Real brains in virtual worlds: Validating a novel oddball paradigm in virtual reality

Jonathan W. P. Kuziek [1,a], Eden X. Redman [1,a], Abdel R. Tayem[a], Jennifer Burrell[a], Jeff Murray[c], Jenna Reinen[c], Aldis Sipolins[c], Kyle E. Mathewson[a,b]

[1] *These authors contributed equally to this work*

[a] *Department of Psychology, Faculty of Science, University of Alberta,*

[b] *Neuroscience and Mental Health Institute, Faculty of Medicine and Dentistry, University of Alberta*

[c] *IBM Research: Education & Cognitive Sciences*

Type of article: Full length research paper

Number of text pages: 28 (including references)

Number of figures: 7

Number of tables: 5

Corresponding Author:

Kyle E. Mathewson, Ph.D.

Assistant Professor of Cognitive Neuroscience

Department of Psychology

P-455 Biological Sciences

University of Alberta

Edmonton, AB, Canada, T6G 2E9

kmathews@ualberta.ca

Electroencephalography (EEG) research is typically conducted in controlled laboratory settings. This limits the generalizability to real-world situations. Virtual reality (VR) sits as a transitional tool that provides tight experimental control with more realistic stimuli. To test the validity of using VR for event-related potential (ERP) research, we used a well-established paradigm, the oddball task. Standard stimuli were presented 80% of the time and target stimuli which were responded to, 20% of the time. For our first study, we compared traditional to VR stimulus presentation using standard visual and auditory oddball tasks. We found that ERPs collected using VR head mounted displays and typical monitors were comparable on measures of latency, amplitude, and spectral composition. In a second study, we implemented a novel depth-based oddball task. We demonstrated that typical oddball ERPs elicited by the presentation of near and far stimuli. Interestingly, we observed significant differences in early ERPs components between near and far stimuli, even after controlling for the effects of the oddball task. Current results suggest that VR can serve as a valid means of stimulus presentation in novel or otherwise inaccessible environments for EEG experimentation. We demonstrated the capability of the depth-based oddball to reliably elicit P3 responses. We also found an interaction between the depth at which objects are presented and early ERP responses. Further research is warranted to better explain this influence of depth on the ERP components.

Keywords: EEG; ERP; Virtual Reality; Depth; 3D

1. Introduction

Electroencephalography (EEG) research is typically conducted in a highly controlled laboratory setting. However, this often limits the generalizability of results to real-world situations. Contemporary research has shown that alternative means of stimulus presentation such as virtual or augmented reality, can yield results comparable to traditional EEG methods. In order to identify ways to make EEG experimentation more accessible and generalizable, here we test the ue of a virtual reality head mounted display to elicit ERP components of interest in an oddball task.

Recently there has been an increase in EEG experimentation moving into more ecologically relevant environments. With this increase, there has been several notable attempts to find alternative means of stimulus presentation. Recent research by Kuziek, Shienh, & Mathewson (2017) and Scanlon, Townsend, Cormier, Kuziek, & Mathewson (2019) has demonstrated methods to increase experimental generalizability by bringing EEG techniques into new environments. These authors showed that EEG can be effectively used outside of a lab. Kuziek et al., (2017) found the event-related potentials (ERP) were comparable to those elicited from a traditional setup. Scanlon et al., (2019) demonstrated that early stage auditory processing waveforms such as the N2 vary across different conditions. Continued use of simple auditory and visual oddball tasks and measurement of the same ERP components across paradigms allow for careful validation of each step outside of the laboratory and into more generalizable environments.

While mobile EEG experimentation is valuable and addresses many challenges facing modern cognitive neuroscience, it can be limited by such practical boundaries as participant safety in potentially hazardous environments and by the presentation of certain types of stimuli remaining unfeasible. Virtual Reality (VR) allows for the presentation of otherwise inaccessible environments while retaining a high level of experimental flexibility and control. VR has over two decades of history in research, having been incorporated into psychology, neuroscience, and medical experimentation (Pugnetti et al., 1998; Sawaragi & Horiguchi, 1999; Lee et al., 2003). However, with recent advances in VR technology, techniques utilizing VR research are rapidly changing. VR has recently seen use in a wide range of research extending from remediation following stroke (Choi, Ku, Lim, Kim, & Paik, 2016; Ho et al., 2019), to

treatment for fear of public speaking (Safir, Wallach, & Bar-Zvi, 2012; Anderson et al., 2013), and PTSD (Rothbaum et al., 2014; Smith et al., 2015). Other studies have begun to employ the use of VR in cognitive neuroscience, allowing for the creation of new paradigms, most notably in embodied cognition to which there are few other methods available to elicit relevant participant states (Argelaguet, Hoyet, Trico, & Lécuyer, 2016). Research involving VR has been constrained to subjective report and behavioural assessments and subsequently would therefore benefit from concomitant brain recording.

There are two primary forms of VR based on its delivery, immersive VR and projective VR. Immersive VR allows for the subject to perceive being physically present in the virtual world; projective VR involves a large screen or monitor that is often concave which imitates the feeling of being surrounded. With VR technology rapidly progressing in recent years, the integration of head mounted displays has delivered the possibility of combining immersive VR with EEG. However, this area of research is still in its infancy, as such there are few paradigms in research that have been appropriately validated. Dankert, Heil and Pfeiffer (2013) studied stereo vision and acuity in VR. In the study, they used projective VR equipment to present animated dots to the participants, and those with stereo vision were able to see a shape emerging. However, the limitations of the screen size allowed for pixilation to occur which significantly decreased the validity of the results. The impact of VR hardware itself on a given experiment remains unclear. Results such as those from Dankert, Heil, & Pfeiffer (2013) demonstrate the lack of systematic testing of the combination of immersive VR, projective VR, and EEG.

Due to the evident lack of systematic testing of VR as a scientific tool, the goal of our first experiment (Exp. 1) was to establish the efficacy of using immersive VR and EEG simultaneously. We did this by having participants engage in identical auditory and visual oddball tasks while being immersed in VR or while using traditional non-VR methods. To further explore the uses of VR we did a second experiment (Exp. 2) in which our goal was to demonstrate that a novel VR-specific depth-based task can elicit the typical oddball P3 response. We were also interested to see whether there are discernable differences in ERPs to stimuli shown at near versus far depths.

2.0 Method – Exp. 1

2.1 Participants

A total of 24 members of the university community participated in the experiment (mean age = 23.33; age range =18-32; 9 males). Each participant completed both an auditory and a visual oddball task. A virtual environment was displayed to the participant with either a ViewPixx monitor (non-VR condition) or an HTC VIVE virtual reality headset (VR condition). All participants gave informed consent and were either given course credit for an introductory psychology course or else were given an honorarium of $10/hour for their time. The experimental procedures were approved by the internal Research Ethics Board of the University of Alberta.

2.2 Materials & Procedure

Participants completed four oddball tasks, an auditory and a visual in each of the VR and non-VR environments. The order of the stimulus presentation, non-VR or VR, was counterbalanced. Task order, auditory or visual, was also counterbalanced. Participants were all right-handed, and all had normal or corrected normal vision and no history of neurological problems. Target stimuli were presented 20% of the time while standard stimuli were presented 80% of the time. Participants were asked to sit still and fixate on a white cross presented directly in front of them that stayed constant throughout the auditory task. In the auditory task two tones at different frequencies were played, 1000 Hz for standard tones and 1500 Hz for target tones. Each tone was sampled at 44100 Hz, presented for a duration of 16 ms and contained a 2 ms linear ramp up and down. For the visual oddball task participants were again told to focus on the fixation cross while also attending to two black orbs located to the left and right of fixation. At the beginning of each trial, these orbs would flash one of two colours, either blue for targets or green for standards. The orbs would remain coloured for approximately 1.4 s until returning to black. Participants were instructed to move only their right hand to press the spacebar on a keyboard placed in front of them each time a target was presented. Following the presentation of a standard, participants were instructed to withhold any response. For both the auditory and visual oddball tasks, the time between each stimulus onset was randomly selected from an evenly distributed array of times between 1.5-2 s.

Response times were collected until 1.4 s following target onset. Figure 1A and 1B show how the auditory and visual oddball tasks were presented, respectively.

Insert Figure 1 About Here

For the VR condition, the fixation cross was approximately 1.238 units away from the participant. The program and source code can be accessed freely from the following repository: github.com/APPLabUofA/VR_Depth. The center of both stimuli presented during the visual oddball task were the same distance from the participant as the fixation cross except each orb was also 0.20 units to the left and right of the fixation. The x, y, and z dimensions of each orb were 0.25 units. Stimuli during the VR condition were displayed using an HTC VIVE headset which contains two PenTile OLED displays with a maximum resolution of 2160 x 1200 pixels and a refresh rate of 90 Hz. For the non-VR condition participants were seated 57-cm away from a ViewPixx/EEG LED monitor running at a resolution of 1920 x 1080 pixels and a refresh rate of 120 Hz with simulated-backlight rastering. Each participant completed 250 trials for each oddball experiment and condition. Stimuli were presented using a Windows 7 PC running Unity 2017.1.0. Auditory stimuli were presented using a pair of Logitech Z130 speakers kept at a constant volume level. Video output was via an Asus Striker GTX1070, and audio was output via an Asus Xonar DSX sound card. Coincident in time with sound and image onset, 8-bit TTL pulses were sent to the EEG amplifier by a parallel port cable connected to the ViewPixx monitor to mark the data for ERP averaging. The experiment was displayed to the ViewPixx monitor in both conditions to allow the experimenter to confirm the task had properly started at the beginning of each condition, and to send TTL pulses to mark stimulus onset within the recorded EEG data. We confirmed in pilot testing the reliability of these triggers compared to the headset stimulus onset using photocells (Mean = 4.8 ms; Max = 11.0ms; Min = 1.0 ms; SD = 3.0 ms).

2.3 EEG recording

During each oddball task, EEG data was collected from participants using passive, wet, low-impedance electrodes (actiCAP passive electrodes kept below 5 kΩ). Inter-electrode impedances were measured at the start of each experiment. All electrodes were arranged in the same 10-20 positions (O1,

P7, T7, P3, C3, F3, Pz, Cz, Fz, P4, C4, F4, P8, T8, O2). A ground electrode was used, positioned at AFz. Ag/AgCl disk electrodes were used, with SuperVisc electrolyte gel and mild abrasion with a blunted syringe tip used to lower impedances. Gel was applied and inter-electrode impedances were lowered to less than 5 kΩ for all electrode sites. EEG was recorded online and referenced to an electrode attached to the left mastoid. Offline, the data were re-referenced to the arithmetically derived average of the left and right mastoid electrodes.

