

UNIVERSITY OF OLDENBURG

Evaluating Stimulus Preceding Negativity with EEG for Neuroadaptive Mixed Reality Systems

Author:

Elnur Imamaliyev
Master Program Neuroscience
Department of Neuroscience

Examiners:

PROF. DR. SVEN MAYER
Ludwig Maximilian University of
Munich

DR. MARTIN G. BLEICHNER
Carl von Ossietzky University of
Oldenburg

Submitted for the degree of
Master of Science
August 1, 2025, Munich

Acknowledgements

Sincere thanks to **Prof. Dr. Sven Mayer** for his supervision and for fostering such a welcoming and inspiring research environment in Munich.

I am also grateful to **Dr. Martin G. Bleichner** for his co-supervision and involvement throughout the project.

I would like to express my deepest gratitude to **Dr. Francesco Chiossi** for his exceptional mentorship, thoughtful feedback, and continuous encouragement throughout this project.

Elnur Imamaliyev

Contents

Acknowledgements	i
Contents	ii
List of Figures	iv
List of Tables	v
Acronyms	vi
Abstract	vii
1 Introduction	1
2 Related Work	3
2.1 Gaze-assisted Neural Interactions	4
2.2 Passive Hybrid BCIs	5
3 Novelty, Contributions, and Hypotheses	7
4 Methods	11
4.1 Experimental Design and Conditions	11
4.2 Block Structure	12
4.3 Participants	13
4.4 Apparatus	13
4.5 Procedure	15
4.6 Task	16
4.7 Trial Structure	18
4.8 Data Processing	19
4.8.1 Ocular Artifact Detection: Saccades and Blinks	19
4.8.2 EEG Processing	20
4.8.3 ERP Extraction	21
5 Results	23
5.1 Mass Univariate Analysis	23
5.2 ERP Analysis	26
5.3 Hypothesis Evaluation	28

6 Discussion	29
6.1 SPN as a Robust Marker in MR	29
6.2 Intention and Feedback: Independent and Interactive Effects . .	29
6.3 SPN Reflects Broader Anticipatory Processes	30
6.4 Implications for MR Interface and hBCI Design	31
6.5 Limitations	31
6.6 Future Directions	32
7 Conclusion	34
Appendix A Supplementary Information	46
Declaration	48

List of Figures

1	Condition Block Design	13
2	Multimodal hBCI Setup	14
3	Mixed Reality and User Interface	17
4	Trial Flow	18
5	Mass Univariate Analysis Results	25
6	ERP and Eye Movement Results	26
7	Topographical scalp activity during SPN window	27
A.1	Single participant ERPs	47

List of Tables

1	Experimental Conditions in a 2×2 Latin Square Design	12
2	Summary of Accepted Trials and Rejection Rates	23
A.1	Trial Counts and Interpolated Bad Channels per Participant . .	46

Acronyms

BCI	Brain-Computer Interface
EEG	Electroencephalography
ERP	Event-Related Potential
hBCI	hybrid Brain-Computer Interface
HCI	Human-Computer Interaction
HMD	Head-Mounted Display
ICA	Independent Component Analysis
ISI	Inter-Stimulus Interval
LSL	LabStreamingLayer
MR	Mixed Reality
MUA	Mass Univariate Analysis
ONF	Observe-No-Feedback
OWF	Observe-With-Feedback
ROI	Region of Interest
SNF	Select-No-Feedback
SPN	Stimulus-Preceding Negativity
SWF	Select-With-Feedback
UI	User interface
VR	Virtual reality
XR	Extended Reality

Abstract

As mixed reality technologies become increasingly embedded in everyday life, the need for intuitive, hands-free interaction methods is growing. Gaze-based interfaces offer a promising solution; however, they are often limited by the so-called "Midas touch" problem, where the system cannot easily distinguish between deliberate selections and passive viewing. Recent studies suggest that a specific slow cortical potential, known as the Stimulus-Preceding Negativity and measurable through electroencephalography, may serve as a neural indicator of user intention in these contexts.

This thesis explores whether this neural signal can reliably differentiate between user intentions to select versus merely observe digital elements in everyday mixed reality interfaces. Additionally, it examines how the expectation of visual feedback influences this neural activity. Twenty-three participants engaged in gaze-based selection and observation tasks across four experimental conditions that manipulated both user intention and feedback availability. Their brain and eye movement data were recorded and analyzed using mass univariate statistical methods and cluster-based permutation testing.

The results revealed that the neural signal was consistently present during intentional selection, regardless of whether participants expected visual feedback. In contrast, observation conditions were highly sensitive to the feedback context, producing a significant interaction effect. Moreover, the neural signal was also observed during tasks that included continuous visual feedback, indicating that it reflects not only intention but also anticipation of system responses.

These findings extend existing research by demonstrating that this anticipatory brain activity captures both action preparation and feedback-driven expectation in realistic mixed reality interactions. This work lays the foundation for future adaptive systems that leverage such brain signals as passive indicators of user intent, enabling more natural, accessible, and confirmation-free gaze-based interactions in immersive environments.

1 Introduction

As digital interactions become an integral part of daily life and immersive Extended Reality (XR) environments continue to advance within Human-Computer Interaction (HCI), the convergence of (Brain-Computer Interface (BCI)s) and these technologies is opening new possibilities for intuitive and accessible digital experiences. These advances are making control technologies more user-friendly and less reliant on physical actions, such as pressing physical buttons, which is especially impactful for users with motor impairments who benefit most from seamless, non-physical interaction methods (Jang et al., 2017; Wolpaw et al., 2002). However, each technology faces its own challenges: in XR, gaze-based selection methods like the Dwell technique often suffer from the "Midas touch" problem, where unintentional glances are misinterpreted as deliberate actions (Jacob, 1990); in BCIs, traditional paradigms frequently require active mental effort or intrusive stimuli, resulting in high cognitive load and limited practicality for everyday use (Banville & Falk, 2016).

To address these limitations, researchers have explored a range of solutions. In XR, manual triggers such as blinks (Lu et al., 2021) or button presses (Fernandes et al., 2023) have been used to improve selection accuracy, but these approaches reintroduce physical effort and disrupt immersion (Jang et al., 2017). In the BCI domain, passive BCIs that leverage slow cortical potentials—such as the Stimulus-Preceding Negativity (SPN)—have emerged as a promising alternative, enabling implicit intent detection without explicit commands or intrusive stimuli (Neumann et al., 2004; Protzak et al., 2013; Shishkin et al., 2016; Zander & Kothe, 2011; Zhao et al., 2021).

Creating interfaces that can sense and respond to user intent without explicit commands could transform accessibility, empower users with motor impairments, and set a new standard for natural interaction in XR (Dybdal et al., 2012; Hong & Khan, 2017; Wolpaw et al., 2002). Integrating BCI with XR has also been shown to enhance immersion, accelerate learning, and improve user performance by increasing motivation (Coogan & He, 2018). While Reddy et al. (Reddy et al., 2024) demonstrated the feasibility of SPN-based intent decoding in a single type of (Virtual reality (VR)) task, the generalizability of this approach to more naturalistic and diverse Mixed Reality (MR) scenarios remains unexplored. In this work, we investigate the dynamics of SPN as a passive neural signal for enhancing gaze-based interaction in realistic, multi-scenario MR environments. We combine high-density Electroencephalography (EEG) and gaze tracking to measure anticipatory neural activity across di-

verse User interface (UI) contexts, including app menus, document editors, and video players.

To our knowledge, our study is the first to demonstrate that SPN can serve as a robust, context-independent marker of selection intent in MR, and to reveal how intention and feedback interact to shape anticipatory neural responses. We introduce a new theoretical framework showing a crossover interaction between intention and feedback, challenging simple additive models and highlighting the complexity of real-world intent decoding. Our results show that SPN is reliably detectable during immersive gaze-based tasks, with its amplitude tracking both user intention and feedback expectation across multiple UI scenarios. We establish the technical feasibility of real-time SPN-based BCI in MR and provide novel evidence that selection intention creates neural stability immune to environmental context.

These findings move us closer to a future where technology can sense user intentions as naturally as they form—enabling more inclusive, adaptive, and effective XR experiences. By integrating gaze tracking with EEG and analyzing SPN responses across conditions, we lay the groundwork for practical, real-time BCI applications in MR and contribute to the development of user-centered interfaces that respond to cognitive intent, not just physical

2 Related Work

This section reviews the landscape of gaze-based interaction, hybrid Brain-Computer Interfaces (BCIs), and the neural correlates of user intent, with a particular focus on the Stimulus-Preceding Negativity (SPN) as a passive neural signal for interaction in Mixed Reality (MR). Our research builds on and extends the findings of Reddy (Reddy et al., 2024), who demonstrated the feasibility of SPN-based intent decoding in a controlled XR (VR) scenario, by systematically investigating SPN in more complex and realistic MR environments.

Gaze has become a central input modality in Human-Computer Interaction (HCI), especially in XR, where it enables hands-free, intuitive, and efficient control (Plopski et al., 2022; Sibert & Jacob, 2000; Tanriverdi & Jacob, 2000). The Dwell technique, which requires users to fixate on a target for a set duration to confirm selection, is the most widely used gaze-based method. However, this approach is prone to the "Midas touch" problem, where unintentional fixations are interpreted as deliberate selections. Jacob (Jacob, 1990) was among the first to identify this issue. Starker and Bolt (Starker & Bolt, 1990) explored Dwell as a measure of user interest. Subsequent research has shown that optimal dwell parameters depend on the specific task and context (Majaranta et al., 2009; Špakov & Miniotas, 2004; Ware & Mikaelian, 1986). Penkar et al. (Penkar et al., 2012) further highlighted the importance of context in dwell parameter selection.

To overcome these limitations, researchers have developed hybrid approaches that combine gaze with other modalities, such as blinks, speech, head or hand movements, and gestures (Lu et al., 2021; Špakov et al., 2014; Whitlock et al., 2018). Other studies have explored the use of handheld devices, tongue-based interfaces, body movements, and hand gestures (Fernandes et al., 2023; Kytö et al., 2018; Wei et al., 2023). Additional modalities include gaze-hand alignment and pinch gestures (Gemicioglu et al., 2023; Klamka et al., 2015; Mutasim et al., 2021).

While these methods can improve accuracy, they often reintroduce physical effort and may disrupt immersion (Lystbæk et al., 2022; Reiter et al., 2022; Wagner et al., 2023). Gaze-based interaction can also cause eye fatigue, though this is lessened when gaze is used only for pointing (Hirzle et al., 2020; Jang et al., 2017; Rajanna & Hammond, 2018). Li et al. (Z. Li et al., 2019) found that eye strain is reduced when gaze is used for pointing rather than selection.

Alternative triggers, such as smooth pursuit or vergence, have also been

explored (Esteves et al., 2017; Khamis et al., 2018; Sidenmark et al., 2023). However, each has its own limitations, such as discomfort or the need for calibration (Kirst & Bulling, 2016). Recent work has focused on predicting user intent from gaze and head features, achieving high accuracy but still facing challenges in real-world use (Bednarik et al., 2012; David-John et al., 2021; Wei et al., 2023).

2.1 Gaze-assisted Neural Interactions

The integration of gaze signals with BCIs has led to the development of hybrid BCIs (hybrid Brain-Computer Interface (hBCI)s), which are commonly used to expand the range of control commands, improve classification accuracy, or reduce intent detection time (Hong & Khan, 2017; Pfurtscheller et al., 2010). Between 2009 and 2016, more than half of the studies in this area investigated the combination of gaze and neural signals (Hong & Khan, 2017). These studies utilized both Electrooculogram (EOG) and eye-tracking data. EOG signals, valued for their temporal precision, are often used to reduce ocular artifacts and minimize false positives in classification (Bashashati et al., 2007; Mingai et al., 2015; Trejo et al., 2006).

Gaze signals have also been leveraged to enhance control in applications such as wheelchair navigation (K.-H. Kim et al., 2006). Kalika et al. (Kalika et al., 2017) demonstrated that combining gaze with EEG in a P300-based speller system outperformed EEG-only setups in both accuracy and bit rate. Eye blinks have been incorporated into both motor imagery (MI) and P300-based BCI paradigms for wheelchair control (H. Wang et al., 2014). Pfurtscheller (Pfurtscheller et al., 2010) showed that a hybrid MI-based BCI could address the Midas touch problem more accurately than gaze-only Dwell techniques, though at the cost of slower activation and increased user effort. Dong et al. (Dong et al., 2015) found that while accuracy can be improved, the requirement for users to imagine motor movements can be challenging.

A notable example by Kim (C.-H. Kim et al., 2016) combined gaze signals from an eye-tracker with the SSVEP paradigm to control a turtle's movement, using SSVEP for directional commands and eye opening/closing for reset and idle states. McCullagh et al. (McCullagh et al., 2013) examined SSVEP with the Dwell technique, noting that users with less technical proficiency struggled with the interface. Putze et al. (Putze et al., 2019) combined gaze with SSVEP for AR-based smart home management, achieving better accuracy than SSVEP alone, though the flashing stimuli disrupted cognitive flow.

2.2 Passive Hybrid BCIs

Passive hybrid BCIs (hBCIs) are designed to enhance interaction by combining signals from multiple modalities and leveraging implicit brain activity, thus reducing the need for explicit user commands or active participation. This approach aims to create more seamless, adaptive, and effortless user experiences, particularly in complex environments such as MR. Early work by Zander (Zander et al., 2016) demonstrated that error-related potentials (ErrPs) could be used for implicit cursor control, where the system detects when a user’s expectations are violated and adapts accordingly. Building on this, Krol (Krol et al., 2017) developed a neuroadaptive Tetris game in which gaze and ErrPs were used together: gaze controlled block movements, while ErrPs allowed the system to infer and correct unwanted actions based on the user’s implicit feedback. Kalaganis (Kalaganis et al., 2018) further applied ErrPs to improve the reliability of gaze-based keyboards, enabling automatic correction of inadvertent selections.

