

# SABER: Spatial Attention, Brain, Extended Reality

Tom Bullock<sup>1,2,3\*</sup> Emily Machniak<sup>1,2,4</sup> You-Jin Kim<sup>5</sup> Radha Kumaran<sup>6</sup> Justin Kasowski<sup>1</sup>  
 Apurv Varshney<sup>6</sup> Julia Ram<sup>1</sup> Melissa M. Hernandez<sup>1</sup> Stina Johansson<sup>1</sup> Neil M. Dundon<sup>1,3</sup>  
 Tobias Höllerer<sup>2,6†</sup> Barry Giesbrecht<sup>1,2‡</sup>

<sup>1</sup> Department of Psychological and Brain Sciences, University of California, Santa Barbara, Santa Barbara, CA, United States.

<sup>2</sup> Institute for Collaborative Biotechnologies, University of California, Santa Barbara, Santa Barbara, CA, United States.

<sup>3</sup> Department of Social and Psychological Sciences, University of Huddersfield, Huddersfield, United Kingdom

<sup>4</sup> Department of Human Systems Engineering, Arizona State University, Mesa, AZ, United States.

<sup>5</sup> School of Computing, University of Nebraska-Lincoln, Nebraska, United States.

<sup>6</sup> Department of Computer Science, University of California, Santa Barbara, Santa Barbara, CA, United States.



Figure 1: A participant is shown superimposed into a virtual environment, standing on a platform wielding a pair of virtual sabers anchored to HTC Vive Controllers. They are required to strike incoming spheres using either saber while their brain activity is recorded at the scalp using electroencephalography (EEG).

## ABSTRACT

Tracking moving objects is a critical skill for many everyday tasks, such as crossing a busy street, driving a car or catching a ball. Attention is a key cognitive function that supports object tracking; however, our understanding of the brain mechanisms that support attention is almost exclusively based on evidence from tasks that present stable objects at fixed locations. Accounts of multiple object tracking are also limited because they are largely based on behavioral data alone and involve tracking objects in a 2D plane. Consequently, the neural mechanisms that enable moment-by-moment tracking of goal-relevant objects remain poorly understood. To address this knowledge gap, we developed SABER (Spatial Attention, Brain, Extended Reality), a new framework for studying the behavioral and neural dynamics of attention to objects moving in 3D. Participants ( $n=32$ ) completed variants of a task inspired by the popular virtual reality (VR) game *Beat Saber*, where they used virtual sabers to strike stationary and moving color-defined target spheres while we recorded electroencephalography (EEG). We first established that standard univariate EEG metrics which are typically used to study spatial attention to static objects presented on 2D screens, can generalize effectively to an immersive VR context involving both static and dynamic 3D stimuli. We then used a computational modeling approach to reconstruct moment-by-moment

attention to the locations of stationary and moving objects from oscillatory brain activity, demonstrating the feasibility of precisely tracking attention in a 3D space. These results validate SABER, and provide a foundation for future research that is critical not only for understanding how attention works in the physical world, but is also directly relevant to the development of better VR applications. The insights gained here can potentially inform the design of more intuitive interfaces, effective training simulations, and immersive experiences optimized for the human attention system.

**Index Terms:** Virtual Reality, VR, Attention, EEG, Object Tracking, Computational Models, Brain Oscillations.

## 1 INTRODUCTION

Attending to moving objects, continuously updating their positions, and predicting their trajectories are critical for goal-directed actions in everyday life. In the physical world, simply walking down a busy street requires us to navigate amongst multiple moving objects, such as other pedestrians, cyclists and vehicles. This challenge extends to the digital world. When playing a VR first-person shooter, one must track and target an enemy player while simultaneously monitoring other players on the battlefield. Similarly, in a rhythm game such as *Beat Saber*, players must track multiple incoming targets and anticipate their positions to prepare and execute precise actions. The human capacity to construct stable mental representations that support the tracking of one, or even multiple, dynamic objects is well-documented [53, 38, 13], while impaired object tracking is often associated with various brain disorders and injuries [2, 35, 41]. Given the fundamental importance of object tracking in both our physical and digital worlds, it is critical that we understand the brain mechanisms that enable this function.

\*e-mail: [tombullock@ucsb.edu](mailto:tombullock@ucsb.edu)

†email: [thollerer@ucsb.edu](mailto:thollerer@ucsb.edu)

‡email: [giesbrecht@ucsb.edu](mailto:giesbrecht@ucsb.edu)

Evidence from cognitive psychology and neuroscience suggests that our ability to track objects is heavily dependent on visual attention and working memory (WM) [4, 21, 38]. These processes support selective sampling, or prioritizing, of behaviorally relevant sensory evidence from the environment, integrating this evidence with expectations about the object, maintaining the representation over time, and then potentially executing a behavioral action [21]. Unfortunately, current accounts of visual attention and working memory are almost exclusively based on evidence from tasks using stable objects at fixed locations [16, 24, 46, 58, 70, 3] and there is limited evidence about the moment-by-moment neural dynamics that support the tracking of dynamic objects. Neural explanations from multiple object tracking (MOT) studies are also limited because these studies are largely based on behavioral data alone and involve tracking objects in a 2D plane [13, 38, 53, 52]. Thus, despite the fundamental importance of tracking dynamic objects in the environment, and the disruption of this capacity by disease and injury, our understanding of the perceptual, cognitive and action-related mechanisms underlying these processes is limited.

To address this knowledge gap, we present a new approach, inspired by *Beat Saber*, which combines VR and EEG to study the neural mechanisms of spatial attention to dynamic objects in an immersive 3D environment. We refer to this approach as SABER (Spatial Attention, Brain, Extended Reality). We validate this approach with results from an experiment where a sample of healthy young adults ( $n=32$ ) used virtual sabers to strike color-defined target spheres while we recorded electrical brain activity at 64 scalp locations using EEG. Targets appeared in one of two ways: they either emerged directly in front of the participants or moved toward them from a distance. These targets were presented either alone or surrounded by an array of distractors (different colored spheres that participants were instructed to ignore).

We first demonstrate that standard univariate EEG methods—typically used to study spatial shifts in attention to static objects presented on 2D screens—can also be successfully applied to analyze brain responses to both static and dynamic objects in immersive 3D VR. We then apply a multivariate computational method to show that both static and moving object locations and trajectories can be reconstructed from patterns of alpha brain activity recorded at the scalp. Together, these results validate SABER, opening up exciting future opportunities to test hypotheses about how the brain supports attention to dynamic objects. The potential insights that can be gained from this new approach to studying attention are critical for two key reasons. First, they provide insights into how attention operates in dynamic scenarios in the physical world. Second, they will aid the development of better virtual and mixed-reality applications, informing the design of more intuitive interfaces, effective training simulations, and immersive experiences that are optimized for the human attention system. Ultimately, the insights gained from this work have broad interdisciplinary relevance, with implications for cognitive neuroscience and psychology, human factors, and the development of extended reality (XR) applications and brain-computer interfaces (BCI).

## 2 RELATED WORK

### 2.1 Multiple Object Tracking

In both the physical and virtual worlds, achieving our goals and avoiding potential dangers often requires that we attend to and track moving objects. This ability relies heavily on spatial attention and WM [4, 21, 38], and the majority of the evidence implicating visual attention and visual working memory comes from studies using the MOT task [4, 21, 38]. In this simple computer task, participants are typically shown multiple objects and cued to track one or more objects as they move on a computer screen. The MOT literature is large and has provided important insights into attention and WM. Spatial attention supports the initial selection of the targets

to be tracked, while WM supports the sustained representation of behaviorally relevant objects [1, 21, 49, 71]. However, accounts of object tracking based on MOT are limited primarily because they largely involve participants tracking objects on a 2D plane (but see [61]) and are typically based on behavioral data alone, with relatively few studies also recording brain data ([21]). Indeed, while the MOT task provides an excellent methodology for studying the behavioral dynamics of object tracking, the task by its very nature is not well-suited for studying the neural dynamics of tracking a specific object.

Contemporary theories of attention and WM provide limited insight on the tracking of dynamic objects because they are primarily based on static objects presented at fixed locations [16, 24, 46, 58, 70]. Consequently, there is a significant gap in our understanding of how the brain handles dynamic stimuli. Indeed, simplifying attention to the perception of still, 2D images overlooks that attention must operate dynamically over time to prioritize relevant information and use it when most needed [19].

SABER is specifically designed with the goal of analyzing both the behavioral and neural dynamics of moment-by-moment object tracking. This will allow for the moment-by-moment modeling of attention to a specific goal-relevant 3D target object that is either tracked within the environment in isolation or amongst multiple other goal-irrelevant objects. The hope is that this study and future work based around this approach will provide insight into how known mechanisms from static object processing extend to contexts involving dynamic interactions with 3D moving objects. This work will inform contemporary models of spatial attention and WM.

### 2.2 EEG Approaches to Studying Spatial Attention

EEG is a method for noninvasive measurement of brain activity from electrodes placed on the scalp [5]. The measured signal captures post-synaptic electrical activity from large populations of neurons that are synchronously activated [50]. The acquired signal has a high temporal resolution and is a direct measure of neural activity. In cognitive neuroscience, there are two main features of the EEG signal that have been used to inform the fundamental understanding of sensory, perceptual, and cognitive phenomena: time-varying amplitude and oscillatory power.