EEG data was recorded with a Brain-Amp 16-channel amplifier (BrainVision) using identical settings across all participants. In addition to the 15 EEG sensors, two reference electrodes, and the ground electrode, vertical and horizontal bipolar EOG was recorded from passive Ag/AgCl easycap disk. EOG electrodes were affixed vertically above and below the left eye and affixed horizontally 1 cm lateral from the outer canthus of each eye. The participant's skin was cleaned using Nuprep (an exfoliating cleansing gel) before the placement of the electrodes, electrolyte gel was used to lower the impedance of these electrodes to under 5 kΩ in the same manner as previously mentioned. These bipolar channels were recorded using the AUX ports of the Brain-Amp amplifier. Data were digitized at 1000 Hz with a resolution of 24 bits. Data were filtered with online hardware with a low cut-off of 1.0 s and a high cut-off of 250 Hz, and additional online software cut-offs of 1.59155 s and 50 Hz. Each experiment was completed in a dimly lit, sound and radio frequency-attenuated chamber from Electromedical Instruments, with copper mesh covering the window. The only electrical devices in the chamber were an amplifier, speakers, keyboard, mouse, and monitor. The monitor and HTC VIVE headset ran on DC power from outside the chamber, the keyboard and mouse were plugged into USB extenders outside the chamber, and the speakers and amplifier were both powered from outside the chamber. The HTC VIVE system's two motion sensors, known as lighthouses, are mounted in the chamber in front of and to the left of the participant, as well as behind and to the right, and are raised 3 feet above the average participant's head level. They were connected to power outside of the chamber, and they emit infrared pulses at 60Hz which are detected by the VIVE headset to locate itself in 3D space. The fan for the chamber was turned on, and

nothing was plugged into the internal power outlets. Any devices transmitting or receiving radio waves (e.g., cell phones) were either turned off or removed from the chamber for the duration of the experiment.

2.4 EEG analysis techniques

Analyses were computed in Matlab R2018a using EEGLAB (Delorme & Makeig, 2004) and custom scripts. The timing of the TTL pulse was marked in the recorded EEG data and used to construct 1200 ms epochs time locked to the onset of standard and target tones with the average voltage in the first 200 ms baseline period, subtracted from the data for each electrode and trial. To remove artifacts due to amplifier blocking and other non-physiological factors, any trials with a voltage difference from baseline larger than +/- 1000 µV on any channel (including eyes) were removed from further analysis. At this time, a regression-based eye-movement correction procedure was used to estimate and remove the artifactual variance in the EEG due to blinks as well as horizontal and vertical eye movements (Gratton, Coles, & Donchin, 1983). After identifying blinks with a template-based approach, this technique computes propagation factors as regression coefficients predicting the vertical and horizontal eye channel data from the signals at each electrode. The eye channel data is then subtracted from each channel, weighted by these propagation factors, removing any variance in the EEG predicted by eye movements. Artifact rejection was again performed except for this second round of artifact rejection removed trials containing a voltage difference of +/- 500µV. Baseline correction was performed a second time following the second artifact rejection.

2.4.1 ERP analysis

Trial-averaged ERPs were derived and the MMN, a negative deflection occurring between 175–275 ms after target onset, and P3 following targets, waveforms were compared. Analysis on the MMN and P3 waveforms was done using difference waveforms, computed by subtracting the ERPs for standard tones from target tones at electrodes Fz and Pz, respectively. A one-tailed t-test was performed to understand if a significant MMN or P3 peak was observed, while a two-tailed t-test was performed to compare peaks across devices, with α set to 0.05 for all analyses. We also employed Bayesian statistics, which has been proposed as a more-informative alternative to statistical tests involving p-values (Jarosz &

Wiley, 2014; Wagenmakers, Love, et al., 2018; Wagenmakers, Marsman, et al., 2018), to directly test our null hypothesis that there is no significant difference in the EEG recordings between our devices. Bayesian analysis was done using version 0.8.6 of the JASP software (JASP Team, 2019), which is a powerful tool for Bayesian statistics (Wagenmakers, Love, et al., 2018; Wagenmakers, Marsman, et al., 2018)

2.4.2 Single trial noise

We then estimated the amount of noise in the data on individual trials in two ways. First, we computed the average frequency spectra of the baseline period in each EEG epoch. For each participant, we randomly selected 180 of their artifact-free standard trials from electrode Pz. For each trial, we computed a Fast Fourier Transform (FFT) by symmetrically padding the 600 time point epochs with zero to make a 1024-point time series for each epoch, providing frequency bins with a resolution of 0.488 Hz. Because the data are collected with an online 30 Hz low-pass filter, we plot only frequencies up to 30-Hz. Each participant's 180 spectra are then averaged together to compute participant spectra, which were then combined to form grand average spectra. To compute a second and related estimate of the noise on single trial EEG epochs, we randomly selected 180 standard tone epochs for each participant and computed the root mean square (RMS) of the baseline period on each trial. We used the 200 ms baseline period (200 time points) prior to trigger onset to avoid the influence of any evoked ERP activity on the RMS measurement. The RMS is a measure of the average absolute difference of the voltage around the baseline and is a good estimate of single trial noise in the EEG data. Smaller RMS values would suggest fewer voltage differences around the baseline and, therefore, less noise present in the data. For each trial, we average the RMS values for each EEG electrode, then averaged over trials for each participant, then computed the grand average RMS across participants (as in Laszlo, Ruiz-Blondet, Khalifian, Chu, & Jin, 2014). To estimate the distribution of RMS in our data for each device, we employed a permutation test in which 180 different epochs were selected without replacement for each participant on each of 10,000 permutations. For each of these random selections, and each electrode, we computed and recorded the grand average single trial RMS. To quantify the level of noise in the participant average ERPs, we again

employed a permutation test of the RMS values in the baseline period. In this ERP version, for each of the 10,000 permutations, we averaged the 180 standard trials that were randomly selected without replacement from the larger pool of that participant's artifact free trials for each device. We then computed the RMS of the resultant 200 time points of the ERP baseline. We averaged these RMS values over all EEG electrodes, then computed a grand average across participants.

2.4.3 ERP power

To compare the ERP statistical power as a function of the number of trials used for both the P3 and MMN, we used another permutation procedure in which we varied the number of trials contributing to the ERP average while keeping the 4 to 1 ratio of standard to target trials (Mathewson, Harrison, & Kizuk, 2017). Trial numbers were varied from 4 standards and 1 target trial, by 20 standard trials, up to 180 standard and 36 target trials, separately for each of the two stimulus types. For each number of trials, 10,000 permutations were randomly selected from the total pool without replacement. For each permutation, the selected single trials were averaged to create participant ERPs separately for target and standard tones. The difference between target and standard tones was then computed at electrode Fz between 175–275 ms (MMN) and electrode Pz between 300–550 ms (P3), and these simulated participant average ERP differences were compared to a null distribution with a standard one-tailed t-test.

3.0 Results – Exp. 1

3.1 Behavioural Results

Following artifact rejection similar trial counts were obtained for targets and standards as shown in Table 1. Figure 1C shows the response times to target stimuli during the auditory and visual oddball tasks. Results of a two-tailed t-test suggest no difference in response times between the auditory VR (Mean = 612.521 ms; SD = 128.375 ms) and non-VR conditions (Mean = 598.212 ms; SD = 123.112 ms; $t(23) = -0.519$; $p = 0.609$) but a significant difference between the visual VR (Mean = 500.293 ms; SD = 124.616 ms) and non-VR conditions (Mean = 479.360 ms; SD = 127.372 ms; $t(23) = -2.476$; $p = 0.021$).

Insert Table 1 About Here

3.2 ERP Analysis

Figure 2 shows the grand average ERPs from electrode Pz and Fz following standard and target tones during the A) auditory and B) visual oddball tasks. A P3 response can be observed following targets in both oddball tasks but an MMN response is only strongly observed following auditory targets. Also shown are the difference waveforms at electrode Fz and Pz for the A) auditory and B) visual oddball tasks. The difference waveforms were calculated by subtracting the ERPs for standards from the ERPs following targets. A large negative deflection between 300-425 ms can be observed for the auditory difference waveform, indicative of an MMN response. By comparison, a much smaller negative deflection is observed between 250-325 ms for the visual difference waveform. A distinct P3 response, signified by a large positive-going deflection, can be observed for both the auditory and visual oddball tasks between 400-550 ms and 300-550 ms respectively. Figure 2 also shows topographies and bar graphs of the averaged ERP effects across the indicated time windows, with similar patterns of activity for both devices in all but the visual MMN region. These electrodes and time windows were chosen because MMN and P3 waveforms are most prominent at these locations (Linden, 2005).

Insert Figure 2 About Here

Table 2A shows the results of the one-tailed *t*-test for the MMN and P3 waveforms at electrode Fz and Pz respectively. Results comparing the waveforms between the VR and Non-VR condition are shown in Table 2B. Hedge's *g* was used to estimate effect size, which was calculated using version 1.5 of the Measures-of-Effect-Size toolbox for Matlab (Hentschke & Stüttgen, 2011).

Insert Table 2 About Here

Results of the Bayesian paired sample *t*-tests, with the electrodes and time windows used for the MMN and P3 waveforms being consistent with those mentioned above, provided Bayes Factors (BF_{10}) of approximately 0.282 and 0.229 for the auditory oddball task while the visual oddball resulted in values of 0.217 and 0.548 respectively.

3.3 Single Trial Noise

Grand average spectra are shown in Figure 3A and 3C. At frequencies below approximately 15Hz, both conditions show similar spectra but start to deviate past 15Hz. This difference could be due to the VR headset itself along with changes induced by participants wearing the headset on top of the EEG cap and inadvertently applying pressure to several electrodes. Both devices showed the expected 1/f frequency structure in the data, as well as the typical peak in the alpha frequency range between 8 and 12 Hz (Mathewson et al., 2011) along with another peak at about 7Hz. Histograms of the single trial RMS values computed for each permutation, along with bar graphs of the mean and standard deviation, are shown in Figure 3B and 3E for the auditory and visual oddball tasks respectively. Also shown are the grand average RMS values computed in each of the 10,000 permutations for both conditions, along with a bar graph of the mean and standard deviation. Both the single trial and grand average RMS distributions suggest less noise was present during the Non-VR condition compared to the VR condition. Given the results of the grand average spectra plots, this difference should not be surprising. The VR headset likely adds electrical, radio, and infrared noise to the testing environment which would not be present in the Non-VR condition since the VR headset would be not be used.

