Experience Matters: Longitudinal Changes in Sensitivity to Rotational Gains in Virtual Reality

ANDREW ROBB, Clemson University, USA
KRISTOPHER KOHM, Clemson University, USA
JOHN PORTER III, Clemson University, USA

Redirected walking techniques use rotational gains to guide users away from physical obstacles as they walk in a virtual world, effectively creating the illusion of a larger virtual space than is physically present. Designers often want to keep users unaware of this manipulation, which is made possible by limitations in human perception that render rotational gains imperceptible below a certain threshold. Many aspects of these thresholds have been studied, however no research has yet considered whether these thresholds may change over time as users gain more experience with them. To study this, we recruited 20 novice VR users (no more than 1 hour of prior experience with an HMD) and provided them with an Oculus Quest to use for four weeks on their own time. They were tasked to complete an activity assessing their sensitivity to rotational gain once each week, in addition to whatever other activities they wanted to perform. No feedback was provided to participants about their performance during each activity, minimizing the possibility of learning effects accounting for any observed changes over time. We observed that participants became significantly more sensitive to rotation gains over time, underscoring the importance of considering prior user experience in applications involving rotational gain, as well as how prior user experience may affect other, broader applications of VR.

CCS Concepts: • **Human-centered computing** → **Virtual reality**.

Additional Key Words and Phrases: rotational gain, perception, longitudinal, virtual reality

1 INTRODUCTION

Rotational gains are commonly used in redirected walking and other similar techniques. A rotational gain alters how much a virtual scene rotates when a user turns their head: an amplification causes the scene to rotate more than the head turns, and a minification causes the scene to rotate less. Steinicke et. al formally define rotational gains as $g_R = \frac{R_{Virtual}}{R_{Real}}$, such that $R_{Virtual}$ equals the amount rotated in VR and R_{Real} the amount rotated in real life [39]. Due to limitations of human perception, rotational gains below a certain threshold are imperceptible. Redirected walking takes advantage of these thresholds to subtly control where users are walking in the real world to allow users to physically walk through virtual spaces that are larger than the available physical space.

While a significant amount of prior work has sought to determine where the threshold for imperceptibility lies under varying circumstances (see Table 3), no research has explored how these thresholds might change as users become more accustomed to them. Prior work in other fields provides examples of how human perception can change after repeated exposure to a sensory phenomenon; improvements in sensory perception are most often documented (e.g. [7, 31]), however reductions in sensitivity have also been observed (e.g. [14]). As many real-world applications of redirected walking would involve repeated use (e.g. a simulation integrated into a

Authors' addresses: Andrew Robb, Clemson University, Clemson, USA, arobb@clemson.edu; Kristopher Kohm, Clemson University, Clemson, USA, kckohm@clemson.edu; John Porter III, Clemson University, Clemson, USA, jjporte@clemson.edu.

1544-3558/2022/9-ART $15.00
https://doi.org/10.1145/3560818

10-week training module) it may be possible for users' perceptual thresholds to shift over time. This potentially poses a challenge to developers if their intention is for the presence of redirection techniques to remain unknown to users, as is often the case. The increased accessibility of VR also raises the possibility that users may be exposed rotational gains in many different applications, even if each application is used only once. As such, perceptual thresholds may vary between users even in a single-use application due to the differing prior experience of each user.

To determine whether sensitivity to rotational gains may evolve as users become familiar with their application, we conducted a longitudinal experiment wherein 20 novice VR users (defined as having less than 1 hour of prior experience with a VR HMD) were loaned Oculus Quest headsets for four weeks and were tasked to complete a short task assessing their sensitivity to rotational gains once each week. Our results showed that participants became more sensitive to the presence of rotational gain as time went on. We also observed that while the time taken by participants to rotate during each trial was affected by the rotation gain, this effect was not moderated by prior experience, which indicates that other effects of rotational gains may not be moderated by prior experience. We also did not observe significant changes in reported sickness scores on either the Simulator Sickness Questionnaire [17] or in response to a single question asked repeatedly at the end of each block of trials.

2 RELATED WORKS

2.1 Longitudinal Research in Virtual Reality

Very few studies have considered how the repeated use of VR impacts user behavior and perception while using a VR application. Baileson and Yee conducted a longitudinal study in 2006 tracking users across fifteen 45 minute sessions as they interacted with each other in VR, finding that presence did not change over time, but that participants did look at each other less over time and that simulator sickness decreased over time as well [1]. More recently, Porter et. al conducted a brief longitudinal study exploring how users played Minecraft on the desktop and in VR, involving three 45 minute sessions on each platform [32]. However, no temporal effects were observed within this timeframe, except for a small effect related to simulator sickness. Porter and Robb later conducted a qualitative analysis of how consumer discussions about VR evolved over the course of two years and found significant changes in how consumers spoke about presence, simulator sickness, and locomotion systems [33]. Ricca et al. found that while performance in a pick-and-place task improved after repeated use of a VR training program, the hand-representation used (gloves holding tools vs. tools only) had no impact on performance over time [35]. Khojasteh and Won examined long-term use of collaborative applications and found that while users reported becoming more adapted to VR over time, workload, emotion recognition, and co-presence did not change significantly over time [18]. Moustafa and Steed found similar evidence that users became more adapted to collaborative VR over time [25]. Unlike earlier work, Han et al. found that presence increased over time in a large-group collaborative environment that participants used for a 10 week span [13]. Oyanagi et. al also found evidence that self-presence may increase over time when users first experience embodiment, but also that this effect may stabilize several weeks after first exposure [27].

Simulator sickness is an exception to the above observation that little research has explored the effects of repeated use of VR. Dużmańska et. al recently published a thorough survey paper regarding what is known about the temporal aspects of simulator sickness [9]. Most relevant to this paper is their discussion of adaptation to simulator sickness. They conclude that, while there is evidence that users can adapt to simulator sickness, more research is necessary. In particular, they note that many of the studies they cite did not report statistical tests or reported non-significant trends. Additionally, Porter and Robb's exploration of how consumer discussions of VR have evolved over time also found evidence that at least some users can adapt to simulator sickness, such that it becomes less of a concern after repeated exposures [33]. However, they also saw some reports of sickness actually increasing with time, though these reports were much less common.

Finally, it should be noted that there have been many longitudinal experiments that *employed* VR without considering whether the effect of the VR system changed over time. The majority of these experiments are best classified as system usability studies, where the research is focused on how repeated use of a VR application impacts an unrelated targeted outcome. Examples are common in the training (e.g. [35, 38, 47]), therapeutic applications (e.g. [8, 11, 12, 16, 19, 42]), education (e.g. [15, 41]), and computer-supported collaborative work (e.g. [18, 25, 43]). Similarly, other studies have considered how the prolonged effects of a single VR intervention evolve over time. These experiments typically involve a single VR intervention paired with multiple follow-up measurements (often post-surveys) administered at spaced-out intervals (e.g. [30]).

2.2 Manipulating Rotational Gain in VR

Steinicke et. al evaluated the perceptability of rotational gain within the context of redirected walking in an HMD [39]. Sensitivity to gain was evaluated in two ways: 1) participants performed a single rotation with a positive or negative gain applied and were then asked if the virtual rotation they had seen was greater than or smaller than the physical rotation they had just performed, and 2) participants performed two rotations back-to-back, where one had a positive or negative rotational gain applied and one had no gain; participants were then asked which of the rotations had had the gain applied. As the second method employed mirrors the method we used, we will focus on these results. Steinicke et. al found that the point of subjective equality was 0.9941, almost indistinguishable from no gain, and lower and upper thresholds of 0.76 and 1.19. A later paper by Steinicke et. al reported a second experiment with results similar to their earlier publication, though they only employed the first evaluation method described above, making it difficult to directly compare their results to ours [40].