Beyond error detection, passive BCIs have also been used to classify user intent. For example, Sharma (Sharma et al., 2022) achieved high accuracy in distinguishing between different types of intent (such as free-viewing, target searching, and target presentation) using EEG features like power spectral intensity and detrended fluctuation analysis. However, these studies often rely on well-known EEG components such as the P300 or reward potentials (Banville & Falk, 2016; Finke et al., 2016), which may be more prominent in visual search tasks than in everyday UI interactions. Other research has explored the use of physiological signals, such as EEG and Electromyography, to detect pre-movement states and model the intent to interact with virtual objects (Nguyen et al., 2023). Some of these approaches likely leverage the contingent negative variation (CNV), a slow cortical potential associated with motor preparation (Neumann et al., 2004). Kato (Kato et al., 2011) used the CNV to develop a BCI master switch, while Xu (Xu et al., 2022) combined the CNV with a P300-based BCI to improve target identification accuracy. Although the CNV is promising as a passive mental ‘click’, its dependence on imagined motor activity can be distracting and may not generalize well to all contexts.

In contrast, the stimulus-preceding negativity (SPN) is a non-motor slow cortical potential linked to the anticipation of feedback or salient events (Neumann et al., 2004). Protzak (Protzak et al., 2013) demonstrated that the SPN could distinguish between intent-to-select and intent-to-observe fixations, but their work was limited to 2D displays and visual search paradigms. (Shishkin et al., 2016) extended this by showing that SPN in the parieto-occipital region

can differentiate intent in gaze-controlled games, although their analysis did not correct for ocular artifacts (Bashashati et al., 2007; Mingai et al., 2015; Trejo et al., 2006). Zhao (Zhao et al., 2021) further validated the role of SPN in distinguishing intent during smooth pursuit tasks, showing that heightened negativity at specific electrodes was associated with intent-to-select actions.

While Reddy (Reddy et al., 2024) demonstrated the feasibility of using SPN for intent decoding in a controlled XR (VR) scenario, the generalizability of SPN as a robust, context-independent marker of intent in more complex and realistic MR environments remains unexplored. Our research addresses this gap by systematically investigating SPN dynamics across diverse MR scenarios, extending the understanding of anticipatory neural processing and informing the development of adaptive, user-centered BCI.

3 Novelty, Contributions, and Hypotheses

Despite significant advances in gaze-based interaction and the identification of the SPN as a neural marker of user intent, important gaps remain. Most previous studies have demonstrated the usefulness of SPN for distinguishing between intent-to-select and intent-to-observe states, primarily in controlled 2D or VR/XR environments (Protzak et al., 2013; Reddy et al., 2024; Shishkin et al., 2016; Zhao et al., 2021). However, these studies were largely limited to simple, single-task settings, leaving open the question of whether SPN is a robust and context-independent marker of intent in more realistic, multi-scenario MR environments.

A major limitation is the frequent conflation of user intention and feedback anticipation. Many studies did not carefully control feedback across conditions, making it difficult to determine whether SPN mainly reflects the user’s goal-directed intention, the expectation of feedback, or some combination of both (Shishkin et al., 2016; Zhao et al., 2021). Early work suggested that anticipating feedback alone might be enough—or even necessary—to trigger SPN, even without explicit selection intent (Protzak et al., 2013; Zhao et al., 2021). More recent findings by Reddy (Reddy et al., 2024) indicate that intention may actually be the stronger driver, as feedback on its own often fails to produce a strong SPN if the user is not actively intending to select.

To clarify these effects, our study uses a latin square design manipulating user intention (select vs. observe) and feedback presence, including a novel control condition where feedback occurs even during passive observation. This approach allows us to disentangle the cognitive mechanisms behind SPN, testing whether feedback is necessary, sufficient, or simply not enough to generate SPN without user intent. We also adopt the terms “intent-to-select” and “intent-to-observe” (Reddy et al., 2024), which more precisely capture user goals than older labels like “intentional” or “non-intentional” fixations. Here, “intent-to-select” means deliberately trying to click a button, while “intent-to-observe” means visually engaging without the goal of triggering a click. This distinction matters because users may remain mentally engaged while observing but without explicit intention to act.

Our work is further motivated by evidence that immersive 3D environments can change neural processing—for example, by altering alpha band power and parieto-occipital connectivity (Keitel et al., 2018; G. Li et al., 2020; Lozano-Soldevilla, 2018)—but it remains unclear if these neural signatures, and SPN itself, generalize to MR, which offers more ecological validity and naturalistic

interaction.

In our MR setup, feedback plays a crucial role. The SPN reflects anticipatory brain activity linked to expecting meaningful, predictable events like feedback following an action (Shishkin et al., 2016; Zhao et al., 2021). Without feedback, or if feedback is random or unpredictable, SPN tends to weaken or disappear because there is no reliable expectation to build on (Reddy et al., 2024). (Reddy et al., 2024) included a “with feedback” condition in their XR study, but this feedback was more like pseudo-feedback—random colors and sounds that did not always connect clearly to the user’s actions or intent. Because this feedback was not consistent or meaningful, their SPN responses were weak or missing, especially during observation without intent to select. This suggests that having some feedback alone is not enough; feedback needs to be meaningful and predictable to reliably drive SPN.

Our MR feedback design improves on this by providing continuous, meaningful feedback tightly linked to gaze interaction. For instance, when participants fixate on an icon for 750 ms during a selection trial, the system triggers clear UI feedback (like a pop-up) and a system action (such as opening an app window). No such feedback or action happens during observation trials. This approach better mimics real-world MR experiences, where feedback is contextually relevant and continuous, boosting anticipation and thus SPN.

The novelty of our work centers on three key contributions: (1) showing robust SPN detection during immersive gaze-based interactions across diverse MR contexts; (2) demonstrating that user intention creates neural stability, with SPN amplitudes remaining consistent during intent-to-select regardless of feedback; and (3) establishing a new framework to understand how intention and feedback interact, revealing that neural anticipation in MR is a complex, context-dependent process.

Hypotheses To investigate the cognitive basis of SPN in MR environments, we tested how different combinations of user **intention** and **feedback expectation** modulate its amplitude.

Our study employed a 2-by-2 Latin square design with four experimental conditions, manipulating the presence of icon pop-up feedback and user click intention as independent variables (see our experimental conditions in Section 4.1 and Table 1). We used the **Observe-No-Feedback (ONF)** condition as our baseline, since it lacks both internal (intention) and external (feedback) drivers of anticipation and is therefore not expected to elicit significant SPN.

H1. This hypothesis builds on findings by Shishkin and Zhao (Shishkin et al., 2016; Zhao et al., 2021), who showed that SPN emerges in response to anticipated feedback in task-relevant scenarios. It also extends Reddy et al.’s attempt to detect SPN in XR contexts, where inconsistent pseudo-feedback likely limited its emergence (Reddy et al., 2024). Accordingly, we expect SPN amplitudes in the Select-With-Feedback (SWF), Select-No-Feedback (SNF), and Observe-With-Feedback (OWF) conditions to be significantly greater than in ONF. Therefore, **Hypothesis 1 (H1):** *SPN will be reliably elicited during gaze-based interactions across all three MR user interface contexts (App, Document, Video), provided that either selection intention or meaningful feedback is present.*

H2. The origin of SPN remains debated: some argue feedback is necessary for anticipation, while others emphasize the role of intention rather than external feedback. To explore this, we perform pairwise comparisons across conditions. Prior work suggests SPN reflects anticipatory activity linked to meaningful and predictable events (Shishkin et al., 2016; Zhao et al., 2021), but the relative influence of internal drivers (intention to act) versus external cues (predictable feedback) is unclear (Reddy et al., 2024). **H2a.** We test whether intention alone can elicit SPN by comparing SNF and SWF (both involve intentional interaction) to ONF. If SPN primarily reflects internal cognitive preparation, then intention—even without feedback—should drive stronger SPN than the passive ONF condition. Therefore, **Hypothesis 2a (H2a):** *Intention can elicit SPN.* **H2b.** We compare SNF (only intention) to OWF (only feedback). If intention has a stronger influence on SPN than passive feedback expectation, then SNF should produce a larger SPN amplitude. This comparison tests whether SPN reflects volitional action more than sensory prediction. Therefore, **Hypothesis 2b (H2b):** *Intention and feedback expectation influence SPN differently.* **H2c.** Finally, we contrast OWF and SWF—both include continuous feedback, but only SWF includes inten-

tion. If feedback alone can elicit SPN, we expect OWF to show a weaker SPN compared to SWF, indicating that feedback contributes to anticipation but intention amplifies it. Therefore, **Hypothesis 2c (H2c):** *Feedback expectation alone can elicit a weaker SPN in the absence of intention.*

4 Methods

Our study used a carefully controlled within-subjects design, employing a Latin square to isolate the effects of intention and feedback during gaze-based interaction in mixed reality (MR). Participants performed realistic Video-through MR tasks that mimic everyday interactions, enhancing ecological validity. Throughout the experiment, we collected high-quality, synchronized EEG and eye-tracking data, and applied a rigorous preprocessing and artifact rejection pipeline to ensure robust physiological analysis. The following sections describe our experimental design, conditions, block structure, participant recruitment, apparatus, procedures, and data processing.

4.1 Experimental Design and Conditions

We employed a within-participants experimental design to explore how intention and feedback influence gaze-based interactions in mixed reality (MR). The two independent variables were **Feedback** (two levels: Feedback / No Feedback) and **Intent** (two levels: Observe / Select), each tested across three different MR scenarios: App, Document, and Video. To control for learning effects, we counterbalanced the order of conditions using a balanced Latin Williams square design with four levels B.-S. Wang et al., 2009. This resulted in a 2×2 experimental design, with **Intent** (Select / Observe) and **Feedback** (With Feedback / No Feedback) as the primary manipulations.

The design aimed to isolate the effects of intention and feedback on gaze-based interactions in MR. Participants completed tasks that closely mimic everyday MR interactions, enhancing ecological validity. Throughout the experiment, we gathered synchronized EEG and eye-tracking data, applying a robust preprocessing and artifact rejection pipeline to ensure high-quality physiological analysis.

Participants performed two main types of tasks: **Active Selection** (clicking icons as buttons) and **Passive Observation** (finding icons without clicking).

Active Intent-to-Select Interactions. In these tasks, participants were instructed to treat the target icons as **buttons to click**, using the **gaze-based dwell mechanism**. If the participant fixated on the icon for 750 ms, this would trigger a click, followed by a simulated **UI action** (e.g., playing a video, opening an app, or editing text). There were two conditions for this task. In the **Intent-to-Select with Feedback (SWF)** condition, the icon popped up after the dwell trigger, followed by the UI action. In the **Intent-to-**

Select No Feedback (SNF) condition, the UI action was triggered directly after the dwell, without any pop-up animation or visual confirmation.

Passive Intent-to-Observe Interactions. For these interactions, participants engaged in a finding task without intending to click the targets, although the gaze-dwell logic still ran silently in the background. These conditions represent instances where a user dwells on targets to find something but has no intention to interact. The **Intent-to-Observe No Feedback (ONF)** condition served as a control, with the dwell mechanism logging when the user fixated on the target for 750 ms but providing no visual feedback or UI action. In contrast, the **Intent-to-Observe with Feedback (OWF)** condition introduced a visual icon pop-up feedback after the dwell threshold was reached, even though participants were not instructed to click but only to find the target.

Table 1: Experimental Conditions in a 2×2 Latin Square Design

	With Feedback	No Feedback
Select	Select-With-Feedback	Select-No-Feedback
Observe	Observe-No-Feedback	Observe-With-Feedback

To summarize, the study included two overarching task types—Intent-to-Select and Intent-to-Observe—with the latter further divided based on the presence of feedback. This resulted in four experimental conditions: (i) Intent-to-Select with ascending feedback, (ii) Intent-to-Select with descending feedback, (iii) Intent-to-Observe without feedback, and (iv) Intent-to-Observe with feedback.

4.2 Block Structure

Each participant completed four blocks—one for each experimental condition. Every block contained all three MR scenes (App Menu, Document Editor, Video Player) presented in a randomized order (see Figure 1). Before each block, participants completed a brief training phase with three sample trials.

Each scene contained 30 trials, totaling 90 trials per block (3 scenes \times 30 trials). With four condition blocks, this resulted in 360 experiment trials per participant. Target icons were randomized but balanced, and the order of conditions was counterbalanced across participants. Each block lasted approximately 20 minutes, for a total session time of 80–90 minutes including breaks and setup.

To ensure adequate statistical power and waveform reliability, we based our trial count on empirical simulations and recent guidelines specific to Event-Related Potential (ERP) research. While participant sample size primarily determines power in repeated-measures designs, increasing the number of trials per condition enhances signal-to-noise ratio and improves measurement precision—a particularly relevant consideration for late ERP components with high intra-individual variability (Boudewyn et al., 2018). Jensen et al. (Jensen & MacDonald, 2022) recommend aiming for 80–150 trials per condition in within-subject ERP paradigms when targeting medium-sized effects with moderate trial-level noise. Accordingly, we presented 90 trials per condition, balancing practical constraints to detect condition-related differences in SPN amplitude with sufficient power and waveform stability.

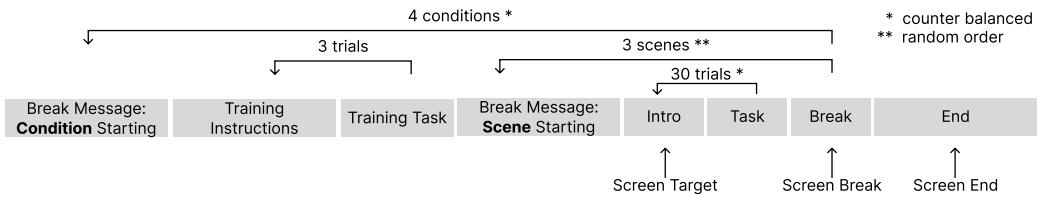


Figure 1: Condition Block Design. Schematic overview of the experimental timeline, depicting the sequence of tasks within each block from left to right.

