

Investigating Noticeable Hand Redirection in Virtual Reality using Physiological and Interaction Data

Martin Feick* Kora P. Regitz* Anthony Tang† Tobias Jungbluth* Maurice Rekrut* Antonio Krüger*
DFKI, Saarland Informatics Campus, Germany* Singapore Management University, Singapore†

ABSTRACT

Hand redirection is effective so long as the introduced offsets are not noticeably disruptive to users. In this work we investigate the use of physiological and interaction data to detect movement discrepancies between a user’s real and virtual hand, pushing towards a novel approach to identify discrepancies which are too large and therefore can be noticed. We ran a study with 22 participants, collecting EEG, ECG, EDA, RSP, and interaction data. Our results suggest that EEG and interaction data can be reliably used to detect visuo-motor discrepancies, whereas ECG and RSP seem to suffer from inconsistencies. Our findings also show that participants quickly adapt to large discrepancies, and that they constantly attempt to establish a stable mental model of their environment. Together, these findings suggest that there is no absolute threshold for possible non-detectable discrepancies; instead, it depends primarily on participants’ most recent experience with this kind of interaction.

1 INTRODUCTION

With Virtual Reality (VR), humans can experience and interact with an immersive simulated environment, opening up a wide range of applications and use cases. VR can enhance interactions with virtual objects—objects that a user cannot physically touch and experience [36]. Recent research has focused on rendering haptic feedback for interactions with the otherwise purely virtual entities through concepts [4, 5, 12, 35, 37, 41, 48, 54, 66], and dedicated controllers and devices [3, 13, 53, 73, 75]. One approach for providing appropriate haptic feedback is the use of illusions combined with proxy objects [4, 5, 12, 41], which is based on multisensory integration theory: when users encounter conflicting sensory information, the most plausible modality dominates over others, which are then suppressed [28]. One common and powerful type of VR illusion is hand redirection, which offsets the visual hand position from the position of the user’s real hand [9, 76]. The result of this manipulation can allow users to grasp objects which are out of reach [60], provide haptic feedback when touching different virtual objects substituted by a single physical “stand-in” (i.e., a proxy) [4, 5, 12, 22, 41] or even simulate virtual objects of different weights [63], weight distributions [79] or dimensions [7].

This method cannot be scaled up infinitely, because if the discrepancy between vision and proprioception becomes too large, it can be noticed by a user (i.e., resulting in a semantic violation). Semantic violations disrupt presence [68], as they are noticeable disruptions in what is “believable” in the simulation. The extent to which illusions may be used without detection is a well-studied area, and the current state of the art is to conduct psychophysical experiments to determine humans’ perceptual boundaries for a certain type of illusion, reporting detection thresholds that future work can build on [1, 6, 7, 14, 22, 22, 23, 26, 27, 69, 73, 76, 79]. However, this approach has three major limitations: (1) the psychophysical experiment needs to be conducted for each new device and illusion technique; (2) contextual factors limit the generalizability of the technique; and (3) the experiments do not account for individual differences, which recent evidence suggests is more meaningful compared to a group mean [23]. The general question remains, how far can we push an

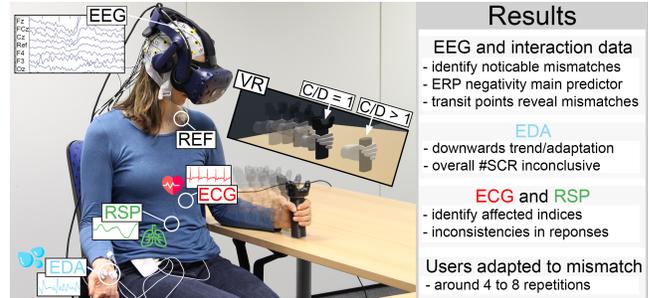


Figure 1: Study setup: a participant is moving a virtual object embodied by a physical proxy. The VR view shows the 1-to-1 mapping and the redirected position of the virtual hand/object.

illusion without an *individual* noticing it? In our work, we consider individuals’ perceptual boundaries of these illusions as a form of personalization. We explore the extent to which a system can automatically detect when these thresholds have been crossed for an individual—i.e., how can we assess whether a semantic violation has occurred without a user telling us? Our long-term vision is to use this information to automatically tailor illusions based on an individual’s perceptual boundary in an *on-the-fly* fashion.

Therefore, the goal of this work is to understand which sensor modalities can be reliably used to detect noticeable visuo-motor discrepancies by investigating electroencephalogram (EEG), electrocardiogram (ECG), electrodermal activity (EDA), respiration (RSP) and interaction data. As part of this work, we also investigate how quickly participants adapt to such discrepancies, and study how long this adaptation process takes. We conducted an experiment with 22 participants performing a simple Fitts’s law style [49] docking task, with and without a noticeable hand redirection offset (tuned to each participant). Our results suggest that EEG is a reliable method to detect that a participant has experienced a semantic violation. In addition to EEG data, we found that interaction data—how participants move during trials—is also a reliable method to detect noticeable discrepancies. On the other hand, ECG and RSP seem to suffer from inconsistencies across participants. Additionally, we observed that participants quickly adapt to (even) larger discrepancies, and found evidence that there seems to be no absolute threshold for the possible discrepancy that can be introduced. Instead, what participants perceived as the “ground truth” depended mostly on their most recent experience with this kind of interaction.

In this work, we make five contributions:

1. We provide a method to study noticeable illusions tailored to an individual’s perceptual boundary.
2. We demonstrate the capability of EEG measures to distinguish between noticeable and unnoticeable discrepancies.
3. We outline the potential of interaction data as a reliable metric to detect noticeable visuo-motor discrepancies.
4. We critically assess the use of physiological data to detect visuo-motor discrepancies and discuss potential limits.
5. We open-source a data set for ECG, EDA, RSP indices and interaction data to the research community.

*e-mail: name.surname@dfki.de

†e-mail: tonyt@smu.edu.sg

2 RELATED WORK

Our work is positioned in the field of (haptic) illusions in VR and neuroscience. We first discuss haptics and illusions in VR, then discuss how physiological data have previously been used in VR. Finally, we outline recent research concerning the detection of semantic violations in VR using EEG.

2.1 Haptics and Illusions in Virtual Reality

Providing adequate haptic feedback when touching or interacting with virtual objects remains a challenge in VR. One approach to address this is called passive haptics, where properties of real-world objects (proxies), such as shape and size [15, 20, 33] or weight [75] are mapped onto virtual counterparts [37]. The ultimate goal would be to have a set of universal proxies that approximate virtual objects to such an extent that users perceive them as “real”.

To achieve this, researchers have proposed the use of illusions alongside proxies [1, 7, 14, 22, 33, 79], taking advantage of the fact that some differences between virtual object and physical proxy may remain unnoticed. This exploits the visual-dominance phenomenon: in cases where information from two senses (e.g. sight and touch) are in conflict, vision usually dominates [10, 25]. For example, [5] used this effect to redirect a user's finger on a simple cylindrical proxy, and by adding various visual overlays atop the proxy, they were able to change the proxy's perceived physical shape. Redirecting users' hands mid-air can be achieved by offsetting the position of the virtual hand from the position in the real world [10, 12, 76]. As a result, users compensate for this and thus, they may touch different virtual objects, but in fact, they have been redirected to the same physical proxy (i.e., haptic retargeting [4]). However, there exist limits to the extent to which such illusions can be used and still remain undetectable for users [1, 14, 23, 76].

A large body of work has looked at how much discrepancy between the real and virtual world may be introduced while remaining unnoticed by a user. These researchers have reported detection thresholds' for illusions such as hand redirection [6, 12, 17, 32, 76], redirected walking [69], bi-manual hand redirection [27] and redirected touch [14]; others have used the technique to simulate object properties, for example weight [63] and size [7]. The consensus is that it is possible to introduce discrepancies; however, two limitations of these works are: (1) the thresholds do not seem to generalize beyond the device and the effect, and (2) the thresholds seem to be highly variable [23]. Our goal is to develop a reliable method to detect when illusions have become noticeable beyond bespoke situations and use cases. To do so, we investigate the potential of physiological data which have previously been used in VR.

2.2 Physiological Data in Virtual Reality

Physiological data such as RSP, EDA and heart rate (HR), and so on, have been used in many systems and situations.

For instance, the study by Egan et al. [16] looked at HR and EDA as an objective evaluation metric to assess the quality of a VR experience. Here, participants experienced the same virtual scene through an HMD vs. a 2D monitor. Differences in participants' HR and EDA were correlated to the display condition (i.e. in VR vs. on a 2D monitor), which was also correlated with their subjective assessments of the experience. HR and heart rate variability (HRV) have also been used to study motion sickness within the context of long-term immersion in VR [31, 51]. Marchiori et al. [52] observed an increase in HR in response to virtual scenes that were perceived as less realistic according to participants' questionnaire responses. EDA is associated with experiencing emotional arousal [11]. Guna et al. [31] studied changes in the skin conductance level (SCL), a measure of EDA, to assess VR sickness, demonstrating a correlation between SCL and participants' subjective responses in a Simulator Sickness Questionnaire [39]. On the other hand, respiratory rate changes with respect to humans' perceived stress, and therefore has

been studied when experiencing stressful scenarios such as flight or roller coaster simulations in VR [18].

The aforementioned studies aim to establish an explicit objective metric, which is in line with our eventual goal of detecting noticeable illusions and consequently allowing us to tailor VR illusions to individuals' perceptual boundaries [44]. To achieve this, we first investigate whether noticeable illusions trigger physiological responses. Most closely related to our work are EEG studies, concerning Error-Related Potentials (ERPs) caused by visual and haptic mismatches, which we discuss below.