Since Steinicke's original papers, various other experiments have measured perceptual thresholds and how they are impacted by different factors, including rotating in a motorized wheelchair [5], acoustic cues [26, 37], visual density [28], 360 degree video [48], gender [46], field of view (FOV) [46], the presence of distractors [46], and dynamic FOV restrictors [6]. The thresholds reported by these papers are summarized in Table 1.

Other research has also considered the question of sensitivity to rotational gain, but has done so in ways that is not as easily integrated into Table 1. Zhang and Kuhl explored sensitivity to dynamic rotational gains, where the gain factor changes during a single rotation, and found that participants were largely unable to distinguish between slow and instantaneous changes in gain factor[49]. Sakono et. al also looked at dynamic gain factors and found that dynamic gain factors could be used to reduce discomfort experienced by participants [36]. Kuhl et. al found that participants could calibrate to both positive and negative rotational gains, such that their rotations in a no-gain scenario were affected after completing multiple rotations where a gain was present [22].

3 METHODS

3.1 Participants

Twenty people participated in this study (5 females, 1 non-binary). Ages ranged from 18 to 33 years old, with a median age of 20 years. All participants had normal or corrected-to-normal vision. Participants were recruited via an email sent to students enrolled at the School of Computing at Clemson University. Given the tendency for women to be underrepresented in VR studies [29], women who responded to the email were preferentially recruited. However, not enough women responded to achieve gender parity, which likely reflects the existing gender gap in most computing programs.

The following selection criteria were used: 1) participants had less than 1 hour of prior experience using VR, 2) agreed to use the HMD regularly for four weeks, 3) would complete our experimental activities each week, 4) had an open floor space at home that could be used for the Quest, and 5) agreed to complete two interviews during the study. Participants received a $50 gift card upon enrolling in the study, and a $75 gift card upon completing the study.

Table 1. Estimated Threshold Parameters from Literature

Paper and Condition	Gain Range	PSE	Lower Threshold	Upper Threshold
Steinicke 2008[39]				
Real vs. Physical rotation	[0.5, 1.5]	0.840	0.59	1.1
Two successive rotations	[0.6, 1.4]	0.994	0.76	1.19
Steinicke 2009[40]	[0.5, 1.5]	0.96	0.67	1.24
Bruder 2012[5]				
Redirected Standing	[0.4, 1.6]	0.954	0.681	1.259
Redirected Driving	[0.4, 1.6]	1.011	0.771	1.262
Serafin 2013[37]	[0.4, 1.6]	1.0	0.82	1.2
Paludan 2016[28]				
Realistic environment	[0.75, 1.25]	1.10	0.93	1.27
Basic environment, 4 objects	[0.75, 1.25]	1.00	0.81	1.19
Basic environment, 16 objects	[0.75, 1.25]	1.01	0.82	1.20
Nilsson 2016[26]				
No Audio	[0.5, 1.5]	0.93	0.77	1.10
Static Audio	[0.5, 1.5]	0.95	0.80	1.11
Moving Audio	[0.5, 1.5]	0.93	0.79	1.08
Zhang 2018 [48]				
Left rotation	[0.6, 1.4]	0.984	0.877	1.092
Right rotation	[0.6, 1.4]	0.972	0.892	1.054
Williams and Peck 2019[46]				
40° FOV, no distractors	[0.6, 1.4]	0.948	0.680	1.217
40° FOV, with distractors	[0.6, 1.4]	0.968	0.681	1.254
110° FOV, no distractors	[0.6, 1.4]	1.055	0.672	1.441
110° FOV, with distractors	[0.6, 1.4]	1.068	0.568	1.568
Brument 2020[6][iii]				
No vingetting (average)	[1.0, 1.4]	1.22	1.135	1.31
Color vignetting (average)	[1.0, 1.4]	1.24	1.135	1.363
Blur vignetting (average)	[1.0, 1.4]	1.22	1.108	1.33
Average (Std. Dev)		0.982 (0.059)	0.754 (0.101)	1.213 (0.129)

[i] Brument et. al only looked at positive gains, which skewed their thresholds compared to other works. As such, these thresholds will be excluded from our analysis.

[ii] Non-significant effects are collapsed due to the many conditions in this experiment.

3.2 Procedure

Each participant was provided with an Oculus Quest to use for the 4 week experiment. Participants were asked to spend ten hours a week using the VR headset, however, this proved infeasible for most participants; the average amount of time spent over the course of the experiment was 17 hours in total across the four weeks. This time could include games, entertainment, or any other activities of interest to participants, as well as three custom activities we developed for the experiment: 1) a "rotational sensitivity" activity (the subject of this paper), 2) a "body awareness" activity, and 3) a "navigation" activity. Participants were instructed to complete the activities in the same order each week; the order in which they were to be completed was randomized for each participant. Participants accessed these activities via an application we developed which enforced the order in which activities were completed, and also reminded participants of when these activities should be completed each week. Participants also completed two semi-structured interviews during the experiment. The first interview took place after the first or the second week; the second interview was completed two weeks after the first interview.

Participants were given detailed instructions on 1) how to use the headset, 2) how to use the Sidequest application [1] to load our custom activities onto the Quest, as well as how to access other free applications available on the Sidequest marketplace, 3) and how to access data stored on the headset about their activities. Our custom activities saved their data as files on the Quest. Accessing the amount of time participants used the headset required the installation of a second application: Activity Watch [2]. Activity Watch is an open-source time tracking application with clients for several operating systems, including Android (the OS used on the Oculus Quest). The Quest is capable of running most Android applications, though this capability is not advertised to users and requires some additional configuration. Activity Watch gathers usage data on Android devices using Android's Usage Stats API, which thankfully functions on the Oculus Quest. This allowed us to access fine-grained information about which applications participants were using, and for how long they were used. Each week, participants were asked to upload logs from our activities and from Activity Watch to a unique Google Drive folder created for each participants, allowing us to keep track of their progress.

The "body awareness" and "navigation" activities are outside of the scope of this paper, however we provide a short description of each activity here. In the "body awareness" activity, participants were instructed to repeatedly stack blocks on top of each other while seated at a virtual table. Blocks were grasped using a pinching gesture, which was tracked using the Oculus Quest's hand tracking system. While stacking blocks, an offset was sometimes applied to the visible location of participants' hand. Participants were asked to identify when an offset was present each time they stacked a set of block. In the "navigation" activity, participants used either teleportation or joystick locomotion to explore twenty maze-like levels. In order to reach the end of the level, they had to first find a keycard located in one of the rooms. After finding the keycard, they could use it to open a locked door leading to the exit. Once arriving at the exit, participants were asked to point back in the direction of where they had started and where they had found the keycard. They also completed a presence and a simulator sickness questionnaire in between levels. Prior to beginning the experiment, participants were randomly assigned to one of the locomotion conditions, which they then used in each of the sessions.

3.3 Apparatus

The rotational gain environment resembled a typical home garage (see Figure 1). This environment was designed to provide a significant amount of optic flow as participants turned their heads, so as to aid in the detection of rotation via visual means. Upon entering the environment, a large TV was visible in front of the participant. This TV was used to provide instructions to the participant prior to the beginning of the experiment and after each

[1] https://sidequestvr.com/
[2] https://github.com/ActivityWatch/aw-android

Fig. 1. The arrows to the left and the right of the TV appeared during a trial to indicate which direction participants should rotate in. After a rotation, the TV was repositioned to appear in front of the user.

trial. Spatial buttons floated in front of and slightly below the TV. Participants could interact with these buttons via raycasting by pulling the trigger on their controllers. Buttons were used to advance through the instructions and to provide their response after each trial. A button was also visible that participants could use to end the experiment early if they began to experience any simulator sickness.

Upon completing the instructions, participants began the first of eight blocks, each of which consisted of eleven trials. During each trial, participants performed two 150° rotations; a rotational gain was applied to the participant's motion during one of these rotations. Participants were instructed to rotate until they saw a floating red sphere in the center of their vision. When starting a trial, participants were instructed to turn to either the left or the right (randomly selected). An arrow provided additional information about which direction to turn in. Upon completing the first rotation, participants were instructed to rotate back in the other direction while completing the second rotation. A different gain factor was used in each of the eleven trials within a block; the order in which these gain factors were presented was randomized within each block. The different gain factors used were: 0.7, 0.8, 0.866, 0.933, 0.966 1.0, 1.033, 1.066, 1.133, 1.2, and 1.3.

