Neural Evidence for Visual Simulation of Object Movement in the Human Visual Cortex

Sixuan Chen (sixuan_chen1@Brown.edu)
Brown University, Providence, RI, 02912, USA

Alekh Karkada Ashok (alekh_karkada_ashok@Brown.edu)
Brown University, Providence, RI, 02912, USA

Anna Kimata (anna_kimata@Brown.edu)
Brown University, Providence, RI, 02912, USA

David Sheinberg (david_sheinberg@Brown.edu)
Brown University, Providence, RI, 02912, USA

Wael Asaad (wael_asaad@Brown.edu)
Brown University, Providence, RI, 02912, USA

Thomas Serre (thomas_serre@Brown.edu)
Brown University, Providence, RI, 02912, USA

Abstract

The ability to visually simulating sensory, motor, and cognitive representations and imagine potential future actions is crucial for our daily life. However, the neural basis underlying the visual simulation of object movement remains partially understood. Past work using human fMRI decoding has provided some evidence for the reactivation of motion-responsive areas of the visual cortex during mental simulation. Here, we tested a human patient with implanted stereoelectroencephalography (sEEG) electrodes while they played the 'Planko' game, where they were asked to predict the trajectory of a ball falling through a field of planks. The patient learned to play the game significantly above chance 68.9%. We use time-frequency analysis to study the similarity between the beta event of mental simulation and video watching. We also used the deep learning method to find a joint embedding between observed and simulated events.

Keywords: mental simulation; human electrophysiology

Introduction

To perform physical reasoning and complex decision-making, we often must use internal models of the world to predict future versions of events. This ability, called "visual simulation", involves picturing upcoming events many steps into the future using understanding of physics. This is in contrast to "visual imagery," which often involves reactivating memories of static precepts. Although previous research has gathered evidence for visual simulation in primates for physical reasoning tasks Ahuja & Sheinberg (2019); Battaglia et al. (2013); Rajalingham et al. (2022), the neural basis underlying the visual simulation of object movement remains poorly understood.

One intriguing hypothesis is that visual imagery "reuses" the same circuits and machinery used for visual perception. In their seminal work, Kosslyn et al. (1995) found that the primary visual cortex is activated when subjects close their eyes and imagine objects in their mind's eye. Surprisingly, Reddy et al. (2010) found that the object category of imagined visual objects could even be decoded from patterns of fMRI activity. Moreover, the same decoder trained to decode category information during perception was able to generalize during imagery, indicating that the visual representation during imagery and perception of objects were remarkably similar.

However, this hypothesis remains poorly studied for tasks involving visual simulation. To test primates on visual simulation, Ahuja & Sheinberg (2019) introduced the "*Planko*" task, which involves judging the final position of a ball (which of two baskets – left or right) as it falls through randomly placed planks (Fig. 1). Previous work on the *Planko* task has found behavioral evidence that both monkeys Ahuja et al. (2024) and humans Ahuja & Sheinberg (2019) resort to a simulation-like strategy to solve this task. Past workAhuja et al. (2021) using human fMRI recordings of subjects performing this task has found reactivation of the same motion-responsive areas
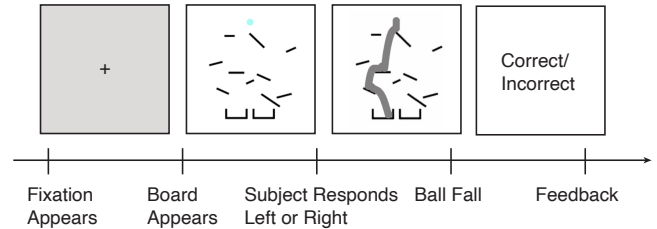

Figure 1: An outline of one complete trial of Planko

of the visual cortex that are active during perception of ball movement through the planks. Here, we tested a human patient with implanted stereoelectroencephalography (sEEG) electrodes while they played *Planko*, to quantify the extent to which the neural representations during visual simulation matched the ones during motion perception.

Method

Participant

A female participant, undergoing epilepsy monitoring at Rhode Island Hospital, participated (IRB protocol #217972.). Eight depth electrodes (87 contacts in total) were implanted in multiple locations, including occipital, lateral, and parietal areas.