2.3 Error-Related Potentials in Virtual Reality

In the HCI and VR community, ERPs have become a useful measure, allowing researchers to detect if participants experience an error without directly asking them [43]. For instance, Si-Mohammed et al. [65] showed that ERPs can be used to detect system errors in VR such as background anomalies or tracking errors while interacting with virtual objects. Gehrke et al. [24] used ERPs to detect visuo-haptic mismatches by comparing three conditions when touching a virtual cube: (1) visual but no haptic feedback, (2) visual feedback + vibrotactile feedback on the fingertip and (3) visual feedback + vibrotactile feedback + electro-muscle stimulation. In 75% of the trials the three feedback conditions matched the performed interaction; however, in the remaining 25% the feedback was presented prematurely. Their results show that it is possible to distinguish matching and mismatched trials using ERPs. In another study by Yazmir et al. [74] ERPs were used in a visuo-haptic error induced task, where participants used a Phantom haptic device to move a sphere horizontally. To get to the target location, it was required to pass an obstacle which momentarily obscured the sphere. In about 60% of the trials they introduced a disturbance, i.e., the sphere was offset, horizontally and/or vertically, 'behind' the obstacle. In a small study, they collected evidence for a strong ERP shortly after the error. Padrao et al. [58] studied the difference between self-generated and externally imposed errors on the sense of agency when users were embodied by a body avatar. Their work provides a strong foundation for our study, because the externally imposed errors were provoked by moving the virtual avatar hand in the opposite direction from the participant's real hand. The results showed a strong similarity to ERP signatures related to semantic or conceptual violations (central cortex area).

This line of research shows that ERPs are promising to detect errors and thus are a candidate to assess whether an illusion exceeds an individual's perceptual boundary—and therefore can be noticed. Please note, illusions usually amplify or alter an interaction that a user is already performing [1, 6, 7, 12–14, 17, 22, 23, 27, 63, 69, 76]—instead of a sudden change in limb trajectory [58], premature haptic feedback [24] or artefacts appearing in the environment [65], leaving it as of yet unknown, when semantic violations occur during such continuous illusions with constantly changing discrepancies.

3 HAND REDIRECTION ILLUSIONS

To investigate continuous illusionary mismatches that are caused by exposing participants to noticeable discrepancies, we used a hand redirection illusion. Here, we apply a continuous offset to the position of the virtual hand from the position in the real world. This can be done by changing the Control-Display (C/D) ratio, introducing a gain factor scaling up participants' virtual movements ($C/D > 1.0$) [12, 76]. In this work, we only consider C/D ratio gains > 1.0 because scaling up participants' virtual movement (i.e. making it faster) is more prevalent than scaling it down [22]. For example, Fig. 1 illustrates the effect when moving a proxy from a start to a target location. As a result, the virtual hand gradually moves further away from the real hand. We included this illusion in our study, since it is widely used in the community, and has been used across a wide application spectrum such as extending interactions [6, 17, 23, 27, 60, 76], increasing the resolution of haptic

proxies [1, 4, 5, 7, 12, 22], creating the sensation of weight [63] or improving realism when using virtual tools [70]. It has also been observed that users adapt to noticeable discrepancies, allowing them to complete tasks effectively [42]. Thus, it may be valuable to understand the process through which this adaptation happens, and when we can reliably understand that participants have accommodated for the discrepancy.

4 PHYSIOLOGICAL DATA & INTERACTIONS

In this work, we investigate if these noninvasive physiological measures may show a unique signature that can be used to reliably detect hand redirection illusions that go beyond individuals’ perceptual boundaries. Additionally, we also incorporate measures that are linked to the interaction itself. Each of these measures has been used in prior related work, and we order them based on whether the measure was likely to provide a strong signal.

EEG: As discussed above, there exists a body of work which uses EEG, specifically ERPs, to detect errors in VR. Commonly, the aforementioned studies report an effect in the frontal cortex area (FCz) between 100 to 360 ms after the stimulus onset [24, 58, 65, 67]. We hypothesize that noticeable illusions show a similar ERP pattern to errors in VR in the frontal cortex area, but in contrast to prior work, it is unclear when the effect occurs (e.g., beginning vs. middle vs. end of the movement phase).

Movement phases: Aimed movements can be separated into two distinct movement phases, ballistic and (an optional) correction [46]. It has been shown that introducing discrepancies between the real and the virtual world influences the execution of targeted movements, because users need to compensate for the offset [22]. Therefore, we expect that noticeable illusions also result in consistent shifts of movement phases, leading to a much longer correction phase when experiencing a noticeable hand offset, because users would need to correct for the unexpected discrepancy.

EDA: Facing noticeable VR illusions may trigger physiological arousal, and thus increases in the number of skin conductance response (SCR) peaks. We formulated this hypothesis based on findings reporting a correlation between greater self-reported VR immersion scores and lower physiological (EDA) responses [16].

RSP: Performing precise interactions while coping with changing visuo-motor discrepancies is a challenging task. Respiration appears to be affected by various environmental stimuli in VR [18, 31]. Therefore, we hypothesize that changes in respiratory rate appear when encountering detectable illusions as a direct response to the unexpected mismatch.

ECG: Changes in HR/HRV measures can be observed when studying long-term immersion [51] and therefore may also be observed when exposing participants to a noticeable illusion, which disrupts presence, resulting in a less realistic experience [68]. However, our work involves a comparatively short interaction, and it is unclear whether this would affect the slow-adapting HRV measure.

5 EXPERIMENT

We conducted an experiment in a quiet room with air conditioning to ensure a room temperature of 22–24°C, which is ideal for high-quality physiological data acquisition. We used a non-distracting virtual environment consisting of two tables, the experimental setup, and an instruction screen. Participants remained seated on a chair throughout the experiment. They wore an HMD, an EEG headset and various sensors on their body. They were told to move the virtual object forward until it matched a target location displayed in the virtual world. After they successfully established the position, they were required to maintain the position for two seconds before moving the object back to the start location. We explicitly showed participants the effect of C/D ratio manipulations multiple times during the warm-up phase to ensure that they understood the effect.

5.1 Research Objectives

In this work, we investigate five research objectives.

R1: Are EDA, RSP and ECG responses triggered when experiencing a noticeable hand redirection illusion?

R2: Can interaction data reveal a noticeable illusion?

R3: Does a noticeable VR illusion show a distinguishable ERP?

R4: When in the interaction does a semantic violation happen?

R5: Do humans adapt to noticeable virtual/real hand discrepancies over time, and if so, how long does it take for them to adapt to these discrepancies?

5.2 Design

In this experiment we use a within-subjects design. We had four study conditions: two Baselines, a Steady and a Mixed, each consisting of 16 trials (see Fig. 2).

Baseline: Uses a 1-to-1 mapping between real and virtual movement corresponding to a C/D ratio of 1.0. These conditions are used to collect ground-truth data [24] about what participants experience to be “normal”, which needs to be captured in VR [16]. By including two baseline conditions, we are able to perform consistency checks on our collected sample.

Steady: A fixed C/D ratio is applied that lies above an individual’s perceptual boundary, i.e. they can detect the visuo-motor discrepancy. To establish this per-participant ratio, we use a pre-calibration described below. This condition enables us to address (R5), because participants may adapt to the pre-calibrated threshold, i.e. they do not notice the manipulation anymore, even though they did initially.

Mixed: Randomly jumps between two C/D ratios (1:1 from Baseline, and the one used in Steady). The system ensures an equal occurrence of both C/D values, and only allows for a max. of three consecutive trials with the same C/D ratio. With this, we want to provoke a situation where a participant fails to adapt to the visuo-motor discrepancy, because of the repeatedly changing C/D ratios.

Each of the Steady and Mixed conditions was paired with a preceding Baseline condition, respectively. This was done to ensure that participants could establish a stable model of the environment, i.e. re-calibrate themselves, before experiencing a condition with visuo-motor discrepancies. To counterbalance the study, we alternated the order of the two blocks A and B for each participant, resulting in a factor two design (see Fig. 2).

5.3 Participants

We recruited 22 right-handed participants (five females, seventeen males), aged 18–31 (mean: 25.05; SD: 3.05) from the general public and the local university. Participants had a range of different educational and professional backgrounds including media informatics, computer science, education, pharmacy, cybersecurity, entrepreneurship, biomedical engineering, data science and artificial intelligence. All participants reported normal or corrected-to-normal vision and did not report any known health issues which might impair their perception or proprioception. Nine participants had never used VR before, ten had used it a few times (one to five times a year), no one reported using it often (6–10 times a year), and three others used it on a regular basis (more than 10 times a year). Participants not associated with our institution received €15 as remuneration for taking part in the experiment. The study was approved by the University’s Ethics and Hygiene Board.



Figure 2: Study design. Even participant#: AB, odd participant#: BA.

5.4 Apparatus

We used the apparatus consisting of an HTC VIVE Pro tracking system with SteamVR (v.1.22) and OpenVR SDK (v.1.16.8). The simple virtual scene was developed in Unity3D (v.2022.1.0). We used an Acer Predator Orion 5000 PO5-615s offering an Intel® Core i9 10900k CPU, 32 GB RAM and an Nvidia® GeForce RTX 3080 for running the experiment. We included an androgynous hand representation without noticeable characteristics [64] to prevent unwanted effects [57]. To avoid mismatches due to error-prone hand tracking [65], the virtual hand was affixed to the virtual object, and initially aligned with a participant's real hand. The experimental logic was implemented using the Unity Experiment Framework (UXF v.2.1.1) [8] and the Unity Staircase Procedure Toolkit [78].

