# Comparing Gamepad and Naturally-mapped Controller Effects on Perceived Virtual Reality Experiences

Monthir Ali*
Rogelio E. Cardona-Rivera
monthir.ali@utah.edu
rogelio@cs.utah.edu
Laboratory for Quantitative Experience Design
School of Computing, University of Utah
Salt Lake City, UT, USA

## ABSTRACT

A virtual reality (VR) system's interface determines its corresponding interaction fidelity. While there exist several naturally-mapped interfaces, these have not been directly compared – in terms of presence, engagement, perceived performance, and real performance – to a more commonly-used interface for interacting with virtual reality: the gamepad. Because gamepads have evolved over several decades and have widespread adoption, it is pragmatic to ask whether it is viable to import them into the design of VR experiences. To study this, we developed a first-person shooter (FPS) Virtual environment and contrasted the VR experience using the Microsoft Xbox 360 controller against the HTC VIVE wand. We assessed the effect of the input device on self-reported presence, engagement, and performance as well as real performance in terms of accuracy and speed, two metrics relevant for our particular environment. A within-subjects FPS shooting task under accuracy and time constraints confirmed our hypotheses in favor of the VIVE wand: presence, engagement, perceived performance, and real performance were deemed significantly higher than the Xbox controller condition. These results are contextualized with participants' comments in a post-experiment debrief, in which (paradoxically) several participants reported *preferring* the gamepad. We discuss the implications of our findings and how what we call *genre fidelity* can impact the design of VR experiences.

## CCS CONCEPTS

• **General and reference** → **Empirical studies**; • **Human-centered computing** → **Virtual reality**; • **Applied computing** → Computer games.

## KEYWORDS

virtual reality, interaction fidelity, naturally-mapped interface, genre fidelity


*SAP '20, September 12–13, 2020, Virtual Event, USA*

ACM ISBN 978-1-4503-7618-1/20/09…$15.00
https://doi.org/10.1145/3385955.3407923

**ACM Reference Format:**
Monthir Ali and Rogelio E. Cardona-Rivera. 2020. Comparing Gamepad and Naturally-mapped Controller Effects on Perceived Virtual Reality Experiences. In *ACM Symposium on Applied Perception 2020 (SAP '20), September 12–13, 2020, Virtual Event, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3385955.3407923

## 1 INTRODUCTION

Fidelity is one of many interesting dimensions to evaluate Virtual Reality (VR) experiences. Ragan et al. [2015] identify three types: (a) **interaction fidelity**, which concerns how real-world interactions are translated inside the virtual environment as well as the accuracy of translation, (b) **display fidelity**, which concerns the realism of the VR's display output as well as how real-world stimuli are reproduced within the virtual environment, and (c) **scenario fidelity**, which concerns the realism of behaviors, attributes, and rules that the simulated scenario reproduces.These different types of fidelity contribute differently to the VR experience and, as they increase, the virtual environment increases in similarity to the real world. This paper focuses on interaction fidelity.

Prior work has explored the effects of varying the level of interaction fidelity on presence, engagement, usability, and task performance. These studies have addressed specific interaction fidelity components such as **view control** (manipulations applied to change the user's current view without them moving) [Barfield et al. 1997; Heineken and Schulte 2007; Narayan et al. 2005; Raja 2006; Ware et al. 1993; Ware and Franck 1996], **locomotion** (moving from one position to another within the virtual environment) [Chance et al. 1998; Chung 1992; Usoh et al. 1999], **object manipulation** in the virtual space [Balakrishnan et al. 1997; Hinckley et al. 1997], **form factor** [Zhai et al. 1996], and **lag** [Arthur et al. 1993; Ware and Balakrishnan 1994]. However, no work has directly compared gamepads and naturally-mapped interfaces *vis-à-vis* presence, engagement, perceived performance, and real performance.

Such a comparison is crucial for designers interested in crafting VR experiences. Bracken and Skalski [2010] defines natural mapping in video games as "*how closely actions represented in the virtual environment match the natural actions used to bring about change in a real environment*." The lack of physical involvement and the use of unrealistic actions that are less similar to real-life actions – via mapped buttons and joysticks – diminishes the ability to induce presence [Skalski et al. 2011]; using naturally-mapped controllers

(NMCs) over non-NMCs has a significant positive effect on naturalness, spatial presence, and enjoyment. However, gamepads[1] have evolved in tandem with games for over forty years [Cummings 2007] and their symbols and corresponding physical arrangements are governed by **design conventions** [Norman 1999] accrued over that time span. These design conventions are not arbitrary, have emerged in a community of practice, and are relied upon to navigate new contexts and manage complexity in virtual environments. Further, it is not obvious that NMCs are always the best candidate for VR experiences: prior work has found support for gamepads being the better alternative in terms of user preference [Hinckley et al. 1997] and target selection performance [Farmani and Teather 2017].

As a result, it is pragmatic to ask whether it is viable to import gamepads controllers as we know them into VR; failing to ask the question would ignore all that we have learned about their use thus far. Furthermore, studying the use of gamepads in virtual worlds may broaden the applicability of existing VR technology.

In this paper, we present a study that compared two input devices; one that is traditionally used to play video games (the Microsoft Xbox 360 controller; hereafter, "Xbox controller") and another which is naturally-mapped (the HTC VIVE wand; hereafter, "VIVE wand"). The experiment compares the perceived and real performances (which correspond to self-reported accuracy and real accuracy in our study), as well as self-reported presence and engagement. We asked participants to complete a target shooting experimental task in a first-person shooter (FPS) virtual experience twice, using a different input device each time while trying to maintain accuracy and speed. We compared participants' data between the two conditions (within subjects) to determine whether the input device has any effect on presence, engagement, perceived performance, or real performance for the given experimental task. Our analysis provides evidence in favor of the VIVE wand: presence, engagement, perceived performance, and real performance were deemed significantly higher than the Xbox controller condition, contradicting earlier related (and different) studies [Farmani and Teather 2017]. These results are contextualized with participant comments in a post-experiment debrief, in which (paradoxically) several participants reported *preferring* the gamepad. We discuss the implications of our findings for the design of VR experiences.

*Contributions.* We address the effects of two heretofore not compared controllers on presence, engagement, perceived performance, and real performance in VR, and discuss the broader impact of our findings on interaction fidelity and scenario fidelity. Our work leads us to introduce a new term, **genre fidelity**, to distinguish a specific kind of interaction fidelity, focused on – not simulation, but – *user expectation.* We believe this distinction can usefully broaden the language with which we approach the design of VR experiences.