Before starting the first block, and after each block, participants were asked a single question, which they rated on a 1 to 10 scale: "How uncomfortable are you right now?". Upon completing all eight blocks, or indicating that they needed to quit early due to VR sickness, participants completed the Simulator Sickness Questionnaire [17]. Participants answered each question in VR using a survey interface designed for this experiment. Questions were shown one at a time with different buttons that participants could select via raycasting in order to provide their answer to the question.

3.4 Dependent Variables

Several dependent variables were recorded during each trial: 1) whether participants correctly identified the manipulated rotation, 2) the time taken to perform each rotation, 3) the time taken to select which rotation participants believed was manipulated (defined as the time from when the question was shown to them and they selected an answer), and 4) general motion data. Participants were asked a single question sickness survey after completing each block of trials, and were asked to complete the Simulator Sickness Questionnaire after upon completing the activity or exiting it early.

4 RESULTS

Analysis was conducted in R [34] and figures were generated using ggplot2[45]. Lme4 [2] was used to fit linear mixed models (LMMs) to the results of this experiment. Buildmer [44] was used to identify the best fitting model, as based on the likelihood-ratio test and BIC criteria. In the event that different models were selected based these criteria, the most parsimonious model was used. P-values were calculated using Satterthwaite's method.

Five participants completed only one full session of this activity during the experiment (a session was considered complete if participants completed at least 50% of the trials). These participants' data were removed from our analysis, leaving a total of 15 participants. Of these 15 participants, all but one completed all four sessions. The remaining participant completed 3 of the 4 sessions. Individual numerical data points were excluded as outliers if their z-score was greater than ±3. Independent numerical variables were grand-mean centered (i.e. the mean of the entire dataset was subtracted from each reading) as this is known to improve the fit of LMMs [20].

4.1 Interpreting Linear Mixed Models

LMMs are rapidly becoming the dominate method used by psychological researchers to analyze quantitative data [24], however they have not yet been widely adopted by VR researchers. As such, we briefly discuss their advantages and how to interpret them here.

LMMs are particularly well suited to analyze data involving correlations between conditions or measurements (e.g. when a single participant's behavior is measured multiple times) [21]. When compared with repeated-measures ANOVAs, LMMs are more robust to violated assumptions and also perform better when analyzing datasets with missing or imbalanced data. Additionally, while repeated-measures ANOVAs require longitudinal data to be collected on a fixed time schedule, LMMs are capable of handling data collected on a variable schedule, making them more practical for use with human-subject research and its natural variability. Taken together, these features make LMMs especially appropriate when analyzing longitudinal data.

LMMs model *fixed* and *random* effects [24]. Fixed effects model the effect of the manipulated variables on the measured quantities, while random effects model the unexplained variability associated with individual differences between participants. To do this, LMMs independently model each participant's response to the data. Random effects can be modeled as *random intercepts*, where the intercept of each participant's individual model is allowed to vary based on the best fit to that participant's data, and as *random slopes*, where the slope of each participant's individual model is also allowed to vary based on the best fit to their particular data. Including random effects in the model improves LLMs ability to precisely model the response of fixed effects independent of the innate variability between participants.

The fitted model includes an intercept, which describes the predicted value of the modeled variable when all predictors are set to 0, and fitted beta values for each fixed effect, which describe how much the modeled variable is predicted to change when the fixed effect increases by 1. Betas are highly sensitive to the scale of a given dependent variable, such that dependent variables spanning a small range of values produce smaller betas, and a larger range of values produce larger betas. This makes it difficult to judge the relative importance of a fixed effect from the size of the beta alone. As such, we also provide standardized betas (abbreviated std. beta in our tables) which are scaled based on the individual distribution of each dependent variable so as to make it easier to determine the relative contribution each fixed effect has on the independent variable. We also provide both marginal and conditional r^2 values; marginal r^2 report the amount of variance explained by the fixed effects alone, and conditional r^2 report the amount of variance explained by both the fixed and random effects.

Table 2. Psychometric Parameters

Session	Goodness of Fit	PSE		Slope	
		Estimate	95% CI	Estimate	95% CI
1	0.940	1.022	[0.999, 1.045]	5.016	[4.257, 5.812]
2	0.947	0.990	[0.975, 1.003]	9.003	[7.973, 10.164]
3	0.995	1.001	[0.987, 1.016]	9.025	[7.947, 10.213]
4	0.961	0.995	[0.983, 1.008]	11.649	[10.398, 13.158]

4.2 Participants' Usage Patterns

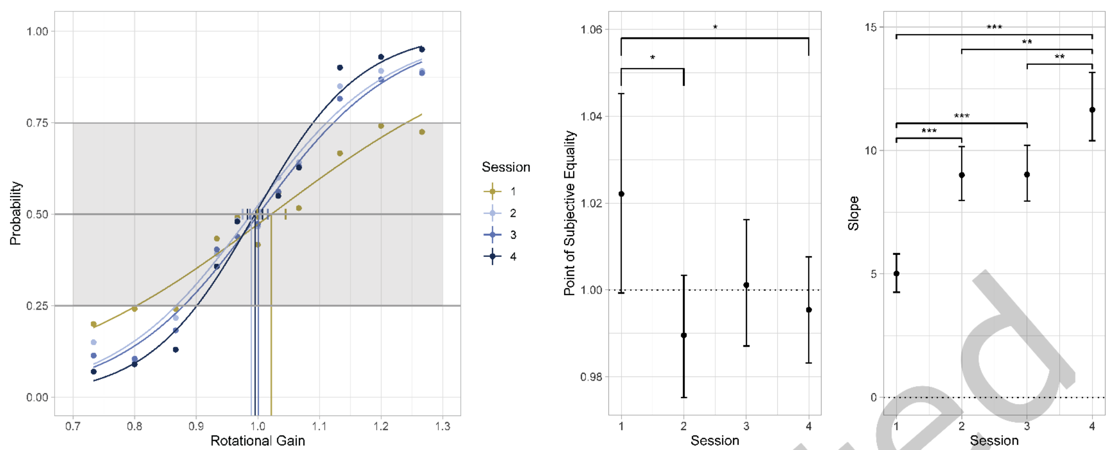
Based on the data recorded by ActivityWatch, participants spent an average of 17.13 hours ($sd = 10.58, min = 7.12, max = 45.24$) using the headset over the four weeks. 119 unique applications were launched by participants, though the majority of these were used for less than 1 hour across all participants. Thirty-three were used for more than 1 hour, with eleven of these being used for more than 5 hours. These applications were our experimental application, the Oculus home application, YouTube VR, Virtual Desktop, Oculus Link, Beat Saber, VRChat, Bait!, SUPERHOT VR, Sports Scramble, and SculptrVR. Discounting the experimental application, participants spent an average of 12.81 hours ($sd = 8.56, min = 3.21, max = 35.46$) in VR over the course of the four weeks.

4.3 Psychometrics of Sensitivity to Rotational Gain

Using R and the quickpsy package [23], we fit the psychometric data with a logistic function where the point of subjective equality and the slope of the function were both treated as free parameters and a lapse rate of 0% was assumed. Maximum liklihood was used as the fitting criteria. A non-parametric bootstrap procedure with 1000 trials was used to estimate the 95% confidence intervals on the parameters. The fitted functions are shown in Figure 2a, and the estimated data for the parameters, with confidence intervals, are reported in Table 2. Each of the estimated functions fitted the data well, as indicated by a non-significant goodness-of-fit.