5.4.1 EEG Setup

EEG data was recorded from 32 actively amplified electrodes using BrainAmp DC amplifiers from BrainProducts. Electrodes were placed according to the international 10–20 system. We applied conductive gel to establish a proper connection between electrodes and scalp, and brought impedance of all active electrodes below 20 kOhm, before continuing in the experiment. Impedance was verified before and after the study. EEG data was recorded with a sampling rate of 500 Hz. The validity of EEG recordings with an HMD has been verified previously [34].

5.4.2 Biosignal Setup

To collect ECG, RSP, and EDA data, we used biosignalsplux's multi-sensor research platform together with the OpenSignal software, allowing for medical-grade data acquisitions. The corresponding sensors were attached to the human body using six pre-gelled & disposable electrodes. Signals were streamed at a 500 Hz sampling rate and 16-bit resolution per channel. Overall, participants wore four sensors consisting of (1) an ECG sensor with an Einthoven Lead I setup, (2) a piezo-electric respiration (PZT) sensor between the 8th and 10th rib, (3) a two electrode EDA sensor attached to participants' palms and (4) a reference on their right collarbone. We synchronized tracking, EEG and biosignal data, and events of the study procedure using labstreaminglayer [47].

5.5 Procedure

Participants were given a general introduction to the study, i.e., we showed them the setup and explained each sensor to make sure that they were comfortable having their physiological data tracked. Next, we gathered participants' consent and asked them to fill in a demographics questionnaire. We then started with the procedure of attaching the physiological sensors. There were always two experimenters available, one identifying as male and another one as female. Participants could choose who should assist during the procedure to ensure proper placement of the sensors. Next, both experimenters fitted the EEG cap on participants' heads. Overall, the preparation time was about 40 minutes.

Following this, participants were placed in the VR environment and guided through an open-ended practice round, showing them the effect of C/D ratio manipulations (i.e., the virtual hand moves faster than their real hand). By doing so, we allowed them to familiarize themselves with the task and the system. Once they felt comfortable, we moved to the second phase, where we calibrated their personal detection threshold as described in the next section.

Participants were instructed to grasp the proxy object as visually indicated and to maintain this pose throughout the experiment. They were told to sit comfortably and only move the object to the target position with a consistent and comfortable speed. The system monitored that they stayed within a reasonable time range. Once they reached the goal position, the object needed to remain in that position for 2 seconds, before the visual start target appeared again. Participants were required to stay within a 5 mm distance (C/D ratio

= 1.0) for the countdown to remain active. To account for increased task difficulty caused by greater C/D ratios, the threshold distance was adjusted using the Fitts's law method [49]. Participant and experimenters were not allowed to talk during each trial, to avoid interrupting the continuous docking task or introducing artifacts in the data. To minimize carry-over effects and cope with proprioceptive fatigue [62], participants took a break after each study conditions, and filled in a questionnaire in VR. On average, the data-collecting process was 25 min long, during which participants were not allowed to remove the VR headset. The total experiment took about 90 min.

5.6 Determining Individuals Perceptual Boundaries

To tailor our study to each participant's individual perceptual boundaries, we calibrated their personal detection thresholds, which were then used in the mixed and steady conditions. This was done because prior work demonstrated large differences between detection thresholds, ranging from 5% to 67%, may be undetected [23]: what is an "obvious" C/D manipulation for one person may not be perceivable by another. To achieve this, we used a 3-up-1-down interleaved staircase procedure, exposing participants to different stimuli (C/D-ratios) repeatedly. Using an unequal step size, we target the Detection Threshold (DT) [38, 40], meaning that a participant can detect the manipulation 75% of the time. Since the procedure can target different probabilities, we can compute the required step-size (Ψ_{target}) for the step Up(Δ^+)/Down(Δ^-) method and $DT = .75$ as follows [38]:

$$\Psi_{target} = \frac{\Delta^+}{\Delta^+ + \Delta^-} \Rightarrow \frac{\Delta^-}{\Delta^+} = \frac{1 - \Psi_{target}}{\Psi_{target}} = \frac{(1 - 0.75)}{0.75} = \frac{1}{3}$$

The interleaved staircase uses a descending and an ascending sequence, and randomly assigns the next trial to one of the sequences. The procedure increases the next stimulus if a participant fails to detect the current stimulus, and decreases the next stimulus if the user detects the manipulation. A directional change within a sequence is noted as a reversal point (see Fig. 3, left). We used the number of reversal points ($r=3$) in each sequence as a convergence criterion. Based on previous studies in this field, we chose 1.0 and 2.0 for our range of manipulation factors with a 0.1 step size [6, 22, 23]. Following our pilot tests, we selected 1.0 (\uparrow asc.) and 1.8 (\downarrow desc.) as the starting values for the procedure to allow for quicker convergence. Finally, we added a relative 25% to the personalized detection threshold and since perception is non-linear [19], this ensures that the C/D gain is always noticeable.

5.7 Data Collection

We collected data from nine sources: a pre-study questionnaire for demographic information; EEG, ECG, RSP, EDA and interaction data; system logs (including trial times, object position and orientation, and velocity using UXF [8]); field notes and observations; and a subset of the avatar embodiment questionnaire [29] after each condition in VR using the VRQuestionnaireToolkit [21].

5.7.1 Avatar Embodiment Questionnaire

To establish a basis for overall analysis, we first need to ensure that participants feel that the avatar is easily controlled, and that the hand redirection is noticeable. Therefore, we use a subset of the complete avatar embodiment questionnaire [29] concerning only body ownership, agency and location, resulting in 10 questionnaire items. Responses were collected on a 7-point Likert scale (-3 = strongly disagree, 0 = neutral, 3 = strongly agree).

5.8 Analysis

For the analysis, our data was split into epochs corresponding to the conditions and the trials within them. Then, we pre-processed, filtered and analyzed the data using the method described below.

5.8.1 EEG

We utilized the EEGLAB [71] and MoBILAB [72] toolboxes inside the MATLAB environment for our analysis. We followed Gehrke et al.’s [24] methodology to pre-process, filter and analyze the EEG data and extract ERPs. To summarize, raw EEG data was re-sampled to 250 Hz, high-pass filtered at 1 Hz and low-pass filtered at 125 Hz. Then, the data was re-referenced to the average of all channels, followed by applying ICA to reject eye and line noise activity.

Extracting ERPs. To obtain the ERPs (shown in Fig. 5 right), we filtered the EEG data with a 0.2 Hz high-pass and 35 Hz low-pass filter. Then we created epochs from -0.2 seconds to 0.9 seconds around the MovementStartEvent, i.e., past the jitter threshold. To guarantee robust data, we rejected 10% of the noisiest epochs [30]. We focused our analysis on one electrode, FCz, located on the forehead, which previous studies found to be a strong predictor for visual and haptic mismatches [24, 65, 67]. Furthermore, we automatically extracted the ERP negativity peaks and their latencies by locating the minimum peak in a 400–700 ms time window after the MovementStartEvent, using a 10 Hz low-pass filter. The time window was derived from visual inspection of the mean difference ERP wave in Fig. 5 left.

5.8.2 Biosignals

We used NeuroKit2 [50] with its automated pipeline for pre-processing ECG, RSP and EDA signals.

ECG: Raw data was cleaned using a fifth order 0.5 Hz high-pass Butterworth filter and powerline filtering at 50 Hz. R peaks were extracted using NeuroKit2’s default method.

RSP: Raw data was cleaned using linear detrending followed by a fifth order 2 Hz low-pass IIR Butterworth filter. To extract RSP peaks, we applied Noto et al.’s [56] analysis method.

EDA: Raw data was cleaned using a fourth order 3 Hz high-pass Butterworth filter. We accepted SCR peaks below the rejection threshold set relative to the largest amplitude in the signal (min 0.1).

Finally, the results were then processed for descriptive and/or statistical analysis.

5.9 Results

In this section, we investigate whether all participants show consistent and systematic responses across the different modalities, seeking to understand whether the monitored measures show a consistent effect across all participants when knowingly exposing them to noticeable virtual/real hand mismatches. Here, the goal is to distinguish between VR experiences that do vs. do not use noticeable illusions.

Our data analysis is split into five parts based on the data source; for each, we discuss which research objective the data source addresses. First, we report the results from the the personalized detection thresholds procedure. Next we analyze the questionnaire responses to ensure that the avatar did not consciously create dissonance for our participants. Then, we look at the EEG data followed by the interaction data, and finally we report our results from the EDA, RSP and ECG data analysis. The biosignals analysis excludes P9 due to data loss, and the EEG analysis excludes participants 1–4 because of an experimenter error with the electrode setup. Statistical tests were chosen based on parametric test assumptions at $\alpha = .05$, and we use outlier removal with the box plot method.

5.9.1 Personalized Detection Thresholds

We collected 277 responses through the interleaved-staircase procedure. On average, it took participants 12.6 (SD: 2.6) trials to reach convergence. For each participant, we obtained one personalized detection threshold by averaging the last three reversal points within the ascending and descending staircase sequence [22, 79]. The results can be found in Fig. 3. The average detection threshold was 1.7 (SD: 0.12). All 22 staircase plots are available in the appendix.

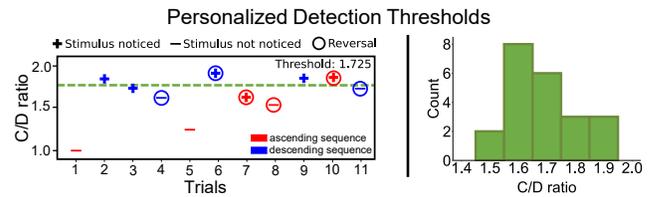


Figure 3: Detection threshold distribution and 3up-1down staircase.