*Time-varying amplitude.* To extract the voltage associated with specific task events, EEG studies have used the Event-Related Potential (ERP) technique [45]. This technique involves averaging EEG activity that has been time-locked to an event, such as a shape appearing on a screen. By averaging across many repetitions of the event (i.e., trials), random background activity in the EEG signal is canceled out, and what remains is a time-varying waveform that represents the neural response to the specific event. Additional analytical steps can then be taken to control for stimulus differences and provide more anatomical specificity by leveraging the known systematic mapping of the visual field in cortical visual areas [67]. In ERP studies of visual attention, one common step leverages the contralateral organization of visual cortex [45] by comparing brain activity over electrodes contralateral (i.e., opposite side) to the location of an attended stimulus versus activity in electrodes ipsilateral (i.e., same side). The difference between these two signals is referred to as *lateralization* and has been leveraged to gain insight into the neural mechanisms that underpin spatial attention. One particularly useful ERP component that is based on this approach is the N2pc, which is a consistently larger negative deflection of the waveform between around 200-350 ms at posterior electrode sites contralateral to an attended stimulus and is thought to reflect the enhancement of the target signal from distracting information [22, 47]. While the ERP technique provides valuable insight into the mechanisms of spatial attention, a key limitation is that it is typically only applied to stimuli that have an abrupt onset. This makes it unsuitable for studying the moment-by-moment fluctua-

tions of attention over time, as it cannot be used with gradual or continuous stimuli.

**Oscillatory power.** The EEG signal can also be converted into the frequency domain to quantify oscillatory activity and this approach is also well-suited to capturing the neural dynamics of spatial attention. In this approach, the EEG signal is converted into the frequency domain using various transformation methods, including Fourier, Hilbert, or Morelet wavelet transforms [15]. It is well-established that activity in the alpha frequency band (~8-13 Hz) tracks both the locus and timing of covert spatial attention [63, 55]. Specifically, alpha builds up at posterior sites ipsilateral to the attended location, which has led to the view that it reflects gating/suppression of irrelevant visual information [27], although it could also reflect signal enhancement [25]. Examining lateralized alpha responses provides insight into moment-by-moment fluctuations in attention over time, which means it can be applied to gradual or continuous stimuli. However, lateralization is still a relatively crude metric in the sense that it limits our understanding of spatial attention to "attend left vs. attend right" and does not provide insight into the quality of the representation.

Amplitude and oscillatory features have traditionally been analyzed using univariate statistical techniques, but more recently, researchers have been using multivariate analytic techniques to investigate patterns of brain activity. One technique used in cognitive neuroscience is the Inverted Encoding Model (IEM). Originally developed for fMRI to recover feature-selective responses in visual cortex [6], this method assumes that the aggregate signal measured at the scalp (EEG) reflects the summed activity of underlying neural populations, each governed by specific tuning properties (e.g., spatial receptive fields). Specifically, the goal of IEM is to determine the extent to which the data conform to an *a priori* formulation of what the response to the stimulus conditions should look like. This method allows researchers to visualize the "quality" or "fidelity" of the population-level representation of the stimulus directly. In the context of attention, IEMs have been used to demonstrate that topographic patterns of alpha-band oscillations can be used to reconstruct spatially selective neural responses to both attended object orientations [29, 8] and spatial locations [54], as well as their encoding and maintenance [26, 48, 30, 12]. Crucially for our study, IEMs applied to scalp EEG have proven robust to noise artifacts that are common in physically active tasks, including motion artifacts and sweating [8, 30].

In this paper, we apply both univariate (ERP, lateralization) and multivariate (IEM) techniques—typically used for stationary objects on 2D screens—to EEG data collected during our task. Our goal is to determine the extent to which these methods can be adapted to study spatial attention to both static and moving objects in an immersive 3D VR environment.

### 2.3 Virtual Reality and Mobile Brain Imaging

A key goal of cognitive neuroscience is to understand the brain mechanisms that support natural human behavior. By building models of how we acquire, process, store and apply information, these findings can help solve real-world problems, such as designing better interfaces for technology, developing therapeutic interventions for cognitive disorders and improving educational techniques. Several decades of laboratory research have provided valuable insights into the human brain and behavior. However, traditional laboratory experiments are often a poor reflection of reality. These experiments typically rely on stationary 2D stimuli presented to participants on a flat screen, requiring simple responses that are disconnected from natural responses (e.g., pressing buttons on a keyboard, clicking a mouse). This approach delivers high internal validity, but limited ecological validity, meaning that the neural mechanisms underlying natural behavior are unclear [62]. Furthermore, traditional brain imaging studies have been largely conducted

on sedentary participants who are required to remain as still as possible to minimize movement artifacts. As a result, current cognitive models largely fail to account for the impact of movement on brain activity, despite increasing evidence that physical activity can have profound effects on multiple aspects of brain and behavior [32, 31]. Indeed, recent work has shown that even the simple act of standing versus sitting while completing a task can impact task-specific brain activity [23, 72]. Together, these limitations in traditional laboratory based tasks highlight a significant knowledge gap in cognitive neuroscience regarding the neural basis of natural human behavior in dynamic, real-world conditions.

A combined VR and mobile brain imaging approach offers a promising solution to this problem, allowing for the study of more naturalistic behavior [60, 62]. VR offers advantages over traditional lab methods because it offers greater ecological validity while maintaining high levels of experimental control. VR uniquely combines environmental immersion with experimental rigor: it allows participants to interact with complex, dynamic 3D stimuli within a standardized framework where events can be precisely replicated across trials. Concurrently, advances in mobile EEG technology allow for non-invasive, wireless recording of brain activity from active humans, with numerous studies successfully demonstrating that movement and environmental noise artifacts can be minimized to achieve a good signal-to-noise ratio (e.g., [68, 17]). Although studies directly comparing VR with traditional 2D displays are sparse, emerging evidence suggests that VR can elicit even more potentiated brain responses [56, 57], indicating that a combined VR-EEG approach is a promising method for studying naturalistic human brain and behavior.

Our new approach, SABER, combines VR and EEG to allow for multimodal data sampling from freely moving participants in an immersive environment. This provides a platform to study the neural mechanisms that underpin tracking and engaging with dynamic 3D objects, in a reproducible, yet more naturalistic manner.

## 3 METHODS

### 3.1 Participants

Thirty-three healthy adults (24 female, 32 right-handed,  $M_{age} = 22.12$ ,  $SD_{age} = 3.10$ ) were recruited. Participants reported normal or corrected-to-normal vision. Informed consent was provided at the beginning of each session. All procedures were approved by the UC Santa Barbara Ethics Committee and the U.S. Army Human Research Protection Office. Participants received \$20/hr for time spent in the lab (~2.5 hrs total per participant).

The sample size was based on a power analysis which revealed that with a significance level of  $\alpha = .05$  and power of .80, we needed a minimum of 34 participants to detect a medium effect and 15 participants to detect a large effect. We aimed to recruit 32 participants as this allowed us to counterbalance the order of the four experimental conditions (see next section), but ended up recruiting an extra participant due to a technical error during one recording resulting in no trigger codes being sent, hence 33 participants total. With the participant with the technical error excluded, the final sample was 32 participants. We further reasoned that this sample size was sufficient for validating our approach, as it meets or exceeds those of foundational studies using these methods. For instance, seminal work on the N2pc [22], alpha lateralization [55], and Inverted Encoding Models [54] relied on sample sizes of 10–11, 22, and 8 subjects, respectively.

### 3.2 Task and Procedure

The experimenter welcomed the participant, providing a detailed explanation of the study and obtaining their informed consent. The participant was then equipped with the VR headset (HTC Vive Eye Pro, HTC Vive, Taoyuan City, Taiwan) and controllers, and guided through single practice blocks of the "Dynamic-Multiple"

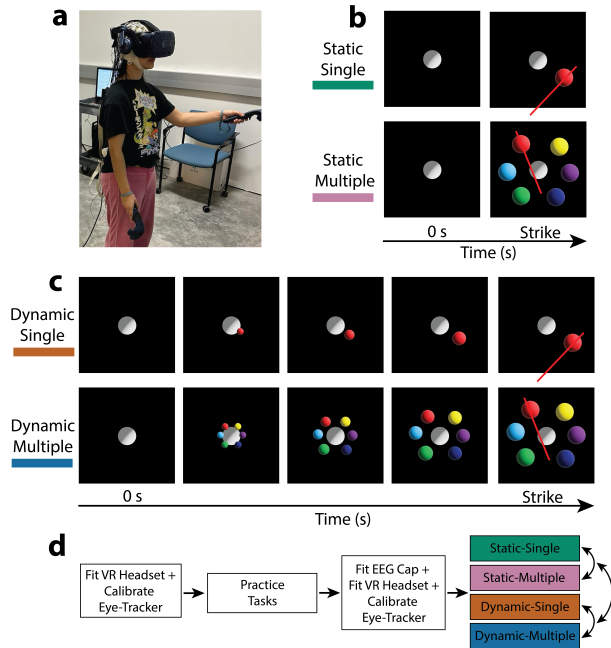


Figure 2: VR-EEG setup, cognitive tasks and protocol. (a) A participant engaged in the task while wearing an EEG cap and VR headset. (b) Schematic examples of trials in the Static-Single and Multiple conditions. A red target sphere appeared in front of the participant either in isolation or flanked by non-targets and the participant used either saber to strike the target as quickly as possible. (c) Schematic examples of trials in the Dynamic-Single and Multiple conditions. On each trial a target sphere appeared at a distance, either in isolation or flanked by non-targets, and moved towards the participant. They used either saber to intercept the target when it came within reach. (d) Experimental workflow. Curved arrows depict the between-participants counterbalancing of the four main conditions.

and "Dynamic-Single" conditions (see below) of the task. All participants successfully completed the practice blocks and proceeded to the full experiment.