## 2 RELATED WORK

As alluded to earlier, the idea of addressing controllers and their design as a component of VR interaction fidelity is not new. However, no study targets both a virtual reality experience that contrasts gamepads to NMCs in terms of real & perceived performance (in terms of speed and accuracy), presence, and engagement. In this paper, we fill this gap.

Farmani and Teather [2017] compared 3D target selection of three different input devices – mouse, gamepad, and VR controller – in an FPS game setting. Task performance was measured with respect to the speed with which targets are acquired and the percentage of missed targets, which they referred to as accuracy. They chose the Razor Hydra as their VR controller and hypothesized that using the Hydra would lead to better performance. Contrary to what they had hypothesized, the VR controller performance came last (mouse was first, gamepad second), suggesting disappointing prospects for VR controllers in VR FPS games. Our study design goes beyond theirs by exploring how the input devices affect – in addition to task performance – immersion aspects of the virtual environment, such as presence and engagement.

Zielinski et al. [2014] compared three input devices (6-DOF wand, Air mouse, Xbox Gamepad) for accuracy, speed, and usability for a VR mining simulation. The simulation involved three tasks: target selection, navigation, and maneuvering. They found the wand was preferred by most users and using it produced better timings for task completion. However, when accounting for target size, the want resulted in higher error rates and less accuracy.

McMahan et al. [2012] demonstrated that both display and interaction fidelity in VR have a significant impact not only on performance, but also on presence, engagement, and usability. Their study measured the combined and independent effects of interaction fidelity and display fidelity using a first-person shooter game, played on a six-sided CAVE in four different conditions: (a) high display (fidelity) / high interaction (fidelity), (b) high display / low interaction, (c) low display / high interaction, and (d) low display / low interaction. Performance results were in favor of configurations that were consistent; high display / high interaction and low display / low interaction settings were favored, which the authors found understandable since these two settings represent the typical ways of interacting in the real world and playing FPS games, respectively. Our work is complementary to theirs: we control for display fidelity by using a different display device and use a gamepad controller instead of a mouse and keyboard for our study's "low interaction fidelity" condition.

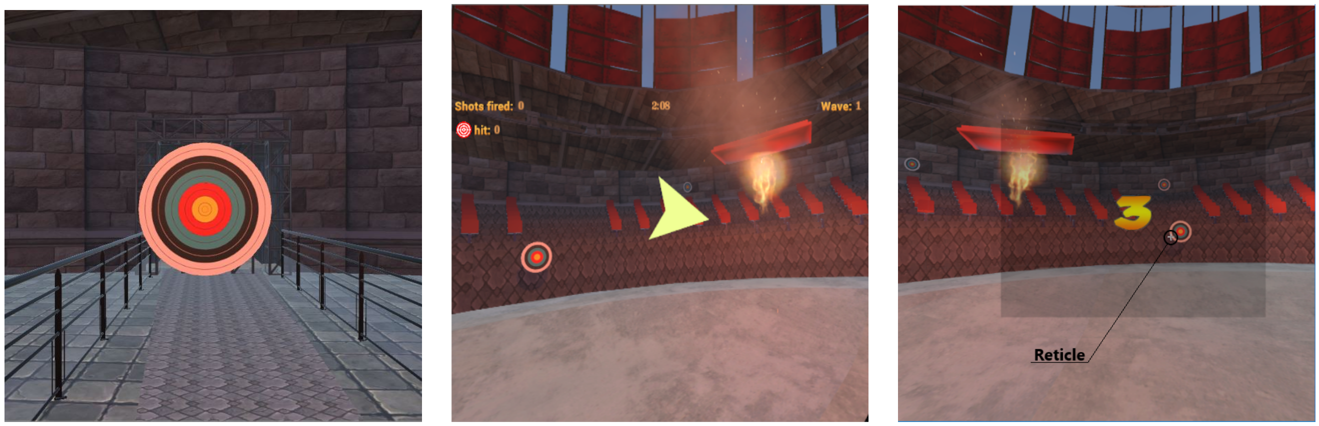
## 3 EXPERIMENT

We address how participants perform with two different input devices (VIVE wand and Xbox controller) when trying to complete the given task in the virtual environment. We also address the impact of those input devices' interactions on presence, engagement, and real and perceived performances, which can directly affect a user's experience. Presence, engagement, and perceived performance are measured through self-assessment, whereas real performance is assessed through timing and accuracy metrics; all measures were taken to better characterize how interaction fidelity actually and perceptually affect a user's experience, which we posit are key considerations for designers of virtual worlds.

### 3.1 Experimental design

The virtual environment was designed as representative of the **first-person shooter** (FPS) genre, a sub-genre of action games in

[1] Alternative names include "joypad", "game controller," and "controller." We use these terms interchangeably.




**(a) Bullseye target: participants need to hit all 24 bullseye targets to finish the experimental task. Hitting anywhere on the target will cause it to be destroyed.**

**(b) Xbox controller reticle finder: whenever the 3D reticle goes offscreen while using the Xbox controller, the HUD will display a yellow arrow pointing towards the reticle position in 3D space.**

**(c) Countdown timer: before the start of each wave, the HUD displays a countdown. Once it is over, the new wave is spawned. The figure also shows the reticle**

**Figure 1: Objects that participants perceived within the virtual environment developed for our experimental task.**

which the most important skill is aiming [Adams 2014]. An FPS is a 3D shooter game that uses a first-person perspective camera, where participants must focus their attention on targets and their immediate vicinity to achieve in-game objectives. We deemed the FPS genre the best fit to conduct our study since it represents an ecologically valid environment within which to evaluate the two input devices we were interested in comparing. On the one hand, modern and commercially available FPS experiences routinely afford gamepads for input [Ng [n.d.]; Polson [n.d.]]. On the other hand, the manipulation of objects in an FPS setting(*i.e.* weapons) lends itself well to **naturalistic control** [McEwan et al. 2014] via the VIVE wand; the wand is an interface that is **directional**, **kinesic**, and **incomplete tangible** [Skalski et al. 2011].

