Designing Coherent Interactions for Virtual Reality

Run Yu

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Computer Science and Applications

Doug A. Bowman. Chair
Hrvoje Benko
Christopher L. North
Denis Gracanin
Nicholas F. Polys

7/17/2019
Blacksburg, VA

Keywords: Virtual reality, plausibility, coherence, user experience,
travel, object manipulation

Designing Coherent Interactions for Virtual Reality

Run Yu

ABSTRACT

Coherence describes the validity of the internal rules that drive the behaviors of a virtual environment (VE) in presenting a credible scenario. A VR system with a high level of coherence could lead to strong plausibility illusion, which is a key component of the sense of presence. There are few existing studies centered around coherence, and they tend to put the user in a passive role when experiencing the VE without emphasizing on their active participation in the interaction. This dissertation makes up this gap by connecting the concept of coherence with fundamental 3D user interface design that focuses on the algorithms that map the user's actions to the VE's behaviors. Specifically, we inspect the design of coherent interactions for two complicated tasks, namely travel and object manipulation. For travel, we propose a family of redirected walking techniques called "narrative driven cell-based redirection", which lets the user traverse a VE that's much larger than the physical space without breaking the coherence of the scenario. For object manipulation, we propose the novel concept of physics coherence to capture whether an interface conforms to the rules of physics and design several novel techniques that try to balance between physics coherence and usability. Together, we provide some useful tools for designing coherent interactions and discuss how coherence affects user experience in VR interaction.

Designing Coherent Interactions for Virtual Reality

Run Yu

GENERAL AUDIENCE ABSTRACT

To create a virtual reality (VR) experience that feels plausible, it's important to consider the validity of the internal rules that drive the behaviors of the virtual environment (VE), which we call "coherence" of a VR system. We discuss how to support coherence in two types of fundamental VR interaction. The first one is travel, which concerns moving the viewpoint around following the user's intention. For this task, we propose a family of novel interaction techniques called "narrative driven cell-based redirection", which lets the user traverse a VE that's much larger than the physical space without breaking the coherence of the scenario. The second one is object manipulation, which is about controlling a virtual object using hand input. For this task, we propose the novel concept of physics coherence to capture whether the interaction conforms to the rules of physics and design several novel techniques that try to balance between physics coherence and controllability. Together, we provide some useful tools for designing coherent interactions and discuss how coherence affects user experience in VR interaction.

Acknowledgement

To my advisor, Dr. Doug A. Bowman, who is a role model and my hero. I learned so much more than research from you. From daily routines like how to talk to another, to bigger matters like how to be a righteous, responsible and caring leader, everyday you teach me something new not by teaching. Thank you for providing such a supportive environment for your students. I'm truly lucky to have a mentor like you, and the days working with you will always make me smile when I look back.

To my committee, Dr. Benko, Dr. Gracanin, Dr. North and Dr. Polys, thank you for believing in me, encouraging me at every milestone and supporting me throughout my (long) PhD program, and thank you for creating the amazing HCI community.

To Dr. Ben Knapp, Dr. Todd Ogle, Dr. Andrea Kavanaugh, Tanner Upthegrove, Mellisa Wyers, Holly Williams, Elizabeth Scharman and other people at ICAT and CHCI, thank you for never giving up on me and helping me till my last day at Virginia Tech. It would be impossible for me to achieve anything without your generous support.

To Dr. Yong Cao, thank you for bringing me to Virginia Tech and helping me start this incredible journey.

To my lab mates at 3DI and CHCI, your company helped me through the hard days.

To Justin Clark, Dr. Sean Keller and people at Facebook Reality Lab, thank you for sharing a great summer with me and all the support you gave.

And I want to thank my parents for the unconditional trust, love and support.

Attribution

Dr. Doug A. Bowman is the Frank J. Maher Professor of Computer Science and the Director for the Center for Human-Computer Interaction (CHCI) at Virginia Tech. He is my PhD advisor and the co-author of all the materials forming this dissertation.

Section 3.1 (except Section 3.1.5) originally appeared in ***Yu, R.***, *Lages, W.S., Nabiyouni, M., Ray, B., Kondur, N., Chandrashekar, V., et al. (2017). "Bookshelf and Bird: Enabling Real Walking in Large VR Spaces through Cell-Based Redirection", in IEEE Symposium on 3D User Interfaces. 116-119. (https://ieeexplore.ieee.org/document/7893327*)

Section 3.2 originally appeared in ***Yu, R.***, *Duer, Z., Ogle, T., Bowman, D.A., Tucker, T., Hicks, D., et al. (2018). "Experiencing an Invisible World War I Battlefield Through Narrative-Driven Redirected Walking in Virtual Reality", in 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). 313-319. (https://ieeexplore.ieee.org/document/8448288*)