The participant then removed the VR headset, sat in a comfortable chair, and was fitted with an EEG cap consisting of 64 scalp electrodes arranged according to the 10-20 system (ActiChamp, Brain Products, Berlin, Germany). Two electrodes were placed directly on the mastoids. Electrode FCz was used as an online reference. A wireless EEG transmitter (MOVE, Brain Products) was then attached. The participant then stood up and the experimenter positioned the VR headset over the EEG cap (see Figure 2A) and handed them a Vive controller for each hand. This VR-EEG setup is consistent with previous studies [17, 42].

The participant then completed the main experimental task while we continuously sampled EEG at 1000 Hz. The task was controlled using custom scripts for Unity (Unity Technologies, San Francisco, USA) that were written with functions from the Simple XR toolbox (<https://github.com/simpleOmnia/sXR>) [40]. A Dell XPS Desktop 8960 PC (NVIDIA RTX4090 GPU, Intel i9 Processor and 64 GB Ram) was used to run the VR setup. Event timing codes were sent from the Unity PC to the EEG amplifier using custom code for Unity. Eye-tracking data from the Vive headset were recorded throughout the experiment at 90 Hz, as were the positions of all relevant objects in the environment. Prior to completing the first condition, the participant ran through the built-in Vive eye-tracker calibration routine, where the physical inter-pupillary distance (IPD)

was manually adjusted and headset position optimized. All participants successfully completed the calibration routine, confirming adequate visual clarity and alignment according to the Vive software (mean IPD = 59.9 mm, SD IPD = 3.1 mm).

The participant was first positioned on a platform in a dark, sparse virtual room (Figure 1). They were handed a pair of Vive controllers, which transformed into virtual sabers measuring 1.5 m in length, the end of each saber anchored to a controller. Both sabers were the same color as the targets (counterbalanced red/blue between participants). They were instructed to position themselves over a white circle on the platform and maintain fixation on a fixation sphere (2 m circumference) positioned at a distance of 30 m and a height of 1 m. They then completed blocks of the four task conditions (order counterbalanced between participants). Each block was initiated with a trigger press, and was comprised of a continuous stream of stimuli.

In the *Static-Single* condition, on each trial a lone target sphere (0.4 m circumference) abruptly appeared in one of six possible pseudo-random locations (centered 1 m from the fixation sphere) and they were instructed to strike the sphere as quickly as possible using either saber (Figure 2B). The target color was either red or blue, counterbalanced between participants. If the target was struck by a saber within the response window (1 s), it disappeared, and the participant received haptic feedback (a short buzz on the striking controller) to signal their successful target strike. The next target then appeared in another pseudo-random location, and the sequence continued. In the *Dynamic-Single* condition, the target sphere emerged from a distant fixation point and moved towards one of the six location bins, and participants were instructed to strike it as it traveled past them (Figure 2C). At the point of saber contact, the moving target sphere was approximately the same distance from fixation as in the Static condition (1 m). Successful saber strikes triggered a short confirmatory buzz on the striking controller. The *Static-Multiple* and *Dynamic-Multiple* conditions were identical to the Static-Single and Dynamic-Single conditions in terms of trial structure and timing, but on each trial five distractor spheres appeared at the same time as the target sphere, one in each of the remaining location bins. The distractor colors were green, magenta, cyan, yellow and either blue if the assigned target color was red, or red if the assigned target color was blue.

The participant was instructed to maintain gaze on the fixation sphere and minimize eye and head movements throughout each trial block. Fixation breaks caused the sphere to turn from white to grey, which was intended to prompt the participant to re-fixate on the sphere. Fixation breaks were defined as gaze deviations exceeding 5 degrees of visual angle from the center of the fixation sphere. In all conditions the target destination was one of six different stimulus locations, which were determined by pseudorandomly selecting one of six locations [30°, 90°, 150°, 210°, 270°, 330°] and then adding a random amount of jitter [+/- 10°]. Target location sequences were pseudo-randomized to ensure equal sampling across bins while preventing repetition of the same bin on consecutive trials. A unique sequence was generated for each participant and condition. Each condition consisted of 6 blocks, and each block contained 102 trials. Condition order was counterbalanced across four possible permutations to control for sequence effects (SS = Static-Single, SM = Static-Multiple, DS = Dynamic-Single, DM = Dynamic-Multiple): [SS,SM,DS,DM; SM,SS,DM,DS; DS,DM,SS,SM; DM,DS,SM,SS]. Participants were distributed across these four orders. They were given the opportunity to take breaks between blocks during each condition, and also longer breaks between conditions, including temporarily removing the VR headset if required. In the static conditions, the 1 s response window was determined via pilot testing. All four conditions took ~1 hr total to complete. Participants tolerated the experimental protocol well, and no sessions were terminated due to simulator sickness

or adverse reactions.

### 3.3 EEG Data Preprocessing

EEG data were processed offline with custom scripts that use functions from the EEGLAB Toolbox [18] and MATLAB (version 2023a, The MathWorks, Inc., Natick, MA, US). First, EEG event markers were shifted to compensate for the latency between the event codes sent from Unity to the amplifier to mark the onset of each stimulus on each trial, and the actual rendering of each stimulus in the VR headset. This correction was based on independent validation using a photodiode, which revealed a mean lag of 25.56 ms ( $SD = 2.02$  ms) across 102 trials. The EEG data were down-sampled to 250 Hz to reduce processing and memory demands. The data were then re-referenced to the average mastoid signal, and low- and high- pass filters were applied at 30 Hz and 0.1 Hz, respectively. The Automatic Artifact Removal (AAR) toolbox [33] was then applied to remove ocular artifacts. Specifically, ocular artifacts were corrected using the AAR toolbox’s conventional recursive least squares (CRLS) regression algorithm, which uses the recorded EOG channels (Fp1 and Fp2) as a reference to adaptively estimate and subtract ocular noise from the EEG data. Next, electrodes with excessive noise were identified using the *clean\_rawdata* function in EEGLAB, whereby an electrode was considered abnormal if it flat-lined for  $>5$  s, had values exceeding 4 standard deviations of the total channel population, or failed to correlate at  $r > .85$  with a reconstruction of it based on other channels. Thresholds were selected based on standard default recommendations for *clean\_rawdata*.

Electrode rejection rates were  $6.5 \pm 2.6$  electrodes (mean  $\pm$  SD). For alpha lateralization analyses, noisy electrodes were interpolated to ensure a consistent electrode array across all participants, enabling the calculation of alpha lateralization indices using standardized electrode clusters (Section 3.5). However, these noisy electrodes were excluded from all IEM analysis to prevent the inclusion of potentially inaccurate reconstructed electrode data in the model (Section 3.6). The data were then epoched from 0.5 s prior to stimulus onset to 2 s post-stimulus offset and the mean pre-stimulus baseline (-0.2 - 0 s) was subtracted. Trials with unsuccessful saber strikes were removed. A threshold-based artifact rejection function was then used to eliminate epochs in the remaining electrodes that exhibited amplitudes surpassing  $\pm 150$   $\mu$ V (Trial rejection rates per condition: C1 =  $5.6 \pm 7.5$  %, C2 =  $6.9 \pm 9.3$  %, C3 =  $4.9 \pm 6.6$  %, C4 =  $6.4 \pm 8.8$  %).

For the lateralization and IEM analyses, we first submitted the cleaned, epoched data to a third-order Butterworth filter to isolate the alpha (8-12 Hz) frequency band. A Hilbert transformation was then applied to extract instantaneous amplitude and phase values. Data were then sorted by location bin. To mitigate bias due to potential trial number discrepancies caused by artifact rejection, trial numbers were equalized across bins. The bin with the fewest trials was identified, and an (n-1) sized subsample was randomly selected from each bin. This ensured every bin contributed a random subsample, including the one with the least trials. Total alpha power (non-phase-locked oscillations) was then computed by squaring the absolute values from the Hilbert transform.

### 3.4 Event-Related Potentials

We used the ERP technique to analyze time-locked neural responses to targets in the Static conditions. This was possible only for the Static conditions because those tasks featured abruptly on-setting lateralized targets. Specifically, we examined the visual N1 and N2pc components - negative deflections occurring at parieto-occipital scalp locations between 150-300 ms after the presentation of a visual stimulus that provide precise time-locked signatures of spatial attention [22, 39]. We isolated these components by averaging brain responses across trials at electrodes both contralateral (i.e., opposite side) and ipsilateral (i.e., same side) to the location

of the target sphere. Consistent with standard guidelines for measuring lateralized visual attention components [45], we analyzed both components at lateral occipito-parietal electrodes PO7/PO8 and P7/P8, where they are typically maximal.

### 3.5 Lateralization

We examined alpha-band lateralization as this provides a moment-by-moment measure of spatially selective attention that, unlike the ERP analyses described in the previous section, is not strictly time-locked to abrupt stimulus onsets. This allows us to measure attention in the static conditions but, also critically in the dynamic conditions, where we want to attend to objects that start at fixation and gradually move into the periphery.

Alpha oscillations are extensively involved in voluntary attentional allocation [27], such that the distribution of alpha power across parieto-occipital scalp regions changes when attention is focused on stimuli located on either side of the visual field [55, 63]. To examine the role of alpha oscillations in our data we computed a lateralization index [63, 36]. We first averaged alpha power across parieto-occipital and occipital electrodes where alpha lateralization is known to be most evident [63, 55]; left ROI: PO3, PO7, O1; right ROI: PO4, PO8, O2]. Then, we computed the degree of lateralization over the trial duration using the following equation:

$$LateralizationIndex = \frac{IpsilateralPower - ContralateralPower}{IpsilateralPower + ContralateralPower}$$

### 3.6 Inverted-Encoding Modeling

Here, we used IEM to estimate spatially selective neural responses based on topographical patterns of alpha activity across the scalp [54, 26, 48, 30]. An IEM attempts to determine whether the input features are carrying information with specific characteristics for each of the stimulus conditions and, in the process, the method can be used to reconstruct a specific feature that the brain is representing at a moment in time. This provides a quantitative estimation of whether the information conforms to an specific expected pattern. A visualization of the IEM pipeline is shown in (Figure 3) and more detail about each step is described below.