In this study we were solely interested in the effect that our two different input devices would have on presence, engagement, perceived performance, and real performance (measured through elapsed time and accuracy). We were not interested in any other parts of the interaction. Thus, we disabled locomotion within the virtual world; the user cannot move their avatar's position throughout the whole experiment session and is limited to changing their orientation. Further, past studies have indicated that self-reported engagement [LaViola, Jr. and Litwiller 2011; Lugrin et al. 2013; McMahan et al. 2012] and self-reported presence [Kim et al. 2014; Kober et al. 2012; McMahan et al. 2012; Zanbaka et al. 2005] are affected by the type of display used. Therefore, we used a single display device for the entirety of this study to control for the effect of display fidelity; all participants used a HMD. This design follows the one used by Farmani and Teather [2017]: it decouples aiming and looking around, with the HMD used to look around the virtual environment and the input device used to aim at targets.

In light of prior work in this area, we formulated three hypotheses for this study:

**H1: Accuracy** – Number of shots fired to complete the experimental task will be lower when using a higher fidelity controller compared to a lower fidelity controller.

**H2: Elapsed Time** – The time measured from the start of the experimental task until it is over, or until the timer runs out, will be less when using a higher fidelity controller compared to a lower fidelity controller.

**H3: Self-reported Presence, Engagement, and Accuracy** – Self-reported presence, engagement, and accuracy will be higher when users complete the experimental task using a higher fidelity controller compared to a lower fidelity controller.

## 3.2 Method

To test our hypotheses, we developed an FPS virtual environment that could be experienced using either an Xbox controller or a VIVE wand. We ranked the VIVE wand as a higher fidelity controller because it is *naturally mapped* to the FPS task domain [Pereira et al. 2018; Schmitt et al. 2019]. The VIVE's afforded interaction is body-centered [Slater and Usoh 1994] and resembles what a person would do when they want to aim at and shoot a target: aligning the hand toward a target and then pulling a trigger.

Its afforded interaction is more similar to what a person would do in real life compared to the gamepad. Participants completed two experimental task sessions inside the virtual environment using a different input device for each session. Both sessions were completely identical with the only difference being the input device used and its associated functions. The session started with the participant placed in the center of a Roman colosseum. Participants were allowed to start shooting after a 3-second countdown (illustrated in Figure 1c). The objective was to destroy all targets in the least amount of time, while refraining from shooting unnecessarily. These constraints help ensure that participants will neither be shooting randomly and hastily to finish fast nor take their time to shoot every target. Participants were made aware of these constraints and that they would be measured on them beforehand, and were taken through a tutorial session that is identical to the experimental task before their main sessions started to ensure they were comfortable with the interface and mechanics.

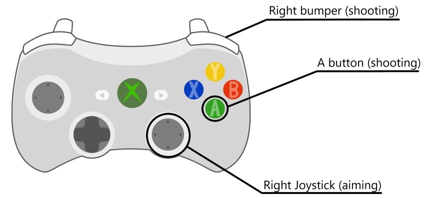
Targets (illustrated in Figure 1a) were grouped into waves and destroyed when hit once. Each session consisted of 3 waves of increasing difficulty. There were 8 targets in each wave (24 targets in total). The first wave had stationary targets. The second wave had moving (some vertically and others horizontally) targets in addition to distant stationary targets. The last wave had targets that moved faster and in two different directions. When a wave was cleared, the heads-up display (HUD) displayed a message informing the participant that they have cleared a wave. Shortly after, a new countdown started and the next wave was spawned when the countdown reached 0. Two metrics were recorded for each participant: (a) number of shots fired and (b) elapsed time.

Participants used a 3D reticle to aim at the targets, which was not fixed to the center of the screen. In the case of the VIVE wand, the reticle was moved around by waving and moving the wand in the desired movement direction. Whereas in the Xbox controller case, it was moved using the right joystick. Orientation and looking around were achieved through the VIVE head mounted display in both cases. Since aiming and orientation were carried out through two different means, the reticle could go off screen. Finding the reticle when it goes off screen is easy when using the VIVE wand because the participant knows exactly where the wand is pointing relative to where he or she is looking (*i.e.* the direction where their hand is oriented toward). Unfortunately, for the Xbox controller it is not as trivial. Players can easily lose track of where they moved the reticle because the joystick physically re-centers itself after each movement. This means that on the joypad, the joystick would return to its original position but the reticle remains at the same last coordinates in the virtual space, making it difficult to track unless the participant can actually see it. To compensate for this, we created a function that is exclusive to the Xbox controller version of the game; whenever the reticle goes off screen, a yellow arrow will appear on screen pointing towards the reticle position in the virtual space making it easier to find (illustrated in Figure 1b). As soon as the reticle is on screen again the arrow disappears. We opted for this design in order to have the interface for both input devices be as similar as possible without adjusting the Xbox controller functionality. We wanted to make sure that when we are comparing those two input devices, we are comparing them in the way they are normally used.

## 3.3 Apparatus

The virtual environment (whose objects are illustrated in Figure 1) used in this experiment was developed using Unity Engine 2017.2.0f3, and ran on a PC with the following specifications: 2.2GHz Intel Core i7-8750H processor, 16GB RAM, and an Nvidia GTX 1060 graphics card. The display device used was an HTC VIVE HMD that has an OLED display with 2160 x 1200 resolution, 110° FOV, and 6-DOF tracker. The two input devices (illustrated in Figure 2) used were: (a) Xbox controller (wired and wireless), and (b) the VIVE wand.

Participants used the wand or the right joystick on the Xbox controller to move the reticle in 3D space. The "A button" or the "right bumper" on the Xbox controller, and the "trigger button" on the VIVE wand functioned as buttons to shoot. Participants




**(a) Configuration for the Xbox controller.**


**(b) Configuration for the VIVE wand.**

**Figure 2: Xbox (top) and VIVE (bottom) controllers: On the Xbox controller, participants use either the "A button" or the "right bumper" to shoot and the "right joystick" to aim. On the VIVE controller, participants use the "trigger button" to shoot and move the wand to aim.**

listened to audio in all cases through Mpow Flame bluetooth noise-cancelling headphones. The HUD displays and starts a countdown timer at the start of each session and before every new wave. in the event of reticle going off screen when using the Xbox controller, the HUD also displays a yellow arrow that points towards the reticle.

## 3.4 Participants

We recruited 28 (23M, 5F) volunteers whose ages ranged from 20-35 (mean value= 25.1, standard deviation = 3.70 ). We asked participants to rate their experience playing FPS video games on a 6-point nominal scale: (1) "I have never played computer or console FPS games," (2) "I only played a few times in my life," (3) "I used to play occasionally but no longer do," (4) "I used to play a lot but no longer do," (5) "I play occasionally," and (6) "I play a lot." Answers ranged from 1 to 6 (Mode = 5), with only one participant reporting no prior FPS game experience. We also asked them if they had used VR before (80% answered no), and if they did, to report their frequency of use on a 4-point nominal scale: (1) "Less than a month," (2) "1 to 6 months," (3) "6–12 months," and (4) "more than a year."