5.9.2 Avatar Embodiment Questionnaire

Here, we analyze the responses from the virtual embodiment questionnaire, namely the items concerning body ownership, agency/motor control and location, by running Friedmann tests on the aforementioned three categories. The results are depicted in Fig. 4; all questionnaire items showed a main effect. Consequently, we ran multiple pairwise comparisons using paired Wilcoxon signed-rank tests and Holm adjustments.

Body ownership. One can have high body ownership over, e.g., a virtual hand that is dislocated from the position of the real hand when visuo-tactile stimuli are provided synchronously [29]. In the case of noticeable hand redirection gains, the initial position of the hand is correctly rendered in place, but once a user starts the movement a gradually changing offset is applied, leading to a break in body ownership. Q1 - “I felt if as my virtual hand was my real hand” supports this by showing significantly greater scores in both baseline conditions compared to the mixed condition.

Agency/motor control. These questions target whether participants can move their virtual body parts in a natural way. Interestingly, the baseline conditions showed significantly greater scores than the mixed condition on the question Q4 - “It felt like I could control the virtual hand as if it was my own hand”, but this was not the case for the steady condition. This shows that the random jumps between gain factors made it difficult for participants to predict and control their movements accurately. The steady condition however, lies between baseline and mixed, suggesting that participants adapted to the (large, but consistent) offset (R5).

Location. Here, the differences between the conditions became most visible in participants’ ratings. For example, the significant difference between baseline and mixed as well as steady and mixed on Q8 - “I felt as if my hand was located where I saw the virtual hand” suggests that participants clearly noticed the positional offset in the mixed condition. Interestingly, participants frequently asked: “I am not sure how to respond, because at the beginning it was faster (the effect of a high C/D ratio), but at the end it felt normal” (P11) after they finished the steady condition, showing that they adapted to even higher C/D ratios (R5).

Summary. The results demonstrate that: (1) we successfully established a solid baseline in our experiment with high body ownership, agency and location scores and (2) the induced hand redirection illusion was noticeable for participants, significantly affecting their questionnaire responses. This ensures that the data analysis below can be linked to the effects we aim to investigate. Next, we report the result from our EEG analysis.

5.9.3 EEG Analysis

The EEG analysis allows us to study research objective (R3), does a noticeable VR illusion show a distinguishable ERP, and (R4), at which point in the interaction does a semantic violation happen. We examined the mean ERP amplitudes across all participants depicted in Fig. 5, which shows a strong prediction error negativity at about 420 ms latency throughout all conditions and a second error negativity around 630 ms after participants’ movements started, but only in the mixed and steady conditions. We observe that main positive components begin to accumulate at 800 ms. Interestingly,

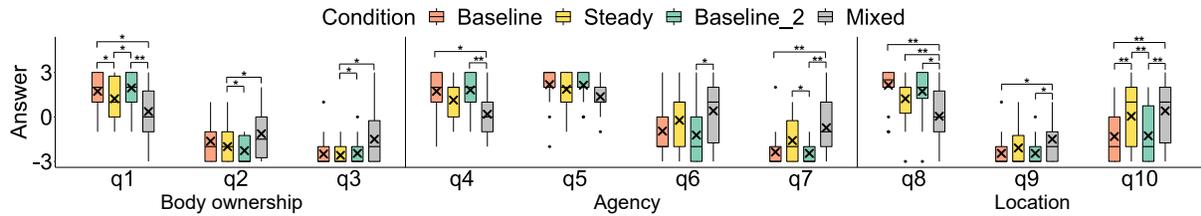


Figure 4: Responses from the avatar embodiment questionnaire items. *** = $p < .001$; ** = $p < .01$; * = $p < .05$ (Holm-adjusted).

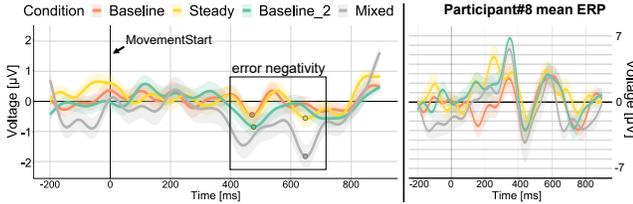


Figure 5: Mean ERP amplitudes for all study conditions across participants (left) and P8 (right). Time window for peak negativity.

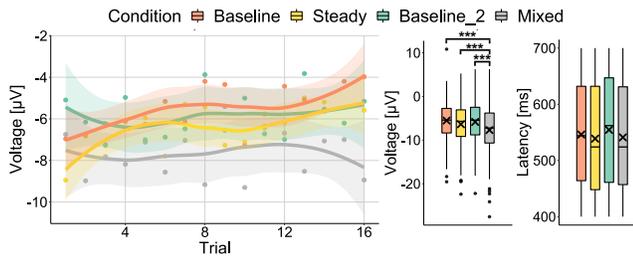


Figure 6: Negativity peak amplitudes trend, severity and latency across conditions. *** = $p < .001$.

there appear to exist two negativity peaks in the mixed and steady condition. Hence, we examined individuals' ERP plots, but found that the second negativity peaks are not a unique ERP signature of these conditions, allowing us to distinguish them from the baseline. Rather, it reflects the temporal differences of participants noticing the discrepancy. This is a reasonable observation, because participants experienced different C/D manipulations and moreover, most likely also differ in the perceptual abilities to detect them. Also, when investigating participants' individual ERP plots, for the vast majority, we observed the greatest prediction error negativity in the time window from 420–630 ms latency (see Fig. 5). **We conclude that participants may detect obvious hand redirection illusions within the first $\frac{1}{3}$ of the total movement time (on average 1.3 seconds) across conditions, taking into account that ERPs occur between 100–360 ms after stimulus onset [24, 58, 65, 67] (R4).**

Since all conditions showed error negativity between 420–630 ms, we investigated the severity of the visuo-motor conflict by computing and analyzing the global minimum prediction error amplitudes and their latencies [24]. We found a main effect for the amplitude of a global minimum $\chi^2(3) = 22.482, p < .0001$ between our four study conditions. The mixed condition showed significantly stronger prediction negativity than all other conditions as illustrated in Fig. 6, right. The strength of global minimum prediction error amplitudes in the mixed condition is comparable to the visual mismatch negative amplitude reported by Gehrke et al. [24], demonstrating the validity of the results. Second, as shown in Fig. 6, right, we observed no significant differences $\chi^2(3) = 3.988, p = 0.262$ for the peak latencies across the four conditions. **We conclude that it is possible**

to distinguish the mixed from the other conditions based on error negativity peak amplitude, i.e. the severity of the mismatch is reflected in the negativity (R3).

Next, we explore ERP negativity in the steady condition to consider research objective (R5), whether humans adapt to noticeable virtual/real hand discrepancies. We investigate the distribution of negativity peaks within our conditions across the 16 consecutive trials per participants. We observed much greater negativity in earlier trials, which decreases throughout the steady condition, quickly matching the baseline after trial 4 and reaching a stable plateau between trial 8 and 12 (see Fig. 6). We ran a Spearman correlation analysis on the negativity peaks along the 16 trials for each condition, and found a small positive correlation in the steady ($peaks_s(2263338) = .17, p = .006$) and the baseline ($peaks_b(2336333) = .15, p = .013$) conditions, but not in the mixed ($p = .778$) and baseline_2 conditions ($p = .606$). **Our results suggest that there is a possibility of almost immediate adaptation to hand redirection i.e., participants quickly re-calibrate themselves to the given discrepancy.**

Finally, we were specifically interested in the mixed condition, because participants reported that this condition was "...all over the place" (P18) and was often perceived as "...movement was either too fast [C/D gain > 1] or too slow [C/D gain = 1]." (P1). We extracted the baseline trials from the mixed condition, and re-ran the analysis on the global minimum prediction error amplitudes $\chi^2(3) = 213, p < .001$. The mixed_baseline showed significantly greater error negativity than baseline ($p < .001$) and baseline_2 ($p < .001$). **Thus, we conclude that even the 1-to-1 mapping can be experienced as disruptive if participants had previously adapted to a different C/D ratio.**

Summary. We found that all conditions (incl. baseline) provoked error negativity consistently between 420–630 ms and thus, we conclude that participants notice errors early in their movements. Additionally, the minimum prediction error amplitudes allowed us to distinguish the mixed from all other conditions. Further, participants seem to adapt quickly to the discrepancy, reflected in decreasing error negativity. In the following we look at the interaction data.

5.9.4 Interaction Data

We analyze how noticeable illusions affect two main movement phases: ballistic and correction. We hypothesized that C/D ratio manipulations increases the relative duration of the correction phase and shortens the ballistic phase, because participants start to compensate for the offset. There exist variations on how these phases are defined, but we follow Nieuwenhuizen et al. [55] to obtain ballistic and correction phases. According to Liu et al. [45]'s recommendation, we analyze movement phases across two dimensions, *distance* and *time*, to achieve a more encompassing view on the interaction.

Movement profiles: Our central interest lies at the transition point between the ballistic and correction phases, which we normalized (by travel distance and time) to account for differences in task completion time and total distance travelled. For analysis, we computed movement profiles for all 1408 interactions (see

appendix). We ran a Friedman test on the normalized (time and distance) transition points. The Friedman test revealed a main effect for both independent variables, distance $\chi^2(3) = 53, p < .001$ and time $\chi^2(3) = 53, p < .001$. Hence, we conducted multiple post-hoc pairwise Wilcoxon rank tests with Holm adjustments (see Fig. 7).