We prepared the data for the IEM by first organizing the spectral data into location bins, ensuring the number of trials per bin was equal. To enhance the signal-to-noise ratio and reduce computational demands, each bin was then randomly split into three sets of trials, which were averaged to create 18 total inputs for the IEM (6 locations x 3 averaged trials). This trial selection process was repeated 10 times, with a new, random set of trials in each iteration, to ensure the model’s results were not influenced by any specific set of trials. This robust approach improved the reliability of our model’s outcomes.

An IEM was run for every time point within the 2.5 s trial epoch, resulting in 625 (250 Hz x 2.5 s) separate models per condition and participant. For model validation, we used an k-fold cross-validation approach, where  $k = 4$ . We divided the averaged trials into four folds, with each fold containing one averaged trial per location bin. The model was trained on three of these folds and then tested on the single remaining "left-out" fold. This process was repeated four times, ensuring that each fold was used once as the test set.

We ran two separate models: one for the static conditions and one for the dynamic conditions. Each model was trained using an equal number of trials from both conditions within its pair. This "fixed" encoding model approach helps to minimize potential bias from variations in signal-to-noise ratio (SNR) between conditions [44, 59]. For each participant and condition, we defined the following variables:  $m$  represents the number of EEG electrodes;  $n_1$  and  $n_2$  represent the number of trials in the training and testing sets, respectively. We then define a basis set. Let  $j$  be the number

of hypothetical response channels ( $C_1, j \times n_1$ ), comprised of half-sinusoidal channels raised to the seventh power. This basis set was comprised of six channels representing equally-spaced locations (i.e.,  $j=6$ ). The training and test sets are represented by  $B_1 (m \times n_1)$  and  $B_2 (m \times n_2)$ , respectively. A general linear model (GLM; standard implementation) was then used to estimate a weight matrix ( $W, m \times j$ ) using the basis set ( $C_1$ ). Specifically:

$$B_1 = WC_1$$

The ordinary least-squared estimate of  $W$  can be computed as follows:

$$\hat{W} = B_1 C_1^T (C_1 C_1^T)^{-1}$$

Using the estimated weight matrix ( $\hat{w}$ ) and the test data ( $B_2$ ), the channel responses ( $\hat{C}_2, j \times n_2$ ) can be estimated by the following:

$$\hat{C}_2 = (\hat{W}^T \hat{W})^{-1} \hat{W}^T B_2$$

After calculating the estimated channel responses  $\hat{C}_2$  for each location bin, we circularly shifted the channel response function (CRF) for each averaged trial to align with a stimulus-centered reference frame. Centered CRFs were then averaged across trials for each participant and condition. This entire modeling process was repeated for every time point in the trial. To ensure the model's output was not biased by any particular trial selection, we ran the model ten times, each with a different random selection of trials. A final centered CRF was then generated by averaging results across the ten iterations.

The IEM sequence was also repeated with randomly permuted location bin labels. We ran this null model ten times, each time using a different set of randomly shuffled bin labels, following the same trial randomization protocol. This entire process was repeated five times using a new set of shuffled labels each time. The resulting CRFs were averaged across all iterations and permutations to generate "permuted" CRFs. This randomization is expected to remove any spatial selectivity, resulting in flat channel response profiles. Permuted CRFs then served as a baseline for hypothesis testing, allowing comparison of "real" CRFs against a null model with no spatial selectivity.

To quantify spatial selectivity, estimated CRFs were then folded around  $0^\circ$  offset. This transformed the responses from  $[-120^\circ, -60^\circ, 0^\circ, 60^\circ, 120^\circ, 180^\circ]$  into  $[0^\circ, 60^\circ, 120^\circ, 180^\circ]$  by averaging responses at corresponding offsets (note that  $\pm 60^\circ$  and  $\pm 120^\circ$  were averaged, but  $0^\circ$  and  $180^\circ$  were not). The slope was then calculated using a linear regression of the resulting alpha power across the offset values. Larger slope values indicate greater spatial selectivity.

### 3.7 Hypothesis Testing

We used a non-parametric permutation-based resampling technique to empirically approximate null distributions for our statistical tests (e.g., [26, 11, 10, 9]). For pairwise analyses, we shuffled condition labels within subjects and ran 1,000 iterations to generate null distributions of  $t$ -statistics. For multivariate techniques, such as IEM, the model was trained and tested for 1,000 iterations, with condition labels shuffled in each iteration to create a matrix of permuted null channel response functions. We then determined statistical significance by calculating the probability ( $p_{null}$ ) of obtaining our observed  $t$ -statistics from these null distributions. The standard  $F$  and  $t$ -statistics are reported alongside the critical  $p$ -value (labeled  $p_{null}$ ). Where appropriate, effect sizes are reported as Cohen's  $d$ . We acknowledge that performing repeated comparisons across the time-course increases the risk of Type I errors. However, our conclusions are not based on isolated significant timepoints, but rather on sustained temporal clusters of significance. The consistency of these effects across contiguous windows mitigates the likelihood of spurious results arising from multiple comparisons.

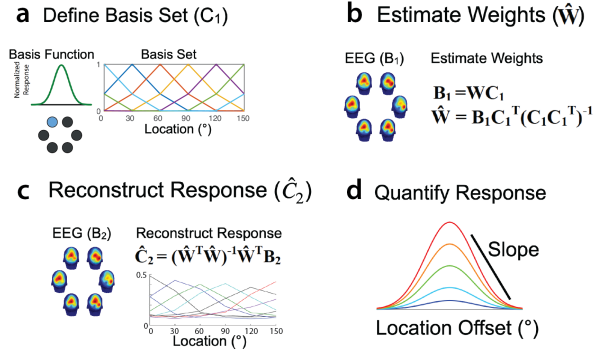


Figure 3: Inverted Encoding Modeling (IEM) Sequence. The IEM assumes brain responses can be modeled as a set of hypothetical information channels [59]. The first step is to estimate to what extent a linear combination of predetermined channel responses (the basis set, (a)) captures the underlying structure of the observed EEG data. This process generates regression weights for each electrode and channel, which essentially represent the contribution of each electrode to each location-specific channel (b). Next, these weights are used to estimate channel responses from the observed EEG data (c). Spatially selective responses are then quantified by circularly shifting the CRFs to a common center ( $0$  degrees), folding them at the center and computing the linear slope (d).

## 4 RESULTS AND DISCUSSION

We present an integrated results and discussion section to create a more coherent narrative. This approach allows us to immediately explain the significance of each key result alongside the data, directly connecting our findings to the existing literature. We then synthesize the results in a General Discussion section and discuss their broader implications.

The static and dynamic tasks differed significantly in nature, timing, and instructions. The static tasks required participants to respond as quickly and accurately to each target as possible, whereas the dynamic tasks required waiting until a target came within reach. Due to these fundamental differences in structure and timing, we analyze the static and dynamic tasks as separate pairs.

### 4.1 Behavior

Accuracy was near ceiling in both Static conditions, hence meaningful statistical comparison of accuracy was not possible due to lack of variance. Responses were faster in the Static-Single compared to Static-Multiple condition [ $t(31) = 9.34, p_{null} < .001$ ] (Figure 4a). Accuracy was not statistically different across Dynamic-Single and Dynamic-multiple conditions [ $t(31) = 1.86, p_{null} > .05$ ], but responses were slightly faster in the Dynamic-Multiple condition, albeit by  $< 1\text{ms}$  [ $t(31) = 3.46, p_{null} < .01$ ] (Figure 4b).

### 4.2 Evoked Brain Responses to Static Targets

We analyzed ERPs time-locked to the onset of static targets to confirm that our 3D VR stimuli produced brain responses consistent with those observed in classic 2D studies of spatial attention. Inspection of contralateral and ipsilateral responses across posterior electrodes to targets presented either in isolation or with distractors revealed substantial differences between conditions (Figure 5). The initial positive deflection around 100 ms post-stimulus (P1 ERP component) in both conditions is thought to reflect early sensory processing and attentional allocation in the brain [37]. In the Static-Single condition, this was a highly lateralized response because the target appeared in isolation, leading to a rapid orientation of attention. In contrast, the Static-Multiple condition, which required a sequential search, showed a non-lateralized P1 because the onset

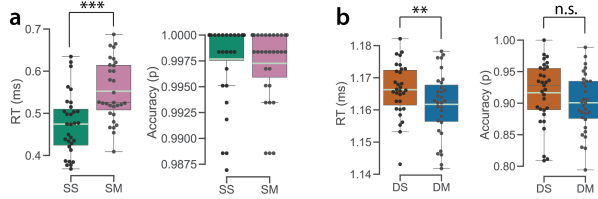


Figure 4: Behavior. Response times and accuracy for (a) Static-Single (SS) and Static-Multiple (SM) conditions, and (b) Dynamic-Single (DS) and Dynamic-Multiple (DM) conditions. Gray and white lines in boxplots represent group medians and means, respectively. \*\* $p_{null} < .01$ , \*\*\* $p_{null} < .001$

of all six stimuli drove the response. These patterns are broadly consistent with previous work [14]. Both conditions then show large subsequent negative-going lateralized deflections (i.e., larger negative-going amplitude in contralateral than ipsilateral). These deflections are broadly consistent with the posterior lateralized N1 and N2pc components, which are thought to reflect sensory gain control processes and the enhancement of the target signal from distracting information [22, 47, 69]. The classic N2pc typically involves separation of the contralateral and ipsilateral waveforms at the second negative-going component in the waveform. In our data, the large negative deflection is the first and only negative component in the Static-Single condition, which is consistent with the visual N1; whereas in the Static-Multiple condition it is the third component. The difference in timing and also order likely reflect additional stages of attentional processing required for selecting a target amongst distractors (volitional, sequential search) versus selecting an isolated target (reflexive, no search). Changes in the timing and order of lateralized negative-going brain responses to spatial stimuli have also been observed in previous work [22, 66]. Furthermore, the notion that search for a target amongst distractors requires additional stages of attentional processing is consistent with slower RTs in the Static-Multiple condition (Figure 4a).