Similarly, answers ranged from 1 to 4 (Mode = 4). Participants were recruited through personal contact, email, and posted flyers. Participants received no compensation and completed 2 sessions (after a training session), using a different input device for each. To balance priming and ordering effects, we had half the participants start with the Xbox controller, and the other half with the VIVE wand. We had 7 of 12 experienced participants (ranked 5 and 6 in the FPS experience question) start with the VIVE controller and 5 with the Xbox controller. 16 of the participants used a wired Xbox controller and 12 used a wireless one. The demographic survey was used to assess whether participants' experience with FPS games and the amount of time they have been using VR had a significant effect on our results.

## 3.5 Measures

To measure the effects of different types of interaction devices we collected data for the following two broad categories of metrics. **Objective metrics** describe assessments that are observable and not prone to interpretation. We use elapsed time and number of shots fired in this category to measure performance after each session. **Subjective metrics** describe assessments that are not observable, but rather obtained from participants' feedback. We asked participants' to self-report presence, engagement and accuracy.

*Self-reported Presence* – To evaluate participants' sense of presence inside the game, We used the Igroup Presence Questionnaire (IPQ) – a 14-item self-report scale [Schubert et al. 2001]. The questionnaire categorizes its items into 4 components: the sense of "being there," experienced realism, involvement and spatial presence. Participants rated their level of agreement with IPQ statements on a 7-point Likert scale, ranging from 1 to 7.

*Self-reported Engagement* – In order to measure the level of engagement participants experienced inside the virtual space, we used the questionnaire administered by Buttussi and Chittaro [Buttussi and Chittaro 2018]. The questionnaire is comprised of a 7-point Likert scale that corresponds to the level of agreement with six statements: (a) "It was boring," (b) "It was engaging," (c) "It aroused emotions in me," (d) "The depicted situation looked real," (e) "I forgot the passing of time," and (f) "I felt immersed in the depicted situation." The first item in the scale was reversed to match the rest of the questionnaire.

*Self-reported Accuracy* – To measure how accurate participants felt the used input device was as well as how comfortable and natural its movement was, we administered a custom 5-point Likert scale with 5 items (Not at all, slightly, moderately, very, extremely) to self-report on. These items were: (a) "The input device movements translated accurately inside the virtual environment," (b) "The input device movements were smooth," (c) "The input device felt comfortable in my hands," (d) "Bullets landed where I expected them to," and (e) "Moving the input device felt like moving a real weapon."

## 3.6 Procedure

Participants were informed about the nature of the experience and that they will play two sessions using two different input devices. They were asked to sign an informed consent form to participate and have their data be recorded. They were also told that they have the right to opt out at any time and were explicitly asked to do so if they felt dizzy or sick. Before the start of the experiment, participants filled a demographic questionnaire. The questionnaire included questions about age, gender, experience with FPS games and experience with VR. Next, participants received their first input device and were given instructions on how to use it. Once they confirmed that they have understood the instructions, they were assisted in putting on the VIVE headset.

After participants had everything equipped, they were informed that they would be measured on speed and accuracy to guarantee they would complete the session to the best of their ability. Participants had a tutorial session before starting each of their real sessions to ensure they understand the mechanics and interface of the game. They were also given the choice of choosing to play with inverted Y-axis when using the Xbox controller if they choose to. After the tutorial, they completed two sessions counter-balanced for ordering effects. Participants filled three surveys about engagement, presence and accuracy after completing the first session. Next, they switched input devices and proceeded to start their second session. Similarly, they were given instructions on how to use the second input device, went through a tutorial session and filled the same three surveys afterwards. At the end of the experiment, participants were asked to give their impression of the experience and the two input devices, what they thought could have been better, and how comfortable the experience was.

# 4 ANALYSIS AND RESULTS

All data was analyzed using standard thresholds for significance ($\alpha = 0.95, \beta = 0.20$). Corresponding power analyses were conducted and are listed next to the tests in question. We discuss the objective and subjective metrics below, in turn.

**H1: Accuracy** – For objective accuracy, we studied the number of shots fired to complete the task. We performed a t-Test to test if the input device had a significant effect, which corresponds to our first hypothesis. Confirming our hypothesis, results showed that participants who used the VIVE wand were significantly more accurate than those who used the Xbox controller (paired-t(df = 27) = $4.2, p < 0.0003$) with effect size $d = 0.79$. Participants shot an average of 27.6 shots to complete the experimental task when using the VIVE wand, and an average of 29.7 when using the Xbox controller (Figure 4a).

**H2: Elapsed Time** – The same t-Test was performed to test our second hypothesis: that the time needed to finish the experimental task will be lower when using the VIVE wand. The test revealed a significant difference (paired-t(df = 27) = $9.5, p < 0.0001$) with effect size $d = 1.8$ in favor of the VIVE wand. When participants used the VIVE wand, it took them an average of 59.14 seconds to finish the experimental task compared to an average of 103.6 when using the Xbox controller (Figure 4b).

Considering that Cohen [Cohen 1992] suggests effect sizes $d$ of 0.2, 0.5 and 0,8 represent small, medium, and large effect sizes respectively for t-Tests, our values[2] of 0.79 and 1.8 for **H1** and **H2** indicate that participants' accuracy and timing benefit significantly – statistically and practically – from using a higher fidelity controller.

To determine if participants' backgrounds had a significant effect on any of the objective metrics, we computed Pearson's correlation coefficient between participants' experience with both FPS games and VR and each of the objective metrics. We found a negative significant correlation between FPS expertise and accuracy using the VIVE wand ($\rho = -0.45, p < 0.02$) and between FPS expertise and the time it took to finish the experimental task using the Xbox Controller ($\rho = -0.65, p < 0.0002$). We did not find any significant correlation between VR experience and participants' accuracy using any

[2] Computed using the **effsize** [Torchiano 2016] library in R.

