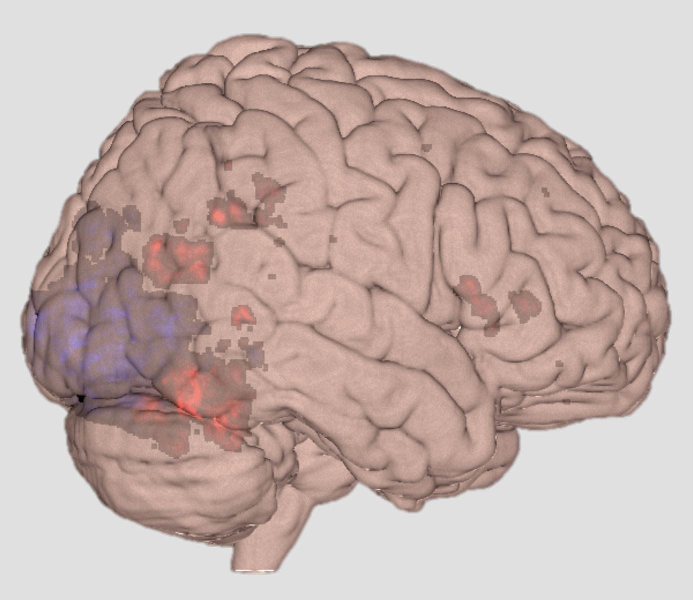
Clockwise vs. Counter-Clockwise Videos



Example Mental Imagery Instruction Screen

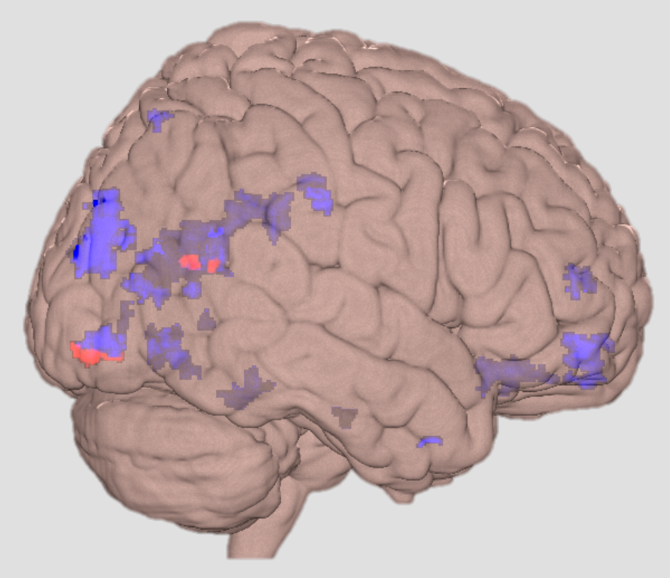
Searchlight Classification Results

Which World?
Video Viewing Mental Imagery



Bilateral: Temporal-Occipital-Fusiform, Lingual Gyrus, Lateral Occipital Complex (LOC), Occipital Pole
Right: Parahippocampal Gyrus

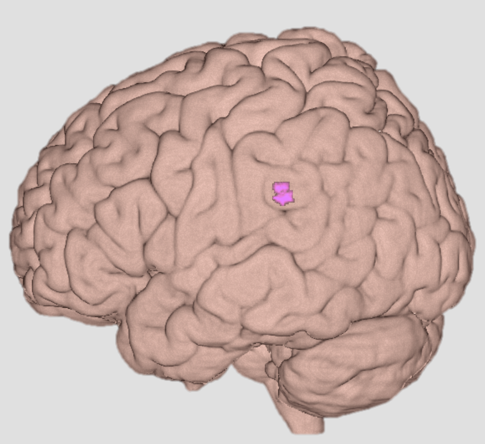
Which Direction?
Video Viewing Mental Imagery



Bilateral: Lingual Gyrus, Cuneus, Inferior Temporal, Retrosplenial, Posterior Cingulate
Right: Paracingulate, LOC, Orbtifrontal, Parietal
Left: Precuneus, Hippocampus

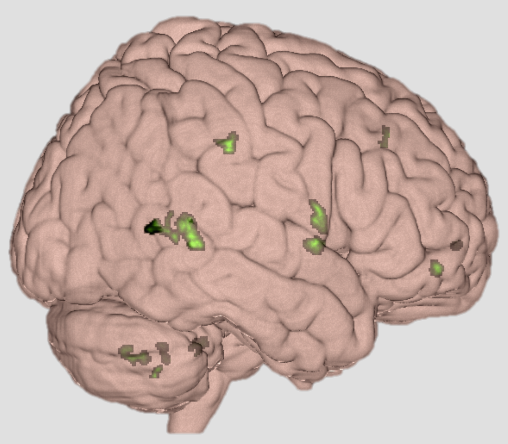
Train on Video Viewing
Test on Mental Imagery

Which World?



Left: Parietal Operculum

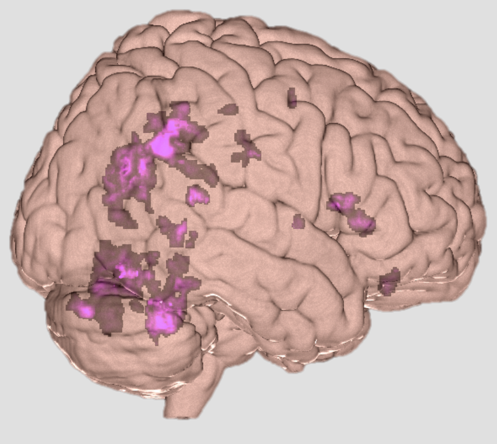
Which Direction?