### 4.3 Lateralized Alpha Tracks Target Locations

We examined lateralized responses in the alpha band to determine if these could be used to track the moment-by-moment allocation of spatial attention to both static and dynamic target objects presented in either left or right visual field locations.

In the Static-Single condition, alpha power increased contralateral to the cue at target onset, then shifted ipsilateral to the cue ( $\sim 0.3$  s) and remained lateralized well beyond the mean RT (Figure 6a). The observation of alpha lateralization appearing slightly before stimulus onset is interpreted as an artifact of temporal smearing caused by the Butterworth filter. The later ipsilateral shift was expected and indicates allocation of attention to the attended location, consistent with previous work (e.g., [63, 55]). The earlier contralateral activation was unexpected and to our knowledge has not been observed previously in screen-based spatial tasks. In the Static-Multiple condition alpha did not lateralize until later in the trial ( $\sim 0.65$  s), indicating lateralization is delayed if search is delayed.

In the Dynamic-Single condition, alpha shifted ipsilateral from the target from  $\sim 0.3$  s and the lateralization index continued to increase until around the time of the response, where it plateaued before starting to decline (Figure 6b). The monotonic increase suggests that alpha tracks the location of the target as it approaches from the center of the visual field and moves towards the periphery. In the Dynamic-Multiple condition alpha lateralized slightly later ( $\sim 0.4$  s to 0.5 s), before dropping, then building up again until the response ( $\sim 0.625$  to 1.175). The delayed lateralization likely reflects the challenge of initially identifying and tracking the target amongst

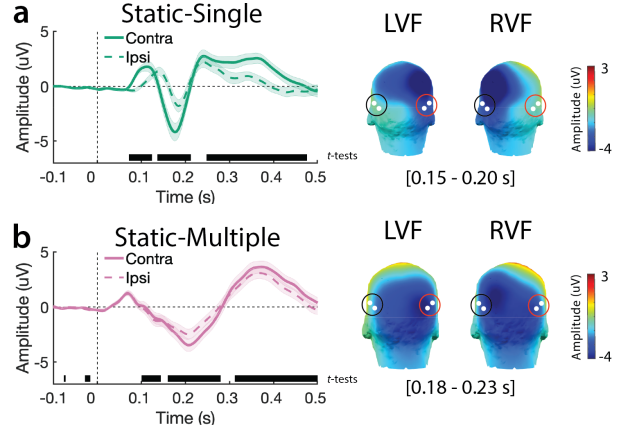


Figure 5: Event-Related Potentials. ERP Waveforms depict averaged contralateral and ipsilateral brain responses to target onsets in the Static-Single (a) and Static-Multiple (b) conditions. Horizontal black bars at the base of each plot mark timepoints where the lateralized responses are statistically different. Topographic maps to the right of the ERP plots show mean activation across posterior scalp regions for the large negative-going deflection averaged across the indicated timepoints. Contralateral and ipsilateral electrodes used to compute waveforms are marked with red and black circles, respectively.

non-targets. The earlier plateau and decline immediately prior to response may reflect attentional capture by the looming non-targets, drawing attention from the target and causing it to be spread elsewhere.

### 4.4 Reconstructing Attended Target Locations

We examined location-selective responses to determine the extent to which the target location was represented in patterns of alpha activity during each time point in the trial. CRF heatmaps and slope plots allow us to quantify and compare the strength of this location-specific coding throughout the trial (Figure 7).

In the Static-Single condition, robust location-selective reconstructions were observed at target onset and endured well beyond the mean saber strike RT. As with our previous lateralization analysis, the observation of a location-selective reconstruction appearing slightly before stimulus onset is interpreted as an artifact of temporal smearing caused by the Butterworth filter. The CRF slopes peaked between  $\sim 0.2$  and 0.3 s before declining and leveling off (Figure 7a). In contrast, in the Static-Multiple condition, location-selective responses emerged much more slowly, with robust reconstructions emerging at  $\sim 0.25$  s and persisting until after the saber strike. Together, these data indicate that it takes participants longer to orient attention to targets surrounded by distractors compared to isolated targets. This is consistent with the previously presented behavioral and EEG data.

In the Dynamic-Single condition, location-selective information was present from  $\sim 0.25$  s post-stimulus and was sustained through the end of the trial (Figure 7b). The CRF slopes initially peaked around 0.45 s and then peaked again just prior to target interception, before dropping off. In the Dynamic-Multiple condition, location-selective information also emerged from  $\sim 0.25$  s, peaked just slightly after the Dynamic-Single condition and then sustained through the end of the trial. Notably, CRF slopes diverged between the two conditions around the time of target interception ( $\sim 1$  to 1.5 s), with a clear peak in the Dynamic-Single relative to the Dynamic-Multiple condition, indicating stronger focus or orientation of attention to a lone target. This divergence may reflect diffusion of attention away from the target when it is surrounded by

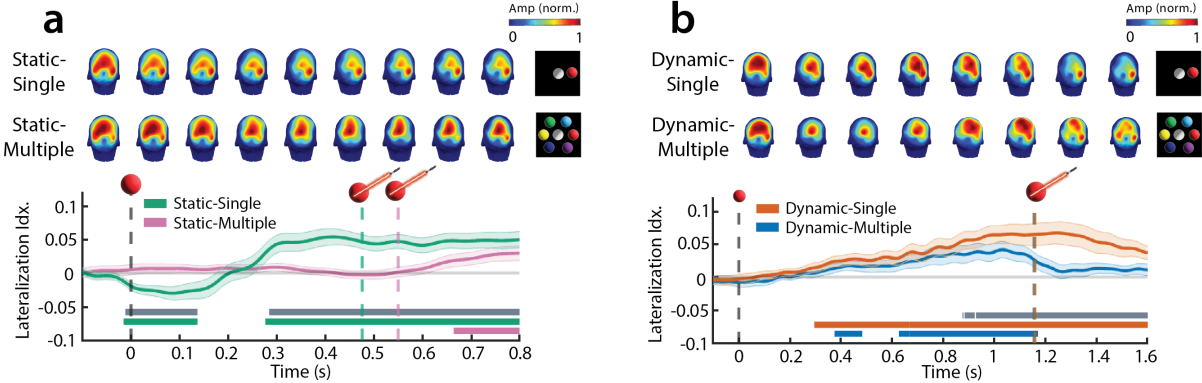


Figure 6: Alpha Lateralization. Plots show lateralization for (a) Static and (b) Dynamic conditions. Topographic maps show the distribution of alpha across posterior scalp sites related to one example target location (red target, 90°). At the base of each plot, condition-specific color-coded bars indicate timepoints where the lateralization index was significantly different from zero, while grey bars mark time points where lateralization differed significantly between conditions. Additionally, condition-specific color-coded vertical dashed lines indicate the mean saber strike reaction times for each condition.

distractors around the time of interception, despite the target already having been identified earlier in the trial. This may reflect participants being unable to disengage attention from the looming distractor stimuli, suggesting these stimuli capture attention despite their task irrelevance. This is consistent with previous behavioral work showing capture by looming stimuli [43, 28].

## 5 GENERAL DISCUSSION

The goal of this paper was to assess the feasibility of the SABER framework for studying the neural mechanisms of object tracking in immersive VR. Participants used virtual sabers to strike target spheres, which appeared either abruptly or loomed from a distance, both alone and with distractors. Participants performed well at the tasks, striking targets with high accuracy. We then demonstrated that methods previously used to study spatial attention to static objects on 2D screens could be applied to study both static and moving 3D objects in immersive VR. Our results validate this approach by demonstrating a consistency in neural mechanisms with previous 2D studies while also providing novel insight into spatial attention to dynamic 3D objects. This paper thus provides a crucial foundation for future research, enabling new hypotheses about how the brain supports attention to dynamic objects to be tested.