Insert Figure 3 About Here

3.4 ERP Power

Figure 4 plots the proportion of the 10,000 permutations in which the *t*-statistic passed the significance threshold, as a function of the number of samples in each permutation. The P3 and MMN waveforms required a similar number of trials to reach significance with 80% power, although as expected the MMN waveform did not reach significance for either condition during the visual oddball task. For the MMN waveform during the auditory task, approximately 37 and 26 trials were needed for the VR and Non-VR conditions respectively. For the P3 waveform during the auditory task about 6 trials were needed to reach significance for both conditions. For the visual oddball task about 4 trials were needed for the P3 waveform to reach significance in both conditions.

Insert Figure 4 About Here

4. Discussion – Exp. 1

The goal of the first experiment was to expand on a recent trend in current literature focusing on EEG generalization and to establish validity in the use of VR technology in basic research in cognitive neuroscience, particularly those involving the oddball paradigm.

Behavioural results indicate there is a small but significant difference between the non-VR and VR in the visual oddball task. Participants slower response to 3D stimuli may present the engagement of more perceptual domains converging, another explanation could be the distraction of wearing a VR headset causes a very slight decrease in response time. However, despite these differences, ERP waveforms derived from stimuli presented in the VR and non-VR conditions were found to be nearly identical during the time windows relevant to the oddball task. Taken together these results demonstrate the versatility of VR and its applicability in cognitive psychology. It is important to note that clarity in the P3 response obtained despite three introduced sources of noise. The first of which is the addition of VR equipment which may introduce electrical noise and presents an added physical challenge to participants with added weight, which may lead to extra muscle artifacts due to neck tension. The second source of noise pertains to the different levels of engagement within the environment, differing levels of processing may be required to interpret a more immersive and realistic scene, however, this may be counteracted by virtue of the VR environment itself presenting as more engaging to the viewer and yielding more readily available attentional resources. Lastly, there are several differences between non-VR and VR stimuli in terms of viewing quality; resolution, refresh rate, and field of view.

Both systems showed the expected ERP, P3 and MMN waveforms along with similar topographies, with the greatest activity focused on electrodes Pz and Fz respectively. Bayesian paired T-tests further support that it is likely the data falls under the null hypothesis suggesting no difference between recording techniques for the MMN and P3. However, we did observe some differences in the amount of noise during both the baseline and within the ERP itself, as indicated by the differences in the RMS distributions for both systems. Despite the apparent difference in baseline and ERP noise between

the VR and non-VR, similar statistical power, frequency distributions, and ERP waveforms were observed between both systems. However, other results have shown that any observed difference between RMS noise has minimal impact on the derived ERPs and their statistical power.

Both our VR and non-VR conditions showed the expected 1/f frequency distribution although we can observe some differences between our conditions. The VR condition shows an increase in frequencies above ~17 Hz. Increased activity in this frequency range is likely due to increased muscle tension in the neck during the VR condition, due to the weight of the VR head mounted displays itself. Electrical artifacts due to additional equipment are likely another source of increased noise in the higher frequencies of the VR condition. It is foreseeable that this could impact findings that reference mid to high beta frequency bands. However, these effects can be mediated with proper experimental design. Furthermore, it is worth noting there is a consistent trend in the commercial VR industry to provide progressively lighter VR head mounted displays; there will likely be less contamination due to muscle artifacts with lighter models.

We were able to demonstrate that a commercial VR headset, a relatively inexpensive and fully immersive VR device can present oddball stimuli comparably to a research-standard non-VR monitor. We were able to show that the ERPs derived from typical auditory and visual oddball tasks are comparable when stimuli are presented using the VR and non-VR. Having confirmed that the ERPs derived from typical auditory and visual oddball tasks are comparable between VR and non-VR, we then looked to explore stimuli relevant to 3D environments. In the next experiment, we explore the addition of a new dimension, depth, to the standard visual oddball task, a task that cannot be replicated on a traditional monitor.

5. Introduction- Exp. 2

Results from Exp.1 suggests that the ERPs from stimuli presented from VR headsets can be comparable to those presented with a traditional non-VR monitor. In the following experiment, we hoped to further demonstrate the utility of the device in an EEG experiment wherein the paradigm domain is more specific to VR itself. We utilized the oddball paradigm of presenting two binary stimulus categories.

We translated the typical oddball paradigm into a novel VR-specific task by specifically modulating the domain of depth. Using the depth-based oddball, we compared ERPs to near and far standards and targets to distinguish a P3 similar to other recent adaptations of the standard oddball task (Raz, Dan, & Zysberg, 2014). The oddball paradigm has been shown to be sensitive to changes in several modalities, ranging from sensitivity to basic visual elements of stimulus, such as colour (Fonteneau & Davidoff, 2007) to higher-level processes such as emotion (Raz & Zysberg, 2014).

Depth has traditionally been studied by moving a stimulus or monitor while covering the subject's eyes which does not allow for all perceived size (retinal size) to remain constant. This effect contributes to the ease in which subjects are able to identify the near or far object due to extraneous depth cues (Bohr & Read, 2013). When using VR objects size can be controlled proportionally to keep retinal size constant at varying depths with respect to the viewer. Naceri, Moscatelli, & Chellali (2015), studied what occurred when size cues were not available as a depth cue, and found that while keeping the size of the objects constant the subjects use convergence, binocular disparity and accommodation cues to allow for depth recognition. Our study followed this principle and kept the perceptual size consistent across different depths.

6. Method – Exp. 2

6.1 Participants

A total of 18 members of the university community participated in experiment two (mean age=21.60 age range=17-43; 8 males). Participants were all right-handed, and all had normal or corrected normal vision and no history of neurological problems. All participants gave informed consent, were given course credit for their time, and the experimental procedures were approved by the internal Research Ethics Board of the University of Alberta.

6.2 Materials & Procedure

Having established the validity of presenting visual stimulus in VR, we focused on modifying the visual task in the VR condition from Exp.1. All hardware remained the same as in the VR condition described in Exp.1. Each participant completed a visual oddball task in VR where orbs were placed either

close to the participant (Near) or further away (Far). The orbs participants responded to were considered targets and orbs that participants were not responding to were considered standards, in-line with the typically oddball paradigm. The experiment was divided into two conditions, a Near condition where participants responded when nearest orbs were presented, and a Far condition where participants responded to the farthest orbs. The conditions henceforth are referred to by the location of the target stimuli Near and Far, respectively. Targets and standards were presented in the same 20% - 80% ratio as described in Exp. 1 for both Near and Far conditions.

Participants completed each condition twice, responding to near and far targets in separate blocks. Both orbs were centred and in line with the central fixation cross and both orbs were perceived to be identical in size. In other words, the farther orb was larger than the near orb but thus orbs were perceived to be the same size by the participant, such that both orbs were detected the same area of the retina, only the relative retinal disparity is modulated. Participants placed their head in a chin rest to reduce noise caused by muscle strain as well as to remove extraneous depth cues due to head movement and position, such that the only distinguishing factor in the VR environment was relative depth. Task order, whether they were responding to near or far targets, was counterbalanced. Participants were instructed to move only their right hand to press the spacebar on a keyboard placed in front of them each time a target stimulus was presented. Following the presentation of a standard stimulus, participants were instructed to withhold any response. During each trial, a fixation cross was presented for a variable duration randomly selected between 500-1000 ms, after which the fixation cross was removed, and an orb was presented. Each orb was kept black and remained on screen for 1400 ms. EEG recording, analysis, artifact rejection, and eye-blink correction procedures were identical to those used in Exp.1. Figure 5A and 5B demonstrate the design of the depth oddball task. To estimate adduction and abduction of the eyes to stimuli of different depths in the task, the position of the horizontal EOGs were modified to over of the medial and lateral rectus muscles of only the left eye.

Insert Figure 5 About Here

7. Results – Exp. 2

7.1 Behavioural Results

Trial counts for targets and standards in each condition, following artifact rejection, are shown in Table 3. Figure 5C shows the mean response times following targets in the Near Targets and Far Targets conditions. Results from a two-tailed t-test suggests no difference between response times for far targets (Mean = 593.566 ms; SD = 175.966 ms) and near targets (Mean = 560.338 ms; SD = 161.841 ms; $t(17) = -1.950$; $p = 0.068$).

Insert Table 3 About Here

7.2 ERP Analysis

Figure 6A shows the grand average ERPs from electrode Pz and Fz following standard and target tones during the Near and Far conditions of the depth oddball task. A noticeable P3 response can be observed following targets in both conditions. Figure 6B shows difference waveforms for standards and targets at electrode Fz and Pz. These waveforms were constructed by subtracting Near standards from Far standards and Near targets from Far targets to observe any potential differences between Near and Far stimuli. Deflections in these difference waveforms suggest differences in ERP activity following Near/Far standards and Near/Far targets. Such deflections are typically occurring early in the difference waveform (up to 500 ms post-stimulus onset) with target stimuli showing the greatest difference in activity between the near and far depths. Figure 6C shows the difference waveforms at electrode Fz and Pz for the Near and Far conditions. These waveforms were calculated by subtracting the ERPs for standards from the ERPs following targets. In the case of the Near condition, ERPs to the far, standard stimuli were subtracted from the ERPs following the near, target stimuli. Difference waveforms for the Far condition were generated by subtracting ERPs following Near, standard stimuli from ERPs generated after the Far, target stimuli. A large P3 response, indicated by a large positive deflection between 300-600 ms at electrode Pz, can be observed in the difference waveforms for both conditions. Interestingly, two distinct, opposing patterns of activity between our conditions can be seen prior to the onset of the P3 response. These patterns of activity at electrode Pz, along with the P3 waveform, were used to generate three time

windows that will be used for all further analysis (175-225 ms; 225-300 ms; 300-600 ms). Electrode Pz was chosen as it was expected that the P3 waveform would be greatest at this electrode and, as such, any differences between our Far and Near targets would also be greatest. Based on the grand-average difference waveforms shown in Figure 6C, we believe these earlier time windows (175-225 ms; 225-300 ms) capture the differences between our Far and Near targets.

Insert Figure 6 About Here

Table 4 shows the results of the one-tailed *t*-tests, effect size measurements, and Bayesian one-tailed *t*-tests for the difference waveforms. Electrode Pz was used for each of the time windows specified above. Table 4A shows the results of a one-tailed t-test comparing the standard and target difference waveforms from Figure 6B to zero. A right or left tailed t-test was performed based on the mean voltage of each time window. The results suggest a significant difference between Near and Far standards at the two early time windows but not the later, 300-600 ms time window. A similar pattern can also be observed between Near and Far targets with significant differences observed during the 175-225 ms and 225-300 ms time windows. Table 4B shows results for the difference waveforms generated from the Near and Far conditions. A left or right tailed *t*-test was used based on the mean voltage of each time window. Hedge's *g* was used to estimate effect size, which was calculated using version 1.5 of the Measures-of-Effect-Size toolbox for Matlab (Hentschke & Stüttgen, 2011). Results suggest that a significant P3 response is obtained during the 300-600 ms time window for both difference waveforms. There is also a significant deviation from 0µV during the 175-225 ms and 225-300 ms time windows for the Near difference waveform.