Time. The results show a significantly shorter ballistic phase in the mixed, compared to both baselines ($p < .001$) and the steady condition ($p < .001$). However, we could not identify an effect between the steady and baseline conditions on the time dimension. The total task completion time $\chi^2(3) = 176, p < .001$ was significantly faster in the steady condition than in all other conditions, which is not surprising, because we compensated for the increased task difficulty using Fitts's law.

Distance. On the other hand, distance results reveal that both mixed and steady conditions have significantly later transition points ($p < .001$). As illustrated in Fig. 7, we observed high levels of participants over- and under- shooting the target due to unexpected slow or fast movements. In contrast, baseline conditions showed consistent transition points, before or right at the required distance, indicating high accuracy. Looking at the total distance travelled $\chi^2(3) = 213, p < .001$ reveals that significantly more distance was covered in the steady and mixed condition than in both baselines. Again, this demonstrates the under- and overshooting in combination with many corrections due to erroneous movements, in contrast to the baseline conditions in which overshooting was nonexistent.

With this in mind, we investigate whether participants adapt to virtual/real hand offsets (R5) by running a Spearman correlation analysis on the normalized transition points alongside the 16 trials for each condition. We found a medium negative correlation in the steady condition on ($distance_r(20) = -.37, p < .001$), and a positive correlation for ($time_r(20) = .25, p < .001$), both indicating that transition points move closer to "normal", i.e., to both baseline conditions, with each trial. In contrast, this correlation was not present in the mixed ($time_r(20) = .08, p = .148; distance_r(20) = -.09, p = .106$) and baseline.2 conditions ($time_r(20) = .07, p = .219; distance_r(20) = -.02, p = .713$). However, baseline showed a weak correlation ($time_r(20) = .12, p = .034; distance_r(20) = .16, p = .006$) which may be attributed to learning effect, because it was the first study condition. Fig. 8 illustrates how P4 adapts to a constant C/D gain in the steady condition with a large correction phase at the beginning (trial 1), and decreasing correction phases in trial 4, 8 and 12 until the end of the study condition. This is in line with our results from the questionnaire and EEG analysis (R5).

Analogously to the ERP analysis, we extracted baseline trials from the mixed condition and re-ran the analysis. Mixed_baseline had shorter and earlier transition points than baseline ($time_p = .004; distance_p = .003$) and baseline.2 ($time_p < .011; distance_p < .001$). Again, this provides supporting evidence that there is no "absolute" baseline, but instead participants quickly adapt to the offsets, and thus, how an illusion is perceived is relative to a user's most recent experience with this kind of interaction (R5).

Summary. *Movement profiles, time and distance appear to be reliable metrics to distinguish unnoticeable vs. noticeable hand redirection offsets, because these are directly affected by the performed interaction (R2). Further, we found that participants adapt to consistent offsets in the steady condition (R5), which was not the case in the mixed condition.*

5.9.5 EDA, RSP and ECG Data Exploration

EDA: Participant 15, 17 and 20 were omitted from the analysis, because it was not possible to compute SCR peaks, possibly due to noisy data. The corresponding raw/cleaned data plots of all participants can be found in the appendix.

To investigate our hypothesis, that the total number of SCR peaks increases in the mixed condition because it creates arousal, we ran a repeated-measures ANOVA. However, we did not find a main

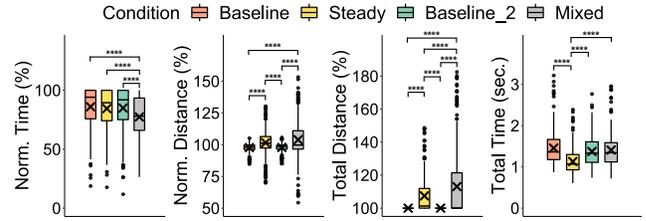


Figure 7: Time and real-world distance needed to reach the virtual target. **** = $p < .0001$ (Holm-adjusted).

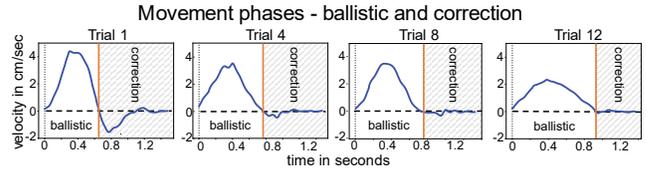


Figure 8: Movement normalizes and becomes more accurate in the steady condition over time. Correction phase becomes shorter.

effect on the total number of SCR peaks across our study conditions ($F_{(3,48)} = 2.025, p = .123$). We further analyzed the data by examining the SCR plots. Overall, we observed that the majority of SCR peaks seem to appear at the beginning of a study block, which might suggest that this phase is used for initial self-calibration—even in the baseline condition, although there was a break and a questionnaire between study conditions. To further investigate this, we split each study condition into quartiles, analogous to an extreme groups approach [61], and then computed the number of SCR peaks for each quartile/condition. The results are depicted in Fig. 9, showing a downward trend in the steady condition similar to the baseline, especially in the 3rd quartile (trial 9 to 12), but an equal distribution of SCR peaks in the mixed condition.

Summary. *Even though we could not identify a significant effect on the overall number of SCR peaks, our results may be interpreted as evidence that participants adapt to even larger offsets (R5), but fail to acclimate to the constantly changing offsets in the mixed condition, reflected in the distribution of SCR peaks.*

To the best of our knowledge, we are the first to explore the relationship between ECG and RSP in the context of VR illusions. Therefore, hypothesis testing using statistical methods is inappropriate; instead, we use a descriptive and exploratory data analysis, splitting the data into four epochs corresponding to our four study conditions. We performed an interval analysis, computing the most-common RSP, time, frequency as well as non-linear ECG indices accordingly to Makowski et al. [50] and Pham et al. [59]. Overall, we obtained 86 indices that may be selected for a statistical analysis or machine-learning approaches. Here, we only report a subset of our results. However, we make the data set publicly available to inform future research hypotheses and studies.

RSP: First, we descriptively analyze respiration profiles, because in our study, we observed that participants' responses differed dramatically. For instance, in the mixed condition (see Fig. 12, bottom), P11 held his breath for several seconds followed by heavily in/exhaling, because he was focused on matching the target and start position while coping with the effects of frequently changing C/D ratios. On the other hand, P5 was "surprised" when the C/D ratio changed, triggering smaller exhales corresponding to the trials, leading to a completely different breathing pattern, demonstrating how different individual participants' reactions can be. However, most participants seem to fall in the latter category, according to RSP

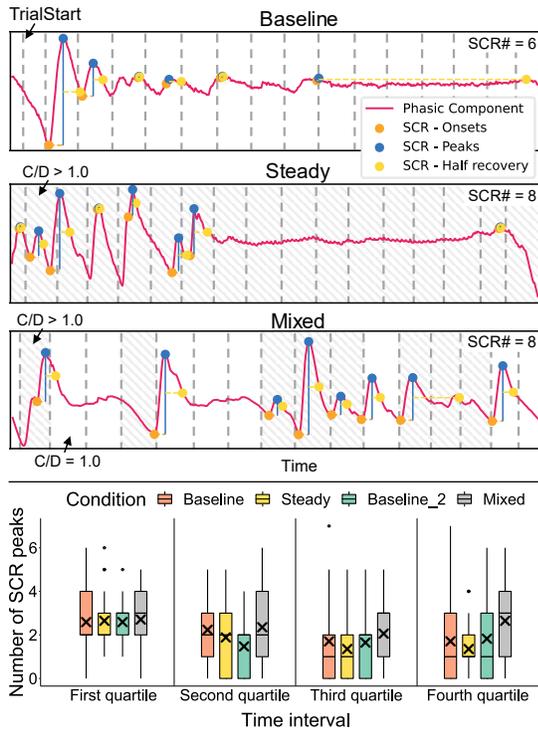


Figure 9: Top: phasic component during baseline, steady and mixed condition. Bottom: number of SCR peaks separated in quartiles.

rate and amplitude measures, when comparing mixed with baseline conditions. Overall, there seems to be a trend towards an increase in RSP rate and more shallow breathing in the mixed condition.

Principal components analysis. Next, we computed 22 common RSP indices, and we performed a dimensional reduction using principal components analysis (PCA) in order to determine the number of components that are needed to describe the variance in the data. In the next step, we check whether the PCs reveal clusters within our collected sample, allowing us to distinguish study conditions. The PCA showed that 65.2% of the variance can be explained by PC1 and PC2. Five additional PCs are needed to reach the often-propagated total variance of $s\%$ for exploratory data analysis (see Fig. 10). However, when inspecting the biplots, we see that even the most powerful principal components PC1 and PC2 do not indicate a distinct separation in clusters that correspond to our study conditions (see Fig. 10). Although this is desirable for both baseline conditions, it shows that there seems to be no consistent pattern within our collected data which can be linked to our study conditions.

Standardized mean difference. Next, we compared standardized mean difference between conditions using paired Cohen's d . A d of .5 is generally considered a medium effect size, meaning that 69.1% of one group will be above the mean of the other group (Cohen's U_3). Six indices between baseline and mixed conditions were above $d = .5$. Additionally, the RSP_Amplitude_Mean metric still shows noticeable mean differences between baseline and steady conditions. Fig. 11 shows the results of the dimensional reduction in combination with d scores above .5 in at least one comparison.