4.3 Participants

A total of 30 participants took part in the study. None reported any history of neurological, psychological, or psychiatric disorders, and all had normal or corrected-to-normal vision. Data from 7 participants were excluded due to poor EEG quality or insufficient valid trials: 3 had high electrode impedance ($> 100 \text{ k}\Omega$), 2 showed a low signal-to-noise ratio, and 2 had flat EEG segments caused by major recording dropouts, resulting in too few usable trials in key conditions. The final analysis included data from 23 participants (9 female; average age ≈ 23).

4.4 Apparatus

We collected two physiological signals during the experiment: EEG and eye tracking, both synchronized with Unity-based interaction events.

Electroencephalography (EEG). EEG data were recorded using a 64-channel BrainVision LiveAmp amplifier (BrainProducts, Germany) with a water-based R-Net cap equipped with Ag/AgCl passive electrodes (Roesler

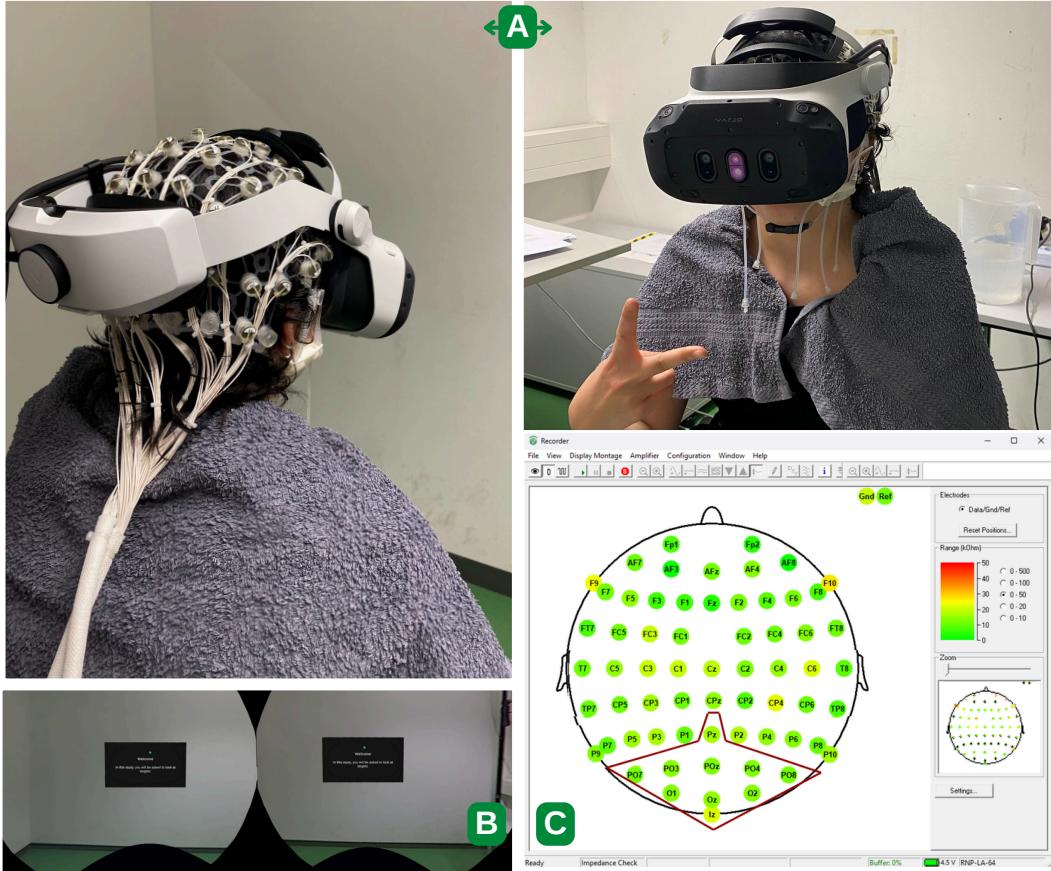


Figure 2: Multimodal hBCI Setup. A) Participant with EEG and Head-Mounted Display (HMD). B) The Video-through MR environment from the participant’s perspective. C) BrainVision Recorder software. Locations of the occipitoparietal electrodes (highlighted in red) used in our analysis.

et al., 2023; Volosyak et al., 2010). Signals were sampled at 500 Hz, referenced to FCz, and grounded at Fz. Electrodes were placed according to the International 10-20 layout. EEG was recorded from the following 64 electrode sites: Fp1, Fz, F3, F7, F9, FC5, FC1, C3, T7, CP5, CP1, Pz, P3, P7, P9, O1, Oz, O2, P10, P8, P4, CP2, CP6, T8, C4, Cz, FC2, FC6, F10, F8, F4, Fp2, AF7, AF3, AFz, F1, F5, FT7, FC3, C1, C5, TP7, CP3, P1, P5, PO7, PO3, Iz, POz, PO4, PO8, P6, P2, CPz, CP4, TP8, C6, C2, FC4, FT8, F6, F2, AF4, AF8. Electrode impedances were kept below $20\text{ k}\Omega$ before recording and monitored during breaks between experimental blocks. The EEG stream was transmitted wirelessly via Bluetooth to a high-performance PC (Windows 11 Pro, Intel Core i9-12900K @ 3.20 GHz, 64 GB RAM, 24 GB total VRAM) and recorded using the LabStreamingLayer (LSL) framework (Kothe & Contributors, 2024). This ensured accurate time synchronization with Unity-based experimental events

Eye Tracking and HMD. Eye tracking was conducted using the built-in

system of the Varjo XR-4 headset (HMD; Varjo, Finland) (Varjo Technologies, 2024), which samples binocular gaze data at 200 Hz. In addition to gaze position, pupillometry was recorded at the same rate. A custom gaze-detection script was implemented in Unity (Unity Technologies, 2024) (v6.0041) to process gaze points in real time and to trigger event markers. Eye tracking calibration was performed using Varjo’s 5-point method at the start of each block and repeated whenever the headset was repositioned or removed. The headset features a 90 Hz refresh rate, 4K-by-4K mini-LED displays (200 nits, 96% DCI-P3 color space), and provides a $120^\circ \times 105^\circ$ field of view. For spatial tracking and immersive interaction, the Varjo XR-4’s built-in 6 degrees-of-freedom (6DoF) tracking was used in combination with two external SteamVR base stations, allowing for room-scale movement and accurate spatial orientation during the tasks.

Event logging. For each trial, both dwell onset and offset timestamps were recorded. When participants fixated on a target icon, a dwell timer was started, and the dwell start time was retrospectively logged. If the gaze remained stable for the required duration (750 ms), the dwell trigger logged the end time along with the retrospectively saved start point. If the gaze deviated before completion, both the timer and dwell start were reset.

4.5 Procedure

Upon arrival, participants signed an informed consent form. EEG cap size was selected based on head measurements, and the cap was soaked in a water-based electrolyte solution for 10 minutes. While the cap was soaking, participants were introduced to the Varjo XR-4 headset (HMD) and shown the eye-tracking calibration sequence. Participants received a brief explanation of the experimental goals along with standardized instructions. They were guided to keep slow head movements, minimize overall movement, maintain a relaxed facial expression, and avoid speaking during the experimental blocks to prevent interference with EEG signals. They were also advised to work at a comfortable pace and focus fully on the tasks.

Once the cap was ready, it was positioned on the participant’s head using tape measurements to ensure correct electrode placement according to the international 10–20 system. Electrode impedance was then reduced to below 50 k Ω using BrainVision Recorder software (Brain Products GmbH, 2024) (Figure 2D), with additional electrolyte solution applied as needed and hair gently adjusted to optimize contact between electrodes and the scalp. The full EEG preparation procedure took approximately 30–40 minutes.

Participants completed the experiment in a series of visual-only blocks. No audio stimuli were presented at any point, and background noise in the testing environment was minimized to reduce distractions. After each block, participants were given a 5–10-minute break, during which they could remove the HMD and relax. During breaks, EEG impedances were re-checked, and more electrolyte solution was applied if needed. Additionally, the eye tracker was recalibrated at the beginning of each block or any time the headset was repositioned.

The entire session lasted approximately 2 hours, including setup and breaks. Participants were compensated at a rate of €12 per hour.

4.6 Task

With the novel challenge of detecting SPN in immersive environments, we developed a Mixed Reality (MR) setup in Unity 3D, where participants could either click icons or passively observe them. The experiment followed a 4-condition \times 3-scene within-subjects design (Figure 1), featuring realistic UIs that reflect everyday MR use cases.

All icons were designed with distinct background and foreground elements and standardized as rounded square buttons with varied colors. This design was informed by prior research showing that varied icon colors combined with a rounded square border significantly improve visual search efficiency, reduce attentional effort, and enhance user satisfaction during icon-based tasks (Liu et al., 2021).

Participants performed a gaze-based icon selection or finding task across three immersive scenes, plus a separate training scene with basic icons to familiarize them with the task and condition structure.

Participants engaged in 3D MR-based icon selection and finding tasks across three everyday scenes. Icon layouts remained consistent across blocks to support learning and familiarity, and only the training scene had different icon positions. This design choice was motivated by evidence that repeated exposure and consistent feedback can foster stronger anticipatory negativity, as practice is known to enhance intentional regulation in slow cortical potentials (Neumann et al., 2004; Reddy et al., 2024).

Each scene contained five distinct icons that represented various functions within the environment. Specifically, the App Menu included icons for Music, TV, Settings, Safari, and Files; the Document Editor had icons for Undo, Redo, Save, Export, and Close; the Video Player featured icons for Rewind, Stop, Pause, Play, and Forward; and the Training scene displayed basic shapes

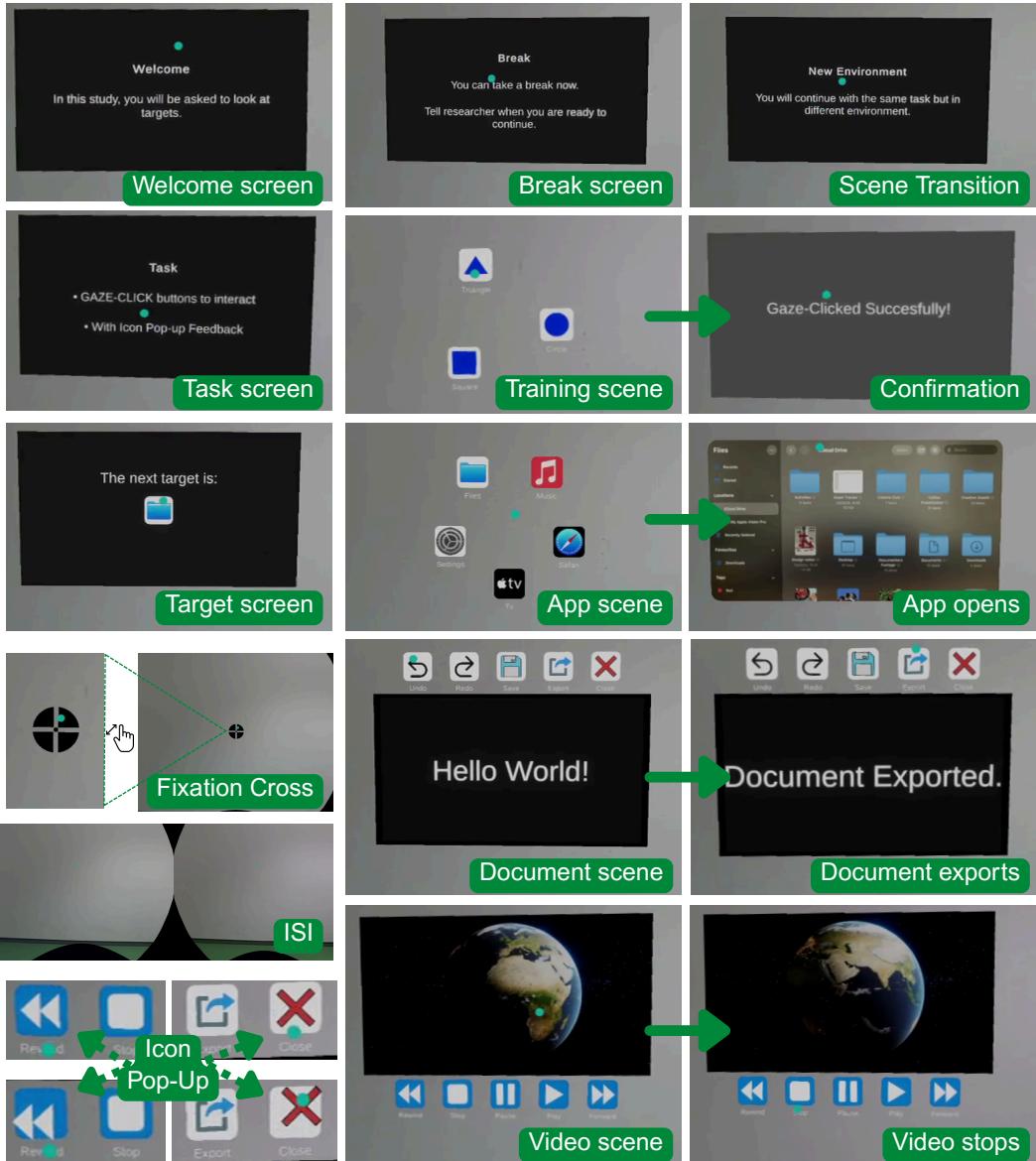


Figure 3: Mixed Reality and User Interface. The top row and left column display the information screens. Below, the left column shows examples of the fixation cross, inter-stimulus interval (Inter-Stimulus Interval (ISI)), and icon pop-up (feedback). The lower four images in the middle column illustrate different MR scenes, with corresponding examples of clicking interactions shown to their right.

such as Square, Triangle, and Circle.

Dwell-Based Gaze trigger. The HMD’s eye-tracker provided real-time coordinates of the gaze origin and the normalized gaze direction. Using Unity’s *Raycast* custom *GazeController* function, we projected a virtual ray from the gaze origin in the direction of this normalized vector. The point at which this ray intersected an object in the environment indicated the user’s current visual focus. When the user’s gaze landed on a target icon, the ray intersected the target’s collider, *IconButton*: custom script starting the countdown for next

interactions.