Insert Table 4 About Here

Results paired sample *t*-tests, estimates of effect size, and Bayesian paired sample *t*-tests are shown in Table 5. These analyses were done to compare our waveforms in each condition. Table 5 shows the results comparing the Near difference waveform to the Far difference waveform. Results suggest that our difference waveforms are not significantly different at the 300-600 ms time window but do differ during the 175-225 ms and 225-300 ms time windows. For the Far Targets – Near Standards difference

waveform a positive deflection is observed during the 175-225 ms time window followed by a negative deflection during the 225-300 ms time window. An opposing pattern of activity is observed for the Near targets – Far standards difference waveform with a negative deflection during the 175-225 ms time window followed by a positive deflection during the 225-300 ms time window. Such patterns of activity can also be observed for the standard and target difference waveforms at electrode Pz, as shown in Figure 6B. Contrary to the ERP plots in Figure 6C, our conditions do not differ during the 225-300 ms time window. These differences are also consistent with the results presented in Table 4, suggesting that these early modulations of the ERP waveform are likely due to the depth of the presented stimuli.

Insert Table 5 About Here

Figure 7 shows a plot of the grand-average ERPs at the HEOG, which is designed to monitor horizontal eye movements. These plots were generated to observe changes in eye movements due to changes in the depths of the stimuli; diverging eye movements were expected for far stimuli and convergent eye movements were expected for near stimuli. Plots of the grand-average standards and targets difference waveforms (as shown in Figure 6) are also shown to compare HEOG changes to ERP changes. A two-tailed t-test was performed at each time point of the HEOG and Pz ERP to understand which time points significantly differ from 0, with significant tests indicated by pink marks at the bottom of each plot. For standards, the earliest, significant point of difference occurs at 212 ms for the HEOG and for the Pz difference waveform the earliest point occurs 146 ms post-stimulus. For target stimuli, the earliest significant point of difference is at 110 ms for the HEOG and 185 ms for the Pz difference waveform.

Insert Figure 7 About Here

8. Discussion – Exp 2.

To further demonstrate the ability of VR headsets as an experimental tool, we had participants complete a third oddball task involving depth. The VR head mounted display provides several advantages compared to a non-VR monitor regarding EEG experimentation such as independent presentation of stimuli to each eye. During the oddball task in the second experiment, participants would complete two

separate blocks, responding either to the near or far target and withholding responses to the respective standard. Separately presenting stimuli to each eye allowed us to manipulate the sense of depth. By maintaining the angle and subjective size of each orb from the participant's point of view, while having a difference in the relative distance to the participant, we were able to selectively manipulate depth. Using this paradigm, we were able to demonstrate that early ERP components are sensitive to changes in stimulus depth.

Beyond a P3 response to targets at both depths, there were several interesting characteristics when comparing the Near and Far conditions. Comparing difference waves at Pz and Fz of the Near and Far conditions depicts an interaction of depth with elicited ERPs. Figure 6B summarizes these differences, particularly at Pz where the early waveform components at 175-225ms (first window) and 225-300ms (second window) vary significantly for both Standard and Target difference waves. Standards show a negative deflection in the first window and a small positive deflection in the second, while targets show the opposite with a large positive deflection in the first window and a negative deflection in the second. This distinction shows that ERPs were distinctly different when the subject perceives a target stimulus, indicating that they were able to perceive the depth of the stimuli in the VR condition. These exciting results also show that ERPs differ with depths, this supports the findings of Liu, Meng, Wu and Dang (2015). Figure 6B compares standards and targets at Fz and Pz showing the oddball effect as it is seen in the two depth conditions of targets and standards. This may not account for all interaction between the oddball stimulus and depth; further study with an extended experimental design is warranted. One such study may be a passive perceptual experiment involving all non-oddball elements at both depths, from this we would be able to gauge variance across ERPs that is only due to depth and not task demands.

Our current results support distinct physiological processes for processing information differentially across varying depths. The HEOG deflections portray an effect where the eye converges to focus on a near object and diverges to focus on a far object, which is expected. What was not previously expected was seeing the timing of this movement consistently appear alongside the ERP changes associated with that object's onset. These novel results are interesting not simply due to the eye

movement that is consistently observed with the onset of these objects which shows that depth cues are being processed by the eyes, they also serve to show that there is a temporal alignment between the brain activity and eye movement. Further investigation of how that depth information processing produces behavioural effects is needed.

However, the differential processing of two neutral stimuli that have been size-matched on the retina is not necessarily intuitive. Henderson, Vo, Chunharas, Sprague and Serences (2019) demonstrated that different depths are processed in different retinotopic areas of the brain, mainly in area V3A, and that an encoding model can be used to map the depth of objects on a z-axis, along with their positions in 2D space along vertical and horizontal axes. This provides an explanation of the different depths producing different ERPs. While this explanation points out a cognitive mapping perspective on where brain activity changes to these stimuli, it does not identify why such a mapping would be designed or beneficial. In that context, a possible explanation of this phenomenon is the evolutionary relevance of an object's distance to the viewer; responding quickly to a stimulus that is near you tends to be more important than responding quickly to stimuli farther away. We may be physiologically poised to react faster to threatening stimuli that are closer, as seen in the earlier onset of eye movements in HEOGs for near stimuli than for far stimuli. The current study may present a generalization of this propensity in detecting and responding to non-threatening stimuli.

Another potential explanation is that differences between Near and Far difference waveforms may depict a perceptual crossing of a distinguishing boundary between peripersonal space (PPS) and extraperipersonal space (EPS), as described by (di Pellegrino & Làdavas, 2015; Làdavas, 2002). PPS refers to a representation of space around an individual that is both malleable and reactive to task demands. There is evidence of distinct regions within this space (de Vignemont & Iannetti, 2015), and known factors demonstrated to change modify the personal space (Canzoneri et al., 2013). Differences between the Near and Far conditions, as shown in the difference waves of Figure 6C, may be interpreted as an interaction with distinct regions of peripersonal space – near objects could be perceived to be within PPS and thus is more reactive than a far object that could be perceived to land in EPS.

With all of these observations and proposed explanations in mind, we must also consider possible interferences and unintended behaviours that could confound these findings. Eye movements that are elicited from other stimuli than the objects presented or participant intent could interact with the effects seen with the HEOGs. Changes in the exact viewing experience of each participant due to potential image blur, visual acuity, and headset positioning on head shapes and sizes that are inherently different between participants are other potential sources of variability. However, measures were taken to minimize these effects, such as adjusting the interpupillary distance of the lenses in the headset, verbal instructions to maintain eye position on the fixation cross, elimination of stimulus shadows and other such depths cues, and readjusting headset size to best fit each participant.

9. Conclusions

Our finding of similar ERPs for traditional and VR tasks in Exp. 1 show the effectiveness of using a VR head mounted display for ERP research. These results also demonstrate the generalizability of the P3 response, as well as its relevance and applicability in VR. We were also able to demonstrate the flexibility of VR in constructing a novel version of the oddball task, with near and far stimuli eliciting a P3 response. This establishes the utility of VR in creating new experimental paradigms within the scope of the P3 response.

If results can be validated with other tasks, movement artifacts, and without participants in a chin rest, there may be a basis for conducting indoor mobile VR experiments. Leveraging recent advances in mobile EEG (Aspinall, Mavros, Coyne, & Roe, 2015; Zink, Hunyadi, Huffel, & Vos, 2016; Scanlon et al., 2019) with advances in the wireless VR industry, there may be space to harness the combination of VR with more mobile experiments within the laboratory.

To investigate the findings from Exp. 2 regarding differences in the Near and Far difference waves, we are currently running a similar experiment in which we present to the participant five size-matched orbs at different depths. The goal is to determine if there is a gradient-like depth effect, such that we expect the three additional depths to fall sequentially between the nearest and farthest orbs.

In future studies, there is room to investigate how other domains interact with the depth oddball. Results may mirror comparable studies of overlapping visual and auditory modalities in the oddball task (Brown, Clarke, & Barry, 2007; Campanella et al., 2010; Robinson, Ahmar, Sloutsky, Robinson, & Sloutsky, 2010). Brown et al. (2007) depicts how auditory and visual oddballs are paired to describe two distinct stages of auditory processing, an early modality-dependent stage and a later context-dependent stage. In addition to auditory dual tasks, we may integrate other visual domains in the depth oddball task such as colour, for example. This method would allow us to ascertain the flow and integration of depth specific information in concert with other modalities.

In line with research into peripersonal space, there is the opportunity to investigate depth-oddball sensitivity to known modulatory factors of peripersonal space. Such as gauging the influence that target threat may have on the relative discrimination of near and far oddballs. In other words, how would different threats affect one's mapping of peripersonal space? Depending on the threat content of the orbs in depth, there is reason to believe we could activate different PPS maps altogether (de Vignemont & Iannetti, 2015). One paradigm to introduce an element of threat while maintaining neutral orb presentation would be to offer a monetary sum at the beginning of the experiment. A loss is then accrued from the initial sum if response times fall below a certain baseline-derived threshold. A comparison of adherence to thresholds and resultant ERPs for each Near and Far conditions both with and without threat may reveal interactions of threat with different regions of personal space. The increased flexibility provided by VR systems has the potential to increase the breadth of EEG studies that can be conducted, greatly improving the external validity of EEG experimentation.

Here we show that a VR head mounted display offers a great tool into a wide variety of pre-existing experimental paradigms. Some topics that may readily benefit from a three-dimensional environment include visual attention, working memory, learning, and will certainly allow for more accurate and realistic models of human navigation. VR as a tool in brain imaging studies offers increased generalizability and flexibility compared to traditional non-VR monitors This is achieved without

sacrificing experimental control or introducing highly variable noise artifacts common with using EEG in environments outside of the lab.