Traditional univariate EEG methods confirmed that our static and dynamic 3D VR stimuli produced reliable brain responses consistent with the existing literature. The ERP analyses served as an important reality check on timing, showing that expected sensory-evoked responses unfolded on a plausible timeline following the abrupt onset of targets. This finding gives us confidence in the temporal precision of our integrated VR-EEG hardware, despite the temporal smearing inherent in oscillatory analyses. Furthermore, lateralization analyses confirmed that alpha power shifted across the posterior scalp as participants attended to targets, a pattern consistent with previous work. A key novel result from this analysis was demonstrating that, in the dynamic conditions, the extent of alpha lateralization tracks the target’s eccentricity from fixation. Additionally, the unexpected rise in contralateral alpha power following target onset in the Static-Single condition represents a novel finding that has not been documented in prior literature. The exact cause of this effect is unclear, but we reason that this is a genuine physiological effect rather than a methodological artifact, as very similar processing pipelines yielded no such patterns in our previous screen-based experiments using comparable single-target paradigms [12, 48]. Replication and further examination of this unexpected effect will be important. Note that while we interpret lat-

eralized alpha activity as a proxy for attentional allocation, the precise functional role of these oscillations remains a subject of debate. Although there is consensus that alpha tracks the spatiotemporal dynamics of covert orienting, current evidence cannot definitively distinguish between mechanisms of distractor suppression and target signal enhancement [25]. Alpha oscillations likely represent just one of several components of a multifaceted attention system, and while alpha does appear to support a causal role in control of spatial attention, the specific processing stages and temporal dynamics governing this effect remain to be fully characterized [51].

Multivariate methods indicated that it was possible to track moment-by-moment spatial attention to 3D static and dynamic objects with greater spatial specificity. Unlike traditional univariate measures which often provide only coarse hemispheric contrasts (attend left vs. right), the IEM approach allowed us to recover precise spatial reconstructions of attended targets from the patterns of alpha oscillations recorded across the scalp. The IEM results are particularly encouraging for cognitive neuroscientists, as they show it is possible to reconstruct dynamic 3D target locations despite the increased visual complexity and potential noise inherent to VR environments. This capability is essential for studying the dynamics of spatial attention, paving the way for future research into how the brain continuously updates priority maps during naturalistic ‘attentional smooth pursuit’ of dynamic targets moving in 3D.

### 5.1 Limitations and Future Directions

While this study maintained high experimental rigor, we acknowledge several limitations that could be addressed in future versions of SABER. First, to isolate covert attention and minimize ocular artifacts known to disrupt alpha-band activity, we required participants to maintain central fixation [48, 12]. We recognize that enforcing fixation limits ecological validity, as natural attention involves overt eye movements; however, this constraint was necessary to minimize noise in this proof-of-concept study and allowed for direct methodological comparison with foundational screen-based paradigms [22, 54, 63]. Aspects of gaze control could be refined depending on future research goals. To ensure stricter control in future versions of SABER aimed at isolating covert attention, a self-paced, gaze-contingent approach will be implemented, where trials automatically terminate upon broken fixation, thereby ensuring participants strictly adhere to the central fixation requirement. Conversely, future adaptations of the task targeting naturalistic, free-viewing behavior will eliminate fixation requirements, thereby enabling participants to overtly attend to targets. Further

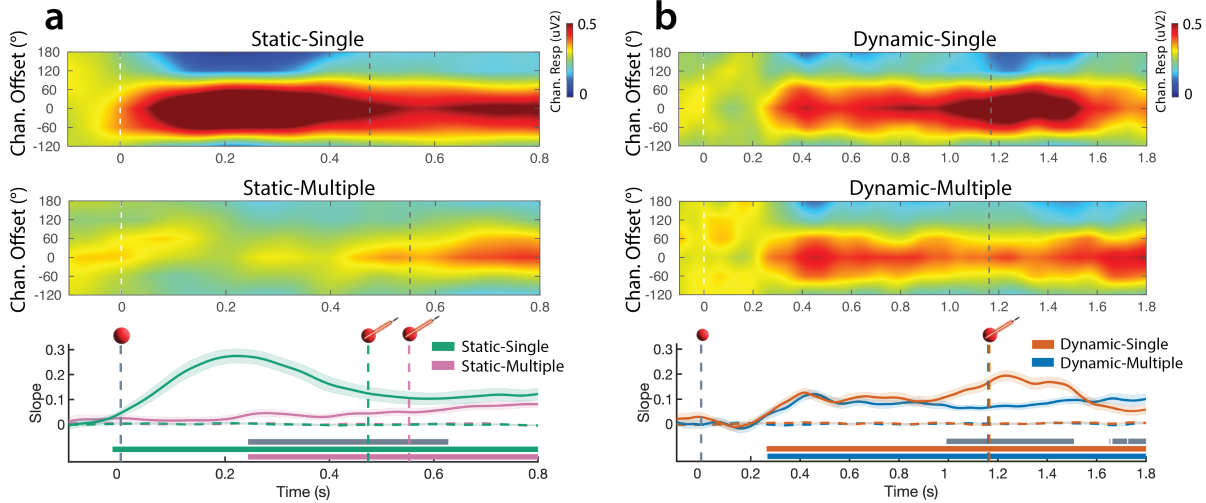


Figure 7: Reconstructing Object Locations from Alpha-Band Power with Inverted Encoding Modeling. CRF heatmaps and slope plots are shown for Static (a) and Dynamic (b) conditions. Heatmaps depict spatially selective responses for targets reconstructed from topographic patterns of alpha, with darker red along the preferred location (0 degs) indicating stronger spatial selectivity. Line plots depict linear slopes of folded CRFs for each condition (solid lines) and their permuted counterparts (dashed lines). Higher slopes indicate more selective responses. At the base of each plot, condition-specific color-coded bars mark timepoints where there is a robust location-selective responses to the target (i.e., CRF slopes are significantly different to their permuted counterparts). Gray bars mark timepoints where CRF slopes differed significantly between conditions. Additionally, condition-specific color-coded vertical dashed lines indicate the mean saber strike reaction times for each condition.

work will also strictly control low-level visual properties by equating stimuli for luminance and contrast. Second, our sample was predominantly composed of young, healthy, female adults, which may limit the generalizability of our findings. Future iterations will prioritize recruiting a demographically representative sample with a balanced sex ratio and broader age range, if possible. Third, despite its immersive nature, VR differs from reality and remains a novel experience for many participants, which may impact behavior. The extended duration of the VR session ( $\sim 1$  hour) likely contributed to participant fatigue, while the standing posture required for the task may have taxed attentional resources and introduced additional variability into the neural data [23, 72]. While we are not yet capturing true real-world behavior, this VR-EEG approach does go a long way towards bridging the gap between traditional lab-based studies and naturalistic human functioning. Fourth, despite the generally clean EEG signals, the VR headset design may have compromised electrode contact in specific areas, notably under the overhead strap and face pad. Fifth, the HTC Vive Pro Eye has a physical minimum IPD of 60.7 mm. As our sample included participants with IPDs slightly below this threshold (mean = 59.9 mm), a subset of participants viewed stimuli with a slight optical misalignment. While this offset (typically  $< .04$  mm per eye) generally remains within the sweet spot of the Fresnel lenses, it may have introduced minor optical aberrations for participants with significantly smaller IPDs. While this is not ideal, since all comparisons were within-subjects, any minor optical discrepancies were consistent across conditions and did not introduce experimental bias. We also obtained verbal confirmation from participants that the stimuli were not blurry each time the headset was repositioned.

Despite limitations, this paper provides proof-of-concept for a new framework for studying attention in dynamic 3D VR. While the primary focus was validating SABER for spatial attention research, this approach also opens the door to investigating the interplay between attention, working memory, and action goals. Recent hypotheses on working memory propose that a neural mechanism exists to track stored items across space and time [3], and our approach provides evidence to support this. Furthermore, considering accumulating evidence that action goals [34, 64, 65], movement,

and physiological state [31, 32, 7] influence attention and working memory, it is critical to adopt more dynamic and embodied tasks to gain a comprehensive understanding of these cognitive functions.

## 6 CONCLUSION

Tracking dynamic objects relies heavily on spatial attention and WM [4, 21, 38] yet contemporary theories of attention and working memory are primarily based on static objects presented at fixed locations [16, 24, 46, 58, 70]. This paper introduces and validates SABER, a new framework that merges a dynamic *Beat Saber* inspired VR task with EEG recording to elucidate the neural mechanisms that underpin real-time tracking of dynamic objects in an immersive 3D environment. We show how EEG techniques typically applied to study spatial attention to 2D static stimuli can be successfully used to study dynamic stimuli moving in 3D. These findings validate the SABER framework and open up new avenues for research into how the brain supports attention to dynamic objects. This approach reveals how known neural mechanisms for static object processing extend and adapt to dynamic contexts, while importantly allowing us to challenge the notion that attentional mechanisms operate identically in static versus dynamic settings [20]. This approach is critical because the insights it can generate will not only advance our understanding of how attention operates in the physical world but will also have practical applications in the design of more intuitive interfaces, effective training simulations, and immersive experiences optimized for the human attention system. Ultimately, this work is relevant across a broad range of fields, including cognitive neuroscience, human factors, and extended reality (XR) and brain-computer interface (BCI) development.

## ACKNOWLEDGMENTS

The authors wish to thank Elise Kormos, Paola Belmontes and Catherine Wen for their assistance with data collection. This research was sponsored by the U.S. Army Research Office and accomplished under contract W911NF-09-D-0001 for the Institute for Collaborative Biotechnologies. Additional support came from NSF award IIS-1911230 and ONR grant N00014-23-1-2118.