Summary. Exposing participants to a noticeable VR illusion seems to have an impact on respiration. That being said, it is unclear whether changes in respiration are consistent following our observations and the PCA analysis. However, a combination of RSP_Amplitude_Mean and RSP_Rate appear to be worth investigating in the future (R1).

ECG: We conducted a similar exploratory analysis for the ECG

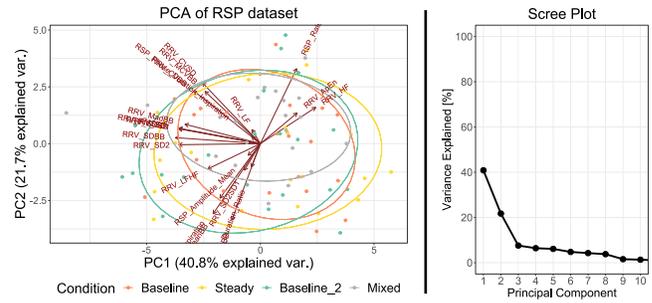


Figure 10: Biplot of PC1 and PC2 describing the greatest variance in the data. Scree plot: number of PCs needed to describe variance.

	Rate Mean	Ampli Mean	RRV CVSD	Inspira tion	Expira tion	Median BB	Mean BB	Cohen's d
Baseline/ Mixed	.333	.743	.502	.493	.162	.507	.400	1.0 0.8 0.6 0.4 0.2
Baseline_2/ Mixed	.437	.459	.477	.384	.547	.527	.524	
Baseline/ Steady	.161	.454	.014	.023	.023	.264	.229	
Baseline_2/ Steady	.307	.307	.014	.218	.323	.264	.293	

Figure 11: RSP Cohen's d in combination with dimensional reduction.

data. We computed 64 common time, frequency, and non-linear ECG indices. Our goal was to narrow down potentially useful indices, providing as many insights into our data as possible.

Principal components analysis. A PCA on this data showed that 51.7% of the variance can be explained by PC1 (40.4%) and PC2 (11.3%). Five additional PCs are needed to reach a total variance of 80%. However, similarly to the RSP data, PC1 and PC2 do not allow a distinct separation in clusters that correspond to our study conditions (similar to Fig. 10; please see plot in appendix).

Standardized mean difference. We again computed Cohen's d scores. The first baseline against the mixed condition showed strong effects across the two indices: HRV_SD1SD2 ($d = .854$), describing long and short-term HRV variance, as well as HRV_CSI_Modified ($d = .835$). Further, we found that the standard deviation of RR-intervals (SDNN) showed consistent d scores across all conditions, especially in the mixed, but also the steady condition, compared to both baselines. Because SDNN is a known stress indicator, we specifically looked at other stress-related HRV indices, since coping with constantly changing C/D ratios in the mixed condition could have an effect on them. However, other well-studied stress indices such as RMSSD remained consistent across all study conditions and therefore did not suggest any differences. Fig. 13 shows the resulting d scores of each variable that contributed to the first seven principal components with at least a medium effect size of $d > .5$.

Summary. We descriptively analyzed ECG data, laying the foundations for future research that aim to detect noticeable VR illusions. We provide an overview of common ECG indices, helping the community to make informed decisions about research hypotheses by outlining the most promising ECG features. Specifically, the stress related index SDNN, as well as the SD1SD2 and CSI_Modified indices, appear to be promising. However, it must be noted that based on our PCA analysis, there exist inconsistencies in participants' ECG responses within our collected sample (R1).

5.9.6 Summary of Results

R1: We investigated EDA, RSP and ECG, providing an overview of the most promising indices that may be selected for future hypothesis-driven studies. Our analysis revealed inconsistencies

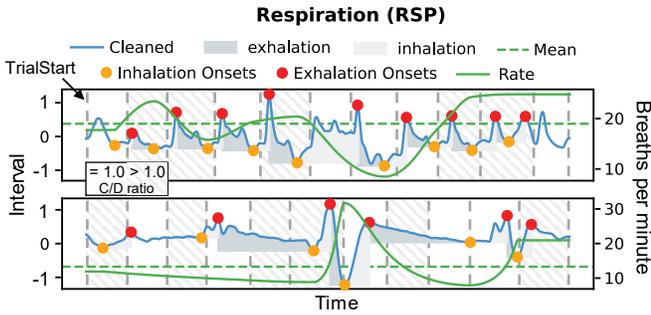


Figure 12: Different RSP responses when exposing P5 (top) and P11 (bottom) to frequently changing C/D ratios in the mixed condition.

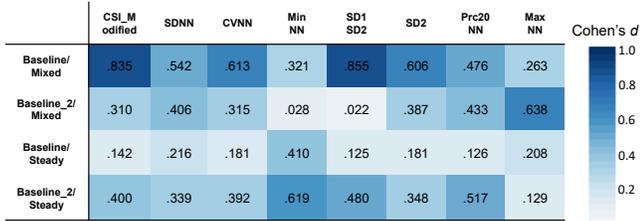


Figure 13: ECG Cohen's *d* in combination with dimensional reduction.

for RSP and ECG, while simultaneously demonstrating substantial individual differences in participants' physiological responses. Hence, at this point, we are unable to conclude if these measures can or cannot be used to distinguish study conditions. Our analysis on the SCR peaks was inconclusive, but we identified a downward trend in SCR peaks except in the mixed condition, which seems to indicate that participants adapted to the discrepancy.

R2: Interaction data, specifically the analysis of movement phases and transition points, showed a significant difference that could be systematically linked to our study conditions.

R3: ERPs were found in all conditions, but differed in their amplitudes and negativity, which allowed us to distinguish the mixed from the other conditions. The steady condition seemed (except for the initial strong negativity during the first four repetitions) mostly indistinguishable from the baseline conditions.

R4: The ERP analysis suggests that noticeable visuo-motor discrepancies were detected early after movement start, which is supported by the movement analysis, showing very early transition points from ballistic to correction phase in the mixed and steady conditions, caused by participants compensating for the discrepancy.

R5: We found overwhelming evidence that participants adapted to the discrepancies in the steady condition through analyzing participants' comments, questionnaire, EDA, movement and EEG data. The exact number of repetitions needed most likely depends on the individual. However, our data suggest that somewhere around repetition 4–8, participants were adapted to the discrepancy. Interestingly, steady and baseline conditions showed very similar patterns across all measures, suggesting that participants used the initial phase for self-calibration and to establish a robust model of their environment.

6 DISCUSSION & LIMITATIONS

6.1 Combining Physiological Modalities & Individuals

In this work, we investigated a set of physiological measures that we intend to expand, including eye blinks, pupil dilation and gaze. The next logical step is to combine multiple modalities for classifying whether a visuo-motor discrepancy was noticed or not. Just as prior research has found that detection thresholds differ between individuals [22], we observed that our participants' individual physiological

responses differed. Similarly, we expect that each user would have their own personalized classifier. Based on our data, it is unclear whether these responses would remain stable—an individual may, for instance, differ from one day to another.

Running an analysis on individuals across time requires substantially more data. This is difficult because of the trade-off between collecting sufficient samples versus proprioceptive fatigue [62]. It is thus possible that the lack of consistency in our data may be attributed to the small sample size; however, our effect-size analysis provides a foundation and a solid starting point for future studies.

6.2 Validity & Applicability of EEG Results

Our results confirm that the frontal cortex area, especially the FCz electrode, and the concept of ERPs [24, 65, 67] can be used to detect noticeable visuo-motor discrepancies. However, previous results show a more consistent ERP mean curve, which can be explained by the distinct events used in their studies. Instead, our work introduces gradually increasing offsets, where the precise time of violation is unclear and thus, temporal shifts in the ERP curves occur. To our surprise, even the baseline conditions showed weak error negativity, which is reduced over time (see Fig. 5) and therefore may be caused by the visualization itself [67] (i.e., the simple floating hands combined with reduced depth perception in VR). Other measures than ERPs to detect a mismatch also exist and could be used for comparison with our results. For instance, a haptic delay has been shown to significantly increase beta and theta band activity [2]. Another interesting aspect is the indistinguishability between the steady and baseline conditions, when compared to the questionnaire responses. We assume that the questionnaire responses were given by participants averaging their experiences in the steady condition, leading to a score in-between those of the baseline and mixed conditions. In contrast, the EEG results perhaps give a more direct estimate of how the illusion was perceived, also showcasing the quick adaptation. Lastly, determining the time (**R4**) when a violation gets noticed is a crucial aspect for many techniques that aim to increase detection thresholds for VR illusions (e.g., by utilizing eye blinks [77]).

6.3 Beyond Noticeable & Hand Redirection

Our long-term goal is to tailor illusions to an individuals' perceptual boundary [23, 44]. In this work, we used discrepancies that are obvious to the user; however, more conservative studies found that the detection thresholds are much smaller [76]. Therefore, we need to investigate discrepancies which are not only above, but also around and below participants' thresholds. Future work needs to investigate whether unnoticeable discrepancies trigger similar effects, allowing us to differentiate them. In our vision, VR designers could already tell when they are reaching perceptual boundary of an individual, by staying just below the threshold. Therefore, we need to study more realistic settings and scenarios to investigate the approach's robustness. Finally, we also plan to apply this method for different types of VR illusions such as redirected touch [42].

7 CONCLUSION

In this work we investigated physiological and interaction data to implicitly detect noticeable movement discrepancies between a user's real and virtual hand by running a study with 22 participants. The results suggest that EEG is a reliable method to detect visuo-motor discrepancies. Furthermore, movement phases appeared to be directly affected by noticeable discrepancies. We provide the first investigation of ECG and RSP in relation to perceptual VR illusions, and outline recommendations for indices that appear to be worth studying in dedicated studies. We also found that participants quickly adapted to larger discrepancies, and whether a discrepancy remains unnoticed primarily depends to their most recent experience with this kind of interaction. Our work marks a first step towards VR experiences that include personalized VR illusions.