References

Anderson, P. L., Price, M., Edwards, S. M., Obasaju, M. A., Schmertz, S. K., Zimand, E., & Calamaras, M. R. (2013). Virtual reality exposure therapy for social anxiety disorder: A randomized controlled trial. Journal of Consulting and Clinical Psychology. https://doi.org/10.1037/a0033559

Argelaguet, F., Hoyet, L., Trico, M., & Lécuyer, A. (2016). The role of interaction in virtual embodiment: Effects of the virtual hand representation. Proceedings - IEEE Virtual Reality. https://doi.org/10.1109/VR.2016.7504682

Aspinall, P., Mavros, P., Coyne, R., & Roe, J. (2015). The urban brain: Analysing outdoor physical activity with mobile EEG. British Journal of Sports Medicine. https://doi.org/10.1136/bjsports-2012-091877

Bohr, I., & Read, J. C. A. (2013). Stereoacuity with Frisby and Revised FD2 Stereo Tests. PLoS ONE, 8(12), e82999. doi:10.1371/journal.pone.0082999

Brown, C. R., Clarke, A. R., & Barry, R. J. (2007). Auditory processing in an inter-modal oddball task: Effects of a combined auditory/visual standard on auditory target ERPs. International Journal of Psychophysiology. https://doi.org/10.1016/j.ijpsycho.2007.04.002

Campanella, S., Bruyer, R., Froidbise, S., Rossignol, M., Joassin, F., Kornreich, C., … Verbanck, P. (2010). Is two better than one? A cross-modal oddball paradigm reveals greater sensitivity of the P300 to emotional face-voice associations. Clinical Neurophysiology. https://doi.org/10.1016/j.clinph.2010.04.004

Canzoneri, E., Ubaldi, S., Rastelli, V., Finisguerra, A., Bassolino, M., & Serino, A. (2013). Tool-use reshapes the boundaries of body and peripersonal space representations. Experimental Brain Research, 228(1), 25–42. https://doi.org/10.1007/s00221-013-3532-2

Choi, Y.-H., Ku, J., Lim, H., Kim, Y. H., & Paik, N.-J. (2016). Mobile game-based virtual reality rehabilitation program for upper limb dysfunction after ischemic stroke. Restorative Neurology and Neuroscience. https://doi.org/10.3233/RNN-150626

Dankert, T., Heil, D., Pfeiffer, T. (2013). Stereo vision and acuity tests within a virtual reality set-up. In M. E. Latoschik, O. Staadt, & F. Steinicke (Eds.), Virtuelle und Erweiterte Realität - 10. Workshop der GI-Fachgruppe VR/AR (pp. 185-188). Shaker Verlag.

de Vignemont, F., & Iannetti, G. D. (2015). How many peripersonal spaces? Neuropsychologia, 70, 327–334. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2014.11.018

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

di Pellegrino, G., & Làdavas, E. (2015). Peripersonal space in the brain. Neuropsychologia, 66, 126–133. https://doi.org/10.1016/j.neuropsychologia.2014.11.011

Fonteneau, E., & Davidoff, J. (2007). Neural correlates of colour categories. NeuroReport. https://doi.org/10.1097/WNR.0b013e3282c48c33

Gratton, G., Coles, M. G. ., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. Electroencephalography and Clinical Neurophysiology, 55(4), 468–484. https://doi.org/10.1016/0013-4694(83)90135-9

Henderson, M., Vo, V., Chunharas, C., Sprague, T., Serences, J. (2019). Multivariate analysis of BOLD activation patterns recovers graded depth representations in human visual and parietal cortex. eNeuro, 6(4), 1-15. https://doi.org/10.1523/ENEURO.0362-18.2019

Hentschke, H., & Stüttgen, M. C. (2011). Computation of measures of effect size for neuroscience data sets. European Journal of Neuroscience, 34(12), 1887–1894. https://doi.org/10.1111/j.1460-9568.2011.07902.x

Ho, T.-H., Yang, F.-C., Lin, R.-C., Chien, W.-C., Chung, C.-H., Chiang, S.-L., … Lee, J.-T. (2019). Impact of virtual reality-based rehabilitation on functional outcomes in patients with acute stroke: a retrospective case-matched study. Journal of Neurology, 266(3), 589–597. https://doi.org/10.1007/s00415-018-09171-2

Jarosz, A. F., & Wiley, J. (2014). What Are the Odds? A Practical Guide to Computing and Reporting Bayes Factors. The Journal of Problem Solving, 7(1). https://doi.org/10.7771/1932-6246.1167

JASP Team (2019). JASP (Version 0.10.2)[Computer software].

Kuziek, J. W. P., Shienh, A., & Mathewson, K. E. (2017). Transitioning EEG experiments away from the laboratory using a Raspberry Pi 2. Journal of Neuroscience Methods. https://doi.org/10.1016/j.jneumeth.2016.11.013

Làdavas, E. (2002). Functional and dynamic properties of visual peripersonal space. Trends in Cognitive Sciences, 6(1), 17–22. https://doi.org/10.1016/S1364-6613(00)01814-3

Laszlo, S., Ruiz-Blondet, M., Khalifian, N., Chu, F., & Jin, Z. (2014). A direct comparison of active and passive amplification electrodes in the same amplifier system. Journal of Neuroscience Methods. https://doi.org/10.1016/j.jneumeth.2014.05.012

Lee, J. H., Ku, J., Kim, K., Kim, B., Kim, I. Y., Yang, B.-H., … Kim, S. I. (2003). Experimental Application of Virtual Reality for Nicotine Craving through Cue Exposure. CyberPsychology & Behavior. https://doi.org/10.1089/109493103322011560

Linden, D. (2005). The P300: Where in the brain is it produced and what does it tell us?. The Neuroscientist, 11(6)**,** 563-576.

Liu, B., Meng, X., Wu, G., & Dang, J. (2013). Correlation between three-dimensional visual depth and N2 component: Evidence from event-related potential study. Neuroscience, 237, 161–169. doi:10.1016/j.neuroscience.2013.01.045

Mathewson, K. E., Harrison, T. J. L., & Kizuk, S. A. D. (2017). High and dry? Comparing active dry EEG electrodes to active and passive wet electrodes. Psychophysiology, 54(1), 74–82. https://doi.org/10.1111/psyp.12536

Mathewson, K. E., Lleras, A., Beck, D. M., Fabiani, M., Ro, T., & Gratton, G. (2011). Pulsed Out of Awareness: EEG Alpha Oscillations Represent a Pulsed-Inhibition of Ongoing Cortical Processing. Frontiers in Psychology, 2, 99. https://doi.org/10.3389/fpsyg.2011.00099

Naceri, A., Moscatelli, A., & Chellali, R. (2015). Depth discrimination of constant angular size stimuli in action space: Role of accommodation and convergence cues. Frontiers in Human Neuroscience, 9. doi:10.3389/fnhum.2015.00511

Pugnetti, L., Mendozzi, L., Attree, E. A., Sc, B., Barbieri, E., Ph, D., … Psychol, C. (1998). Probing Memory and Executive Functions with Virtual Reality : Past and Present Studies. CyberPsychology.

Raz, S., Dan, O., & Zysberg, L. (2014). Neural correlates of emotional intelligence in a visual emotional oddball task: An ERP study. Brain and Cognition. https://doi.org/10.1016/j.bandc.2014.09.003

Robinson, C., Ahmar, N., Sloutsky, V., Robinson, C. W., & Sloutsky, V. M. (n.d.). UC Merced Proceedings of the Annual Meeting of the Cognitive Science Society Title Evidence for auditory dominance in a passive oddball task Publication Date Evidence for auditory dominance in a passive oddball task. (32), 32. Retrieved from https://escholarship.org/dist/wJD5zTgC2vrImRR/dist/prd/content/qt9gg173d6/qt9gg173d6.pdf

Rothbaum, B. O., Price, M., Jovanovic, T., Norrholm, S. D., Gerardi, M., Dunlop, B., … Ressler, K. J. (2014). A randomized, double-blind evaluation of D-cycloserine or alprazolam combined with virtual reality exposure therapy for posttraumatic stress disorder in Iraq and Afghanistan war veterans. American Journal of Psychiatry. https://doi.org/10.1176/appi.ajp.2014.13121625

Safir, M. P., Wallach, H. S., & Bar-Zvi, M. (2012). Virtual Reality Cognitive-Behavior Therapy for Public Speaking Anxiety. Behavior Modification, 36(2), 235–246. https://doi.org/10.1177/0145445511429999

Sawaragi, T., & Horiguchi, Y. (n.d.). Analysis of task morphologies for networking behavioral skills via virtual reality. Proceedings. IEEE 8th International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WET ICE'99), 200–207. https://doi.org/10.1109/ENABL.1999.805202

Scanlon, J. E. M., Townsend, K. A., Cormier, D. L., Kuziek, J. W. P., & Mathewson, K. E. (2019). Taking off the training wheels: Measuring auditory P3 during outdoor cycling using an active wet EEG system. Brain Research. https://doi.org/10.1016/j.brainres.2017.12.010

Smith, M. J., Boteler Humm, L., Fleming, M. F., Jordan, N., Wright, M. A., Ginger, E J., … Bell, M. D. (2015). Virtual reality job interview training for veterans with posttraumatic stress disorder. Journal of Vocational Rehabilitation. https://doi.org/10.3233/JVR-150748

Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., … Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. Psychonomic Bulletin & Review, 25(1), 58–76. https://doi.org/10.3758/s13423-017-1323-7

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., … Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. Psychonomic Bulletin & Review, 25(1), 35–57. https://doi.org/10.3758/s13423-017-1343-3

VIVE Virtual Reality System. (2011). Retrieved from https://www.vive.com/us/product/vive-virtual-reality-system/

Zink, R., Hunyadi, B., Huffel, S. Van, & Vos, M. De. (2016). Mobile EEG on the bike: Disentangling attentional and physical contributions to auditory attention tasks. Journal of Neural Engineering. https://doi.org/10.1088/1741-2560/13/4/046017

Acknowledgements

This work was supported by a discovery grant to KEM from the Natural Sciences and Engineering Research Council (NSERC) of Canada and start-up funds from the Faculty of Science. Thank you to all members and volunteers of the Mathewson lab for assisting with data collection and experimental setup. Thank you to our collaborators at IBM Research for providing help and guidance designing the virtual reality task and environment.

Tables

Condition	Oddball Task	Stimulus Type	Mean Trial Count	Standard Deviation	Trial Count Range	Minimum Trials Remaining (%)
Non-VR	Auditory	Target	50.63	7.48	37-70	94.12
Non-VR	Auditory	Standard	197.04	7.78	180-213	94.47
VR	Auditory	Target	49.71	6.07	37-61	94.12
VR	Auditory	Standard	198.33	6.88	187-213	94.47
Non-VR	Visual	Target	50.71	5.07	43-60	95.65
Non-VR	Visual	Standard	197.13	6.36	182-206	91.46
VR	Visual	Target	49.38	4.98	39-60	97.92
VR	Visual	Standard	199.92	4.90	190-208	97.01

Table 1: Trial count information for both target and standard trials across condition (Non-VR; VR) and oddball modality (Auditory; Visual).