4.3.1 Point of Subjective Equality. The fitted point of subjective equality (PSE) was not significantly different from zero in any of the Sessions (see Figure 2b). The PSE in the first session was significantly different from the PSE in Session 2 ($p = 0.0150$) and Session 4 ($p = 0.0486$), but no other differences were observed. These results suggest that while a small effect of prior experience may be present for PSE, any effect is fairly negligible and appears to quickly disappear as users gain experience.

4.3.2 Slope. Fitting a psychometric function yields two parameters, the PSE and the slope. For whatever reason, slope has not often been discussed in previous literature exploring rotational gains, in favor of a focus on reporting the thresholds (which we report below). In contrast to the PSE, a significant effect of Session was consistently observed for the fitted slope (see Figure 2c). Slope increased monotonically across sessions, from a starting value of 5.02 to a final value of 11.65. Pairwise-significant differences were observed between almost all Sessions. Session 2 and 3 were the only exception, where the slope was not observed to change significantly between these sessions. This suggests that Session had a meaningful effect on the slope of the detection curve, indicating that participants became more sensitive to detecting rotational gain as they increased in experience. The largest difference was observed between Session 1 and Session 2, which may indicate that participants increased in sensitivity to rotation gain rapidly at first, and that this slowed as they progressed. However, it is important to note that participants were still improving in the final session. It is unclear whether participants would have continued to improve if they had completed additional sessions.



(a) The fitted psychometric functions (b) PSE by session (c) Slope by session

Fig. 2. Session can be seen to have a clear effect on the slope of the fitted curve, suggesting that participants grew more sensitive to rotation gain as their experience with VR increased.

Table 3. Detection Threshold Values. The percentage change is computed relative to a threshold of 1.0, corresponding to a perfect detection rate.

Session	Lower Threshold	Percent Change	Upper Threshold	Percent Change
1	0.803	–	1.242	–
2	0.862	29.9%	1.117	51.7%
3	0.874	36.0%	1.128	47.1%
4	0.894	46.1%	1.095	60.7%

4.3.3 Detection Thresholds. By convention, the detection threshold is defined as when there was a 75% probability of participants detecting the presence of a rotational gain. These regions are shown with a white background in Figure 2a. Detection thresholds decreased consistently as Session increased; the specific threshold values are reported in Table 3. Participants were able to detect gain factors that were up to 60.7% smaller by the end of the experiment.

4.4 Accuracy by Block and Session

Given the observation that participants' became more sensitive to rotational gains across Sessions, we also wanted to explore whether sensitivity increased over time *within* a Session, as measured by the accuracy of their responses to each trial within a given block. To do this, we first computed the accuracy within each block of trials (i.e. the number of trials where the correct rotation was selected), which was then used to fit a linear mixed model.

The best fit model included Session as a fixed effect and Session and Participant ID as random effects. The model's total explanatory power is substantial ($conditional r^2 = 0.35$) and the part related to the fixed effects

Table 4. Accuracy Model Parameters

Fixed Effect	beta [95% CI]	Std. beta [95% CI]	t(9415)	p	m. r^2
Intercept	-0.05 [-0.10, -2.73e-04]	–	-1.64	0.119	–
Session	0.03 [0.2, 0.05]	0.23 [-0.10, -0.35]	3.14	0.002	0.049

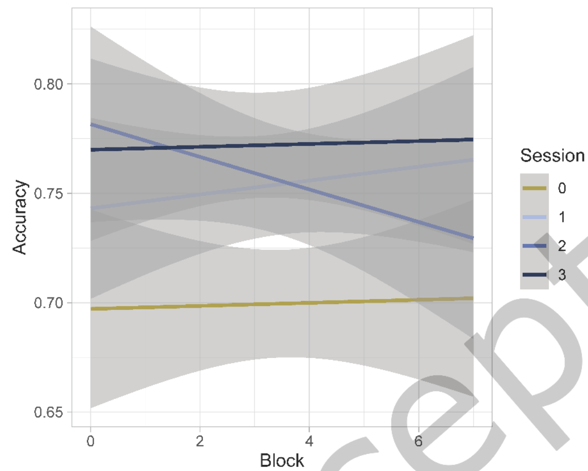


Fig. 3. Session had a clear impact on accuracy, but Block had no effect.

alone ($marginal r^2$) is of 0.06. The model parameters can be seen in Table 4, and a graph of the data can been seen in Figure 3. 1.31% of data points were removed as outliers.

Session was observed to have a significant effect on accuracy, such that accuracy improved across sessions. As Block was not included in the best fitting model, it can be safely assumed to not have significantly affected participants' accuracy. We confirmed this by running an additional model that included Block as a fixed effect, in addition to Session. Block's lack of an effect on accuracy can also be seen when examining the interaction-plot for Session by Block, shown in Figure 3. It can be observed that, on average, the slope of the fit lines tends to be relatively flat, suggesting that Block had little effect on participants' accuracy, especially when compared to the large change in average accuracy between Sessions (most notably Sessions 1 and 4). These results suggest that short-term learning effects were largely absent within any given session, as participants' accuracy did not improve across blocks.

4.5 Response Times

We next consider the amount of time taken by participants while performing the rotation tasks, and while indicating which rotation they believed to have a gain applied once a trial was completed.

4.5.1 Time to Rotate. In this model, we sought to understand how the duration of each rotation within a task was affected by Session, Block, and the Rotational Gain present. The best fit model included Session, Block, Rotational Gain, and the interaction between Session and Block as fixed effects and Session and Participant ID as random effects. The model's total explanatory power is moderate (conditional r^2 = 0.17) and the part related to the fixed

Table 5. Rotation Time Model Parameters

Fixed Effect	beta [95% CI]	Std. beta [95% CI]	t(9415)	p	m. r^2
Intercept	0.43 [0.23, 0.62]	–	4.21	<0.001	–
Session	-0.22 [-0.33, -0.11]	-0.17 [-0.27, -0.07]	-3.69	<0.001	0.027
Amplification	-1.45 [-1.64, -1.26]	-0.14 [-0.16, -0.12]	-14.64	<0.001	0.019
Block	-0.05 [-0.07, -0.04]	-0.07 [-0.09, -0.05]	-6.51	<0.001	3.7e-03
Session:Block	0.01 [3.23e-3, 0.02]	0.03 [7.2e-03, 0.05]	2.70	.007	4.7e-04

Table 6. Response Time Model Parameters

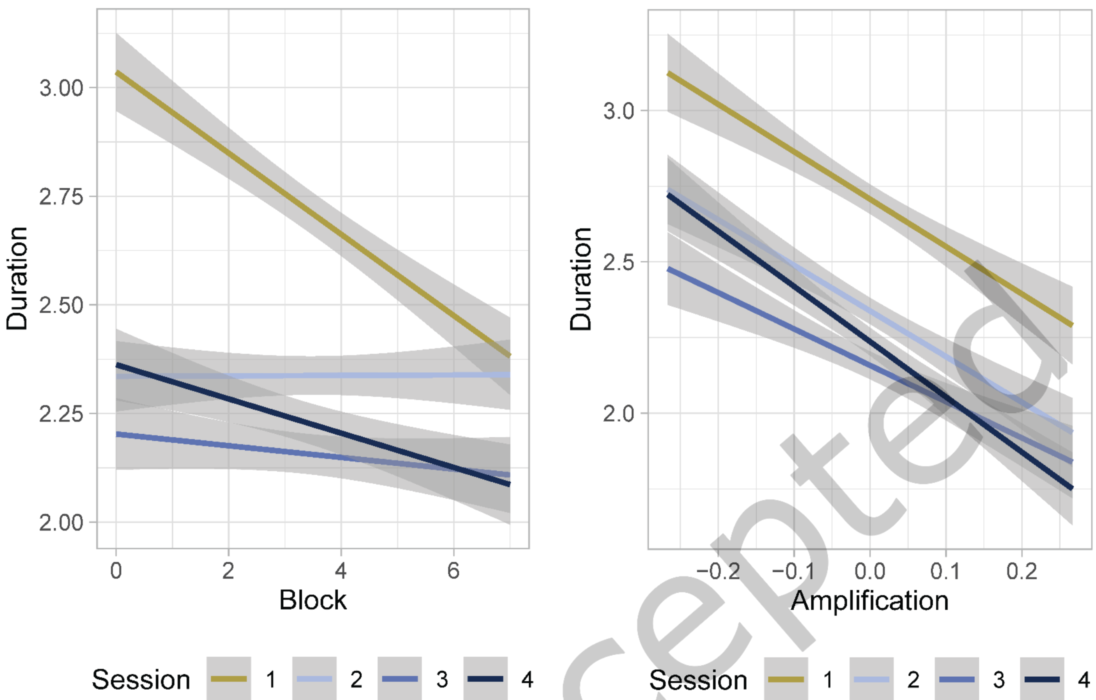
Fixed Effect	beta [95% CI]	Std. beta [95% CI]	t(4424)	p	m. r^2
Intercept	0.032 [0.14, 0.49]	–	3.56	<0.001	–
Session	-0.14 [-0.20, -0.07]	-0.17 [-0.26, -0.07]	-4.17	<0.001	0.026
AbsAmplif.	-0.24 [-0.45, -0.03]	-0.03 [-0.06, -3.8e-03]	-2.23	0.026	9.7e-04
Block	-0.03 [-0.05, -0.02]	-0.06 [-0.09, -0.03]	-4.49	<0.001	3.4e-03
Session:Block	9.5e-03 [1.7e-03, 0.02]	0.03 [6.3e-03, 0.06]	2.40	0.017	8.0e-04