We adopted a 750 ms dwell threshold, in the line of Reddy et al. (Reddy et al., 2024) as a balanced point between previously used 500 ms and 1000 ms thresholds in previous research (Shishkin et al., 2016).

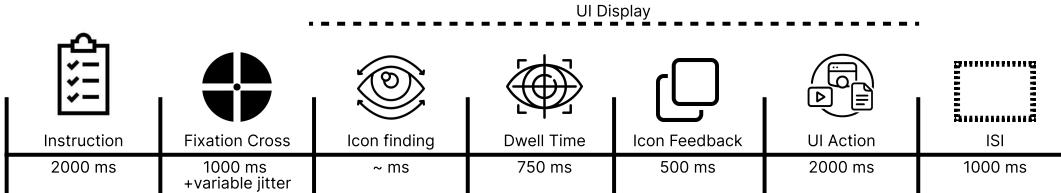


Figure 4: Trial Flow. Full sequence of the trial flow, presented from left to right.

4.7 Trial Structure

The full trial sequence followed these steps (see Figure 4):

1. **Instruction Phase:** The target icon was shown for 2000 ms.
2. **Fixation Cross:** Presented at the center for 1000 ms plus a random jitter (250, 500, or 750 ms). Fixation on the cross was required to continue. The design of the fixation cross was adapted from best practices outlined by Thaler et al. (2012).
3. **Icon Search Phase:** Participants scanned the scene to locate the instructed target icon.
4. **Interaction Phase:** Once the participant’s gaze first landed on the correct target icon, a dwell countdown of 750 ms began. If the gaze remained stable on the target for the full duration, the interaction was triggered. Up to this point, the trial was identical across all conditions. The following phases varied depending on condition:
 5. *Icon Feedback Phase (Conditional):* In ‘with feedback’ conditions, the target icon’s foreground part visibly animated out (PopUp), lasting 500 ms after the participant’s dwell finished. The pop-up feedback design was inspired by VisionOS.
 6. *UI Action Phase (Conditional):* In ‘Select’ interaction conditions, a simulated UI action was triggered following the pop-up (e.g., document editing, app opening, or video playback). This phase lasted 2000 ms.

7. **Inter-Stimulus Interval (ISI):** A blank screen was shown for 1000 ms before the next trial began.

If participants failed to fixate during the cross or icon phases for more than 5 seconds, a timeout was triggered and the trial restarted from the instruction phase to complete the same trial.

4.8 Data Processing

All EEG analysis was performed offline using the data collected during the experiments. All data streams, including EEG, eye tracking, and gaze events, were synchronized through the LSL and dwell trigger times (Kothe & Contributors, 2024).

4.8.1 Ocular Artifact Detection: Saccades and Blinks

Eye-tracking data were collected using Varjo XR headsets with built-in eye trackers, yielding an average sampling rate of \sim 66 Hz. Gaze data were exported as CSV files per participant, which were merged, sorted by timestamp, and cleaned by removing duplicates and rows where gaze tracking was invalid (`gaze_status` = 0). Additionally, the `left_status` and `right_status` fields were used to identify periods where eyes were visible but not reliably tracked (e.g., during blinks or saccades, marked with status = 1).

Blinks were detected using a combination of signal dropout and eye openness metrics. Specifically, gaps in valid gaze data ranging from 50 to 700 ms were labeled as blinks. Shorter gaps between 25 and 50 ms were treated as noise and interpolated linearly, whereas gaps longer than 700 ms were classified as signal dropouts. An additional blink detection method relied on eye openness, calculated from the average openness of both eyes. Blinks were inferred when this value exceeded a threshold of 0.9.

Saccades and fixations were identified using a velocity-based detection algorithm adapted from the EYE-EEG toolbox (Dimigen et al., 2011; Engbert & Kliegl, 2003; Engbert & Mergenthaler, 2006). Gaze velocities were computed using the projected gaze coordinates (`gaze_projected_to_right_view_x/y`) after excluding frames with unrealistic temporal resolution (i.e., steps < 1 ms). Saccades were defined as time windows where gaze velocity exceeded six times the participant's median velocity. To remove implausible saccade detections, additional constraints were applied: durations between 12 and 100 ms, amplitudes between 1 and 30° , and peak velocities between 30 and $1000^\circ/\text{s}$. Fixations were identified as all remaining gaze samples that were not labeled as

blinks or saccades and that fell within a valid spatial window. Fixation durations were restricted to a biologically plausible range of 100 to 2000 ms.

4.8.2 EEG Processing

EEG data were processed using MATLAB (version 2024b, MathWorks, Natick, MA) with the EEGLAB toolbox (Delorme & Makeig, 2004) (v2024.1), ERPLAB (Lopez-Calderon & Luck, 2014), alongside custom scripts tailored to our experiment. Raw EEG signals were imported from LSL recorded CSV files.

First, we cleaned the data by removing any long gaps (longer than 5 seconds) that included breaks or dropouts. Those gaps and their corresponding events were deleted to avoid corrupting the analysis. After this initial cleanup, the data was downsampled to 250 Hz to speed up later processing using `pop_resample`, and converted to double precision using `double`. We ran `eeg_checkset` to ensure integrity and then the cleaned datasets were saved in EEGLAB’s .set format. To get rid of large artifacts like muscle bursts or noise bursts, we specifically checked Region of Interest (ROI) channels (O1, Oz, O2) by first filtering the data with a bandpass between 0.1 and 20 Hz then calculating RMS in 500 ms windows and removing segments exceeding a robust threshold (median + 8 times MAD). These bursts were padded by 1 second on each side, and any overlapping task events were removed to keep the data clean.

ICA calculation. Ocular artifacts, such as those arising from eye movements and extraocular muscles, can contaminate neural EEG signals during complex visual search tasks (Plöchl et al., 2012). To address this, we applied independent component analysis (Independent Component Analysis (ICA)) using the extended Infomax algorithm (Delorme & Makeig, 2004).

Cleaned .set files were loaded with `pop_loadset`, and a copy was made for ICA computation. Channel locations were preserved as `urchanlocs` to maintain spatial reference. To mitigate slow drifts—especially prevalent in water-based EEG systems (Volosyak et al., 2010)—and to reduce high-frequency noise, we applied a high-pass filter at 1 Hz and a low-pass filter at 42 Hz using zero-phase FIR filters with Hann windows (`pop_firws`). Noisy or bad channels were automatically detected and removed using EEGLAB’s `clean_channels`, and excluded channels were marked accordingly.

The continuous EEG was segmented into 1-second epochs using `eeg_regepochs`. Epochs with artifacts were rejected based on joint probability metrics with a 3 SD threshold (`pop_jointprob`). This preprocessing improved ICA decom-

position by reducing contamination from artifacts. ICA weights were computed using the extended Infomax algorithm (`pop_runica` with parameters `icatype = 'runica'`, `extended = 1`, `interrupt = 'on'`, `concatenate = 'on'`), and the resulting weights, along with bad channel indices, were back-projected into the original dataset.

ICA Artifact rejection. The dataset with back-projected ICA components was automatically labeled using `pop_iclabel` with the '`default`' parameter (Pion-Tonachini et al., 2019) to classify components as brain, eye, muscle, heart, line noise, channel noise, or other. While many prior studies conservatively reject components with a 0.7–0.8 threshold for the “brain” label (H.-Z. Li et al., 2025; Pion-Tonachini et al., 2019), or retain ICs only if “brain” probability exceeds 0.5 (Callan et al., 2024), we adopted more lenient thresholds to maximize neural signal retention in our noisy MR and water-based EEG environment (Mihajlović & Peuscher, 2012; Roesler et al., 2023; Volosyak et al., 2010). Specifically, components were retained if the “brain” probability was at least 0.2, and rejected if the probability for muscle, eye, heart, line noise, channel noise, or other exceeded 0.5. These thresholds were applied uniformly across all participants to ensure a fully automated and reproducible pipeline, even though this sometimes resulted in suboptimal artifact removal for certain cases (e.g., participants 8 and 30; see Appendix).

Flagged components were removed using `pop_subcomp`, resulting in an average of 14.0 ICs removed (SD = 4.26) per participant. The remaining ICs, including ICA weights and labels, were saved to the new pre-processed file.

4.8.3 ERP Extraction

The preprocessing for ERP analysis was performed after ICA cleaning. We first removed 60 Hz line noise using the *Zapline-plus* extension for *EEGLAB* (de Cheveigné, 2020; Klug & Kloosterman, 2022). Bad channels were identified and spherically interpolated using `pop_interp`.

To preserve low-frequency activity relevant to anticipatory ERPs, we applied a bandpass filter from 0.1 Hz (high-pass) to 15 Hz (low-pass) using `pop_firws`. The 0.1 Hz high-pass filter removes slow drifts (e.g., baseline shifts) but retains slow cortical potentials such as the SPN, following recommendations for low-order filters with cutoffs around 0.1–0.5 Hz (de Cheveigné, 2020). The 15 Hz low-pass filter attenuates muscle and high-frequency noise while minimizing waveform distortion, and matches the settings used in previous ERP studies with our EEG system (Mitrevska et al., 2025).

Then, we used ERPLAB to epoch the data, extracting ERPs time-locked

to the dwell ending event, epoch window: -1000 ms to $+500$ ms. We avoided baseline correction to prevent unstable ERP results.

Epochs were created in ERPLAB, time-locked to the dwell end event, using a window from -1000 ms to $+500$ ms. We skipped baseline correction to avoid distorting low-frequency components. The ERP analysis focused on the occipitoparietal ROI (O1, Oz, O2, Iz, PO7, PO3, POz, PO4, PO8, Pz, electrodes as depicted in Figure 2), in line with prior work highlighting the SPN’s emergence in these regions (Protzak et al., 2013; Reddy et al., 2024; Shishkin et al., 2016; Zhao et al., 2021).

Event Rejection. Since Unity logs data at a fixed 60 Hz frame rate (~ 16.7 ms/frame) and lacks sub-millisecond precision, in practice, the dwell durations often landed slightly below the target 750 ms—averaging around 734 ms, consistent with being one frame short. To reduce logging base artifact, we kept only fixations within ± 17 ms (one frame) of the average dwell duration.

We included all trials that successfully triggered the interaction (i.e., 750 ms dwell), as task accuracy wasn’t relevant to our analysis. Training trials and timeout trials (when participants either forgot the target or failed to dwell long enough) were excluded.

We also removed trials where no EEG data was available within a 1-second window around the dwell trigger. To finalize artifact rejection, epochs with strong artifacts were automatically rejected using a $50\ \mu\text{V}$ sample-to-sample voltage threshold. This resulted in an average rejection rate of approximately 17.5% (SD = 5.2%) per condition across all participants. The average number of trials used for the grand average ERP for each task is shown in Table 2.

Table 2: Summary of Accepted Trials and Rejection Rates. Mean number of accepted trials used in analysis and rejection rates (%) per condition and scene, averaged across 23 participants.

Condition	Scene	Accepted	Rejected
SWF	App	528 (85.3%)	91 (14.7%)
	Document	543 (84.8%)	98 (15.2%)
	Video	515 (85.1%)	90 (14.9%)
SNF	App	523 (83.1%)	106 (16.9%)
	Document	485 (83.5%)	96 (16.5%)
	Video	523 (82.7%)	109 (17.3%)
OWF	App	538 (82.2%)	117 (17.8%)
	Document	503 (81.6%)	114 (18.4%)
	Video	521 (81.9%)	115 (18.1%)
ONF	App	478 (80.7%)	114 (19.3%)
	Document	509 (81.3%)	117 (18.7%)
	Video	481 (81.0%)	113 (19.0%)

5 Results

This section presents the main findings from our analyses of anticipatory EEG activity during gaze-based interactions in mixed reality (MR), focusing on how intention and feedback modulate the Stimulus Preceding Negativity (SPN). We evaluate our results in relation to the study’s hypotheses (H1, H2a, H2b, H2c).

5.1 Mass Univariate Analysis

To examine anticipatory EEG differences across experimental conditions, we performed one-way within-subject Mass Univariate Analyses (Mass Univariate Analysis (MUA)) using the Factorial Mass Univariate Toolbox (FMUT) (E. C. Fields, 2024). The analysis focused on the pre-stimulus dwell period (-750 to 0 ms relative to stimulus onset), sampled at 250 Hz, across ten posterior electrodes typically associated with SPN topography (O1, Oz, O2, PO7, PO3, POz, PO4, PO8, Pz, Iz) (Protzak et al., 2013; Reddy et al., 2024; Zhao et al., 2021). This resulted in 1,850 spatiotemporal comparisons (185 time points × 10 channels; n = 23).

We intentionally avoided defining the time window based on visual inspec-

tion to minimize experimenter bias (Luck & Gaspelin, 2017), allowing the analysis to reveal when and where anticipatory effects emerged. This broad, data-driven temporal scope enabled exploratory identification of condition differences without presupposing their onset or duration (Groppe et al., 2011; Maris & Oostenveld, 2007).

To correct for multiple comparisons, we employed a cluster-based permutation test using the cluster mass statistic (Bullmore et al., 1999), with the family-wise alpha set to .05. An ANOVA was computed for each spatiotemporal point using both the observed data and 10,000 random within-subject permutations (Manly, 2018). For each permutation, data points with F-values corresponding to uncorrected p-values ≤ 0.05 were grouped into clusters based on spatial adjacency (electrodes within 5.44 cm) and temporal continuity (E. Fields & Kuperberg, 2020). The cluster mass was defined as the sum of F-values within each cluster. From the 10,001 iterations (original + permutations), the largest cluster mass was used to build a null distribution. Clusters in the observed data were deemed significant if their mass exceeded the 95th percentile of the null distribution (family-wise error corrected, $p < 0.05$).

This mass univariate approach was preferred over traditional mean-amplitude ANOVAs due to its improved temporal and spatial specificity and its robustness against multiple comparisons (Groppe et al., 2011). Cluster-based MUA enables detailed localization of when and where anticipatory SPN activity differentiates conditions and is particularly effective for detecting widespread ERP effects such as the P300 (Groppe et al., 2011; Maris & Oostenveld, 2007).