A

Condition	Oddball Task	Electrode/ Time Window	Mean Voltage (µV)	Standard Deviation (µV)	*t* (23)	*p*	Confidence Interval	Hedge's *g*	BF_{10}
Non-VR	Auditory	Fz/MMN	-2.65	2.89	-4.50	<0.01	-Inf:-1.64	-0.89	353.0
Non-VR	Auditory	Pz/P3	5.51	4.24	6.36	<0.01	*4.02:Inf*	1.26	21741.0
VR	Auditory	Fz/MMN	-2.22	2.45	-4.44	<0.01	-Inf:-1.36	-0.88	306.1
VR	Auditory	Pz/P3	5.34	4.64	5.64	<0.01	*3.72:Inf*	1.11	4513.0
Non-VR	Visual	Fz/MMN	-0.77	3.73	-1.01	0.16	-Inf:0.54	-0.20	0.562
Non-VR	Visual	Pz/P3	6.90	3.87	8.74	<0.01	*5.55:Inf*	1.73	2.815e+6
VR	Visual	Fz/MMN	-0.86	3.82	-1.11	0.14	-Inf:0.47	-0.22	0.629
VR	Visual	Pz/P3	6.10	3.76	7.95	<0.01	*4.79:Inf*	1.57	594100

B

Oddball Task	Electrode/ Time Window	Mean Voltage (µV)	Standard Deviation (µV)	*t* (23)	*p*	Confidence Interval	Hedge's *g*	BF_{10}
Auditory	Fz/MMN	-0.43	2.75	-0.56	0.58	-1.99:1.12	-0.16	0.282
Auditory	Pz/P3	0.16	2.17	0.13	0.90	-2.42:2.75	0.04	0.229
Visual	Fz/MMN	0.09	3.23	0.08	0.93	-2.10:2.28	0.02	0.217
Visual	Pz/P3	0.80	2.68	0.73	0.47	-1.42:3.01	0.20	0.548

Table 2: Results of one-tailed and two-tailed *t*-tests on the ERP difference waves. A) One tailed *t*-test and Bayes Factor results to determine if a significant MMN and P3 response was observed for each condition and oddball task. Bayes factors were calculated by testing if the MMN/P3 waveforms for each condition/oddball task were less than or greater than zero, respectively. B) Two-tailed *t*-test and Bayes paired-t-test results comparing the MMN and P3 waveforms for each condition and oddball task. Bayes Factors (BF_{01}) were calculated by comparing the MMN and P3 waveforms between Non-VR and VR conditions.

Stimulus Type	Mean Trial Count	Standard Deviation	Trial Count Range	Minimum Trials Remaining (%)
Far Targets	94.44	9.72	80:112	72.10
Near Standards	394.72	38.73	244:416	62.72
Near Targets	100.33	7.51	87:113	96.67
Far Standards	397.06	8.39	377:412	95.93

Table 3: Trial count information for both target and standard trials across both the Far and Near conditions.

Condition	Electrode/ Time Window	Mean Voltage (µV)	Standard Deviation (µV)	*t* (17)	*p*	Confidence Interval	Hedge's *g*	BF_{10}
A	Pz/175-225ms	-1.17	2.07	-2.40	0.014	-Inf:-0.32	-0.54	2.28
A	Pz/225-300ms	1.15	1.93	2.52	0.011	0.35:Inf	0.57	2.77
A	Pz/300-600ms	-0.18	1.29	-0.61	0.28	-Inf:0.35	-0.14	0.29
B	Pz/175-225ms	2.70	2.52	4.54	<0.01	1.67:Inf	1.02	111.00
B	Pz/225-300ms	-1.44	2.91	-2.10	0.25	-Inf:-0.25	-0.47	1.43
B	Pz/300-600ms	-0.53	3.25	-0.70	0.25	-Inf:0.80	-0.16	0.30

A = Far Standards – Near Standards; B = Far Targets – Targets

Condition	Electrode/ Time Window	Mean Voltage (µV)	Standard Deviation (µV)	*t* (18)	*p*	Confidence Interval	Hedge's *g*	BF_{10}
C	Pz/175-225ms	1.37	2.26	2.57	<0.01	0.44:Inf	0.58	3.02
C	Pz/225-300ms	-0.62	3.29	-0.79	0.22	-Inf:0.73	-0.18	0.32
C	Pz/300-600ms	5.77	3.11	7.88	<0.01	4.50:Inf	1.77	3.73e+4
D	Pz/175-225ms	-2.51	1.94	-5.49	<0.01	-Inf:-1.71	-1.24	641.08
D	Pz/225-300ms	1.97	2.68	3.12	<0.01	0.87:Inf	0.70	7.87
D	Pz/300-600ms	6.12	2.48	10.47	<0.01	5.10:Inf	2.36	1.52e+6

C = Far Targets – Near Standards; D = Near Targets – Far Standards

Table 4: Results of one-tailed *t*-tests, Bayesian one-tailed *t*-tests, and estimated measures of effect size for the ERP waveforms at three time windows. Hedge's *g* is used to estimate effect size. Bayes factors were calculated by testing if the waveforms at each time window was less or greater than zero, based on the mean voltage. BF_{10} indicates support for H1 over H0.

Oddball Tasks Compared	**Electrode/ Time Window**	**Mean Voltage (µV)**	**Standard Deviation (µV)**	***t* (17)**	***p***	**Confidence Interval**	**Hedge's *g***	**BF_{10}**
C:D	Pz/175-225ms	3.88	3.68	4.47	<0.01	2.05:5.70	1.80	97.33
C:D	Pz/225-300ms	-2.59	4.03	-2.72	0.014	-4.59:-0.58	-0.84	3.91
C:D	Pz/300-600ms	-0.35	3.44	-0.43	0.67	-2.06:1.36	-0.12	0.26

C = Far Targets – Near Standards; D = Near Targets – Far Standards

Table 5: Results of two-tailed *t*-tests, Bayesian paired *t*-tests, and estimates of measure of effect size between the ERP waveforms at three time windows for electrode Pz. Hedge's *g* is used to estimate effect size. Bayes factors were calculated by comparing both waveforms at each time window. BF_{10} indicates support for H1 over H0.

Figure Captions

Figure 1. Design of the experiments used for the A) auditory oddball and the B) visual oddball tasks. C) Response times (ms) following targets during the auditory and visual oddball tasks.

Figure 2: Grand-averaged ERP plots for target and standard stimuli at electrode Fz and Pz. Also shown are the difference waveforms for the A) auditory oddball and B) visual oddball tasks at electrode Fz and Pz. Difference waveforms were generated by subtracting the ERPs following standard stimuli from the ERPs following target stimuli. Yellow shaded regions represent the MMN and P3 time windows used for analyses. Topographies represent activity across the scalp while the bar graphs show activity at either electrode Fz or Pz, averaged across the MMN and P3 time windows. Error bars represent the standard error of the mean.

Figure 3: Spectra and RMS analysis for the auditory and visual oddball tasks. A) Spectra and B) single-trial and grand-averaged RMS distributions for the auditory oddball task. Error bars in the spectra plots represent the standard deviation at each frequency. C = Spectra and D = single-trial and grand-averaged RMS distributions for the visual oddball task. Error bars on the mean RMS plots represent the standard deviation of the respective RMS distribution.

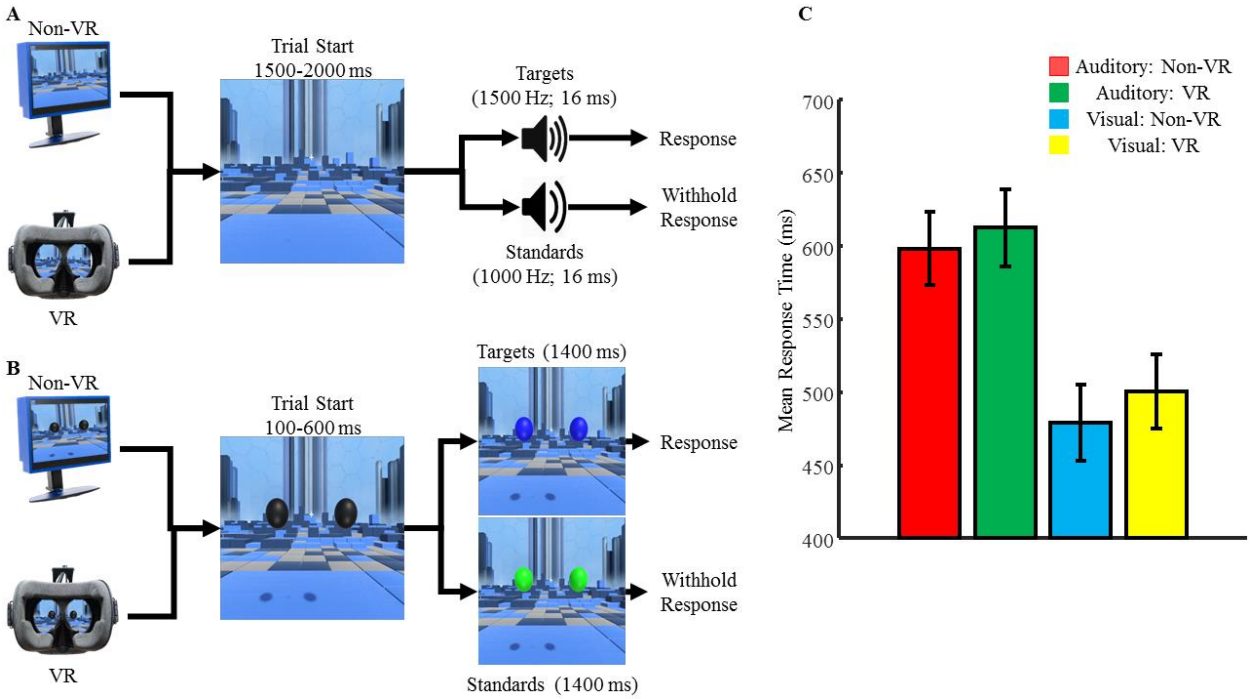
Figure 4: Plots showing the trials counts needed to obtain a significant MMN or P3 with 80% power for the A) auditory oddball and B) visual oddball tasks. Analyses were run using electrode Fz for the MMN waveform and electrode Pz for the P3 waveform.