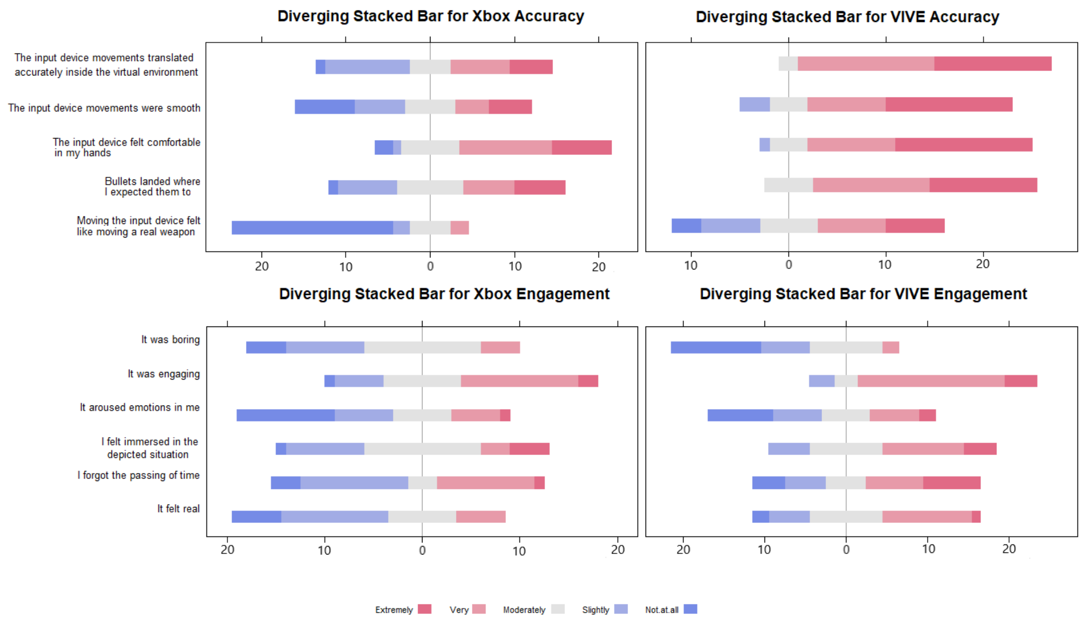



**Figure 3: Diverging Stacked Bars showing Likert scale distributions for (a) Xbox Accuracy, (b) VIVE Accuracy (c) Xbox Engagement, and (d) VIVE Engagement. Wilcoxon Signed-Rank tests yielded statistically significant differences with medium to large effect sizes, all in favor of the VIVE wand.**

input device, nor between FPS expertise and the time it took to finish the experimental task using the VIVE wand. We also performed another t-Test to see using a wired or wireless controller made a difference. We took a random sample of 12 participants who used the wired controller and another random sample of 12 who used the wireless one and performed the test on their data. The results were insignificant for both accuracy (paired t(df = 11) = 1.62, $p > 0.05$), and elapsed time (paired t(df = 11) = 0.63, $p > 0.05$).

**H3: Self-reported Presence, Engagement, and Accuracy** – In Figure 3, we illustrate the Likert scale data distribution of self-reported presence, engagement and accuracy using a Diverging Stacked Bar [Heiberger and Robbins 2014]. We performed the non-parametric Wilcoxon Signed-Rank Test to compare difference in survey scores. Effect sizes for these tests were obtained using Rosenthal's [Rosenthal 1994] formula to calculate effect size for non-parametric tests.

For engagement, our test indicated that participants' engagement ranks mean values when using the VIVE wand (M = 3.34) were statistically significantly higher (Z-value= $-3.11, p < 0.002$) than ranks mean values when using the Xbox controller (M = 2.90), with effect size $r = 0.42$. We used a scale with an acceptable reliability (Cronbach's Alpha = 0.73) that was computed from averaging the six engagement ratings.

For presence, our test revealed statistically significant differences (Z-value = $-3.35, p < 0.001$) in presence ranks total score values in favor of the VIVE wand, with effect size $r = 0.44$. The IPQ individual components ranks values were also analyzed using Wilcoxon Signed-Rank Test to reveal the following statistically significant differences between participants when using the VIVE wand and the Xbox controller (all in favor of the VIVE wand); results for these components are reported in Table 2.

For accuracy, we used a separate test for each individual item in the scale since it was a custom one. Results (see Table 1) indicated that for 5 out of 6 items, participants' accuracy when they used the VIVE wand (M = 4.04) was statistically significantly higher than when they used the Xbox controller (M = 2.93).

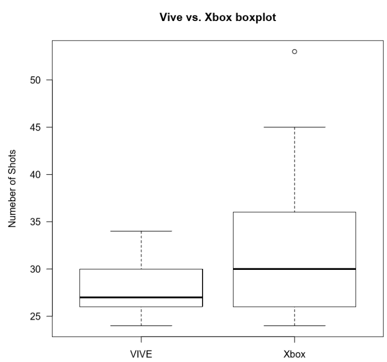
Cohen [1988] suggests that r values of 0.1, 0.3, and 0.5 represent small, medium, and large effect sizes respectively. Our values fall mostly between large and medium effect sizes, meaning that our results are statistically *and* practically significant. Finally, we performed additional tests and concluded that there was no significant effect when switching from wired to wireless on presence (Z-value= $-1.84, p > 0.05$), engagement (Z-value= $-0.13, p > 0.05$) or perceived accuracy (Z-value= $-1.26, p > 0.05$).

## 5 DISCUSSION

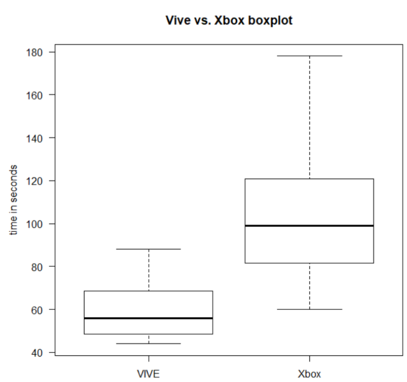
*Summary of Findings.* Our results confirmed all three of our hypotheses. There was a statistical and practical significant difference in accuracy and elapsed time (real performance) when comparing the outcome of the experimental task using the higher-fidelity VIVE wand against the lower-fidelity Xbox controller. This indicates that *in VR experiences that have accuracy and time constraints, increasing the level of interaction fidelity may result in better performance in accuracy and elapsed time, as well as with greater (self-reported) senses of presence, engagement, and accuracy.* However, it is not necessarily the case that every individual would prefer to use higher-fidelity controllers even if they perform better with it.