Pairwise comparisons across the four experimental conditions—Select With Feedback (SWF), Select No Feedback (SNF), Observe With Feedback (OWF), and Observe No Feedback (ONF)—revealed the following:

Full Window (-750 to 0 ms): A significant negative cluster was found for SNF vs. ONF ($p = .032$) between -364 ms and -324 ms. For SWF vs. ONF, an earlier significant cluster appeared ($p = .0135$) from -752 ms to -664 ms, though this falls outside the typical SPN latency. OWF vs. ONF had a relatively low p-value ($p \geq .0819$) but no significant clusters. All other pairwise comparisons (SWF vs. SNF, SWF vs. OWF, SNF vs. OWF) were non-significant (all $p \geq .5498$).

Focused Window (-500 to -200 ms): Significant negative clusters were found for SWF vs. ONF ($p = .0388$, -376 to -324 ms), SNF vs. ONF ($p = .0135$, -364 to -324 ms), and OWF vs. ONF ($p = .0297$, -368 to -316 ms), as visualized in Figure 5. All three showed the strongest effects at occipital electrodes, particularly O1 and Oz. No significant differences were found for

SWF vs. SNF, SWF vs. OWF, or SNF vs. OWF (all $p \geq .4954$).

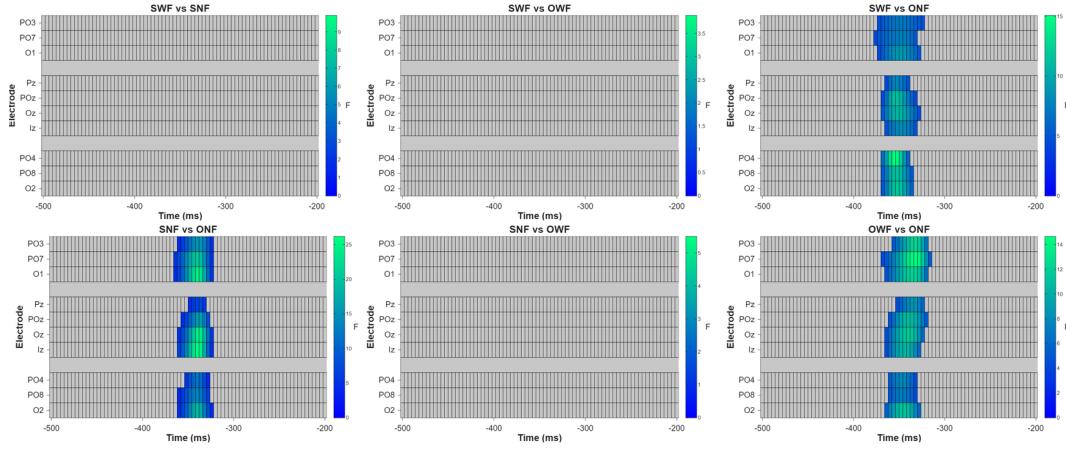


Figure 5: Mass Univariate Analysis Results. The top raster plot shows the main effect of trial condition across time and electrodes, followed by raster plots for each pairwise condition comparison. Significant clusters ($p < 0.05$) are highlighted using a color gradient scaled by the F-value, where more intense colors indicate stronger effects.

5.2 ERP Analysis

Grand average ERPs ($n = 23$), time-locked to the interaction trigger and averaged over the occipital ROI (O1, Oz, O2), revealed a clear negative-going potential in the -500 to -200 ms interval for the SWF, SNF, and OWF conditions, peaking around -300 to -400 ms (see Figure 6). The ONF condition showed no such negativity, indicating the absence of anticipatory activity.

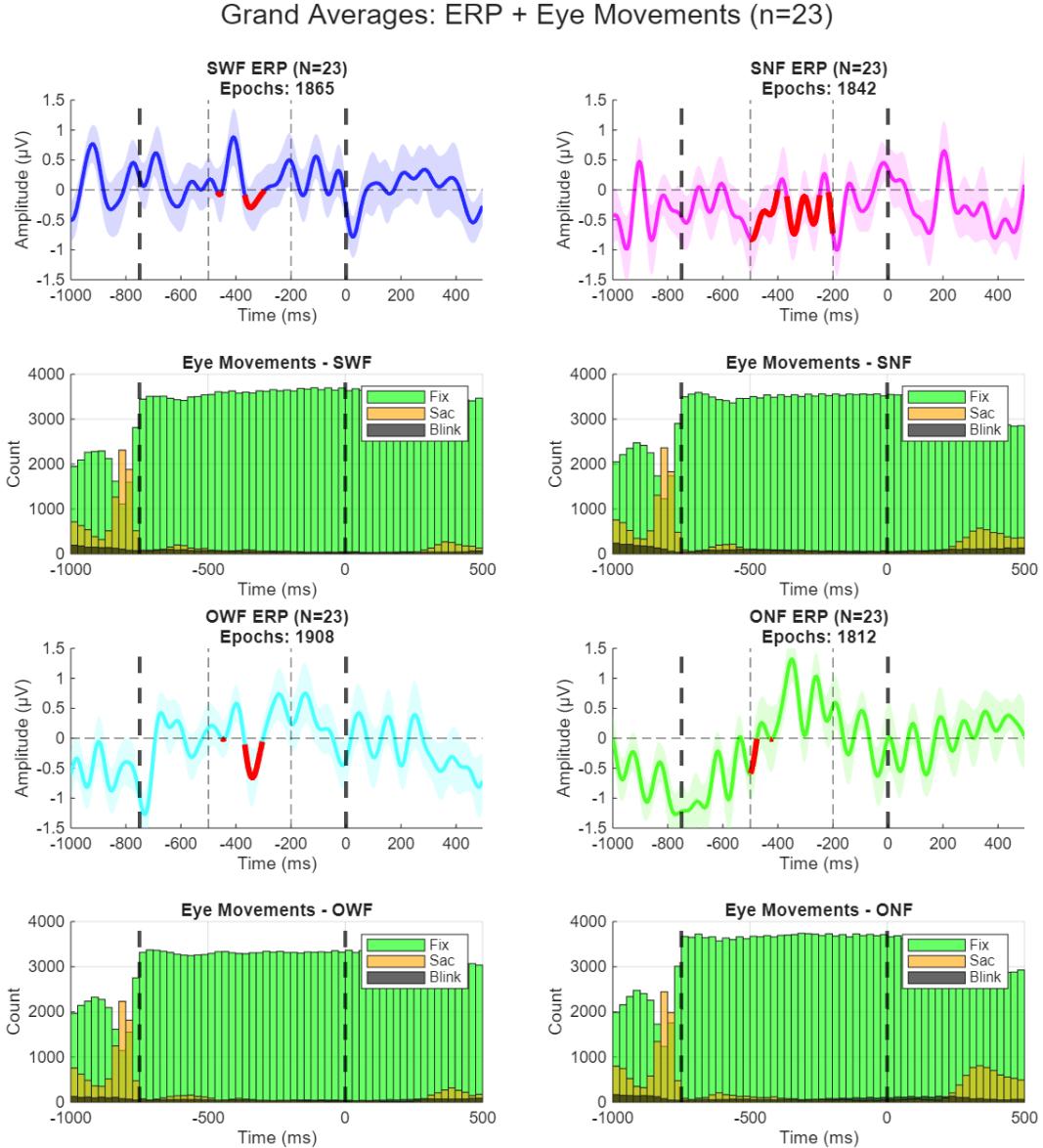


Figure 6: ERP and Eye Movement Results. Time-locked to the dwell-end trigger. The first and third rows show grand average ERP waveforms ($n = 23$), averaged over the occipital ROI (O1, Oz, O2) for each condition, with shading representing ± 1 standard error. The red-highlighted area marks the negative-going activity in the -500 to -200 ms interval.

The second and fourth rows show histograms of eye-related activity: blinks (black), saccades (yellow), and fixations (green), all time-locked to the dwell duration. Bold dashed lines indicate the -750 ms and 0 ms window used for SPN analysis.

The topographical map results (Figure 7) showed that the SNF condition elicited the broadest negative distribution over the occipital lobe among all conditions, not only for the -500 to -200 ms window but also for almost the entire dwell time. Nevertheless, consistent negative activity was observed around the -300 to -400 ms window in the occipital and slightly left-lateralized regions (around O1, P03) with the ONF and OWF conditions. The ONF condition showed strong negative activity when the dwell started, followed by a positive potential from -400 to -150 ms.

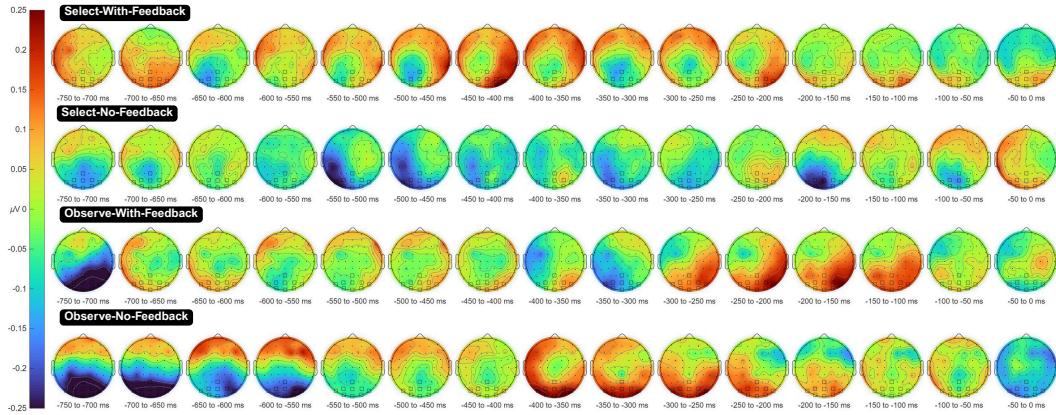


Figure 7: Topographical Distributions of Scalp Activity.

Grand average topographical distributions ($n = 23$) of mean EEG amplitude in 50 ms intervals during the -750 ms to 0 ms window, time-locked to the interaction trigger. The maps are shown for the four experimental conditions: intent-to-select with feedback (top row), intent-to-select without feedback (second row), intent-to-observe with feedback (third row), and intent-to-observe without feedback (bottom row). Channels within the region of interest are marked with square indicators. Squares indicate ROI channel locations.

These ERP and topographical map findings were consistent with the MUA results: significant SPN clusters were found in comparisons of SWF vs. ONF, SNF vs. ONF, and OWF vs. ONF, but not between SWF vs. SNF, SWF vs. OWF, or SNF vs. OWF. Topographical maps confirmed that the strongest effects were localized to occipital regions.

To assess potential confounds from eye movements, we analyzed the timing of saccades and blinks relative to the dwell period. As shown in Figure 6, saccade and blink activity mostly occurred outside the critical analysis window, suggesting minimal interference from eye movements during the gaze-based dwell period.

5.3 Hypothesis Evaluation

Hypothesis 1 (H1): *SPN will be elicited in any condition where either intention or feedback is present, across all MR scenes.* **Supported.** In the focused analysis window (-500 to -200 ms), significant SPN effects were found for Select With Feedback (SWF), Select No Feedback (SNF), and Observe With Feedback (OWF) compared to the baseline Observe No Feedback (ONF) condition. This indicates that either intention or feedback alone is sufficient to elicit SPN. The effect was consistent across all MR scenes (App, Document, Video), supporting the generalizability of SPN as a marker in naturalistic MR environments.

Hypothesis 2a (H2a): *Intention alone is sufficient to elicit SPN.* **Supported.** Both SWF and SNF conditions, which included intent to select, showed significant SPN activity compared to ONF, even when feedback was absent (in SNF). This demonstrates that internal volitional states reliably elicit anticipatory neural activity.

Hypothesis 2b (H2b): *Intention has a stronger influence on SPN than feedback expectation alone.* **Not statistically supported, though trends align.** Comparing SNF (intention only) to OWF (feedback only), no significant difference was found in the MUA results. However, ERP waveforms and topographical maps showed a trend toward stronger and earlier SPN effects in SNF compared to OWF. While this suggests intention may have a stronger effect, the lack of statistical significance means this hypothesis remains inconclusive.

Hypothesis 2c (H2c): *Feedback expectation alone can elicit a weaker SPN compared to when intention is also present.* **Not supported.** Comparing OWF (feedback only) to SWF (intention + feedback), no significant difference was found. SPN amplitude was not reliably stronger in SWF compared to OWF. Thus, while feedback alone can elicit SPN (as shown in H1), there is no evidence for additive or amplifying effects when combined.

6 Discussion

This thesis set out to investigate the detectability of Stimulus-Preceding Negativity (SPN) and its potential use as a trigger for hands-free technologies, addressing the "Midas touch" problem, where gaze leads to unintended clicks. We explored how user intention and feedback expectation modulate anticipatory neural activity during gaze-based interactions in mixed reality (MR) environments. By systematically manipulating both intent (select vs. observe) and feedback (with vs. without feedback), and measuring SPN as a neural marker of anticipation, we aimed to clarify the cognitive drivers underlying SPN in realistic MR interfaces. Our group level findings offer a potential solution to the "Midas touch" problem by demonstrating that SPN can serve as a neural trigger to distinguish between a user's intent to observe an object and their intent to select it, thereby enabling a more natural and intuitive gaze-based interface. This work extends prior findings from XR research (Protzak et al., 2013; Reddy et al., 2024; Zhao et al., 2021) to more realistic MR contexts.

6.1 SPN as a Robust Marker in MR

Our results provide clear evidence that SPN is reliably detectable at the group level in MR settings whenever either selection intent or feedback expectation is present. In the focused analysis window (-500 to -200 ms), significant SPN effects were found for Select With Feedback (SWF), Select No Feedback (SNF), and Observe With Feedback (OWF) compared to the baseline Observe No Feedback (ONF) condition. This supports Hypothesis 1 (H1), demonstrating that SPN is a robust neural signature of anticipatory processing in MR environments, and that its detectability generalizes across diverse MR scenes (App, Document, Video). The ability to elicit SPN in both selection and feedback conditions suggests that MR interfaces can leverage this neural marker for implicit intent detection and adaptive interaction, even in complex, naturalistic scenarios.