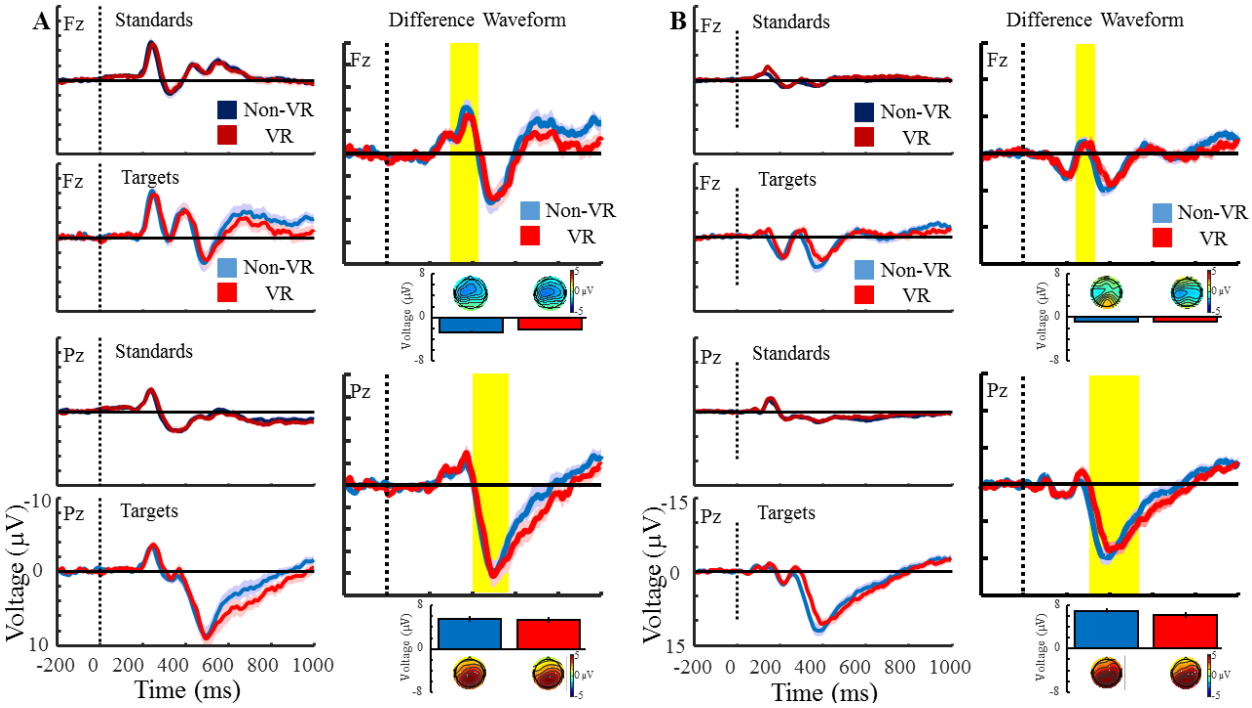
Figure 5: Design and setup of the depth oddball task. A) Task setup for each participant and the location of the near and far orbs with respect to the participant. B) Stimuli and virtual environment from the perspective of the participant. C) Response times for target stimuli at both near and far depths. Error bars represent the standard error of the mean.

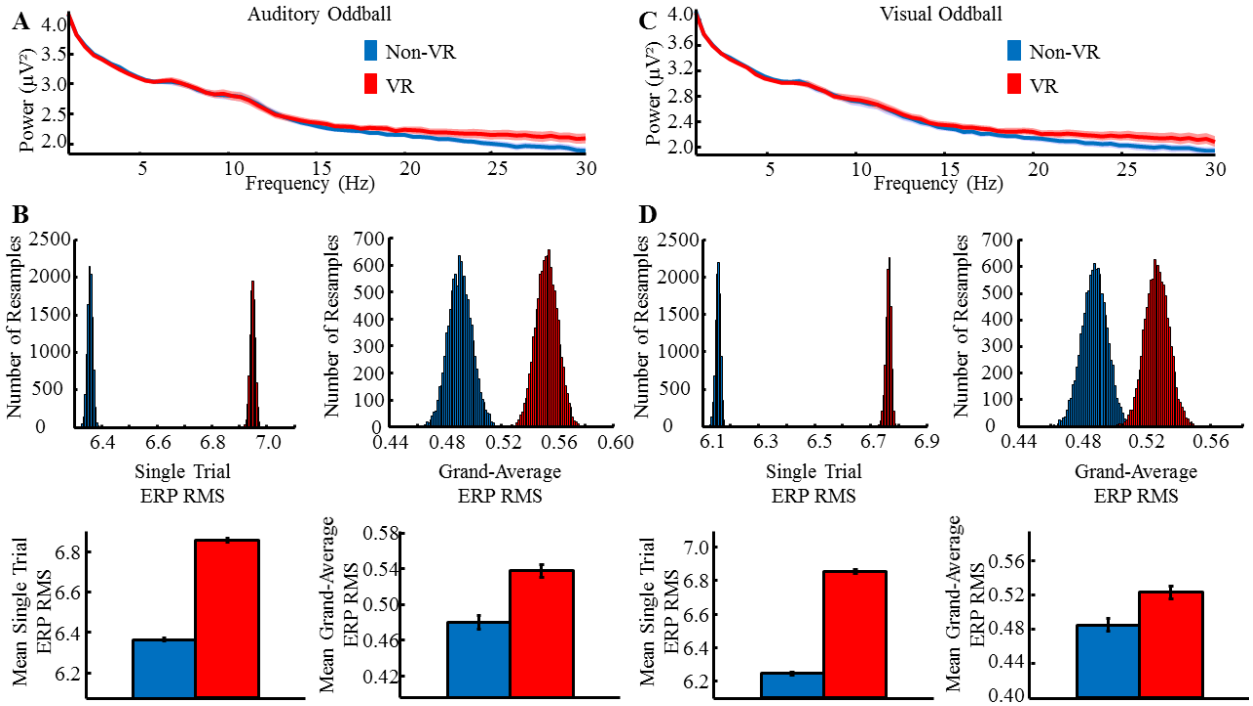
Figure 6: A) Grand-averaged ERP plots for target and standard stimuli at electrode Fz and Pz. B) Difference waveforms for near and far targets and standards. These waveforms were created by subtracting near standards from far standards, and near targets from far targets, at electrodes Fz and Pz. C) Difference waveforms for the Far Targets and Near Targets conditions. The left-most plots were generated by subtracting ERPs to standards from the ERPs to targets in each condition. Far Targets Difference Wave = Far Targets – Near Standards; Near Targets Difference Waveform = Near Targets – Far Standards. Coloured shaded regions in each difference waveform represent the three time windows used for analyses. These windows are based on two opposing patterns of activity, and the P3 waveform, observed in the Far Targets and Near Targets difference waveforms at electrode Pz. Yellow = 175-225 ms; Orange = 225-300 ms; Red = 300-600 ms. Topographies represent activity across the scalp while bar graphs represent activity at either electrode Fz or Pz, averaged across these time windows. All error bars in each plot represent the standard error of the mean.

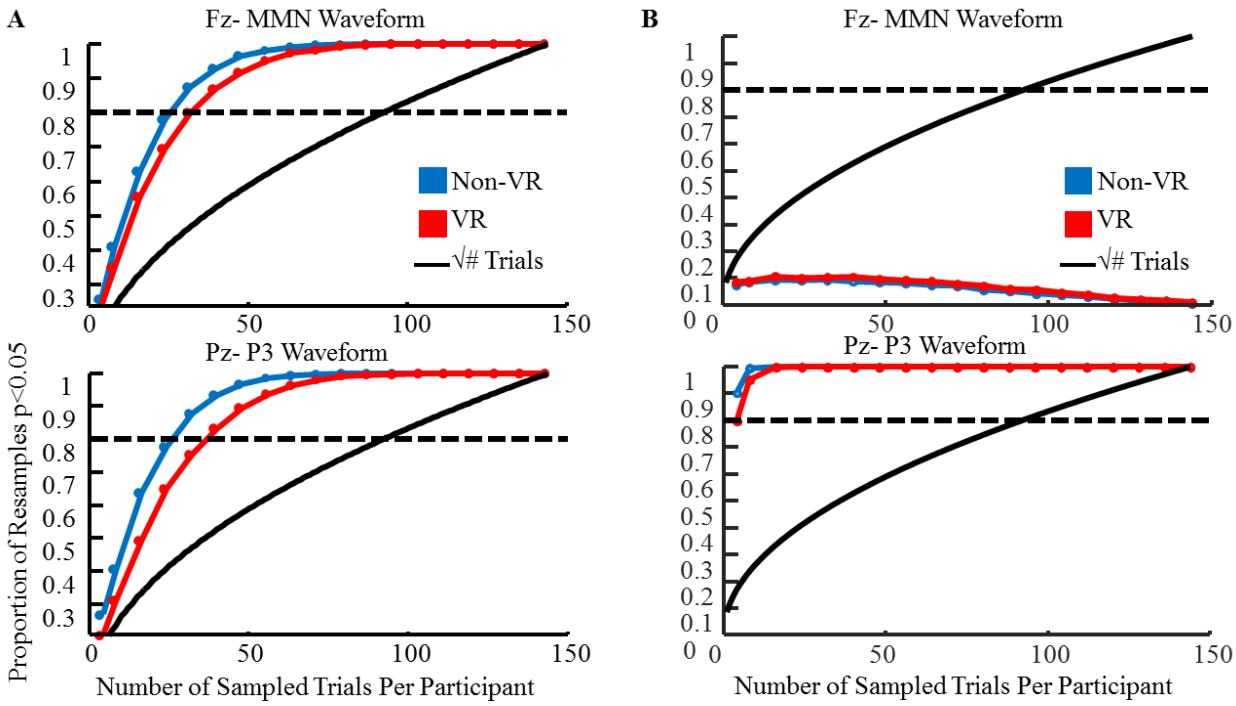
Figure 7: A) Spectra along with B) single-trial and grand-averaged RMS analysis for the Far Targets and Near Targets difference waveforms. Error bars on the mean RMS plots represent the standard deviation of the respective RMS distribution.