Section 4.2.1.4 and Section 4.2.2 originally appeared in the poster, ***Yu, R.***, *and Bowman, D.A. (2018). "Force Push: Exploring Expressive Gesture-to-Force Mappings for Indirect 3D Object Manipulation", in 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). 733-734. (https://ieeexplore.ieee.org/document/8446453).* It also appeared in the journal paper ***Yu, R.***, *and Bowman, D.A. (2018). "Force Push: Exploring Expressive Gesture-to-Force Mappings for Remote Object Manipulation in Virtual Reality". Frontiers in ICT 5, 25. (https://www.frontiersin.org/articles/10.3389/fict.2018.00025/full*)

Emmet Hobgood was an undergraduate student conducting research at the 3D Interaction (3DI) group at Virginia Tech. He conducted the study in Section 3.1.5 and directly interacted with the participants under the guidance of our advisor Dr. Doug A. Bowman. This study is used in this dissertation with the permission from Dr. Doug A. Bowman. Please refer to the permission letter in Appendix A.1.3.

Contents

List of Figures

List of Tables

1. INTRODUCTION

1.1 Motivation

The sense of presence is one of the key outcomes of virtual reality (VR) and there has been a tremendous amount of work in understanding and creating this sensation. It is believed that the capability of creating a strong sense of presence provides unique opportunities for VR applications such as psychological treatment, training and entertainment (Skarbez et al., 2017b). To understand and dissect this important concept, Mel Slater proposed a theoretical framework that breaks presence into two orthogonal components: Place Illusion (PI) and Plausibility Illusion (Psi) (Slater, 2009). On top of this, previous research also identifies the two corresponding characteristics of a VR system that lead to these illusions, namely Immersion and Coherence (Skarbez, 2016). We start this dissertation by introducing these concepts and the relationship between them.

Among the two components of presence, PI describes the "sense of being there" (Slater, 2009), "the qualia (individual instances of subjective, conscious experience) of having a sensation of being in a real place" (Slater, 2009), which is similar to the traditional definition of overall presence in telepresence systems (Minsky, 1980). PI is a subjective psychological response to what is presented to the user, and the objective characteristic of this presentation that leads to such illusion is called "immersion", namely the "valid actions that are possible within the system." Immersion is constrained by "the sensorimotor contingencies afforded by the virtual reality system". For example, a display device that has head tracking capability could support more valid actions (turning and moving the head to naturally look at the virtual world from different perspectives) than one that doesn't, so the former one provides more immersion. Hence, immersion depicts an objective capability of the system usually described by hardware/software parameters (e.g. field of view, resolution, tracking accuracy and range) (Slater, 2009).

Psi, on the other hand, describes how reasonable the scenario feels to the user. It is determined by "the extent to which the system can produce events that directly relate to the participant, the overall credibility of the scenario being depicted in comparison with expectations." (Slater, 2009). Just like immersion to PI, Psi is supported by the "coherence" of the VR application, as an objective criterion that measures the consistency and validity of what drives the behaviors of a virtual world (Skarbez, 2016; Skarbez et al., 2017a; Skarbez et al., 2017c). For example, a virtual environment (VE) with characters talking with you in a socially appropriate manner is likely to be more coherent than one that's with lifeless characters, which could in turn results in stronger Psi.

We can use these two components of presence (and the corresponding two system characteristics) to inspect the design and engineering of a system. For example, a VR system could have a very high level of immersion by using high-resolution display and accurate full-body tracking, but a very low level of coherence by showing an unreasonable scenario in which objects are teleporting around. This could lead to a strong place illusion but a weak plausibility illusion. The opposite of this case could also happen, and a typical example is a book (or text adventure)—

the scenario and story could be highly coherent, which leads to a high level of plausibility illusion. But text is not an immersive media, so one doesn’t experience place illusion.

There has been a lot of effort in enhancing immersion and place illusion, but much less work focusing on VR coherence and plausibility illusion (Skarbez, 2016). On top of this, the few existing studies mostly treat coherence as how the environment is presented to the user instead of emphasizing one’s active participation. To investigate coherence, the experiments usually put the user in a passive role to “view” the environment under different conditions, and how the user’s input could affect the environment’s action is often not the major focus. For example, in the recent work by Skarbez et al. that inspects the components of plausibility illusion, four coherence characteristics were included: the behavior of other virtual human, the appearance of the environment, the avatar of the user’s body, and how a ball would react to the user’s kicking it (Skarbez et al., 2017c). Among these four, only the last one emphasizes how the environment reacts to the user’s active participation, and only three simple conditions were created to vary its level of coherence. The one for low coherence would make the ball follow the foot when they are in contact but “does not roll or maintain momentum once out of contact”; the one for high coherence would simulate its movement driven by inertia using a physics engine; there is also a third condition with mid-level coherence, which makes the ball randomly behave like one of the previous two conditions with equal likelihood. These three conditions are very limited and can hardly capture the nuances in designing the interaction with the ball. We feel there is a missing link between coherence and fundamental interaction design that emphasizes the user’s participation.