Task

Each trial featured a board with a single ball at the top, a series of semi-randomly positioned planks in the middle, and two baskets at the bottom. The sequence of events for a trial included a 1000ms fixation dot followed by the board display, with the positions of the ball and baskets remaining constant throughout. Participants were instructed to predict which basket the ball would land in when dropped from its central position. Responses were recorded by pressing one of two buttons corresponding to the baskets on the screen. The ball dropped once a response was made, providing visual feedback (9000ms). Participants were instructed to track the falling ball visually until it landed in the selected basket.

The task comprised 3 practice trials followed by 70 experimental trials. During each trial, sEEG data were recorded across three main phases: fixation, mental simulation, and passive video viewing.

Defining beta events and features

Recent research in both humans and non-human primates has revealed that neural activity at the trial level occurs in burst-like events are believed to play a crucial role in supporting higher-order cognitive functions (Jones, 2016; Shin et al., 2017). One of them is the beta frequency activity (15–29 Hz) events. They are usually associated with cognitive demands or shifts in spatial attention Shin et al. (2017); Sacchet et al. (2015). Using the toolbox provided byMcKeon et al. (2023), beta events were defined as local maxima in
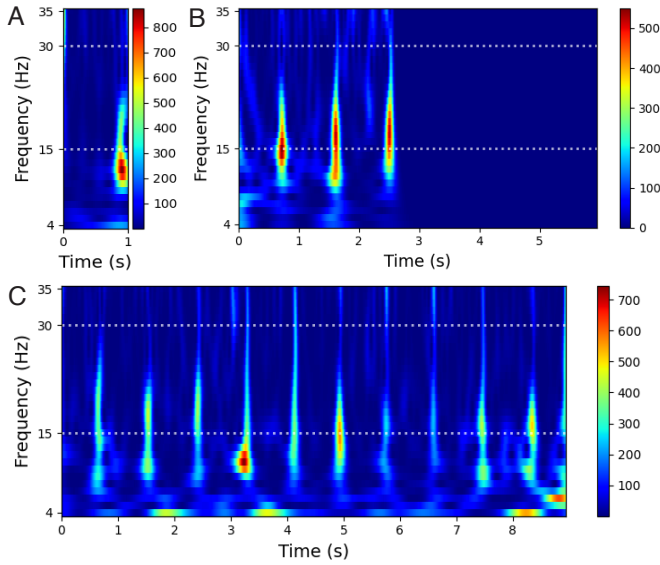
Figure 2: A: Time-frequency representation (TFR) of fixation, B: TFR of simulation, C: TFR of passive viewing

Figure 3: Contrastive learning model architecture.

the time-frequency representation (TFR) matrix for each frequency value within the beta band, with power exceeding a cutoff of 6 times the median power Shin et al. (2017).(Fig. 2)

Spectral analysis

The sEEG preprocessing was done following revised pipeline of EEGlab toolbox (https://eeglab.org/tutorials/, retrieved Feb 19, 2024). Spectral data were computed for every electrode in the range 4–35 Hz by convolving the neural signals with a complex Morlet wavelet Qin & Shen (2000).

Learning a neural embedding with contrastive learning

We used a contrastive loss (Défossez et al. (2023)) to learn a neural embedding in order to associate a mental simulation sEEG segment X_k from trial k with its corresponding sEEG segment during passive viewing of the ball dropping, Y_k. To calculate this loss, we first selected one positive example Y_k, which is the true match for X_k, and then sample N-1 negative examples, $Y_{1,..N-1}$. These negative samples are sEEG segments during viewing of ball drop of different trials. By minimizing this loss, the model learns to match sEEG segments during simulation and passive video watching from the same trial while making sEEG segments from different trials as different as possible. (Fig. 3)