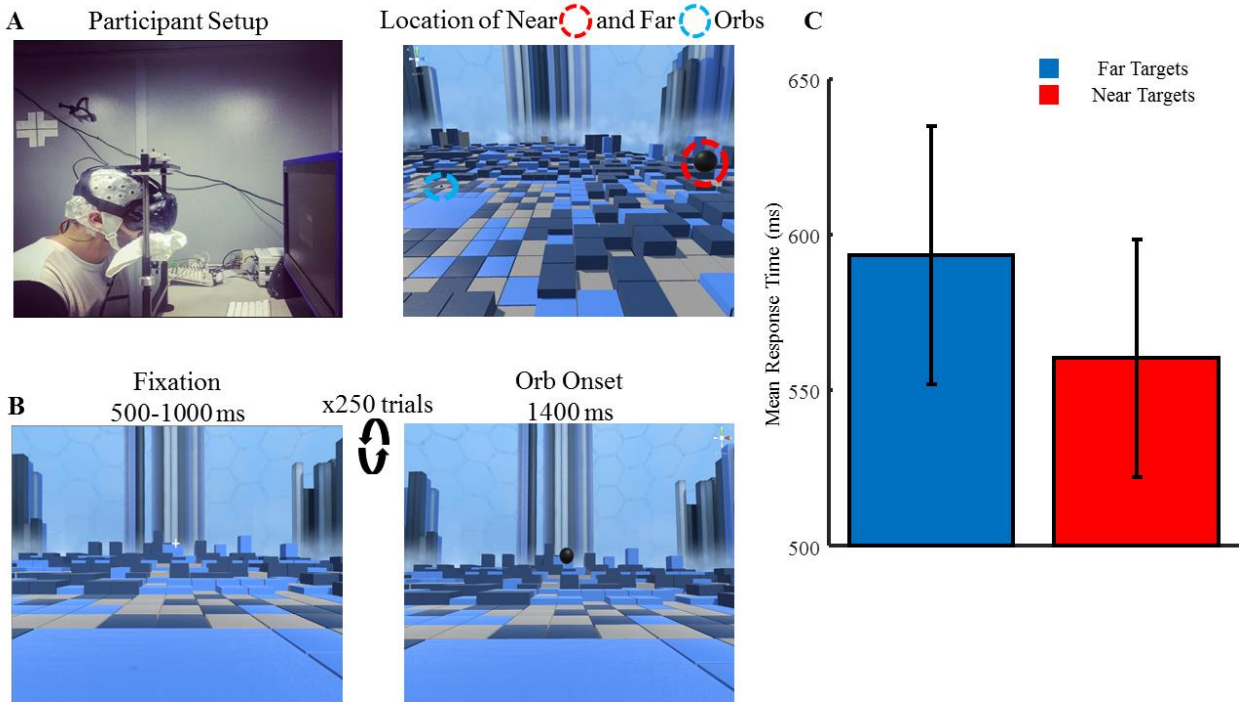
Figures

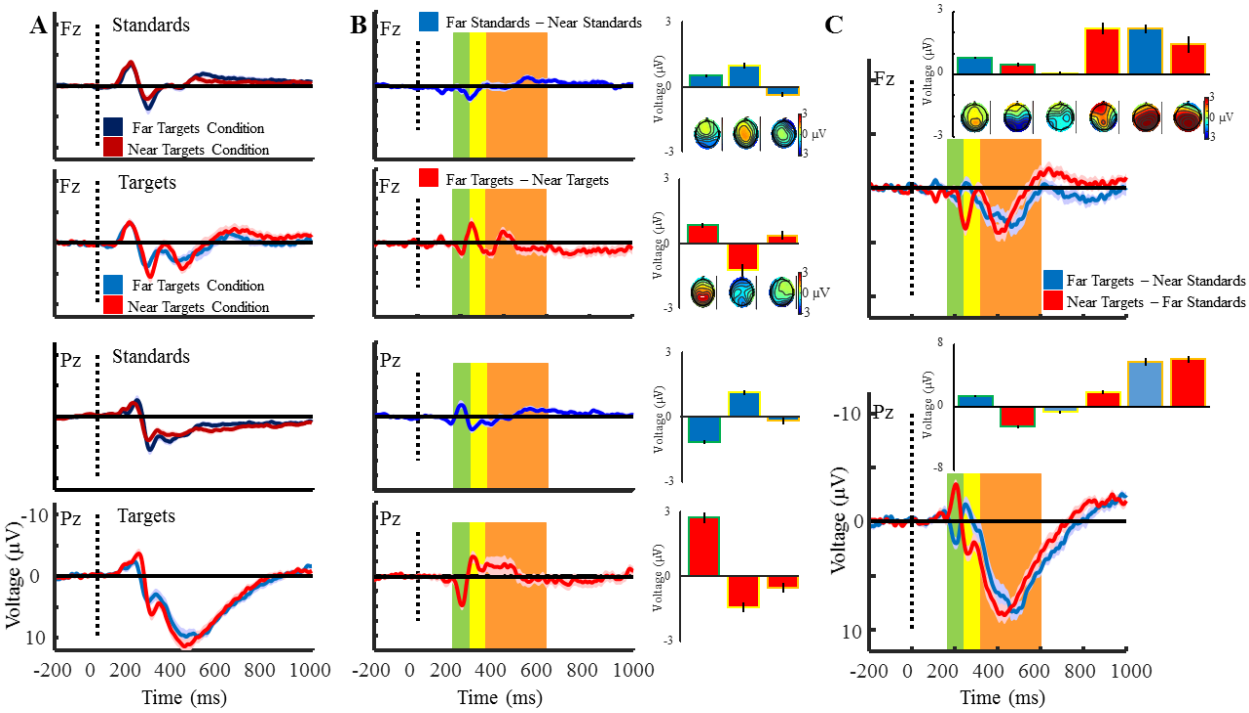

A
Non-VR
VR
Trial Start
1500-2000 ms
Targets
(1500 Hz; 16 ms)
Response
Withhold
Response
Standards
(1000 Hz; 16 ms)
B
Non-VR
VR
Trial Start
100-600 ms
Targets (1400 ms)
Response
Withhold
Response
Standards (1400 ms)
C
Auditory: Non-VR
Auditory: VR
Visual: Non-VR
Visual: VR
Mean Response Time (ms)
700
650
600
550
500
450
400

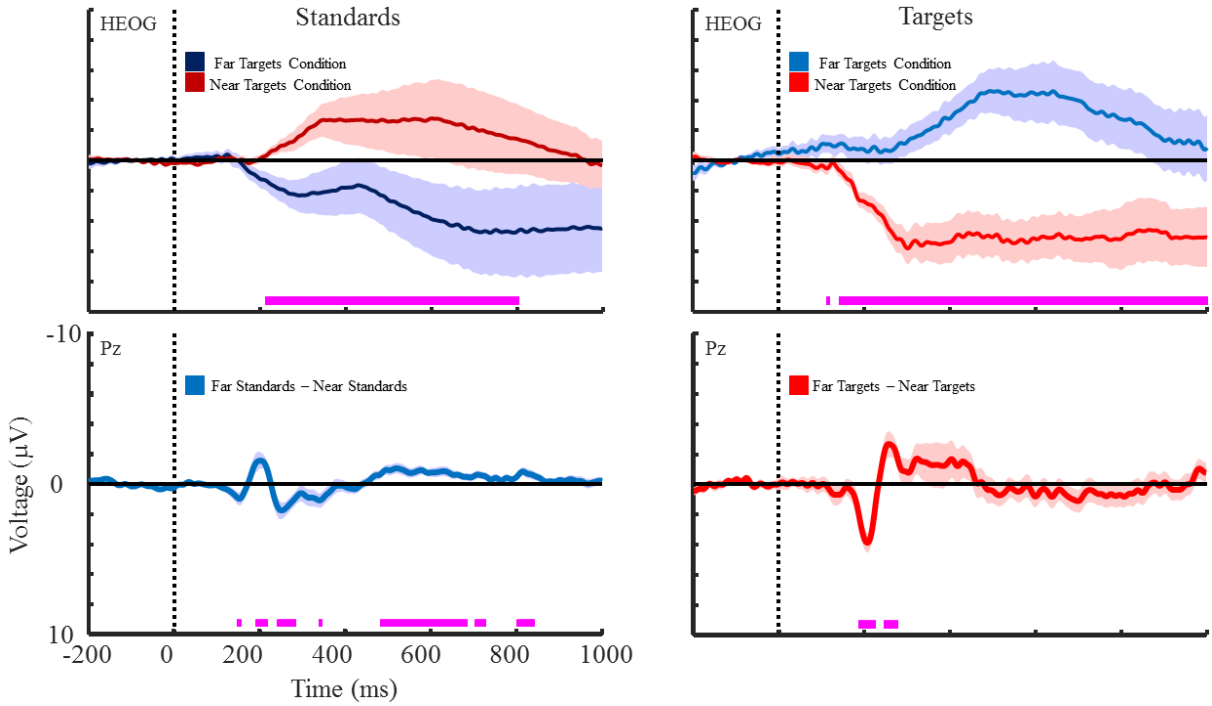

A
Fz
Standards
Non-VR
VR
Fz
Targets
Non-VR
VR
Pz
Standards
Pz
Targets
Voltage (µV)
-10
0
10
-200 0 200 400 600 800 1000
Time (ms)
Difference Waveform
Fz
Non-VR
VR
Pz
B
Fz
Standards
Non-VR
VR
Fz
Targets
Non-VR
VR
Pz
Standards
Pz
Targets
Voltage (µV)
-15
0
15
-200 0 200 400 600 800 1000
Time (ms)
Difference Waveform
Fz
Non-VR
VR
Pz


A
Auditory Oddball
Non-VR
VR
Power (μV²)
Frequency (Hz)
C
Visual Oddball
Non-VR
VR
Power (μV²)
Frequency (Hz)
B
Number of Resamples
Single Trial ERP RMS
Grand-Average ERP RMS
Mean Single Trial ERP RMS
Mean Grand-Average ERP RMS
D
Number of Resamples
Single Trial ERP RMS
Grand-Average ERP RMS
Mean Single Trial ERP RMS
Mean Grand-Average ERP RMS


A
Fz- MMN Waveform
Pz- P3 Waveform
B
Fz- MMN Waveform
Pz- P3 Waveform
Non-VR
VR
√# Trials
Proportion of Resamples p<0.05
Number of Sampled Trials Per Participant
Number of Sampled Trials Per Participant


A
Participant Setup
Location of Near and Far Orbs
B
Fixation
500-1000 ms
x250 trials
Orb Onset
1400 ms
C
Far Targets
Near Targets
Mean Response Time (ms)
650
600
550
500


A
Fz
Standards
Far Targets Condition
Near Targets Condition
Fz
Targets
Far Targets Condition
Near Targets Condition
Pz
Standards
Pz
Targets
Voltage (µV)
-10
0
10
-200 0 200 400 600 800 1000
Time (ms)
B
Fz
Far Standards – Near Standards
Fz
Far Targets – Near Targets
Pz
Pz
-200 0 200 400 600 800 1000
Time (ms)
C
Fz
Far Targets – Near Standards
Near Targets – Far Standards
Pz
Voltage (µV)
-10
0
10
-200 0 200 400 600 800 1000
Time (ms)


HEOG
Standards
Far Targets Condition
Near Targets Condition
HEOG
Targets
Far Targets Condition
Near Targets Condition
Pz
Far Standards – Near Standards
Pz
Far Targets – Near Targets
-10
0
10
Voltage (µV)
-200
0
200
400
600
800
1000
Time (ms)