Bilateral: Posterior Middle Temporal Gyrus
Right: Frontal Pole (BA11), Planum Polare, Orbital (BA47), Central Operculum, Cerebellum (Crus I/ VIIIa)
Left: Hippocampus

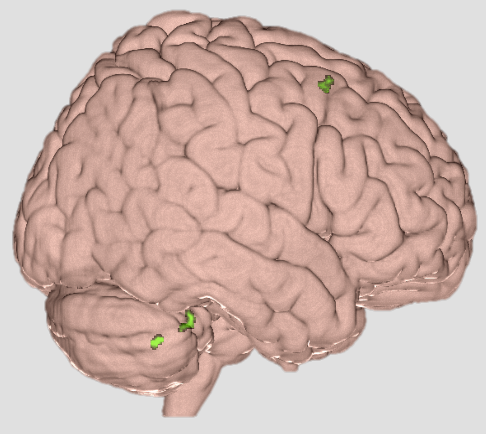
Train on Video Viewing & Mental Imagery
Test on Mental Imagery

Which World?



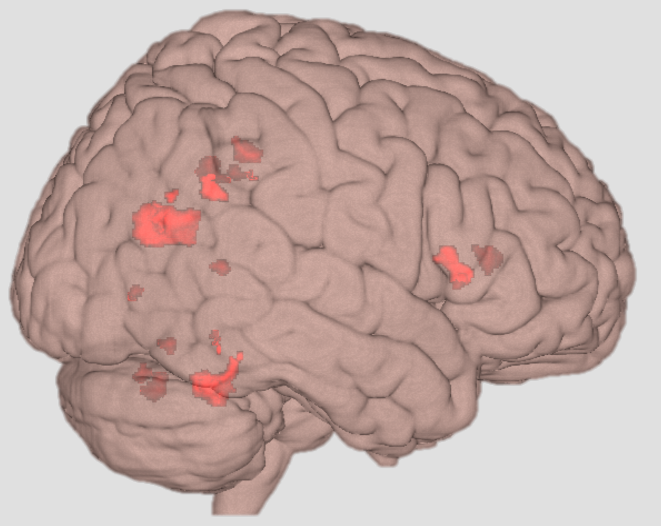
Bilateral: Precuneus, Angular Gyri, LOC, Superior Occipital Gyri
Right: Lingual Gyrus, Posterior Hippocampus, Fronto-Operculum, Cerebellum (V)
Left: Frontal Cortex (BA11) , Putamen, Occipirotemporal, Cerebellum (VI)

Which Direction?



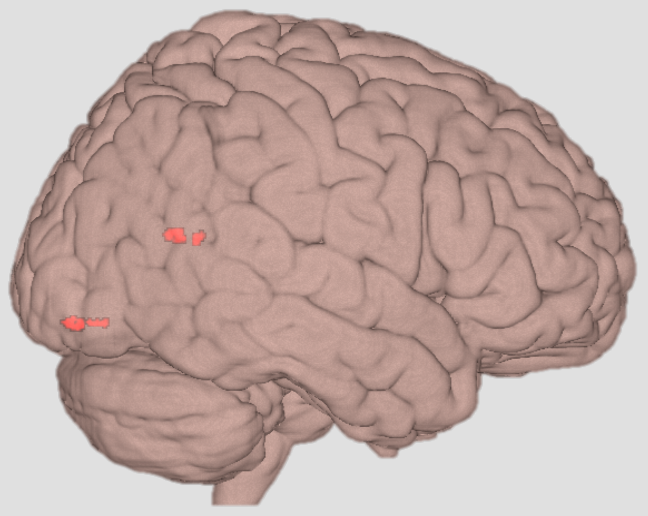
Bilateral: Superior Frontal Gyri
Right: Cerebellum (Crus II)
Left: Pontocerebellar area

Which World? [Unique to Imagery]



Bilateral: Cerebellar lobule VI, Lingual Gyrus
Right: Superior Parietal Lobule, Insular Cortex
Left: Angular Gyrus, Retrosplenial Cortex, Caudate

Which Direction? [Unique to Imagery]



Right: Angular Gyrus, Inferior LOC

Conclusions

- These preliminary results showcase that neural representations of environment identity—evoked during both viewing and mental imagining distinct virtual worlds—can be decoded from fMRI activity patterns across a number of cortical regions.
- Furthermore, cortical activity patterns appear to be sensitive to one's navigational trajectory (clockwise vs. counter-clockwise), during both perception and mental imagery, irrespective of the specific environment.
- While visual cortices were most informative for decoding analyses conducted on the video viewing data, a largely non-overlapping set of regions were diagnostic for mental imagery decoding (including higher level visual association areas and several frontoparietal regions). This likely indicates that imagery-trained classifiers learn neural representations tied to more abstract features of the respective worlds.
- Classifiers trained on viewing data struggled to decode Which World but showed some success at decoding Which Direction. Regions showing such transfer include those known to be involved in sequence-based navigation (Cerebellum Crus I and the Left Hippocampus)[1], and likely recapitulate the same direction-selective neural patterns evoked during initial processing of the video clips.
- Interestingly, a model trained on data from both perception and imagery trials showed a far better ability to transfer its learned representations of environment identity to held-out imagery trials. By pooling across trial types during training, the classifier was likely forced to acquire more generalized neural representations of navigation about the three unique worlds.

This study was funded by:
1) Defense Advanced Research Projects Agency award (D13AP00057) to Dr. Jesse Rissman
2) National Science Foundation Graduate Research Fellowships awarded to both Nicco Reggente and Joey Essoe

References:
1) Iglói, K., et al. (2014). Interaction between hippocampus and cerebellum Crus I in sequence-based but not place-based navigation. *Cerebral Cortex*, bhu132.