## REFERENCES

- [1] R. Allen, P. McGeorge, D. G. Pearson, and A. Milne. Multiple-Target Tracking: A Role for Working Memory? *Quarterly Journal of Experimental Psychology*, 59(6):1101–1116, June 2006. doi: 10.1080/02724980543000097 2
- [2] M. M. Alnawmasi and S. K. Khuu. Deficits in multiple object-tracking and visual attention following mild traumatic brain injury. *Scientific Reports*, 12(1):13727, Aug. 2022. doi: 10.1038/s41598-022-18163-2 1
- [3] E. Awh and E. K. Vogel. Working memory needs pointers. *Trends in Cognitive Sciences*, p. S1364661324003279, Jan. 2025. doi: 10.1016/j.tics.2024.12.006 2, 9
- [4] M. Belledonne, E. Butkus, B. J. Scholl, and I. Yildirim. Adaptive computation as a new mechanism of dynamic human attention. *Psychological Review*, June 2025. doi: 10.1037/rev0000572 2, 9
- [5] H. Berger. Über das Elektroencephalogramm des Menschen. *Archiv für Psychiatrie und Nervenkrankheiten*, 87(1):527–570, Dec. 1929. doi: 10.1007/BF01797193 2
- [6] G. J. Brouwer and D. J. Heeger. Decoding and reconstructing color from responses in human visual cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 29(44):13992–14003, 2009. doi: 10.1523/JNEUROSCI.3577-09.2009 3
- [7] T. Bullock, H. Cecotti, and B. Giesbrecht. Multiple stages of information processing are modulated during acute bouts of exercise. *Neuroscience*, 307:138–150, Oct. 2015. doi: 10.1016/j.neuroscience.2015.08.046 9
- [8] T. Bullock, J. C. Elliott, J. T. Serences, and B. Giesbrecht. Acute Exercise Modulates Feature-selective Responses in Human Cortex. *Journal of cognitive neuroscience*, 29(4):605–618, Apr. 2017. doi: 10.1162/jocn.a.01082 3
- [9] T. Bullock, B. Giesbrecht, A. E. Beaudin, B. G. Goodyear, and M. J. Poulin. Effects of changes in end-tidal PO<sub>2</sub> and PCO<sub>2</sub> on neural responses during rest and sustained attention. *Physiological Reports*, 9(21):1–25, Nov. 2021. doi: 10.14814/phy2.15106 6
- [10] T. Bullock, E. Machniak, J. Passananti, K. Yu, R. Kumaran, Y.-J. Kim, J. Garrett, M. Viswesvaran, J. Ram, T. Höllerer, and B. Giesbrecht. Electrophysiological evidence for differential semantic processing of words and objects presented in augmented reality. *Computers in Human Behavior Reports*, p. 100677, Apr. 2025. Publisher: Elsevier BV. doi: 10.1016/j.chbr.2025.100677 6
- [11] T. Bullock, M. H. MacLean, T. Santander, A. P. Boone, V. Babenko, N. M. Dundon, A. Stuber, L. Jimmons, J. Raymer, G. N. Okafor, M. B. Miller, B. Giesbrecht, and S. T. Grafton. Habituation of the stress response multiplex to repeated cold pressor exposure. *Frontiers in Physiology*, 13:752900, Jan. 2023. doi: 10.3389/fphys.2022.752900 6
- [12] T. Bullock, K. Pickett, A. Salimian, C. Gregory, M. H. MacLean, and B. Giesbrecht. Eye movements disrupt EEG alpha-band coding of behaviorally relevant and irrelevant spatial locations held in working memory. *Journal of Neurophysiology*, 129(5):1191–1211, May 2023. doi: 10.1152/jn.00302.2021 3, 8
- [13] P. Cavanagh and G. Alvarez. Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences*, 9(7):349–354, July 2005. doi: 10.1016/j.tics.2005.05.009 1, 2
- [14] V. P. Clark, S. Fan, and S. a. Hillyard. Identification of early visual evoked potential generators by retinotopic and topographic analyses. *Human Brain Mapping*, 2(3):170–187, 1995. doi: 10.1002/hbm.460020306 7
- [15] M. X. Cohen. *Analyzing Neural Time Series Data: Theory and Practice*. The MIT Press, Jan. 2014. doi: 10.7551/mitpress/9609.001.0001 3
- [16] M. Corbetta and G. L. Shulman. Control of goal-directed and stimulus-driven attention in the brain. *Nature reviews. Neuroscience*, 3(3):201–15, Mar. 2002. doi: 10.1038/nrn755 2, 9
- [17] A. Delaux, J. De Saint Aubert, S. Ramanoël, M. Bécu, L. Gehrke, M. Klug, R. Chavarriaga, J. Sahel, K. Gramann, and A. Arleo. Mobile brain/body imaging of landmark-based navigation with high-density EEG. *European Journal of Neuroscience*, 54(12):8256–8282, Dec. 2021. doi: 10.1111/ejn.15190 3, 4
- [18] A. Delorme and S. Makeig. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*, 134(1):9–21, Mar. 2004. doi: 10.1016/j.jneumeth.2003.10.009 5
- [19] R. N. Denison. Visual temporal attention from perception to computation. *Nature Reviews Psychology*, 3(4):261–274, Mar. 2024. doi: 10.1038/s44159-024-00294-0 2
- [20] Z. Djebbara, D. C. Huynh, A. Koselevs, Y. Chen, L. B. Fich, and K. Gramann. Turning corners in built environments shifts spatial attention costs. *NeuroImage*, 322:121549, Nov. 2025. doi: 10.1016/j.neuroimage.2025.121549 9
- [21] T. Drew and E. K. Vogel. Neural Measures of Individual Differences in Selecting and Tracking Multiple Moving Objects. *The Journal of Neuroscience*, 28(16):4183–4191, Apr. 2008. doi: 10.1523/JNEUROSCI.0556-08.2008 2, 9
- [22] M. Eimer. The N2pc component as an indicator of attentional selectivity. *Electroencephalography and clinical Neurophysiology*, 99:225–234, 1996. 2, 3, 5, 7, 8
- [23] S. Fakorede, F. Alkhameys, K. Liao, L. Martin, and H. Devos. Neural evidence for attentional resource allocation to postural control using brain-body imaging. *Behavioural Brain Research*, 494:115716, Oct. 2025. Publisher: Elsevier BV. doi: 10.1016/j.bbr.2025.115716 3, 9
- [24] I. C. Fiebelkorn and S. Kastner. Functional Specialization in the Attention Network. *Annual Review of Psychology*, 71(1):221–249, Jan. 2020. doi: 10.1146/annurev-psych-010418-103429 2, 9
- [25] J. J. Foster and E. Awh. The role of alpha oscillations in spatial attention: limited evidence for a suppression account. *Current Opinion in Psychology*, 29:34–40, 2019. Publisher: Elsevier Ltd. doi: 10.1016/j.copsyc.2018.11.001 3, 8
- [26] J. J. Foster, D. W. Sutterer, J. T. Serences, E. K. Vogel, and E. Awh. The topography of alpha-band activity tracks the content of spatial working memory. *Journal of Neurophysiology*, 115(1):168–177, Jan. 2016. ISBN: 6135336215. doi: 10.1152/jn.00860.2015 3, 5, 6
- [27] J. J. Foxe and A. C. Snyder. The Role of Alpha-Band Brain Oscillations as a Sensory Suppression Mechanism during Selective Attention. *Frontiers in Psychology*, 2, July 2011. Publisher: Frontiers. doi: 10.3389/fpsyg.2011.00154 3, 5
- [28] S. L. Franconeri and D. J. Simons. Moving and looming stimuli capture attention. *Perception & Psychophysics*, 65(7):999–1010, Oct. 2003. doi: 10.3758/BF03194829 8
- [29] J. O. Garcia, R. Srinivasan, and J. T. Serences. Near-real-time feature-selective modulations in human cortex. *Current biology : CB*, 23(6):515–22, Mar. 2013. Publisher: Elsevier Ltd. doi: 10.1016/j.cub.2013.02.013 3
- [30] J. Garrett, T. Bullock, and B. Giesbrecht. Tracking the contents of spatial working memory during an acute bout of aerobic exercise. *J Cogn Neurosci*, 33:1271–1286, 2021. doi: 10.1162/jocn.a.01714 3, 5
- [31] J. Garrett, C. Chak, T. Bullock, and B. Giesbrecht. A systematic review and Bayesian meta-analysis provide evidence for an effect of acute physical activity on cognition in young adults. *Communications Psychology*, 2(1):1–21, Aug. 2024. Publisher: Nature Publishing Group. doi: 10.1038/s44271-024-00124-2 3, 9
- [32] B. Giesbrecht, T. Bullock, and J. Garrett. Physically activated modes of attentional control. *Trends in Cognitive Sciences*, 29(3):295–307, Mar. 2025. doi: 10.1016/j.tics.2024.11.006 3, 9
- [33] G. Gomez-Herrero, W. Clercq, H. Anwar, O. Kara, K. Egiazarian, S. Huffel, and W. Paesschen. Automatic Removal of Ocular Artifacts in the EEG without an EOG Reference Channel. *Proceedings of the 7th Nordic Signal Processing Symposium - NORSIG 2006*, pp. 130–133, June 2006. doi: 10.1109/NORSIG.2006.275210 5
- [34] T. C. Handy, S. T. Grafton, N. M. Shroff, S. Ketay, and M. S. Gazzaniga. Graspable objects grab attention when the potential for action is recognized. *Nature Neuroscience*, 6(4):421–427, Apr. 2003. doi: 10.1038/nn1031 9
- [35] S. Harenberg, J. St. Onge, J. Robinson, O. Eguakun, A. Feinstein, K. Dorsch, R. S. Kakar, R. Abdulhakim, Z. Rehman, M. Shawush, and V. Pillay. Effectiveness of Three-Dimensional Multiple-Object Tracking in Patients with Multiple Sclerosis. *International Journal of MS Care*, 23(4):143–149, July 2021. doi: 10.7224/1537-2073.2020-007 1