**Table 1: Self-reported Accuracy Questionnaire items and their corresponding Wilcoxon Signed-Rank Test values. We found statistically significant differences ($p < 0.05$) between controllers for 5 out of 6 items (marked with *).**

| Self-reported Accuracy Question | Z-value | p-value | Effect size (r) |
|---|---|---|---|
| The input device movements translated accurately inside the virtual environment | -3.24 | 0.001* | 0.43 |
| The input device movements were smooth | -2.95 | 0.003* | 0.39 |
| The input device felt comfortable in my hands | -1.90 | 0.057 | 0.25 |
| Bullets landed where I expected them to | -3.10 | 0.002* | 0.41 |
| Moving the input device felt like moving a real weapon | -4.02 | 0.00001* | 0.53 |




**(a) Number of shots fired in the experimental task.**




**(b) Time (seconds) taken in the experimental task.**

**Figure 4: Boxplots of objective metrics recorded in the VIVE wand and the Xbox controller experiment settings. t-Tests for both number of shots fired and time taken in the experimental task yielded statistically significant differences with medium to large effect sizes, all in favor of the VIVE wand. Notably, the difference in time taken was dramatic; Xbox controller users took an average of 44.46 seconds more than VIVE wand users to complete the experimental task.**


**Figure 5: Means of self-reported engagement, accuracy, and presence using Xbox controller v. VIVE wand.**

**Table 2: Igroup Presence Questionnaire sub-scale scores and their corresponding Wilcoxon Signed-Rank Test values. We found statistically significant differences ($p < 0.05$) between controllers for all sub-scales as well as for the overall questionnaire score (marked with *).**

| Igroup Presence Sub-scale | Z-value | p-value | Effect size (r) |
|---|---|---|---|
| Sense of being there | -2.66 | <0.008* | 0.36 |
| Spatial presence | -2.90 | <0.004* | 0.39 |
| Involvement | -3.28 | <0.002* | 0.44 |
| Experienced realism | -2.09 | <0.04* | 0.28 |
| Total score | -3.35 | <0.001* | 0.44 |

We found negative correlations between FPS expertise and the VIVE controller accuracy, as well as Xbox controller timing. The first correlation can possibly be explained by the fact that experienced participants have expectations of prototypical FPS games [Adams 2014] and therefore may have attempted to map their prior knowledge onto a new interaction device resulting in degraded (*i.e.* less accurate) performance. However, for the controller case, experienced participants are used to the interaction device, resulting in a better (*i.e.* faster) performance. Further, during the experiment debrief, three participants indicated that they in fact preferred the Xbox controller over the VIVE wand because they have "shaky hands," and thus found it easier to aim using the Xbox controller's joystick. Five other participants stated that they prefer the Xbox controller overall because they are used to it. This extends prior findings that have demonstrated that form factor of a controller influences a user's acceptance of it [Hinckley et al. 1997]; beyond form factor, we can say that in some cases, *preference* affects acceptance as well.

In general, most participants offered negative feedback about using the Xbox controller with a VR HMD. Four of these expressed discomfort due to the weight of the controller. More than ten noted – although they felt it did not affect them – that others who are not used to playing with gamepads will find it difficult to figure out the button positions. It is possible that discomfort due to having to remember button positions may be mediated by the individual's prior experience in playing games, although this will need to be verified in a subsequent study. Out of all participants, only one elected to use an inverted y-axis during the gamepad condition. We had only one participant who had no prior FPS game experience; they explained how it was much easier for them to use the VIVE wand because it "felt more natural to aim with." This (once again) suggests that the participant's knowledge is a potential determinant of their experience in VR.

Interaction fidelity played a role in self-reported presence, engagement and perceived accuracy. Our results suggest that participants had a higher feeling of engagement and presence when they were using the VIVE wand. Participants' feedback was also consistent with this finding; the majority stated that they felt it was more fun and natural to do the shooting task with the wand because it closely resembles shooting targets in real life; *i.e.* it was an NMC [McEwan et al. 2014; Skalski et al. 2011]. There was a significant difference between the two controllers for all sub-scales of presence (including the total score), with the highest difference being in the "sense of being there." This is consistent with prior findings that body-centered interactions elicit the strongest feeling of presence in the virtual world [Slater and Usoh 1994]. For self-reported accuracy, we found a significant difference between the VIVE wand and the Xbox controller in favor of the wand.

*Theoretical Implications.* These findings support the idea that people, overall, prefer and perform better when engaging in time and accuracy restricted activities in the virtual world using higher interaction fidelity interfaces. We only used one display device (the VIVE HMD), which is considered a high-fidelity display, for the entirety of this experiment to control for the display fidelity effect. Previous studies found that higher display fidelity leads to stronger feeling of presence [Buttussi and Chittaro 2018; Kim et al. 2014; McMahan et al. 2012]. Results from a study by McMahan et al. [2012] suggest that consistency between interaction and display fidelity results in better performance: low-interaction, low-display and high-interaction high-display settings can result in a better performance than a combination of a high-display, low-interaction or low-display, high-interaction settings. The authors attributed this to familiarity: people are used to playing on different types of consoles where low-display low-interaction settings are the default, but they did not test this conjecture directly with an input device typical of consoles (the joypad). In this work we have confirmed their intuition.

In summary, we found that changing the type of interaction interface had a significant impact on participants' sense of presence and engagement, as well as on their real and perceived performances. For tasks that require accuracy and/or speed, providing users with higher-fidelity equipment will result in better performance.

*Limitations.* For our study, we assessed participants' experience with VR and FPS games in general, and took that into account when analyzing the results. We did not directly assess their experience with the two input devices they used specifically, which might have had an effect on the results. Additionally, randomizing the order in which participants used input devices allowed us to balance ordering effects. Lastly, our study considered an FPS task as its experimental task to compare the effects of gamepads and NMCs on VR experiences. Hence, our findings are specifically applicable to this genre of virtual experiences.

Consistent with the findings in [Brown et al. 2010], nine participants mentioned sensitivity as a problem, and discussed how they felt that the Xbox controller's joystick sensitivity negatively affected their performance. These comments came from participants who had experienced with FPS games. Our experiment design intentionally controlled for interface consistency between the conditions, and so we designed our experiment to not vary in terms of joystick sensitivity. Future work may control for joystick sensitivity by ensuring that every participant's setting is within their **comfort zone** [Kölsch et al. 2003], but to our knowledge there exist no methods for measuring comfort with joypads in a manner that would afford controlling for this experimentally.