REFERENCES

- [1] P. Abtahi and S. Follmer. Visuo-Haptic Illusions for Improving the Perceived Performance of Shape Displays. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, pp. 1–13. ACM, Montreal, QC, Canada, 2018.
- [2] H. Alsuradi, W. Park, and M. Eid. Midfrontal Theta Power Encodes the Value of Haptic Delay. *Scientific Reports*, 12(1):1–11, 2022.
- [3] B. Araujo, R. Jota, V. Perumal, J. X. Yao, K. Singh, and D. Wigdor. Snake Charmer: Physically Enabling Virtual Objects. In *Proceedings of the TEI '16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*, TEI '16, p. 218–226. Association for Computing Machinery, New York, NY, USA, 2016.
- [4] M. Azmandian, M. Hancock, H. Benko, E. Ofek, and A. D. Wilson. Haptic Retargeting: Dynamic Repurposing of Passive Haptics for Enhanced Virtual Reality Experiences. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2016.
- [5] Y. Ban, T. Narumi, T. Tanikawa, and M. Hirose. Displaying Shapes with Various Types of Surfaces using Visuo-Haptic Interaction. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*, VRST '14, pp. 191–196. Association for Computing Machinery, Edinburgh, Scotland, Nov. 2014.
- [6] B. Benda, S. Esmacili, and E. D. Ragan. Determining Detection Thresholds for Fixed Positional Offsets for Virtual Hand Remapping in Virtual Reality (ISMAR). In *2020 IEEE International Symposium on Mixed and Augmented Reality*, ISMAR '20, pp. 269–278, 2020.
- [7] J. Bergström, A. Mottelson, and J. Knibbe. Resized Grasping in VR: Estimating Thresholds for Object Discrimination. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, UIST '19, p. 1175–1183, 2019.
- [8] J. Brookes, M. Warburton, M. Alghadier, M. Mon-Williams, and F. Mushtaq. Studying Human Behavior with Virtual Reality: The Unity Experiment Framework. *Behavior Research Methods*, 52, 2020.
- [9] E. Burns, S. Razzaque, A. Panter, M. Whitton, M. McCallus, and F. P. Brooks. The Hand is Slower than the Eye: A Quantitative Exploration of Visual Dominance over Proprioception. In *IEEE Proceedings. VR 2005. Virtual Reality*, 2005, pp. 3–10, 2005.
- [10] E. Burns, S. Razzaque, A. T. Panter, M. C. Whitton, M. R. McCallus, and F. P. Brooks. The Hand Is More Easily Fooled than the Eye: Users Are More Sensitive to Visual Interpenetration than to Visual-Proprioceptive Discrepancy. *Presence: Teleoperators and Virtual Environments*, 15(1):1–15, 2006.
- [11] D. Caruelle, A. Gustafsson, P. Shams, and L. Lervik-Olsen. The Use of Electrodermal Activity (EDA) Measurement to Understand Consumer Emotions – A Literature Review and a Call for Action. *Journal of Business Research*, 104:146–160, Nov. 2019.
- [12] L.-P. Cheng, E. Ofek, C. Holz, H. Benko, and A. D. Wilson. Sparse Haptic Proxy: Touch Feedback in Virtual Environments Using a General Passive Prop. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, p. 3718–3728, 2017.
- [13] I. Choi, H. Culbertson, M. R. Miller, A. Olwal, and S. Follmer. Grability: A Wearable Haptic Interface for Simulating Weight and Grasping in Virtual Reality. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, UIST '17, pp. 119–130. Association for Computing Machinery, Québec City, QC, Canada.
- [14] X. de Tinguy, C. Pacchierotti, M. Emily, M. Chevalier, A. Guignardat, M. Guillaudeux, C. Six, A. Lécuyer, and M. Marchal. How Different Tangible and Virtual Objects Can Be While Still Feeling the Same? In *2019 IEEE World Haptics Conference*, WHC, pp. 580–585, 2019.
- [15] X. de Tinguy, C. Pacchierotti, M. Marchal, and A. Lécuyer. Toward Universal Tangible Objects: Optimizing Haptic Pinching Sensations in 3D Interaction. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 321–330. IEEE, Osaka, Japan.
- [16] D. Egan, S. Brennan, J. Barrett, Y. Qiao, C. Timmerer, and N. Murray. An evaluation of Heart Rate and Electrodermal Activity as an Objective QoE Evaluation Method for Immersive Virtual Reality Environments. In *2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX)*, pp. 1–6, 2016.
- [17] S. Esmacili, B. Benda, and E. D. Ragan. Detection of Scaled Hand Interactions in Virtual Reality: The Effects of Motion Direction and Task Complexity. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 453–462. IEEE, 2020.
- [18] K. A. Fadeev, A. S. Smirnov, O. P. Zhigalova, P. S. Bazhina, A. V. Tumialis, and K. S. Golokhvast. Too Real to Be Virtual: Autonomic and EEG Responses to Extreme Stress Scenarios in Virtual Reality. *Behavioural Neurology*, 2020:e5758038, Mar. 2020. Publisher: Hindawi.
- [19] G. T. Fechner. *Elemente der Psychophysik*, vol. 2. Breitkopf and Härtel, 1860.
- [20] M. Feick, S. Bateman, A. Tang, A. Miede, and N. Marquardt. TanGi: Tangible Proxies for Embodied Object Exploration and Manipulation in Virtual Reality. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 195–206. ISSN: 1554-7868.
- [21] M. Feick, N. Kleer, A. Tang, and A. Krüger. The Virtual Reality Questionnaire Toolkit. In *Adjunct Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, AP UIST 2020, UIST '20 Adjunct, pp. 68–69. ACM, 2020.
- [22] M. Feick, N. Kleer, A. Zenner, A. Tang, and A. Krüger. Visuo-Haptic Illusions for Linear Translation and Stretching Using Physical Proxies in Virtual Reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, pp. 1–13, 2021.
- [23] M. Feick, R. Kora, A. Tang, and A. Krüger. Designing Visuo-Haptic Illusions with Proxies in Virtual Reality: Exploration of Grasp, Movement Trajectory and Object Mass. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, number 220, pp. 1–13. Association for Computing Machinery, New York, NY, USA.
- [24] L. Gehrke, S. Akman, P. Lopes, A. Chen, A. K. Singh, H.-T. Chen, C.-T. Lin, and K. Gramann. Detecting Visuo-Haptic Mismatches in Virtual Reality using the Prediction Error Negativity of Event-Related Brain Potentials. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow, Scotland, UK.
- [25] J. J. Gibson. Adaption, After-effect and Contrast in the Perception of Curved Lines. *Journal of Experimental Psychology*, 16(1):1–31, 1933.
- [26] E. J. Gonzalez, P. Abtahi, and S. Follmer. REACH+: Extending the Reachability of Encountered-type Haptics Devices through Dynamic Redirection in VR. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*. Virtual Event USA, 2020.
- [27] E. J. Gonzalez and S. Follmer. Investigating the Detection of Bimanual Haptic Retargeting in Virtual Reality. In *25th ACM Symposium on Virtual Reality Software and Technology*, VRST '19, pp. 1–5, 2019.
- [28] M. Gonzalez-Franco and J. Lanier. Model of Illusions and Virtual Reality. *Frontiers in Psychology*, 8, 2017.
- [29] M. Gonzalez-Franco and T. C. Peck. Avatar Embodiment. Towards a Standardized Questionnaire. *Frontiers in Robotics and AI*, 5, 2018.
- [30] K. Gramann, F. U. Hohlefeld, L. Gehrke, and M. Klug. Human Cortical Dynamics during Full-body Heading Changes. *bioRxiv*, 2021.
- [31] J. Guna, G. Geršak, I. Humar, M. Krebl, M. Oreš, H. Lu, and M. Pogačnik. Virtual Reality Sickness and Challenges Behind Different Technology and Content Settings. *Mobile Networks and Applications*, 25(4):1436–1445, Aug. 2020.
- [32] J. Hartfill, J. Gabel, L. Kruse, S. Schmidt, K. Riebandt, S. Kühn, and F. Steinicke. Analysis of Detection Thresholds for Hand Redirection during Mid-Air Interactions in Virtual Reality. In *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology*, VRST '21, pp. 1–10. ACM, New York, NY, USA.
- [33] S. Heo, J. Lee, and D. Wigdor. PseudoBend: Producing Haptic Illusions of Stretching, Bending, and Twisting Using Grain Vibrations. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, UIST '19, p. 803–813, 2019.
- [34] S. Hertweck, D. Weber, H. Alwanni, F. Unruh, M. Fischbach, M. E. Latoschik, and T. Ball. Brain Activity in Virtual Reality: Assessing Signal Quality of High-Resolution EEG While Using Head-Mounted Displays. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 970–971, 2019.
- [35] A. Hettiarachchi and D. Wigdor. Annexing Reality: Enabling Opportunistic Use of Everyday Objects as Tangible Proxies in Augmented Reality. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 1957–1967. San Jose California USA.
- [36] K. Hinckley, R. Pausch, J. C. Goble, and N. F. Kassell. Passive Real-World Interface Props for Neurosurgical Visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*,