6.2 Intention and Feedback: Independent and Interactive Effects

Hypothesis 2a (H2a) posited that intention alone is sufficient to elicit SPN. Our findings support this: both SWF and SNF conditions, which included intent to select, showed significant SPN activity compared to ONF, even when feedback was absent (in SNF). This indicates that internal volitional states can

reliably elicit anticipatory neural activity, consistent with prior work in 2D and XR environments (Shishkin et al., 2016; Zhao et al., 2021). The implication is that SPN can serve as a neural indicator of goal-directed engagement, even in the absence of explicit system feedback.

Hypothesis 2b (H2b) proposed that intention has a stronger influence on SPN than feedback expectation alone. While SNF (intention only) showed a trend toward stronger and earlier SPN effects than OWF (feedback only), no statistically significant difference was found between these conditions. Thus, although the descriptive data suggest intention may have a stronger effect, the lack of statistical significance means this hypothesis remains inconclusive. This finding highlights the nuanced interplay between cognitive intention and feedback anticipation, suggesting that both factors independently contribute to SPN, but neither is strictly dominant in MR contexts.

Hypothesis 2c (H2c) suggested that feedback expectation alone can elicit a weaker SPN compared to when intention is also present. However, comparing OWF (feedback only) to SWF (intention + feedback) yielded no significant difference, indicating that SPN amplitude was not reliably stronger in SWF compared to OWF. Therefore, while feedback alone can elicit SPN, there is no evidence for additive or amplifying effects when combined with intention. This result challenges the notion of a simple summative relationship between intention and feedback, and instead points to a more complex, context-dependent interaction.

6.3 SPN Reflects Broader Anticipatory Processes

These findings challenge the traditional view that SPN is exclusively tied to selection intent. Both SNF (select, no feedback) and OWF (observe, with feedback) exhibited significant SPN activity, similar to SWF (select, with feedback), while ONF showed none. This pattern suggests that SPN is not strictly a marker of intent-to-select, but is also sensitive to the anticipation of system feedback or meaningful post-dwell events—even in the absence of explicit selection intent. In the OWF condition, participants may have experienced the feedback as a form of confirmation or reward for successfully locating the target, thereby triggering anticipatory neural activity akin to that seen in active selection trials. This aligns with recent perspectives that SPN reflects a broader class of anticipatory processes, including reward prediction and feedback expectation (Chwilla & Brunia, 1991; Donkers & van Boxtel, 2005; Donkers et al., 2005; Zander & Kothe, 2011).

The mass univariate analysis approach used in this study enabled precise

spatial and temporal localization of anticipatory effects, overcoming the limitations of mean amplitude ANOVAs used in previous work (E. Fields & Kuperberg, 2020; Groppe et al., 2011; Shishkin et al., 2016; Zhao et al., 2021). By avoiding experimenter bias in time window selection and leveraging cluster-based permutation statistics (Bullmore et al., 1999; Manly, 2018; Maris & Oostenveld, 2007), we were able to identify the earliest neural differentiation between conditions, which is crucial for rapid and reliable intent detection in real-world MR BCI applications.

6.4 Implications for MR Interface and hBCI Design

The robust detection of SPN in MR at the group level supports its generalizability and potential as a practical signal for real-world BCI applications, including accessibility solutions for users with motor impairments. SPN can serve as a reliable, implicit neural indicator of user engagement and anticipation in naturalistic MR settings, supporting the development of hands-free, confirmation-free interaction techniques. By including conditions that decoupled feedback from selection intent, our study addresses a confound overlooked in earlier research (Protzak et al., 2013; Shishkin et al., 2016; Zhao et al., 2021) and demonstrates that feedback alone, as continuous confirmation, can elicit SPN. This is consistent with evidence that temporal predictability and learned associations are necessary for SPN emergence (Chwilla & Brunia, 1991; Donkers & van Boxtel, 2005; Donkers et al., 2005), and reinforces the link between SPN and visual feedback processing (Crispin et al., 2020; Zander & Kothe, 2011).

For MR interface designers, these results suggest that both intention and feedback should be considered when developing adaptive systems that rely on neural signals for implicit interaction. SPN-based BCIs could monitor user engagement and dynamically adjust feedback to enhance user experience, as proposed by Crispin et al. (Crispin et al., 2020). The ability to detect SPN in response to both selection intent and feedback expectation opens new possibilities for confirmation-free, context-aware interaction paradigms in MR.

6.5 Limitations

While our study advances understanding of SPN in mixed reality, several limitations should be acknowledged.

The Dwell technique used for gaze-based selection may induce user fatigue (Hirzle et al., 2020; Rajanna & Hammond, 2018). While prior research

has explored optimal dwell durations (Majaranta et al., 2009; Špakov & Miniotas, 2004), a definitive threshold for minimizing fatigue remains unclear. Additionally, some participants reported discomfort due to the heaviness of the HMD, which could have impacted their experience. To address user comfort, lighter alternatives to current HMDs could be explored, such as emerging smart AR glasses, provided they can achieve 6DoF (Degree of Freedom). Furthermore, the use of more ergonomic EEG systems with localized 3-5 occipital channels should be investigated.

The analysis was performed at the group level, which may obscure important individual variability. While our group-level p-values suggest statistically reliable effects, they do not fully capture the nuances of how different participants respond to the experimental conditions. A larger and more diverse sample would also enhance the generalizability of our findings, as individual differences in cognitive processing and motivation can affect SPN amplitude and timing (Zhao et al., 2021).

We also did not develop or evaluate a classification model for single-trial SPN detection. This limits our understanding of the potential for real-time intent detection and means that the practical accuracy of distinguishing between intent-to-select and intent-to-observe interactions remains unverified.

Finally, while our design decouples feedback from selection intent, we did not investigate how other factors, such as task difficulty, user motivation, or reward value, might modulate SPN. Previous studies suggest these factors could influence SPN (Chwilla & Brunia, 1991; Donkers & van Boxtel, 2005).

6.6 Future Directions

In this work, we generalized the scenes and only investigated the condition-specific SPN detectability. Future research should address these limitations by focusing on several key areas.

First, we will conduct scene-specific ERP analysis to understand how different MR scenes modulate neural activity and to find which use cases work better for different contexts. We will analyze different scene-specific ERPs to determine which one initiates the strongest SPN. With this information, we could calibrate the system to work better in specific contexts. For example, if a scene has a higher amplitude SPN, we can calibrate the trigger to be more sensitive for that specific context. This would enable individual calibration in the system itself, allowing the trigger for a specific context to be adjusted based on the higher amplitude SPN observed in that scene.

In addition to this, we will also conduct more granular analyses at the single-

participant and single-trial levels to better understand individual variability in SPN responses. Future studies should examine neural signatures when intent-to-select and intent-to-observe trials are intermixed, reflecting more natural MR usage patterns. This will also involve exploring how different interaction scenes and UI actions influence anticipatory neural processes.

If we want to use SPN as a trigger, a crucial next step is to develop and validate classification models for single-trial SPN-based intent detection, using techniques such as LSTM or UNFOLD regression. This will enable us to assess the practical accuracy of distinguishing between intent-to-select and intent-to-observe in real-time. For robust real-world applications, we also need to develop real-time preprocessing and ocular artifact correction methods. While offline analyses can benefit from methods like OPTICAT-based ICA (Jafari-farmand et al., 2017), online applications will require more practical solutions, such as SGEYESUB (Kobler et al., 2020).

Furthermore, we plan to investigate how cognitive and contextual factors like task difficulty and reward value interact with feedback and selection intent to modulate anticipatory neural activity. We also aim to investigate the integration of natural gaze dynamics (e.g., ambient vs. focal fixations (Krejtz et al., 2016)) and explore more complex interaction paradigms, such as object manipulation or continuous control.

Finally, the extension of SPN-based intent detection to more ecologically valid MR scenarios, such as mobile and outdoor environments, will be crucial for practical deployment. As SPN-based BCIs become more capable of decoding user states, research into privacy-preserving techniques will be essential to address important ethical considerations.

7 Conclusion

In conclusion, this study demonstrates that the Stimulus-Preceding Negativity (SPN) is a reliable neural marker of anticipatory processing in mixed reality (MR), sensitive to both selection intent and feedback expectation. We showed that SPN can be reliably detected during gaze-based interactions in MR (supporting H1), but it also emerges in response to anticipated feedback, indicating that it is not solely tied to selection intent (partially supporting H2). These findings extend existing VR/XR research and provide a foundation for developing passive BCI systems that enable more intuitive and natural gaze-based interactions. The results emphasize the importance of considering both user intent and feedback in MR interface design, offering potential solutions to challenges such as the "Midas touch" problem. Moving forward, future work should focus on developing and validating classification models for real-time SPN-based intent detection, investigating scene-specific neural signatures, and addressing practical challenges in real-time preprocessing. The integration of SPN-based systems into lightweight, real-world applications, along with ethical considerations related to user privacy, will be critical for broader adoption in immersive environments.

References

- Banville, H., & Falk, T. H. (2016). Recent advances and open challenges in hybrid brain-computer interfacing: A technological review of non-invasive human research. *Brain-Computer Interfaces*, 3(1), 9–46. <https://doi.org/10.1080/2326263X.2015.1134958>
- Bashashati, A., Nouredin, B., Ward, R., Lawrence, P., & Birch, G. (2007). Effect of eye-blinks on a self-paced brain interface design. *Clinical Neurophysiology*, 118(7), 1639–1647. <https://doi.org/10.1016/j.clinph.2007.03.020>
- Bednarik, R., Vrzakova, H., & Hradis, M. (2012). What do you want to do next: A novel approach for intent prediction in gaze-based interaction. *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '12)*, 83–90. <https://doi.org/10.1145/2168556.2168569>
- Boudewyn, M. A., Luck, S. J., Farrens, J. L., & Kappenman, E. S. (2018). How many trials does it take to get a significant erp effect? it depends. *Psychophysiology*, 55(6), e13049. <https://doi.org/10.1111/psyp.13049>
- Brain Products GmbH. (2024). *Brainvision recorder user manual* [Version 1.24]. Brain Products GmbH. <https://www.brainproducts.com/>
- Bullmore, E., Suckling, J., Overmeyer, S., Rabe-Hesketh, S., Taylor, E., & Brammer, M. (1999). Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural mr images of the brain. *IEEE Transactions on Medical Imaging*, 18(1), 32–42. <https://doi.org/10.1109/42.750253>
- Callan, D. E., Torre-Tresols, J. J., Laguerta, J., & Ishii, S. (2024). Shredding artifacts: Extracting brain activity in eeg from extreme artifacts during skateboarding using asr and ica [PMCID: PMC11233536, PMID: 38989056]. *Frontiers in Human Neuroscience*.
- Chwilla, D., & Brunia, C. (1991). Event-related potentials to different feedback stimuli. *Psychophysiology*, 28(2), 123–132. <https://doi.org/10.1111/j.1469-8986.1991.tb00400.x>
- Coogan, C. G., & He, B. (2018). Brain-computer interface control in a virtual reality environment and applications for the internet of things. *IEEE Access*, 6, 10840–10849. <https://doi.org/10.1109/ACCESS.2018.2809453>
- Crispin, S., Yildiz, I., & Mulliken, G. (2020). Biofeedback method of modulating digital content to invoke greater pupil radius response [WIPO Patent (PCT) WO2020159784A1].

- David-John, B., Peacock, C., Zhang, T., Murdison, T., Benko, H., & Jonker, T. (2021). Towards gaze-based prediction of the intent to interact in virtual reality. *ACM Symposium on Eye Tracking Research and Applications (ETRA '21 Short Papers)*, Article 2. <https://doi.org/10.1145/3448018.3458008>
- de Cheveigné, A. (2020). Zapline: A simple and effective method to remove power line artifacts. *NeuroImage*, 207, 116356. <https://doi.org/10.1016/j.neuroimage.2019.116356>
- Delorme, A., & Makeig, S. (2004). Eeglab: An open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>
- Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A., & Kliegl, R. (2011). Coregistration of eye movements and eeg in natural reading: Analyses and review. *Journal of Experimental Psychology: General*, 140(4), 552. <https://doi.org/10.1037/a0023885>
- Dong, X., Wang, H., Chen, Z., & Shi, B. (2015). Hybrid brain computer interface via bayesian integration of eeg and eye gaze. *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, 150–153. <https://doi.org/10.1109/NER.2015.7146582>
- Donkers, F., Nieuwenhuis, S., & Van Boxtel, G. (2005). Mediofrontal negativities in the absence of responding. *Cognitive Brain Research*, 25(3), 777–787. <https://doi.org/10.1016/j.cogbrainres.2005.09.007>
- Donkers, F., & van Boxtel, G. (2005). Mediofrontal negativities to averted gains and losses in the slot-machine task: A further investigation. *Journal of Psychophysiology*, 19(4), 256–262. <https://doi.org/10.1027/0269-8803.19.4.256>
- Dybdal, M., San Agustin, J., & Hansen, J. (2012). Gaze input for mobile devices by dwell and gestures. *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '12)*, 225–228. <https://doi.org/10.1145/2168556.2168601>
- Engbert, R., & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research*, 43(9), 1035–1045. [https://doi.org/10.1016/S0042-6989\(03\)00084-1](https://doi.org/10.1016/S0042-6989(03)00084-1)
- Engbert, R., & Mergenthaler, K. (2006). Microsaccades are triggered by low retinal image slip. *PNAS*, 103(18), 7192–7197. <https://doi.org/10.1073/pnas.0509557103>