*Future Work.* Previous limitations point to avenues for future work. A good first step would be to explore the effects of other interfaces, in addition to the ones we used, on FPS and other virtual environment genres. There is a variety of controllers that were not addressed in our study and are used to play all kinds of genres, including FPS games. Developing the right environment where these controllers can be used and tested should provide a suitable means to measure relevant metrics, and eventually, help improve interface technology. Furthermore, a study similar to ours can be conducted, with means to assess participants' prior experience with the provided controllers, to evaluate if participants expertise with controllers had a significant effect on presence, engagement, or performance.

# 6 CONCLUSION

In this paper, we studied the role of interaction fidelity in a VR FPS virtual environment by contrasting the VR experience with the Microsoft Xbox 360 controller against the HTC VIVE wand. Based on the results of our work, both the type of interaction device used, as well as users' prior experience with FPS games had an effect on their VR experience. Aspects affected include task performance, presence, and engagement. Our results when manipulating interaction fidelity are inconsistent with previous related (and different) literature [Farmani and Teather 2017] in terms of real/perceived performance, but we remain consistent with respect to presence and engagement [Cummings and Bailenson 2016; McMahan et al. 2012]. The comparison we made is crucial to further our understanding of design and interaction interfaces in VR: while the VIVE wand seems to be preferred in terms of experience, the Xbox controller seems to be preferred in terms of familiarity. Designers of interaction devices for VR experiences would do well to reconcile these opposing forces: how might we retain the experience of one while remaining familiar to our expectations of the other?

The point of familiarity is related to the broader role that we feel expectations play in our experience of VR. We conjecture that we cannot cleanly separate interaction fidelity from scenario fidelity. In our case, we do not necessarily mean that our interaction conforms to expectations of how our non-virtual reality is (the original sense of scenario fidelity), but that our interaction conforms to expectations of how other VR experiences – in our case FPS experiences – are; perhaps better termed **genre fidelity**. This is supported by prior work that has looked at interaction fidelity *outside* of VR, which found that an interface's functionality evaluation depends on the virtual environment's portrayed setting [Brown et al. 2010].

Further, the VIVE wand makes sense in the context of our FPS virtual environment: it represents a naturally-mapped interface [McEwan et al. 2014]. However, in the case of other genres, this might not hold true; other genres should be investigated and our results should be interpreted with this in mind. The fact that the VIVE is naturally mapped onto VR experiences of a certain type (*e.g.* FPS games) begs the question as to whether the VIVE controllers are implicitly constraining the space of potential designs in virtual worlds that make sense in that platform. In other words, interaction fidelity might only be assessable if scenario and/or genre fidelity exceeds a certain threshold. It is *vital* to understand how much we can evaluate interaction fidelity independent from scenario fidelity in order to be able to address and design for individual genres effectively.

## REFERENCES


Ernest Adams. 2014. *Fundamentals of Game Design*. Pearson Education.

Kevin W. Arthur, Kellogg S. Booth, and Colin Ware. 1993. Evaluating 3D Task Performance for Fish Tank Virtual Worlds. *ACM Transactions on Information Systems* 11, 3 (jan 1993), 239–265.

Ravin Balakrishnan, Thomas Baudel, Gordon Kurtenbach, and George Fitzmaurice. 1997. The Rockin'Mouse: Integral 3D Manipulation on a Plane. In *Proc. CHI*. ACM, 311–318.

Woodrow Barfield, Claudia Hendrix, and Karl Bystrom. 1997. Visualizing the structure of virtual objects using head tracked stereoscopic displays. In *Proc. VR*. IEEE, Piscataway, NJ, USA, 114–120.

Cheryl Campanella Bracken and Paul Skalski. 2010. *Immersed in media: Telepresence in everyday life*. Routledge.

Michael Brown, Aidan Kehoe, Jurek Kirakowski, and Ian Pitt. 2010. Beyond the Gamepad: HCI and Game Controller Design and Evaluation. In *Evaluating User Experience in Games*, Regina Bernhaupt (Ed.). Springer, 209–219.

Fabio Buttussi and Luca Chittaro. 2018. Effects of Different Types of Virtual Reality Display on Presence and Learning in a Safety Training Scenario. *IEEE Transactions on Visualization and Computer Graphics* 24, 2 (2018), 1063–1076.

Sarah S. Chance, Florence Gaunet, Andrew C. Beall, and Jack M. Loomis. 1998. Locomotion Mode Affects the Updating of Objects Encountered During Travel: The Contribution of Vestibular and Proprioceptive Inputs to Path Integration. *Presence: Teleoperators and Virtual Environments* 7, 2 (apr 1998), 168–178.

James C. Chung. 1992. A comparison of head-tracked and non-head-tracked steering modes in the targeting of radiotherapy treatment beams. In *Proc. I3D*. ACM, New York, NY, USA, 193–196.

Jacob Cohen. 1988. *Statistical Power Analysis for the Behavioral Sciences* (2$^{\text{nd}}$ ed.). Erlbaum Associates, Hillsdale, Routledge.

Jacob Cohen. 1992. A Power Primer. *Psychological Bulletin* 112, 1 (1992), 155–159.

Alastair H. Cummings. 2007. The Evolution of Game Controllers and Control Schemes and their Effect on their Games. In *Proc. of the 17th Annual University of Southampton Multimedia Systems Conference*.

James J Cummings and Jeremy N Bailenson. 2016. How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology* 19, 2 (2016), 272–309.

Yasin Farmani and Robert J Teather. 2017. Player performance with different input devices in virtual reality first-person shooter games. In *Proceedings of the 5th symposium on spatial user interaction*. ACM, 165–165.

Richard M. Heiberger and Naomi B. Robbins. 2014. Design of Diverging Stacked Bar Charts for Likert Scales and Other Applications. *Journal of Statistical Software* 57, 5 (2014), 1–32.

Edgar Heineken and Frank P. Schulte. 2007. Seeing Size and Feeling Weight: The Size-Weight Illusion in Natural and Virtual Reality. *Human Factors* 49, 1 (2007), 136–144.

Ken Hinckley, Joe Tullio, Randy Pausch, Dennis Proffitt, and Neal Kassell. 1997. Usability Analysis of 3D Rotation Techniques. In *Proc. UIST*. ACM, New York, NY, USA, 1–10.