- CHI '94, p. 452–458, 1994.
- [37] B. E. Insko. *Passive Haptics Significantly Enhances Virtual Environments*. phd, The University of North Carolina at Chapel Hill, 2001.
- [38] C. Kaernbach. Simple Adaptive Testing with the Weighted Up-Down Method. *Perception & Psychophysics*, 49(3):227–229, 1991.
- [39] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology*, 3(3):203–220, July 1993.
- [40] F. A. Kingdom and N. Prins. Chapter 5 - Adaptive Methods. In *Psychophysics (Second Edition)*, pp. 119–148. San Diego, 2016.
- [41] L. Kohli. Redirected Touching: Warping Space to Remap Passive Haptics. In *2010 IEEE Symposium on 3D User Interfaces (3DUI)*.
- [42] L. Kohli, M. C. Whitton, and F. P. Brooks. Redirected Touching: The Effect of Warping Space on Task Performance. In *2012 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 105–112, Mar. 2012.
- [43] E. Krokos and A. Varshney. Quantifying VR Cybersickness using EEG. *Virtual Reality*, 26(1):77–89, Mar. 2022.
- [44] A. Lécuyer, J.-M. Burkhardt, S. Coquillart, and P. Coiffet. “Boundary of Illusion”: An Experiment of Sensory Integration with a Pseudo-haptic System. In *Proceedings IEEE Virtual Reality 2001*, pp. 115–122. IEEE Comput. Soc, Yokohama, Japan, 2001.
- [45] L. Liu and R. van Liere. Designing 3D Selection Techniques Using Ballistic and Corrective Movements. In *Proceedings of the 15th Joint Virtual Reality Eurographics Conference on Virtual Environments, JVRC'09*, p. 1–8. Eurographics Association, Goslar, DEU, 2009.
- [46] L. Liu, R. van Liere, C. Nieuwenhuizen, and J.-b. Martens. *Comparing Aimed Movements in the Real World and in Virtual Reality*. Mar. 2009. Proceedings - IEEE Virtual Reality.
- [47] LSL. The Lab Streaming Layer (LSL). <https://github.com/sccn/labstreaminglayer>, 2022. Accessed: 2022-10-05.
- [48] A. Lécuyer. Simulating Haptic Feedback Using Vision: A Survey of Research and Applications of Pseudo-Haptic Feedback. *Presence: Teleoperators and Virtual Environments*, pp. 39–53, 2009. MIT Press.
- [49] I. S. MacKenzie. Fitts' Law as a Research and Design Tool in Human-Computer Interaction. *Human-Computer Interaction*, 1992.
- [50] D. Makowski, T. Pham, Z. J. Lau, J. C. Brammer, F. Lespinasse, H. Pham, C. Schölzel, and S. H. A. Chen. NeuroKit2: A Python Toolbox for Neurophysiological Signal Processing. *Behavior Research Methods*, 53(4):1689–1696.
- [51] M. Malińska, K. Zużewicz, J. Bugajska, and A. Grabowski. Heart Rate Variability (HRV) during Virtual Reality Immersion. *International Journal of Occupational Safety and Ergonomics*, 2015.
- [52] E. Marchiori, E. Niforatos, and L. Preto. Analysis of Users' Heart Rate Data and Self-reported Perceptions to Understand Effective Virtual Reality Characteristics. *Information Technology & Tourism*, 2018.
- [53] T. H. Massie and J. K. Salisbury. The PHANToM Haptic Interface: A Device for Probing Virtual Objects. In *Proceedings of the ASME Dynamic Systems and Control Division*, pp. 295–301, 1994.
- [54] W. A. McNeely. Robotic Graphics: A New Approach to Force Feedback for Virtual Reality. In *Proceedings of IEEE Virtual Reality Annual International Symposium*, pp. 336–341. IEEE, Seattle, WA, USA, 1993.
- [55] K. Nieuwenhuizen, L. Liu, R. van Liere, and J.-B. Martens. Insights from Dividing 3D Goal-Directed Movements into Meaningful Phases. *IEEE Computer Graphics and Applications*, 29(6):44–53, 2009.
- [56] T. Noto, G. Zhou, S. Schuele, J. Templer, and C. Zelano. Automated Analysis of Breathing Waveforms using BreathMetrics: A Respiratory Signal Processing Toolbox. *Chemical Senses*, 43(8):583–597, 07 2018.
- [57] N. Ogawa, T. Narumi, and M. Hirose. Effect of Avatar Appearance on Detection Thresholds for Remapped Hand Movements. *IEEE Transactions on Visualization and Computer Graphics*, 2021.
- [58] G. Padrao, M. Gonzalez-Franco, M. V. Sanchez-Vives, M. Slater, and A. Rodriguez-Fornells. Violating Body Movement Semantics: Neural Signatures of Self-generated and External-generated Errors. *NeuroImage*, 124:147–156.
- [59] T. Pham, Z. J. Lau, S. H. A. Chen, and D. Makowski. Heart Rate Variability in Psychology: A Review of HRV Indices and an Analysis Tutorial. *Sensors*, 21(12), 2021. doi: 10.3390/s21123998
- [60] I. Poupyrev, M. Billinghurst, S. Weghorst, and T. Ichikawa. The Go-go Interaction Technique: Non-linear Mapping for Direct Manipulation in VR. In *Proceedings of the 9th Annual ACM Symposium on User Interface Software and Technology*, UIST '96, pp. 79–80. Association for Computing Machinery, Seattle, Washington, USA, Nov. 1996.
- [61] K. J. Preacher, D. D. Rucker, R. C. MacCallum, and W. A. Nicewander. Use of the Extreme Groups Approach: A Critical Reexamination and New Recommendations. *Psychological Methods*, 10(2):178, 2005.
- [62] F. Ribeiro and J. Oliviera. Factors Influencing Proprioception: What Do They Reveal? *Psychophysical Reviews*, 2011.
- [63] M. Samad, E. Gatti, A. Hermes, H. Benko, and C. Parise. Pseudo-Haptic Weight: Changing the Perceived Weight of Virtual Objects by Manipulating Control-Display Ratio. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–13. Association for Computing Machinery, Glasgow, Scotland, UK.
- [64] V. Schwind, P. Knierim, C. Tasci, P. Franczak, N. Haas, and N. Henze. “These Are Not My Hands!”: Effect of Gender on the Perception of Avatar Hands in Virtual Reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017.
- [65] H. Si-Mohammed, C. Lopes-Dias, M. Duarte, F. Argelaguet, C. Jeunet, G. Casiez, G. R. Müller-Putz, A. Lécuyer, and R. Scherer. Detecting System Errors in Virtual Reality Using EEG Through Error-Related Potentials. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 653–661, Mar. 2020. ISSN: 2642-5254.
- [66] A. L. Simeone, E. Velloso, and H. Gellersen. Substitutional Reality: Using the Physical Environment to Design Virtual Reality Experiences. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pp. 3307–3316. Association for Computing Machinery, Seoul, Republic of Korea, Apr. 2015.
- [67] A. K. Singh, H.-T. Chen, Y.-F. Cheng, J.-T. King, L.-W. Ko, K. Gramann, and C.-T. Lin. Visual Appearance Modulates Prediction Error in Virtual Reality. *IEEE Access*, 6:24617–24624, 2018.
- [68] M. Slater. Place Illusion and Plausibility Can Lead to Realistic Behaviour in Immersive Virtual Environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535), 2009.
- [69] F. Steinicke, G. Bruder, J. Jerald, H. Frenz, and M. Lappe. Estimation of Detection Thresholds for Redirected Walking Techniques. *IEEE Transactions on Visualization and Computer Graphics*, Jan. 2010.
- [70] P. L. Strandholt, O. A. Dogaru, N. C. Nilsson, R. Nordahl, and S. Serafin. Knock on Wood: Combining Redirected Touching and Physical Props for Tool-Based Interaction in Virtual Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Honolulu, HI, USA.
- [71] Swartz Center for Computational Neuroscience. EEGLab. <https://sccn.ucsd.edu/eeglab/index.php>. Accessed: 2022-10-05.
- [72] Swartz Center for Computational Neuroscience. MoBILAB. <https://github.com/sccn/mobilab>, 2022. Accessed: 2022-10-05.
- [73] J. J. Yang, H. Horii, A. Thayer, and R. Ballagas. VR Grabbers: Ungrounded Haptic Retargeting for Precision Grabbing Tools. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, UIST '18, p. 889–899. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3242587.3242643
- [74] B. Yazmir, M. Reiner, H. Pratt, and M. Zacksenhouse. Brain Responses to Errors During 3D Motion in a Hapto-Visual VR. In F. Bello, H. Kajimoto, and Y. Visell, eds., *Haptics: Perception, Devices, Control, and Applications*, vol. 9775, pp. 120–130. Springer International Publishing, Cham, 2016. Series Title: Lecture Notes in Computer Science.
- [75] A. Zenner and A. Krüger. Shifty: A Weight-Shifting Dynamic Passive Haptic Proxy to Enhance Object Perception in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*, 2017.
- [76] A. Zenner and A. Krüger. Estimating Detection Thresholds for Desktop-Scale Hand Redirection in Virtual Reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces*, VR, pp. 47–55, 2019.
- [77] A. Zenner, K. P. Regitz, and A. Krüger. Blink-Suppressed Hand Redirection. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 75–84, Mar. 2021. ISSN: 2642-5254.
- [78] A. Zenner and K. Ullmann. The Unity Staircase Procedure Toolkit. <https://github.com/AndreZenner/staircase-procedure>, 2021.
- [79] A. Zenner, K. Ullmann, and A. Krüger. Combining Dynamic Passive Haptics and Haptic Retargeting for Enhanced Haptic Feedback in Virtual Reality. *IEEE Transactions on Visualization and Computer*