Result

Beta event

Beta event with peak frequency at 15-16Hz showed increased events during mental simulation ($M = 0.209$, $SD = 0.066$) compared to fixation ($M = 0.154$, $SD = 0.112$, $t(72) = -3.52$, $p < 0.001$), with similar levels to passive viewing ($M = 0.19$,
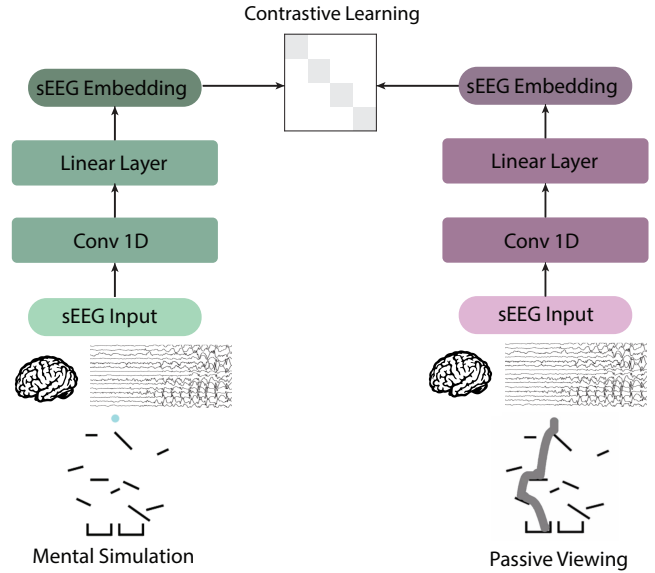
$SD = 0.042$, $t(72) = 2.01$, $p = 0.0485$), suggesting engagement of similar neural processes.

Contrastive learning

We performed five times five-fold cross-validation confirmed the model's efficacy in distinguishing between correct and random data sequences, validating our approach. Our top 1 validation accuracy using time-domain data ($M = 0.125$, $SD = 0.091$) is significantly higher than randomly shuffled data ($M = 0.062$, $SD = 0.058$, $p < 0.001$) .

Discussion

In this study, we presented some early evidence for the "reuse" of machinery involved in motion perception during visual simulation. This paves the way for future work involving larger participant cohorts and a broader range of brain regions where ball positions could be decoded during visual simulation using a decoder trained during passive ball drop viewing. This will allow for understanding the nature of the time course of such simulations as well the neural circuits involved.

References

Ahuja, A., Desrochers, T. M., & Sheinberg, D. L. (2021). A role for visual areas in physics simulations. *Cognitive neuropsychology*, *38*(7-8), 425–439.

Ahuja, A., & Sheinberg, D. L. (2019). Behavioral and oculomotor evidence for visual simulation of object movement. *Journal of vision*, *19*(6), 13–13.

Ahuja, A., Yusif Rodriguez, N., Ashok, A. K., Serre, T., Desrochers, T., & Sheinberg, D. (2024). Monkeys engage in visual simulation to solve complex problems. *bioRxiv*, 2024–02.

Battaglia, P. W., Hamrick, J. B., & Tenenbaum, J. B. (2013). Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences*, *110*(45), 18327–18332.

Défossez, A., Caucheteux, C., Rapin, J., Kabeli, O., & King, J.-R. (2023). Decoding speech perception from non-invasive brain recordings. *Nature Machine Intelligence*, *5*(10), 1097–1107.

Jones, S. R. (2016). When brain rhythms aren't 'rhythmic': implication for their mechanisms and meaning. *Current opinion in neurobiology*, *40*, 72–80.

Kosslyn, S. M., Thompson, W. L., Klm, I. J., & Alpert, N. M. (1995). Topographical representations of mental images in primary visual cortex. *Nature*, *378*(6556), 496–498.

McKeon, S. D., Calabro, F., Thorpe, R. V., de la Fuente, A., Foran, W., Parr, A. C., . . . Luna, B. (2023). Age-related differences in transient gamma band activity during working memory maintenance through adolescence. *NeuroImage*, *274*, 120112.

Qin, X., & Shen, L. (2000). Wavelet transform and its application in spectral analysis. *Guang pu xue yu Guang pu fen xi= Guang pu*, *20*(6), 892–897.

Rajalingham, R., Piccato, A., & Jazayeri, M. (2022). Recurrent neural networks with explicit representation of dynamic latent variables can mimic behavioral patterns in a physical inference task. *Nature Communications*, *13*(1), 5865.

Reddy, L., Tsuchiya, N., & Serre, T. (2010). Reading the mind's eye: decoding category information during mental imagery. *Neuroimage*, *50*(2), 818–825.

Sacchet, M. D., LaPlante, R. A., Wan, Q., Pritchett, D. L., Lee, A. K., Hämäläinen, M., . . . Jones, S. R. (2015). Attention drives synchronization of alpha and beta rhythms between right inferior frontal and primary sensory neocortex. *Journal of neuroscience*, *35*(5), 2074–2082.

Shin, H., Law, R., Tsutsui, S., Moore, C. I., & Jones, S. R. (2017). The rate of transient beta frequency events predicts behavior across tasks and species. *Elife*, *6*, e29086.