- Esteves, A., Verweij, D., Suraiya, L., Islam, R., Lee, Y., & Oakley, I. (2017). Smoothmoves: Smooth pursuits head movements for augmented reality. *UIST '17*, 167–178. <https://doi.org/10.1145/3126594.3126616>
- Fernandes, A., Murdison, T., & Proulx, M. (2023). Leveling the playing field: A comparative reevaluation of unmodified eye tracking as an input and interaction modality for vr. *IEEE Transactions on Visualization and Computer Graphics*, 29(5), 2269–2279. <https://doi.org/10.1109/TVCG.2023.3247058>
- Fields, E., & Kuperberg, G. (2020). Having your cake and eating it too: Flexibility and power with mass univariate statistics for erp data. *Psychophysiology*, 57(2), e13468. <https://doi.org/10.1111/psyp.13468>
- Fields, E. C. (2024). Factorial mass univariate toolbox (fmut) [Accessed: July 2025].
- Finke, A., Essig, K., Marchioro, G., & Ritter, H. (2016). Toward frp-based brain-machine interfaces—single-trial classification of fixation-related potentials. *PLoS ONE*, 11(1), e0146848. <https://doi.org/10.1371/journal.pone.0146848>
- Gemicioğlu, T., Winters, R., Wang, Y.-T., Gable, T., Paradiso, A., & Tashev, I. (2023). Gaze & tongue: A subtle, hands-free interaction for head-worn devices. *CHI EA '23*, Article 456. <https://doi.org/10.1145/3544549.3583930>
- Groppe, D., Urbach, T., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields i: A critical tutorial review. *Psychophysiology*, 48(12), 1711–1725. <https://doi.org/10.1111/j.1469-8986.2011.01273.x>
- Hirzle, T., Cordts, M., Rukzio, E., & Bulling, A. (2020). A survey of digital eye strain in gaze-based interactive systems. *ETRA '20 Full Papers*, Article 9. <https://doi.org/10.1145/3379155.3391313>
- Hong, K.-S., & Khan, M. (2017). Hybrid brain–computer interface techniques for improved classification accuracy and increased number of commands: A review. *Frontiers in Neurorobotics*, 11, 35. <https://doi.org/10.3389/fnbot.2017.00035>
- Jacob, R. (1990). What you look at is what you get: Eye movement-based interaction techniques. *CHI '90*, 11–18. <https://doi.org/10.1145/97243.97246>
- Jafarifarmand, A., Badamchizadeh, M.-A., Khanmohammadi, S., Nazari, M., & Mozaffari Tazehkand, B. (2017). Real-time ocular artifacts removal of

- eeg data using a hybrid ica-anc approach. *Biomedical Signal Processing and Control*, 31, 199–210. <https://doi.org/10.1016/j.bspc.2016.08.006>
- Jang, S., Stuerzlinger, W., Ambike, S., & Ramani, K. (2017). Modeling cumulative arm fatigue in mid-air interaction based on perceived exertion and kinetics of arm motion. *CHI '17*, 3328–3339. <https://doi.org/10.1145/3025453.3025523>
- Jensen, D., & MacDonald, K. (2022). Towards thoughtful planning of erp studies: How participants, trials, and effect sizes influence statistical power. *Psychophysiology*, 59(12), e14208. <https://doi.org/10.1111/psyp.14208>
- Kalaganis, F., Chatzilari, E., Nikolopoulos, S., Kompatsiaris, I., & Laskaris, N. (2018). An error-aware gaze-based keyboard by means of a hybrid bci system. *Scientific Reports*, 8, 13176. <https://doi.org/10.1038/s41598-018-31425-2>
- Kalika, D., Collins, L., Caves, K., & Throckmorton, C. (2017). Fusion of p300 and eye-tracker data for spelling using bci2000. *Journal of Neural Engineering*, 14(5), 056010. <https://doi.org/10.1088/1741-2552/aa776b>
- Kato, Y., Yonemura, T., Samejima, K., Maeda, T., & Ando, H. (2011). Development of a bci master switch based on single-trial detection of contingent negative variation related potentials. *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 4629–4632. <https://doi.org/10.1109/IEMBS.2011.6091146>
- Keitel, C., Benwell, C., Thut, G., & Gross, J. (2018). No changes in parieto-occipital alpha during neural phase locking to visual quasi-periodic theta-, alpha-, and beta-band stimulation. *European Journal of Neuroscience*, 48(7), 2551–2565. <https://doi.org/10.1111/ejn.13935>
- Khamis, M., Oechsner, C., Alt, F., & Bulling, A. (2018). Vrpursuits: Interaction in virtual reality using smooth pursuit eye movements. *AVI '18*, Article 18. <https://doi.org/10.1145/3206505.3206522>
- Kim, C.-H., Choi, B., Kim, D.-G., Lee, S., Jo, S., & Lee, P.-S. (2016). Remote navigation of turtle by controlling instinct behavior via human brain-computer interface. *Journal of Bionic Engineering*, 13(3), 491–503. [https://doi.org/10.1016/S1672-6529\(16\)60322-0](https://doi.org/10.1016/S1672-6529(16)60322-0)
- Kim, K.-H., Kim, G., Kim, J.-S., Son, W., & Lee, S.-Y. (2006). A biosignal-based human interface controlling a power-wheelchair for people with motor disabilities. *ETRI Journal*, 28(1), 111–114. <https://doi.org/10.4218/etrij.06.0205.0069>

- Kirst, D., & Bulling, A. (2016). On the verge: Voluntary convergences for accurate and precise timing of gaze input. *CHI EA '16*, 1519–1525. <https://doi.org/10.1145/2851581.2892307>
- Klamka, K., Siegel, A., Vogt, S., Göbel, F., Stellmach, S., & Dachselt, R. (2015). Look & pedal: Hands-free navigation in zoomable information spaces through gaze-supported foot input. *ICMI '15*, 123–130. <https://doi.org/10.1145/2818346.2820751>
- Klug, M., & Kloosterman, N. (2022). Zapline-plus: A zapline extension for automatic and adaptive removal of frequency-specific noise artifacts in m/eeg. *Human Brain Mapping*, 43(9), 2743–2758. <https://doi.org/10.1002/hbm.25832>
- Kobler, R., Sburlea, A., Lopes-Dias, C., Schwarz, A., Hirata, M., & Müller-Putz, G. (2020). Corneo-retinal-dipole and eyelid-related eye artifacts can be corrected offline and online in electroencephalographic and magnetoencephalographic signals. *NeuroImage*, 218, 117000. <https://doi.org/10.1016/j.neuroimage.2020.117000>
- Kothe, C. A., & Contributors, L. (2024). Lab streaming layer (lsl) [Accessed: July 2025].
- Krejtz, K., Duchowski, A., Krejtz, I., Szarkowska, A., & Kopacz, A. (2016). Discerning ambient/focal attention with coefficient k. *ACM Transactions on Applied Perception*, 13(3), Article 11. <https://doi.org/10.1145/2896452>
- Krol, L., Freytag, S.-C., & Zander, T. (2017). Meyendtris: A hands-free, multimodal tetris clone using eye tracking and passive bci for intuitive neuroadaptive gaming. *Proceedings of the 19th ACM International Conference on Multimodal Interaction (ICMI '17)*, 433–437. <https://doi.org/10.1145/3136755.3136805>
- Kytö, M., Ens, B., Piumsomboon, T., Lee, G., & Billinghurst, M. (2018). Pinpointing: Precise head- and eye-based target selection for augmented reality. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*, 1–14. <https://doi.org/10.1145/3173574.3173655>
- Li, G., Anguera, J., Javed, S., Khan, M., Wang, G., & Gazzaley, A. (2020). Enhanced attention using head-mounted virtual reality. *Journal of Cognitive Neuroscience*, 32(8), 1438–1454. https://doi.org/10.1162/jocn_a_01560
- Li, H.-Z., Yang, J.-J., Lv, Z., Wan, L.-Y., Wang, W., Li, D.-Q., Zhou, D.-D., & Kuang, L. (2025). Eeg dataset from playing multiplayer online battle

- arena games in natural settings. *Scientific Data*, 12, 1129. <https://doi.org/10.1038/s41597-025-05435-5>
- Li, Z., Akkil, D., & Raisamo, R. (2019). Gaze augmented hand-based kinesthetic interaction: What you see is what you feel. *IEEE Transactions on Haptics*, 12(2), 114–127. <https://doi.org/10.1109/TOH.2019.2896027>
- Liu, W., Cao, Y., & Proctor, R. W. (2021). How do app icon color and border shape influence visual search efficiency and user experience? evidence from an eye-tracking study. *International Journal of Industrial Ergonomics*, 84, 103160. <https://doi.org/10.1016/j.ergon.2021.103160>
- Lopez-Calderon, J., & Luck, S. (2014). Erplab: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, 8, 213. <https://doi.org/10.3389/fnhum.2014.00213>
- Lozano-Soldevilla, D. (2018). On the physiological modulation and potential mechanisms underlying parieto-occipital alpha oscillations. *Frontiers in Computational Neuroscience*, 12, 23. <https://doi.org/10.3389/fncom.2018.00023>
- Lu, X., Yu, D., Liang, H.-N., & Goncalves, J. (2021). Itext: Hands-free text entry on an imaginary keyboard for augmented reality systems. *UIST '21*, 815–825. <https://doi.org/10.1145/3472749.3474788>
- Luck, S., & Gaspelin, N. (2017). How to get statistically significant effects in any erp experiment (and why you shouldn't). *Psychophysiology*, 54(1), 146–157. <https://doi.org/10.1111/psyp.12639>
- Lystbæk, M., Rosenberg, P., Pfeuffer, K., Grønbæk, J., & Gellersen, H. (2022). Gaze-hand alignment: Combining eye gaze and mid-air pointing for interacting with menus in augmented reality. *Proc. ACM Hum.-Comput. Interact.*, 6, Article 145. <https://doi.org/10.1145/3530886>
- Majaranta, P., Ahola, U.-K., & Špakov, O. (2009). Fast gaze typing with an adjustable dwell time. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*, 357–360. <https://doi.org/10.1145/1518701.1518758>
- Manly, B. (2018). *Randomization, bootstrap and monte carlo methods in biology*. Chapman; Hall/CRC. <https://doi.org/10.1201/9781315273075>
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of eeg- and meg-data. *Journal of Neuroscience Methods*, 164(1), 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>
- McCullagh, P., Galway, L., & Lightbody, G. (2013). Investigation into a mixed hybrid using ssvep and eye gaze for optimising user interaction within a virtual environment. In *Universal access in human-computer interaction*

- tion. design methods, tools, and interaction techniques for einclusion: 7th international conference, uahci 2013 (pp. 530–539). https://doi.org/10.1007/978-3-642-39188-0_57
- Mihajlović, V., & Peuscher, J. (2012). To what extent can dry and water-based eeg electrodes replace conductive gel ones? a steady state visual evoked potential brain-computer interface case study. *Proceedings of the 34th Annual International Conference of the IEEE EMBS*. https://www.researchgate.net/publication/254762634_To_what_extent_can_dry_and_water-based_EEG_electrodes_replace_conductive_gel_ones_A_Steady_State_Visual_Evoked_Potential_Brain-Computer_Interface_Case_Study
- Mingai, L., Shuoda, G., Guoyu, Z., Yanjun, S., & Jinfu, Y. (2015). Removing ocular artifacts from mixed eeg signals with fastkica and dwt. *Journal of Intelligent & Fuzzy Systems*, 28(6), 2851–2861. <https://doi.org/10.3233/JIFS-151564>
- Mitrevska, T., Chiossi, F., & Mayer, S. (2025). Erp markers of visual and semantic processing in ai-generated images: From perception to meaning. *CHI EA '25: Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, Article 209, 1–7. <https://doi.org/10.1145/3706599.3719907>
- Mutasim, A., Batmaz, A., & Stuerzlinger, W. (2021). Pinch, click, or dwell: Comparing different selection techniques for eye-gaze-based pointing in virtual reality. *ETRA '21 Short Papers*, Article 15. <https://doi.org/10.1145/3448018.3457998>
- Neumann, N., Hinterberger, T., Kaiser, J., Leins, U., Birbaumer, N., & Kübler, A. (2004). Automatic processing of self-regulation of slow cortical potentials: Evidence from brain-computer communication in paralysed patients. *Clinical Neurophysiology*, 115(3), 628–635. <https://doi.org/10.1016/j.clinph.2003.10.030>
- Nguyen, W., Gramann, K., & Gehrke, L. (2023). Modeling the intent to interact with vr using physiological features. *IEEE Transactions on Visualization and Computer Graphics*, 1–8. <https://doi.org/10.1109/TVCG.2023.3308787>
- Penkar, A., Lutteroth, C., & Weber, G. (2012). Designing for the eye: Design parameters for dwell in gaze interaction. *OzCHI '12*, 479–488. <https://doi.org/10.1145/2414536.2414609>
- Pfurtscheller, G., Allison, B., Bauernfeind, G., Brunner, C., Solis Escalante, T., Scherer, R., Zander, T., Mueller-Putz, G., Neuper, C., & Birbaumer,

- N. (2010). The hybrid bci. *Frontiers in Neuroscience*, 4, 1283. <https://doi.org/10.3389/fnpro.2010.00003>
- Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). Iclabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*, 198, 181–197. <https://doi.org/10.1016/j.neuroimage.2019.05.026>
- Plöchl, M., Ossandón, J., & König, P. (2012). Combining eeg and eye tracking: Identification, characterization, and correction of eye movement artifacts in electroencephalographic data. *Frontiers in Human Neuroscience*, 6, 278. <https://doi.org/10.3389/fnhum.2012.00278>
- Plopski, A., Hirzle, T., Norouzi, N., Qian, L., Bruder, G., & Langlotz, T. (2022). The eye in extended reality: A survey on gaze interaction and eye tracking in head-worn extended reality. *ACM Computing Surveys*, 55(3), Article 53. <https://doi.org/10.1145/3491207>
- Protzak, J., Ihme, K., & Zander, T. (2013). A passive brain-computer interface for supporting gaze-based human-machine interaction. In *Universal access in human-computer interaction. design methods, tools, and interaction techniques for einclusion: 7th international conference, uahci 2013* (pp. 662–671). https://doi.org/10.1007/978-3-642-39188-0_71
- Putze, F., Weiß, D., Vortmann, L.-M., & Schultz, T. (2019). Augmented reality interface for smart home control using ssvep-bci and eye gaze. *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 2812–2817. <https://doi.org/10.1109/SMC.2019.8914390>
- Rajanna, V., & Hammond, T. (2018). A fitts' law evaluation of gaze input on large displays compared to touch and mouse inputs. *Workshop on Communication by Gaze Interaction*, 1–5.
- Reddy, G., Proulx, M., Hirshfield, L., & Ries, A. (2024). Towards an eye-brain-computer interface: Combining gaze with the stimulus-preceding negativity for target selections in xr. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)*, Article 376, 1–17. <https://doi.org/10.1145/3613904.3641925>
- Reiter, K., Pfeuffer, K., Esteves, A., Mittermeier, T., & Alt, F. (2022). Look & turn: One-handed and expressive menu interaction by gaze and arm turns in vr. *ETRA '22*, Article 66. <https://doi.org/10.1145/3517031.3529233>
- Roesler, D., Terry, J., & Kappenman, E. (2023). Comparing the performance of a water-based electrode eeg system with a conventional gel-based