Kwanguk Kim, M. Zachary Rosenthal, David J. Zielinski, and Rachael Brady. 2014. Effects of virtual environment platforms on emotional responses. *Computer Methods and Programs in Biomedicine* 113, 3 (mar 2014), 882–893.

Silvia Erika Kober, Jürgen Kurzmann, and Christa Neuper. 2012. Cortical correlate of spatial presence in 2D and 3D interactive virtual reality: An EEG study. *International Journal of Psychophysiology* 83, 3 (mar 2012), 365–374.

Mathias Kölsch, Andrew C. Beall, and Matthew Turk. 2003. An Objective Measure for Postural Comfort. In *Proc. HFES*. 725–728.

Joseph J. LaViola, Jr. and Tad Litwiller. 2011. Evaluating the Benefits of 3D Stereo in Modern Video Games. In *Proc. CHI*. ACM, New York, NY, USA, 2345–2354.

Jean-Luc Lugrin, Marc Cavazza, Fred Charles, Marc Le Renard, Jonathan Freeman, and Jane Lessiter. 2013. Immersive FPS Games: User Experience and Performance. In *Proc. ImmersiveMe*. ACM, New York, NY, USA, 7–12.

Mitchell W. McEwan, Alethea L. Blackler, Daniel M. Johnson, and Peta A. Wyeth. 2014. Natural mapping and intuitive interaction in videogames. In *Proc. CHIPlay*. ACM, New York, NY, USA, 191–200.

Ryan P. McMahan, Doug A. Bowman, David J. Zielinski, and Rachael B. Brady. 2012. Evaluating Display Fidelity and Interaction Fidelity in a Virtual Reality Game. *IEEE Transactions on Visualization and Computer Graphics* 18, 4 (apr 2012), 626–633.

Michael Narayan, Leo Waugh, Xiaoyu Zhang, Pradyut Bafna, and Doug Bowman. 2005. Quantifying the Benefits of Immersion for Collaboration in Virtual Environments. In *Proc. VRST*. ACM, New York, NY, USA, 78–81.

Kevin Ng. [n.d.]. 'First-Person Control Design for Dual Analog Stick Controllers'. *Gamasutra Article*. 14-Jul-2004 [Online] Available: http://www.gamasutra.com/view/feature/2119/firstperson_control_design_for_.php [Accessed: 23-Nov-2018].

Donald A. Norman. 1999. Affordance, Conventions, and Design. *Interactions* 6, 3 (may/jun 1999), 38–43.

Michael Pereira, Ferran Argelaguet, José del R Millán, and Anatole Lécuyer. 2018. Novice shooters with lower pre-shooting alpha power have better performance during competition in a virtual reality scenario. *Frontiers in psychology* 9 (2018), 527.

John Polson. [n.d.]. First-person shooter design: What to save, and what to frag. *Gamasutra Article*. 22-May-2012 [Online] Available: https://www.gamasutra.com/view/news/170721/Firstperson_shooter_design_What_to_save_and_what_to_frag.php [Accessed: 23-Nov-2018].

Eric D. Ragan, Doug A. Bowman, Regis Kopper, Cheryl Stinson, Siroberto Scerbo, and Ryan P. McMahan. 2015. Effects of field of view and visual complexity on virtual reality training effectiveness for a visual scanning task. *IEEE Transactions on Visualization and Computer Graphics* 21, 7 (feb 2015), 794–807.

Dheva Raja. 2006. *The Effects of Immersion on 3D Information Visualization*. Ph.D. Dissertation. Virginia Polytechnic Institute and State University.

Robert Rosenthal. 1994. Parametric Measures of Effect Size. In *The Handbook of Research Synthesis*, Harris Cooper and Larry V. Hedges (Eds.). Russell Sage Foundation, 231–244.

Johann Schmitt, Carolin Wienrich, Marc Erich Latoschik, and Jean-Luc Lugrin. 2019. Investigating Gesture-based Commands for First-Person Shooter Games in Virtual Reality. *Mensch und Computer 2019-Workshopband* (2019).

Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. 2001. The Experience of Presence: Factor Analytic Insights. *Presence: Teleoperators and Virtual Environments* 10, 3 (jun 2001), 266–281.

Paul Skalski, Ron Tamborini, Ashleigh Shelton, Michael Buncher, and Pete Lindmark. 2011. Mapping the road to fun: Natural video game controllers, presence, and game enjoyment. *New Media & Society* 13, 2 (2011), 224–242.

Mel Slater and Martin Usoh. 1994. Body Centred Interaction in Immersive Virtual Environments. In *Artificial Life and Virtual Reality*, Nadia Magnenat Thalmann and Daniel Thalmann (Eds.). John Wiley and Sons Ltd., 125–148.

Marco Torchiano. 2016. Effsize - a package for efficient effect size computation. https://doi.org/10.5281/zenodo.196082

Martin Usoh, Kevin Arthur, Mary C. Whitton, Rui Bastos, Anthony Steed, Mel Slater, and Frederick P. Brooks Jr. 1999. Walking > Walking-in-Place > Flying, in Virtual Environments. In *Proc. SIGGRAPH*. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 359–364.

Colin Ware, Kevin Arthur, and Kellogg S. Booth. 1993. Fish Tank Virtual Reality. In *Proc. CHI/INTERACT*. ACM, New York, NY, USA, 37–42.

Colin Ware and Ravin Balakrishnan. 1994. Reaching for Objects in VR Displays: Lag and Frame Rate. *ACM Transactions on Computer-Human Interaction* 1, 4 (dec 1994), 331–356.

Colin Ware and Glenn Franck. 1996. Evaluating Stereo and Motion Cues for Visualizing Information Nets in Three Dimensions. *ACM Transactions on Graphics* 15, 2 (apr 1996), 121–140.

Catherine A. Zanbaka, Benjamin C. Lok, Sabarish V. Babu, Amy C. Ulinski, and Larry F. Hodges. 2005. Comparison of path visualizations and cognitive measures relative to travel technique in a virtual environment. *IEEE Transactions on Visualization and Computer Graphics* 11, 6 (nov/dec 2005), 694–705.

Shumin Zhai, Paul Milgram, and William Buxton. 1996. The Influence of Muscle Groups on Performance of Multiple Degree-of-Freedom Input. In *Proc. CHI*. ACM, New York, NY, USA, 308–315.

David Zielinski, Brendan Macdonald, and Regis Kopper. 2014. Comparative study of input devices for a VR mine simulation. In *2014 IEEE Virtual Reality (VR)*. IEEE, 125–126.