effects alone (marginal r^2) is of 0.06. The model parameters can be seen in Table 5, and graphs of the data can been seen in Figures 4a and 4b. 1.94% of data points were removed as outliers.

Main effects of Session, Block, and Rotational Gain were all observed to affect the duration of participants' rotations, as well as an interaction between Session and Block. Interestingly, there was not an interaction between Session and Rotational Gain, which suggests that the manner in which Rotational Gain impacts head rotation speed is not influenced by prior experience with VR. The interaction between Session and Block appears to be driven by participants turning their heads more slowly in the early blocks of the first Session, potentially due to still becoming familiar with the task, or with VR in general.

4.5.2 Time to Respond After Each Trial. In this model, we sought to understand how the speed at which participants indicated which rotation had a gain applied was affected by Session, Block, and the Absolute Rotational Gain. We examined this variable as a proxy for participants' certainty, where a decision made more quickly suggests that participants were more confident about which rotation involved a rotational gain. Absolute Rotational Gain was used here as the magnitude of the gain likely influences how detectable it is, rather than whether it was a positive or negative gain. The best fit model included Session, Block, Absolute Rotational Gain, and an Interaction between Session and Block. The model included Session and Participant ID as random effects. The model's total explanatory power is moderate (conditional r^2 = 0.14) and the part related to the fixed effects alone (marginal r^2) is of 0.03. The model parameters can be seen in Table 6, and graphs of the data can been seen in Figures 5a and 5b. 4.07% of data points were removed as outliers.

Session had the largest impact on time to respond, as indicated by both the standardized beta value and the marginal r^2. We again see an interaction between Session and Block, such that the effect of Block diminished as Session increased, and observed no interaction between Session and Absolute Amplification. Surprisingly, Absolute Amplification had the smallest impact on time to respond, which may indicate that our assumption connection between time to respond and participants' certainty was overstated.



(a) Session by Block

(b) Session by Amplification

Fig. 4. The effects of Session, Block, and Amplification on Time to Rotate.

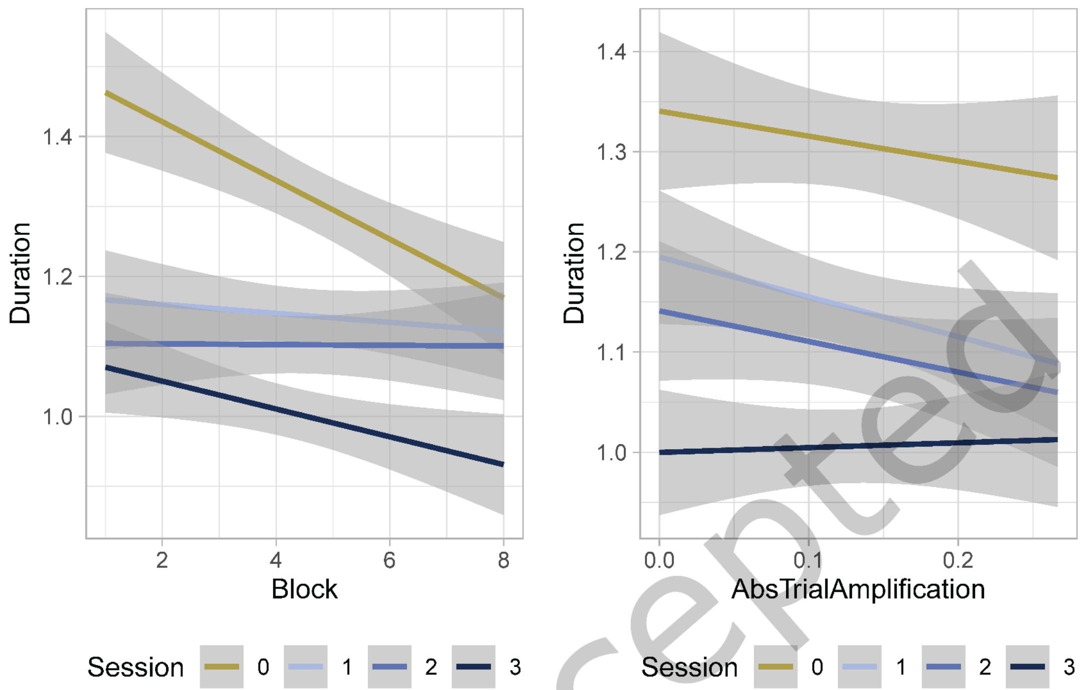
4.6 Simulator Sickness

Finally, we consider two perspectives on simulator sickness during the experiment. Participants completed the SSQ after each Session, and also answered a single question survey (which we refer to as the Short Sickness Questionnaire in this paper) after completing each block. Participants were asked to give their agreement to the following statement, on a scale from 1 to 10: “On a scale from 1 to 10, how uncomfortable are you right now?”

4.6.1 Simulator Sickness Questionnaire. We fitted individual linear mixed models to each of the four SSQ factors. The factor scores were grand-mean centered prior to fitting. Each of these models included Session as a fixed effect and Participant ID and Session as random effects. The individual model parameters can be seen in Table 7, and a graph of the data can be seen in Figure 6. 1.72% of data points were removed as outliers.

4.6.2 Short Sickness Questionnaire. We fitted a linear mixed model to predict participants’ Short Sickness Score with Session and Question. Short Sickness Score was grand-mean centered prior to analysis. The model included Session and Participant ID as random effects. The model’s total explanatory power is substantial (conditional r^2 = 0.67) and the part related to the fixed effects alone (marginal r^2) is of 0.08. The model parameters can be seen in Table 8, and a graph of the data can been seen in Figure 7. 9.39% of data points were removed as outliers.

Participants’ Short Sickness Scores increased as they completed the activity, which suggests some simulator sickness was experienced. Once again, we see no effect of Session on participants’ self-reported simulator sickness.



(a) Session by Block

(b) Session by Amplification

Fig. 5. The effects of Session, Block, and Amplification on Response Time after a trial was completed.