- [36] A. M. Harris, P. E. Dux, C. N. Jones, and J. B. Mattingley. Distinct roles of theta and alpha oscillations in the involuntary capture of goal-directed attention. *NeuroImage*, 152:171–183, May 2017. doi: 10.1016/j.neuroimage.2017.03.008 5
- [37] S. A. Hillyard, E. K. Vogel, and S. J. Luck. Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 353(1373):1257–70, Aug. 1998. doi: 10.1098/rstb.1998.0281 6
- [38] A. Holcombe. *Attending to Moving Objects*. Cambridge University Press, 1 ed., Feb. 2023. doi: 10.1017/9781009003414 1, 2, 9
- [39] S. Johannes, T. Münte, H. Heinze, and G. Mangun. Luminance and spatial attention effects on early visual processing. *Cognitive Brain Research*, 2(3):189–205, July 1995. doi: 10.1016/0926-6410(95)90008-X 5
- [40] J. Kasowski and M. Beyeler. SimpleXR: An open-source Unity toolbox for simplifying vision research using augmented and virtual reality. *Journal of Vision*, 23(9):5518, Aug. 2023. doi: 10.1167/jov.23.9.5518 4
- [41] O. Kelemen, O. Nagy, A. Mátyássy, I. Bitter, G. Benedek, Z. Vidnyánszky, and S. Kéri. How well do patients with schizophrenia track multiple moving targets? *Neuropsychology*, 21(3):319–325, May 2007. doi: 10.1037/0894-4105.21.3.319 1
- [42] M. Liang, M. J. Starrett, and A. D. Ekstrom. Dissociation of frontal-midline delta-theta and posterior alpha oscillations: A mobile EEG study. *Psychophysiology*, 55(9):e13090, Sept. 2018. doi: 10.1111/psyp.13090 4
- [43] J. Y. Lin, S. Franconeri, and J. T. Enns. Objects on a Collision Path With the Observer Demand Attention. *Psychological Science*, 19(7):686–692, July 2008. Publisher: SAGE Publications Inc. doi: 10.1111/j.1467-9280.2008.02143.x 8
- [44] T. Liu, D. Cable, and J. L. Gardner. Inverted encoding models of human population response conflate noise and neural tuning width. *The Journal of Neuroscience*, 38(2):2453–17, 2017. doi: 10.1523/JNEUROSCI.2453-17.2017 5
- [45] S. J. Luck. *An introduction to the event-related potential technique*. MIT press, 2014. 2, 5
- [46] S. J. Luck, N. Gaspelin, C. L. Folk, R. W. Remington, and J. Theeuwes. Progress toward resolving the attentional capture debate. *Visual Cognition*, 29(1):1–21, Jan. 2021. doi: 10.1080/13506285.2020.1848949 2, 9
- [47] S. J. Luck and S. A. Hillyard. Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, 31(3):291–308, May 1994. doi: 10.1111/j.1469-8986.1994.tb02218.x 2, 7
- [48] M. H. MacLean, T. Bullock, and B. Giesbrecht. Dual Process Coding of Recalled Locations in Human Oscillatory Brain Activity. *The Journal of Neuroscience*, 39(34):6737–6750, Aug. 2019. doi: 10.1523/JNEUROSCI.0059-19.2019 3, 5, 8
- [49] H. S. Meyerhoff, F. Papenmeier, and M. Huff. Studying visual attention using the multiple object tracking paradigm: A tutorial review. *Attention, Perception, & Psychophysics*, 79(5):1255–1274, July 2017. doi: 10.3758/s13414-017-1338-1 2
- [50] P. L. Nunez and R. Srinivasan. *Electric fields of the brain: The neurophysics of EEG*. Oxford University Press, Jan. 2006. doi: 10.1093/acprof:oso/9780195050387.001.0001 2
- [51] C. Peylo, Y. Hilla, and P. Sauseng. Cause or consequence? Alpha oscillations in visuospatial attention. *Trends in Neurosciences*, 44(9):705–713, Sept. 2021. doi: 10.1016/j.tins.2021.05.004 8
- [52] Z. W. Pylyshyn. Perception, Representation, and the World: The FINST That Binds. In D. Dedrick and L. Trick, eds., *Computation, Cognition, and Pylyshyn*, pp. 3–48. The MIT Press, June 2009. doi: 10.7551/mitpress/8135.003.0004 2
- [53] Z. W. Pylyshyn and R. W. Storm. Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, 3(3):179 – 197, 1988. Place: Leiden, The Netherlands Publisher: Brill. doi: 10.1163/156856888X00122 1, 2
- [54] J. Samaha, T. C. Sprague, and B. R. Postle. Decoding and Reconstructing the Focus of Spatial Attention from the Topography of Alpha-band Oscillations. *Journal of Cognitive Neuroscience*, 28(8):1090–1097, Aug. 2016. arXiv: 1511.04103 ISBN: 9780192880512. doi: 10.1162/jocn.a.00955 3, 5, 8
- [55] P. Sauseng, W. Klimesch, W. Stadler, M. Schabus, M. Doppelmayr, S. Hanslmayr, W. R. Gruber, and N. Birbaumer. A shift of visual spatial attention is selectively associated with human EEG alpha activity. *The European journal of neuroscience*, 22(11):2917–26, Dec. 2005. doi: 10.1111/j.1460-9568.2005.04482.x 3, 5, 7
- [56] D. Schubring, M. Kraus, C. Stolz, N. Weiler, D. A. Keim, and H. Schupp. Virtual Reality Potentiates Emotion and Task Effects of Alpha/Beta Brain Oscillations. *Brain Sciences*, 10(8):537, Aug. 2020. Number: 8 Publisher: Multidisciplinary Digital Publishing Institute. doi: 10.3390/brainsci10080537 3
- [57] B. Schöne, J. Kisker, L. Lange, T. Gruber, S. Sylvester, and R. Osinsky. The reality of virtual reality. *Frontiers in Psychology*, 14, Feb. 2023. Publisher: Frontiers. doi: 10.3389/fpsyg.2023.1093014 3
- [58] S. Shomstein, X. Zhang, and D. Dubbelde. Attention and platypuses. *WIREs Cognitive Science*, 14(1), Jan. 2023. doi: 10.1002/wcs.1600 2, 9
- [59] T. C. Sprague, K. C. S Adam, J. J. Foster, M. Rahmati, D. W. Suterer, and V. A. Vo. Assaying population-level stimulus representations Title: Inverted encoding models assay population-level stimulus representations, not single-unit neural tuning. pp. 0–10, 2018. doi: 10.1523/ENEURO.0098-18.2018 5, 6
- [60] M. Stangl, S. L. Maoz, and N. Suthana. Mobile cognition: imaging the human brain in the ‘real world’. *Nature Reviews Neuroscience*, Apr. 2023. doi: 10.1038/s41583-023-00692-y 3
- [61] L. E. Thomas and A. E. Seiffert. Self-motion impairs multiple-object tracking. *Cognition*, 117(1):80–86, Oct. 2010. doi: 10.1016/j.cognition.2010.07.002 2
- [62] K. Thurley. Naturalistic neuroscience and virtual reality. *Frontiers in Systems Neuroscience*, 16:896251, Nov. 2022. doi: 10.3389/fnsys.2022.896251 3
- [63] G. Thut, A. Nietzel, S. a. Brandt, and A. Pascual-Leone. Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 26(37):9494–502, Sept. 2006. doi: 10.1523/JNEUROSCI.0875-06.2006 3, 5, 7, 8
- [64] C. Trentin, L. Falanga, J. Jeske, C. N. Olivers, and H. A. Slagter. Action similarity warps visual feature space in working memory. *Proceedings of the National Academy of Sciences*, 121(49):e2413433121, Dec. 2024. doi: 10.1073/pnas.2413433121 9
- [65] C. Trentin, H. A. Slagter, and C. N. Olivers. Visual working memory representations bias attention more when they are the target of an action plan. *Cognition*, 230:105274, Jan. 2023. doi: 10.1016/j.cognition.2022.105274 9
- [66] R. Verleger, B. Żurawska Vel Grajewska, and P. Jaśkowski. Time-course of hemispheric preference for processing contralateral relevant shapes: P1pc, N1pc, N2pc, N3pc. *Advances in Cognitive Psychology*, 8(1):19–28, Mar. 2012. doi: 10.5709/acp-0098-9 7
- [67] B. a. Wandell, S. O. Dumoulin, and A. a. Brewer. Visual field maps in human cortex. *Neuron*, 56(2):366–383, 2007. ISBN: 0896-6273 (Print)\n0896-6273 (Linking). doi: 10.1016/j.neuron.2007.10.012 2
- [68] S. Wang, Z. Djebbara, G. Sanches De Oliveira, and K. Gramann. Human brain dynamics dissociate early perceptual and late motor-related stages of affordance processing. *European Journal of Neuroscience*, p. ejn.16461, July 2024. doi: 10.1111/ejn.16461 3
- [69] E. Wascher, S. Hoffmann, J. Sängler, and M. Grosjean. Visuo-spatial processing and the N1 component of the ERP. *Psychophysiology*, 46(6):1270–1277, Nov. 2009. doi: 10.1111/j.1469-8986.2009.00874.x 7
- [70] J. M. Wolfe. Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review*, 28(4):1060–1092, Aug. 2021. doi: 10.3758/s13423-020-01859-9 2, 9
- [71] H. Zhang, Y. Xuan, X. Fu, and Z. W. Pylyshyn. Do objects in working memory compete with objects in perception? *Visual Cognition*, 18(4):617–640, Apr. 2010. doi: 10.1080/13506280903211142 2
- [72] A. Zinchenko, N. Busch, G. Dodwell, and T. Geyer. Withstand Context: Standing Posture Improves Contextual Cueing in Challenging Visual Search. 2025. 3, 9