- system. *Brain and Behavior*, 13(2), e2889. <https://doi.org/10.1002/brb3.2889>
- Sharma, M., Rekrut, M., Alexandersson, J., & Krüger, A. (2022). Towards improving eeg-based intent recognition in visual search tasks. *International Conference on Neural Information Processing*, 604–615. <https://doi.org/10.1145/3577190.3614166>
- Shishkin, S., Nuzhdin, Y., Svirin, E., Trofimov, A., Fedorova, A., Kozyrskiy, B., & Velichkovsky, B. (2016). Eeg negativity in fixations used for gaze-based control: Toward converting intentions into actions with an eye-brain-computer interface. *Frontiers in Neuroscience*, 10, 528. <https://doi.org/10.3389/fnins.2016.00528>
- Sibert, L., & Jacob, R. (2000). Evaluation of eye gaze interaction. *CHI '00*, 281–288. <https://doi.org/10.1145/332040.332445>
- Sidenmark, L., Clarke, C., Newn, J., Lystbæk, M., Pfeuffer, K., & Gellersen, H. (2023). Vergence matching: Inferring attention to objects in 3d environments for gaze-assisted selection. *CHI '23*, Article 257. <https://doi.org/10.1145/3544548.3580685>
- Špakov, O., Isokoski, P., & Majaranta, P. (2014). Look and lean: Accurate head-assisted eye pointing. *ETRA '14*, 35–42. <https://doi.org/10.1145/2578153.2578157>
- Špakov, O., & Miniotas, D. (2004). On-line adjustment of dwell time for target selection by gaze. *NordiCHI '04*, 203–206. <https://doi.org/10.1145/1028014.1028045>
- Starker, I., & Bolt, R. (1990). A gaze-responsive self-disclosing display. *CHI '90*, 3–10. <https://doi.org/10.1145/97243.97245>
- Tanriverdi, V., & Jacob, R. (2000). Interacting with eye movements in virtual environments. *CHI '00*, 265–272. <https://doi.org/10.1145/332040.332443>
- Trejo, L., Rosipal, R., & Matthews, B. (2006). Brain-computer interfaces for 1-d and 2-d cursor control: Designs using volitional control of the eeg spectrum or steady-state visual evoked potentials. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(2), 225–229. <https://doi.org/10.1109/TNSRE.2006.875578>
- Unity Technologies. (2024). *Unity documentation*. <https://docs.unity.com>
- Varjo Technologies. (2024). *Varjo xr-4 eye tracking specifications*. <https://varjo.com>
- Volosyak, I., Valbuena, D., Malechka, T., Peuscher, J., & Gräser, A. (2010). Brain–computer interface using water-based electrodes. *Journal of Neu-*

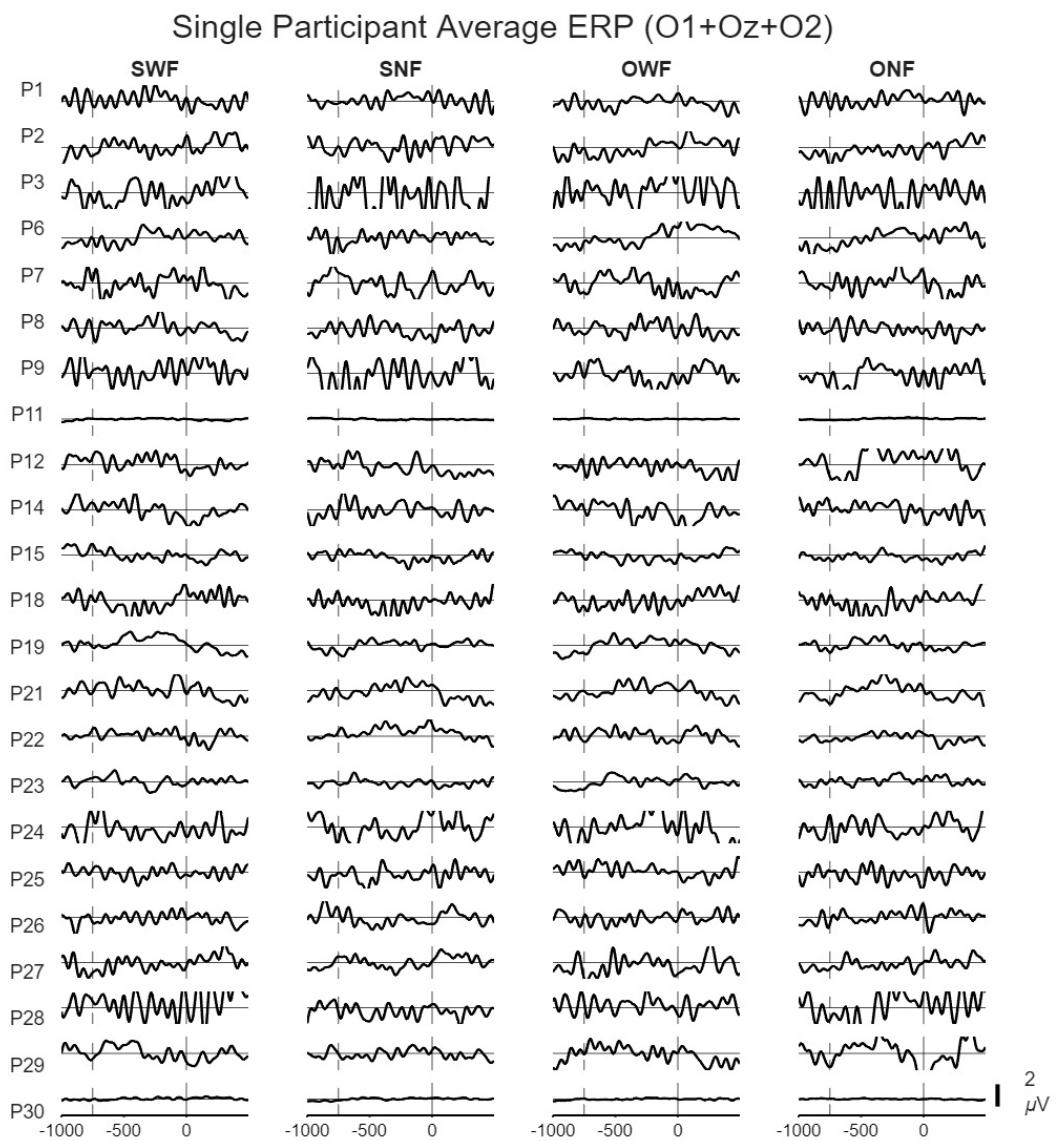
- ral Engineering, 7(6), 066007. <https://doi.org/10.1088/1741-2560/7/6/066007>
- Wagner, U., Lystbæk, M., Manakhov, P., Grønbæk, J., Pfeuffer, K., & Gellersen, H. (2023). A fitts' law study of gaze-hand alignment for selection in 3d user interfaces. *CHI '23*, Article 252. <https://doi.org/10.1145/3544548.3581423>
- Wang, B.-S., Wang, X.-J., & Gong, L.-K. (2009). The construction of a williams design and randomization in cross-over clinical trials using sas. *Journal of statistical software*, 29, 1–10. <https://doi.org/10.18637/jss.v029.c01>
- Wang, H., Li, Y., Long, J., Yu, T., & Gu, Z. (2014). An asynchronous wheelchair control by hybrid eeg–eog brain–computer interface. *Cognitive Neurodynamics*, 8, 399–409.
- Ware, C., & Mikaelian, H. (1986). An evaluation of an eye tracker as a device for computer input. *CHI '87*, 183–188. <https://doi.org/10.1145/29933.275627>
- Wei, Y., Shi, R., Yu, D., Wang, Y., Li, Y., Yu, L., & Liang, H.-N. (2023). Predicting gaze-based target selection in augmented reality headsets based on eye and head endpoint distributions. *CHI '23*, Article 283. <https://doi.org/10.1145/3544548.3581042>
- Whitlock, M., Harnner, E., Brubaker, J., Kane, S., & Szafir, D. (2018). Interacting with distant objects in augmented reality. *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 41–48. <https://doi.org/10.1109/VR.2018.8446381>
- Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clinical Neurophysiology*, 113(6), 767–791. [https://doi.org/10.1016/S1388-2457\(02\)00057-3](https://doi.org/10.1016/S1388-2457(02)00057-3)
- Xu, W., Gao, P., He, F., & Qi, H. (2022). Improving the performance of a gaze independent p300-bci by using the expectancy wave. *Journal of Neural Engineering*, 19(2), 026036. <https://doi.org/10.1088/1741-2552/ac4e6c>
- Zander, T., & Kothe, C. (2011). Towards passive brain–computer interfaces: Applying brain–computer interface technology to human–machine systems in general. *Journal of Neural Engineering*, 8(2), 025005. <https://doi.org/10.1088/1741-2560/8/2/025005>
- Zander, T., Krol, L., Birbaumer, N., & Gramann, K. (2016). Neuroadaptive technology enables implicit cursor control based on medial prefrontal cortex activity. *PNAS*, 113(52), 14898–14903. <https://doi.org/10.1073/pnas.1605155114>

- Zhao, D., Vasilyev, A., Kozyrskiy, B., Melnichuk, E., Isachenko, A., Velichkovsky, B., & Shishkin, S. (2021). A passive bci for monitoring the intentionality of the gaze-based moving object selection. *Journal of Neural Engineering*, 18(2), 026001. <https://doi.org/10.1088/1741-2552/abdfb1>

Appendix A Supplementary Information

Supplementary Table A.1: Trial Counts and Interpolated Bad Channels per Participant. Number of accepted epochs for each participant, condition, and scene, along with interpolated EEG channels (ROI channels in **bold**). Used for descriptive purposes (Appendix).

Participant	SWF			SNF			OWF			ONF			Interpolated Bad Channels
	App	Doc	Vid										
1	21	22	30	30	25	30	30	30	28	27	28	12	AF4, AF8, CPz, F5, FC2, FT7, FT8, P10, P9
2	26	28	0	28	29	28	30	26	29	26	30	26	C2, C4, CP2, Cz, FC1, FC2, FC5, T8
3	29	29	28	27	0	0	26	29	28	30	30	30	AF4, AF7, CP4, CP6, CPz, F7, FC2, FT8, I _Z , P5, P6, P9, TP7
6	29	30	30	29	9	30	30	30	29	29	29	30	C1, C6, CP6, Cz, FC2, FC5, FT7, FT8, P10, P5, P6, P7, P8, P9, PO3 , PO8 , T7, T8, TP7, TP8
7	28	30	29	29	30	30	30	28	29	30	30	30	C2, P8
8	30	30	30	28	29	28	29	29	29	29	29	29	FC2, FT8, T8
9	30	29	30	29	1	30	28	30	29	28	28	28	F3, F9, FC5, FT8
11	28	29	29	29	29	30	29	15	30	28	30	30	F8, FT7
12	30	29	28	29	30	30	30	30	30	29	28	29	C2, CP2, Cz, F7, FC1
14	25	26	26	30	30	22	25	28	29	29	28	29	F1, P10
15	30	30	30	27	29	27	30	19	30	30	29	28	AF7, C6, CPz, Cz, F10, FC6, FT7, I _Z , P6, P8, P9, TP8, PO3
18	28	29	28	28	29	29	29	29	30	30	15	28	AF8, C3, C5, F1, FC5, I _Z , T8
19	29	30	15	29	30	29	30	30	30	29	30	30	FT8, PO3 , T8
21	30	30	30	29	30	30	30	29	28	29	30	30	AF8, CP4, F1, F2, F6, F8, FC2, FC3, FC4, FC5, Fp1, P10
22	27	27	27	30	29	30	27	30	28	30	30	30	F1, F3
23	30	30	30	27	30	30	30	29	27	30	30	30	C2, Cz, F1, FC2, Fz, P10, P8, TP8, PO3
24	28	29	29	26	28	29	28	26	30	28	27	29	AF4, AF7, AF8, C1, C3, F10, F8, FC3, FT7, FT8, P7, T7, TP7
25	30	30	30	30	30	29	29	30	30	1	30	30	C4, Cz, PO3
26	30	28	24	29	28	28	25	24	29	29	26	2	AF8, C1, C2, CP3, CP5, CPz, Cz, F1, F3, FC1, FC2, FT8, Fp1, Fz, T7, TP8, PO3 , I _Z
27	26	28	28	30	30	26	28	12	0	27	28	30	CPz, F1, F4, F8, Fp2, Fz, P10, PO8
28	25	18	17	0	22	29	24	28	28	15	25	29	C2, F4, T8, TP8
29	30	30	30	29	29	30	30	28	30	0	30	6	PO3
30	0	20	27	27	25	28	28	28	26	29	30	19	



Supplementary Figure A.1: ERPs for each single participant, condition averaged.

Declaration of Originality

I hereby confirm that this thesis is entirely my own work. I confirm that no part of the document has been copied from either a book or any other source – including the internet – except where such sections are clearly shown as quotations and the sources have been correctly identified within the text or in the list of references. Moreover I confirm that I have taken notice of the ‘Leitlinien guter wissenschaftlicher Praxis’ of the University of Oldenburg.

01.08.2025, Munich

Date, Place



Signature