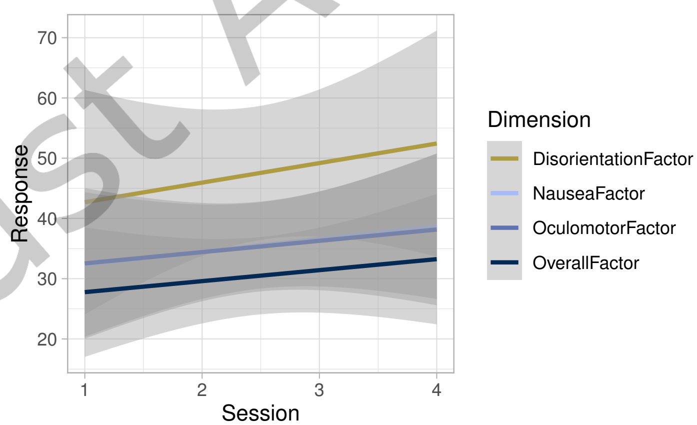


Fig. 6. While the graph suggest that sickness increased across Session, this apparent trend was not significant ($p > 0.599$). The trend lines for Oculomotor and Nausea are overlapping, making it appear that Nausea factors is missing.

Table 7. SSQ Models Parameters

Fixed Effect	beta [95% CI]	Std. beta [95% CI]	t(59)	p	m. r^2
Nausea	conditional r^2 = 0.82	marginal r^2 = 7.17e-03			
Intercept	-4.589 [-21.71, 12.53]	–	-0.538	0.599	–
Session	2.01 [-2.99, 7.02]	0.08 [-0.12, 0.29]	0.81	0.423	7.17e-03
Disorientation	conditional r^2 = 0.71	marginal r^2 = 1.42e-03			
Intercept	-3.26 [-26.17, 19.64]	–	-0.286	0.766	–
Session	1.34 [-6.56, 9.24]	0.04 [-0.18, 0.26]	0.34	0.735	1.42-e-03
Oculomotor	conditional r^2 = 0.70	marginal r^2 = 8.10e-04			
Intercept	-1.95 [-15.15, 11.26]	–	-0.296	0.772	–
Session	0.75 [-4.54, 6.04]	0.03 [-0.17, 0.23]	0.28	0.777	8.10e-04
Overall	conditional r^2 = 0.72	marginal r^2 = 1.20e-03			
Intercept	-1.85 [-14.14, 10.43]	–	-0.303	0.766	–
Session	0.75 [-3.70, 5.20]	0.03 [-0.17, 0.24]	0.34	0.737	1.20e-03

Table 8. Short Sickness Model Parameters

Fixed Effect	beta [95% CI]	Std. beta [95% CI]	t(455)	p	m. r^2
Intercept	-0.22 [-1.13, 0.70]	–	-0.46	0.648	–
Session	-0.12 [-0.41, 0.17]	-1.03e-03 [-0.18, 0.18]	-0.82	0.410	-0.0017
Question	0.11 [0.02, 0.20]	0.30 [0.24, 0.37]	2.39	0.017	3.59e-03
Session:Question	0.03 [-2.12e-03, 0.07]	0.06 [-3.76e-03, 0.12]	1.84	0.066	2.52e-03

It should be noted that the amount of simulator sickness reported was significantly less than expected, given the strong connection between large rotational gains and simulator sickness. It may be that the five participants who only completed one Session of this activity (and thus were dropped from the analysis) experienced more simulator sickness, which motivated them to avoid this particular activity. If so, then this analysis may underestimate the degree of simulator sickness experienced during this activity and could have potentially failed to identify trends associated with Session.

5 DISCUSSION

The psychometric model suggests that the PSE did not change substantially over time, or if it did, the change stabilized quickly after the first Session. The observed PSE values were well within the range observed in other studies investigating sensitivity to rotational gain, as shown in Table 1. In contrast, the slope of the psychometric curve, and thus the detection thresholds, changed substantially over the course of the four Sessions. The fitted values of 0.802 and 1.242 from the first Session compare reasonably well with the average thresholds reported in other studies, namely 0.754 ± 0.101 and 1.213 ± 0.129 [3]. However, the values found in the final Session, 0.894

[3]Ideally, a full meta-analysis would have been conducted to definitively place our results in the context of prior work, however most earlier papers provided insufficient information for this to be possible, namely a 95% confidence interval for the thresholds.

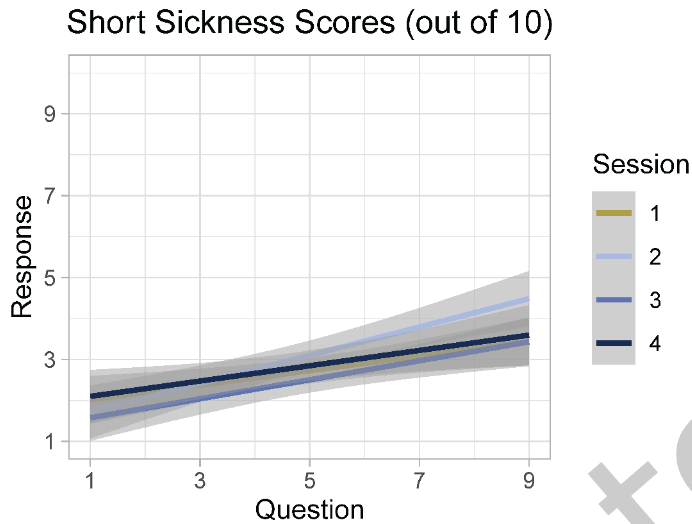


Fig. 7. The results from the short sickness questionnaire.

and 1.095, deviate by roughly a full standard deviation from these averages. An alternative way at considering the change in sensitivity is to consider the number of studies that reported more sensitive thresholds than we observed in the first and final Session. For the first Session, 33% of the studies reported a more sensitive lower threshold, and 33% of the studies reported a more sensitive upper threshold. By the final Session, only 5.55% of the studies reported a more sensitive lower threshold, and only 16.7% of studies reported a more sensitive upper threshold.

The paper which reported threshold averages closest to our own was published by Zhang et. al in 2018 [48], with average threshold values of 0.885 and 1.073. It is worth noting that the Zhang et. al specifically reports in this paper that 14 of their 15 participants had used a HMD before, and that most had used a stereoscopic 3D display before. This suggests that the population used in their study likely had significant prior experience using VR, which positions them much more closely to our participants in Session 4 than when they started the experiment. The similarity of their findings to our own further suggests that the observed change in detection thresholds corresponds to a real change linked to the amount of prior experience a user has with an HMD.

In additional to psychometric modeling, we also considered the question of sensitivity to rotational gain by examining how participants' accuracy within a block of trials evolved both within and between Sessions. We again observed a substantial effect of Session on accuracy, such that accuracy (and thus sensitivity) increased significantly across Sessions. However, we did *not* observe an effect of Block on accuracy. This indicates that the observed increase in sensitivity is unlikely to be solely attributable to learning effects that occurred during each Session. If so, we would expect participants to perform substantially better by the end of the Session than they did at the beginning. Instead, we see substantial increases in accuracy *between* Sessions, but little change within a Session. This suggests that something more than an immediate learning effect is responsible for the observed increase in sensitivity, potentially reflection a slower cognitive process. Taken together, these findings strongly suggest that a user's sensitivity to rotational gain will be influenced by the amount of prior exposure they have had to systems employing a rotational gain, either for redirected walking or for some other purpose. If developers wish to hide rotational gain from users, it may be necessary to use a smaller rotational gain than was indicated by previous studies.

Session also influenced how much time participants took to rotate during trials and to give their response after each trial. However, we did not observe an interaction effect between Session and Amplification on time to rotate or time to respond. This underscores that some aspects of participants' behavior may not change as sensitivity increases. In particular, we did not observe participants varying the speed of rotation of their head based on the amount of amplification present; instead, the baseline speed increased with time, likely due to familiarity with the task. We see a similar lack of effect when considering participants' simulator sickness during the experiment. No significant differences were observed for the SSQ factors across Sessions, suggesting that participants' simulator sickness during this activity did not change substantially over the course of four weeks. The results from the short sickness questionnaire similarly showed no effect of Session, though it did show participants' began to experience more discomfort over time within a single Session.

As to why these changes were observed, one possible explanation is the phenomenon of perception-action calibration. According to the perception-action theory of calibration, there is a joint relationship between our perceptions and our actions whereby acting allows us to refine our perceptions of the world around us, which in turn allows to us to act with more precision [4]. Kuhl et. al [22] have already shown that users can calibrate to rotational gains, though they did not consider whether this effect might persist with time. While not using VR, Bingham and Romack have explored how repeated calibration to a system affects performance over time; they had participants repeatedly calibrate to wearing prism goggles (which distort one's view of the world) across multiple sessions and found participants' baseline performance improved over time, indicating that the process of repeated calibration improved participants ability to adapt to perceptual distortions [3]. There are several distinctions between this experiment and our own, but they share a common theme: participants performance in a task involving perceptual distortions improved over time. In Bingham and Romack's case, reaching towards a target while wearing prism goggles, and in our case, identifying when a distortion was present while rotating towards a target. Unlike Bingham and Romack, and most other perception-action work, our participants were not provided with explicit feedback about their performance (they never knew if they selected the correct rotation where a gain was present). However, they still received feedback from their sensory systems while performing the rotations, most notably when larger gains were used which produced a clear conflict with participants' vestibular system. This perceptual information may have been sufficient to enable some degree of calibration to occur during the activities.

5.1 Limitations

As this experiment was conducted in the wild, we did not have control over what other applications participants used between activities which increases the variability between participants' experiences. We chose to allow participants to engage in other activities besides ours so as to increase the ecological validity of this experiment, as real users are most likely being exposed to several applications with different locomotion methods simultaneously. This increased the ecological validity of our results, however there is a tradeoff in terms of increased variability between participants other experiences in VR.

While we believe this work has applications for developers using redirected walking in real settings, participants did not actually engage in redirected walking in this research due to the limited amount of space available in most users' homes. As such, sensitivity to rotational gains in real world applications may be less susceptible to prior experience. Additionally, most real applications would involve additional activities that would serve as distractors which could further limit users ability to detect small rotational gains.

Finally, it should be noted that this experiment took part as a batch of three activities which raises the possibility of spillover between the activities. We sought to carefully separate the effects that were the focus of these studies (sensitivity to rotational gain, sensitivity to proprioceptive offsets, and spatial awareness and navigation). We made this decision so as to maximize the useful data collected from this experiment, as longitudinal experiments of this nature are difficult and expensive to run.

6 CONCLUSION

Our primary question in this paper was whether a user's sensitivity to rotational gain is influenced by their prior experience in VR. We found clear indications of this, including an increase in sensitivity to rotational gain of up to 60%, a reduction in the duration of their rotations during the tasks, and a reduction in the time spent indicating which rotation in a trial had a rotational gain applied. However, other aspects of our data showed no indication of being impacted by a user's prior experience in VR, namely that the effect of rotational gain on turning time and response time was not moderated by Session, and that simulator sickness scores also showed no sign of being affected by Session.

As many aspects of a user's perception can be altered by VR (e.g. depth compression in distance estimation [10]), this work has broader implications for many other aspects of VR research and development. Future work is needed to assess what other aspects of perception may be influenced not only by the VR application and hardware, but also by users' previous experience in similar VR scenarios. Similarly, more work is needed to understand how, if at all, increases in sensitivity to rotational gain may impact user behavior. This is especially critical for the application of rotation gains to redirected walking. Finally, this work underscores the importance of collection and reporting information about participants' prior experience with VR to better contextualize the results of an experiment.

ACKNOWLEDGMENTS

This work was supported in part by US National Science Foundation (CISE HCC) grant #1717937.

REFERENCES

[1] Jeremy N Bailenson and Nick Yee. 2006. A longitudinal study of task performance, head movements, subjective report, simulator sickness, and transformed social interaction in collaborative virtual environments. *Presence: Teleoperators and Virtual Environments* 15, 6 (2006), 699–716.

[2] Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67, 1 (2015), 1–48. https://doi.org/10.18637/jss.v067.i01

[3] Geoffrey Bingham and Jennifer L Romack. 1999. The rate of adaptation to displacement prisms remains constant despite acquisition of rapid calibration. *Journal of Experimental Psychology: Human Perception and Performance* 25, 5 (1999), 1331.

[4] Geoffrey P Bingham and Christopher C Pagano. 1998. The necessity of a perception–action approach to definite distance perception: Monocular distance perception to guide reaching. *Journal of Experimental Psychology: Human Perception and Performance* 24, 1 (1998), 145.

[5] Gerd Bruder, Victoria Interrante, Lane Phillips, and Frank Steinicke. 2012. Redirecting walking and driving for natural navigation in immersive virtual environments. *IEEE transactions on visualization and computer graphics* 18, 4 (2012), 538–545.

[6] Hugo Brument, Maud Marchal, Anne-Hélène Olivier, and Ferran Argelaguet. 2020. Influence of Dynamic Field of View Restrictions on Rotation Gain Perception in Virtual Environments. In *International Conference on Virtual Reality and Augmented Reality*. Springer, 20–40.

[7] Pamela Dalton, Nadine Doolittle, and Paul AS Breslin. 2002. Gender-specific induction of enhanced sensitivity to odors. *Nature neuroscience* 5, 3 (2002), 199–200.

[8] Felipe Augusto dos Santos Mendes, José Eduardo Pompeu, Alexandra Modenesi Lobo, Keyte Guedes da Silva, Tatiana de Paula Oliveira, Andrea Peterson Zomignani, and Maria Elisa Pimentel Piemonte. 2012. Motor learning, retention and transfer after virtual-reality-based training in Parkinson's disease–effect of motor and cognitive demands of games: a longitudinal, controlled clinical study. *Physiotherapy* 98, 3 (2012), 217–223.

[9] Natalia Dużmańska, Paweł Strojny, and Agnieszka Strojny. 2018. Can simulator sickness be avoided? A review on temporal aspects of simulator sickness. *Frontiers in psychology* 9 (2018), 2132.

[10] Elham Ebrahimi, Andrew Robb, Leah S Hartman, Christopher C Pagano, and Sabarish V Babu. 2018. Effects of anthropomorphic fidelity of self-avatars on reach boundary estimation in immersive virtual environments. In *Proceedings of the 15th ACM Symposium on Applied Perception*. 1–8.

[11] AL Faria, J Couras, MS Cameirão, T Paulino, GM Costa, and S Bermúdez i Badia. 2016. Impact of combined cognitive and motor rehabilitation in a virtual reality task: an on-going longitudinal study in the chronic phase of stroke. *Proceedings of the 11th ICDVRAT* (2016).

[12] Daniel Freeman, C Thompson, Natasha Vorontsova, G Dunn, L-A Carter, P Garety, E Kuipers, Mel Slater, Angus Antley, Ed Glucksman, et al. 2013. Paranoia and post-traumatic stress disorder in the months after a physical assault: a longitudinal study examining shared and differential predictors. *Psychological medicine* 43, 12 (2013), 2673–2684.

[13] Eugy Han, Mark Roman Miller, Nilam Ram, Kristine L Nowak, and Jeremy N Bailenson. 2022. Understanding Group Behavior in Virtual Reality: A Large-Scale, Longitudinal Study in the Metaverse. In *72nd Annual International Communication Association Conference, Paris, France*.

[14] MA Hernandez-Mocholi, FJ Dominguez-Muñoz, H Corzo, SCS Silva, JC Adsuar, and N Gusi. 2016. Whole body vibration training improves vibration perception threshold in healthy young adults: A randomized clinical trial pilot study. *Journal of Musculoskeletal & Neuronal Interactions* 16, 1 (2016), 12.

[15] Tanya Hill and Hanneke du Preez. 2021. A Longitudinal Study of Students' Perceptions of Immersive Virtual Reality Teaching Interventions. In *2021 7th International Conference of the Immersive Learning Research Network (iLRN)*. IEEE, 1–7.

[16] Xieyining Huang, Kensie M Funsch, Esther C Park, Paul Conway, Joseph C Franklin, and Jessica D Ribeiro. 2021. Longitudinal studies support the safety and ethics of virtual reality suicide as a research method. *Scientific reports* 11, 1 (2021), 1–12.

[17] Robert S Kennedy, Norman E Lane, Kevin S Berbaum, and Michael G Lilienthal. 1993. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology* 3, 3 (1993), 203–220.

[18] Negar Khojasteh and Andrea Stevenson Won. 2021. Working Together on Diverse Tasks: A Longitudinal Study on Individual Workload, Presence and Emotional Recognition in Collaborative Virtual Environments. *Frontiers in Virtual Reality* 2 (2021), 53.

[19] Hyun-Jin Kim, Seulki Lee, Dooyoung Jung, Ji-Won Hur, Heon-Jeong Lee, Sungkil Lee, Gerard J Kim, Chung-Yean Cho, Seungmoon Choi, Seung-Moo Lee, et al. 2020. Effectiveness of a participatory and interactive virtual reality intervention in patients With social anxiety disorder: longitudinal questionnaire study. *Journal of medical Internet research* 22, 10 (2020), e23024.

[20] Ita GG Kreft, Jan De Leeuw, and Leona S Aiken. 1995. The effect of different forms of centering in hierarchical linear models. *Multivariate behavioral research* 30, 1 (1995), 1–21.

[21] Charlene Krueger and Lili Tian. 2004. A comparison of the general linear mixed model and repeated measures ANOVA using a dataset with multiple missing data points. *Biological research for nursing* 6, 2 (2004), 151–157.

[22] Scott A Kuhl, Sarah H Creem-Regehr, and William B Thompson. 2008. Recalibration of rotational locomotion in immersive virtual environments. *ACM Transactions on Applied Perception (TAP)* 5, 3 (2008), 1–11.

[23] Daniel Linares and Joan López-Moliner. 2016. quickpsy: An R Package to Fit Psychometric Functions for Multiple Groups. *The R Journal* 8, 1 (2016), 122–131. https://journal.r-project.org/archive/2016-1/linares-na.pdf

[24] Lotte Meteyard and Robert AI Davies. 2020. Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language* 112 (2020), 104092.

[25] Fares Moustafa and Anthony Steed. 2018. A longitudinal study of small group interaction in social virtual reality. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*. 1–10.

[26] Niels Christian Nilsson, Evan Suma, Rolf Nordahl, Mark Bolas, and Stefania Serafin. 2016. Estimation of detection thresholds for audiovisual rotation gains. In *2016 IEEE Virtual Reality (VR)*. IEEE, 241–242.

[27] Akimi Oyanagi, Takuji Narumi, Kazuma Aoyama, Kenichiro Ito, Tomohiro Amemiya, and Michitaka Hirose. 2021. Impact of Long-Term Use of an Avatar to IVBO in the Social VR. In *International Conference on Human-Computer Interaction*. Springer, 322–336.

[28] Anders Paludan, Jacob Elbaek, Mathias Mortensen, Morten Zobbe, Niels Christian Nilsson, Rolf Nordahl, Lars Reng, and Stefania Serafin. 2016. Disguising rotational gain for redirected walking in virtual reality: Effect of visual density. In *2016 IEEE Virtual Reality (VR)*. IEEE, 259–260.

[29] Tabitha C Peck, Laura E Sockol, and Sarah M Hancock. 2020. Mind the gap: The underrepresentation of female participants and authors in virtual reality research. *IEEE transactions on visualization and computer graphics* 26, 5 (2020), 1945–1954.

[30] Carina Peckmann, Kyra Kannen, Max C Pensel, Silke Lux, Alexandra Philipsen, and Niclas Braun. 2022. Virtual reality induces symptoms of depersonalization and derealization: A longitudinal randomised control trial. *Computers in Human Behavior* 131 (2022), 107233.

[31] Uri Polat, Tova Ma-Naim, Michael Belkin, and Dov Sagi. 2004. Improving vision in adult amblyopia by perceptual learning. *Proceedings of the National Academy of Sciences* 101, 17 (2004), 6692–6697.

[32] John Porter III, Matthew Boyer, and Andrew Robb. 2018. Guidelines on successfully porting non-immersive games to virtual reality: a case study in minecraft. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*. 405–415.

[33] John Porter III and Andrew Robb. 2019. An analysis of longitudinal trends in consumer thoughts on presence and simulator sickness in VR games. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. 277–285.

[34] R Core Team. 2017. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

[35] Aylen Ricca, Amine Chellali, and Samir Otrnane. 2021. The influence of hand visualization in tool-based motor-skills training, a longitudinal study. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, 103–112.

[36] Hiroaki Sakono, Keigo Matsumoto, Takuji Narumi, and Hideaki Kuzuoka. 2021. Redirected walking using continuous curvature manipulation. *IEEE Transactions on Visualization & Computer Graphics* 27, 11 (2021), 4278–4288.

[37] Stefania Serafin, Niels C Nilsson, Erik Sikstrom, Amalia De Goetzen, and Rolf Nordahl. 2013. Estimation of detection thresholds for acoustic based redirected walking techniques. In *2013 IEEE Virtual Reality (VR)*. IEEE, 161–162.

[38] Sherrill J Smith, Sharon Farra, Deborah L Ulrich, Eric Hodgson, Stephanie Nicely, and William Matcham. 2016. Learning and retention using virtual reality in a decontamination simulation. *Nursing education perspectives* 37, 4 (2016), 210–214.

[39] Frank Steinicke, Gerd Bruder, Jason Jerald, Harald Frenz, and Markus Lappe. 2008. Analyses of human sensitivity to redirected walking. In *Proceedings of the 2008 ACM symposium on Virtual reality software and technology*. 149–156.

[40] Frank Steinicke, Gerd Bruder, Jason Jerald, Harald Frenz, and Markus Lappe. 2009. Estimation of detection thresholds for redirected walking techniques. *IEEE transactions on visualization and computer graphics* 16, 1 (2009), 17–27.

[41] Tuukka M Takala, Lauri Malmi, Roberto Pugliese, and Tapio Takala. 2016. Empowering students to create better virtual reality applications: A longitudinal study of a VR capstone course. *Informatics in Education* 15, 2 (2016), 287–317.

[42] Ioannis Tarnanas, Winfried Schlee, Magda Tsolaki, René Müri, Urs Mosimann, and Tobias Nef. 2013. Ecological validity of virtual reality daily living activities screening for early dementia: longitudinal study. *JMIR serious games* 1, 1 (2013), e2778.

[43] Viswanath Venkatesh and Philip Johnson. 2002. Telecommuting technology implementations: a within-and between-subjects longitudinal field study. *Personnel psychology* 55, 3 (2002), 661–687.

[44] Cesko C Voeten. 2020. Using 'buildmer'to automatically find & compare maximal (mixed) models.

[45] Hadley Wickham. 2009. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. http://ggplot2.org

[46] Niall L Williams and Tabitha C Peck. 2019. Estimation of rotation gain thresholds considering fov, gender, and distractors. *IEEE transactions on visualization and computer graphics* 25, 11 (2019), 3158–3168.

[47] Alexander Winkler-Schwartz, Khalid Bajunaid, Muhammad AS Mullah, Ibrahim Marwa, Fahad E Alotaibi, Jawad Fares, Marta Baggiani, Hamed Azarnoush, Gmaan Al Zharni, Sommer Christie, et al. 2016. Bimanual psychomotor performance in neurosurgical resident applicants assessed using NeuroTouch, a virtual reality simulator. *Journal of surgical education* 73, 6 (2016), 942–953.

[48] Jingxin Zhang, Eike Langbehn, Dennis Krupke, Nicholas Katzakis, and Frank Steinicke. 2018. Detection thresholds for rotation and translation gains in 360 video-based telepresence systems. *IEEE transactions on visualization and computer graphics* 24, 4 (2018), 1671–1680.

[49] Ruimin Zhang and Scott A Kuhl. 2013. Human sensitivity to dynamic rotation gains in head-mounted displays. In *Proceedings of the ACM Symposium on Applied Perception*. 71–74.