This dissertation intends to make up this gap by investigating how to *design for coherence* in complex, active interaction by the user in the virtual environment. We specifically focus on two complex and fundamental interaction tasks, namely travel and object manipulation (Bowman et al., 2004). The reason we picked walking and manipulation out of the five fundamental 3D interaction tasks (selection, manipulation, travel, system control and symbolic input) (Bowman et al., 2004) is that they are “physical” by nature, and such nature makes them particularly valuable to our exploration in connecting coherence to 3D UI design. On the one hand, they have real-world roots, which intuitively gives some basis in how coherence could be created (e.g., physical rules for object manipulation). On the other hand, they are complicated in the temporal dimension, which naturally demands the design of complex rules to drive the VE’s behaviors (e.g., redirected walking algorithms and transfer functions in manipulation). As we’ll see, creating coherent interactions for either task is not trivial and requires careful design.

1.1.1 Coherent Redirected Walking

In the case of travel, real walking is a highly coherent interaction technique by nature. It’s what we are used to do in real life, and previous research shows that it could enhance the sense of presence (Usoh et al., 1999). However, it’s not always feasible to directly use walking in a real VR application, since the presented virtual environment is usually larger than the available physical play area. Various variations of redirected walking (RDW) techniques have been proposed to address this issue, with the idea of making the user walk within the smaller physical space while presenting the illusion that she is traversing a much larger virtual world (Langbehn and Steinicke, 2018). These methods rely on manipulating the mapping between the virtual and real space, and such manipulation can only be hidden from the user when the physical space is large enough (Steinicke et al., 2010; Grechkin et al., 2016).

For room-scale VR that only has a play area with the size of a living room, it's almost inevitable to use explicit, overt manipulation to reconstruct the user's walking path when she approaches the real-world boundaries in order to make sure her movement is confined in the physical area (Suma et al., 2012). This manipulation needs to be carefully designed for it to feel reasonable in the presented scenario, so that the user's perception of the virtual world is kept consistent with what they would expect. Otherwise the coherence of the experience could be fundamentally broken. For example, in the experiment presented by Williams et al., their locomotion techniques insert reset events that freeze the visual display and ask the user to turn around (Williams et al., 2007). This brute force approach breaks the narrative of the experience without an appropriate reason, which is less coherent than the technique by Peck et al., which uses a visual distraction that's part of the VE to keep the experience seamless while changing the virtual environment in the background (Peck et al., 2009).

This dissertation aims to expand on the effort of maintaining coherence in RDW for room-scale VR. We propose a family of techniques called "narrative driven cell-based redirection (CBR)", which turns the manipulation of redirection into built-in features of the VE and narrative itself. We'll show several implementations of this idea, and how it's applied to a real-world educational VR application.

1.1.2 Physically Coherent Object Manipulation

As for object manipulation, current VR technology cannot fully replicate the real-world interaction yet. Different from walking, which is straightforward to implement (to the extent that's within the range of the physical area), an authentic re-creation of how we manipulate an object in reality would require extremely accurate tracking of individual fingers, an advanced physics engine that supports simulation of friction and deformation, and a high-fidelity haptic rendering system to display subtle forces. To our knowledge, such a system still doesn't exist to date.

Even if we could perfectly replicate real-world object manipulation, we have plenty of reasons not to do so, mainly because it could be desirable to go beyond reality. One example is that we want to facilitate remote control in many situations—if moving around in VR is already challenging, it is desirable for the user to be able to manipulate the object without necessarily walking towards it first, especially when the VE is larger than the physical space. Besides that, we may also want the user to be able to move large, heavy items that would be difficult to hold in reality.

Lacking an authentic replication of real-world interaction while intending to perform more than that, the most common existing interfaces for object manipulation use a "simple virtual hand" metaphor. The user holds a physical proxy as a representation of the virtual object, and a zero-order mapping is applied between the proxy's movement and the virtual object's movement. Such a design is easy to understand and results in high usability. It could also be easily extended to remote control—the user remotely “grabs and drags” an object just like how we move an icon in a desktop environment using a mouse.

However, despite its high usability, the coherence of interaction is potentially broken when judged by the rules of physics, which fundamentally govern the process of object manipulation. For example, a large and heavy item wouldn’t go through the proper acceleration/deceleration according to its weight. Instead, it would demonstrate “jumpy” movement as if it is as heavy as the hand-held proxy. When the user perceives that an object that’s supposed to be heavy feels much lighter than its proper weight, the physical coherence of the interaction is compromised, which could in turn negatively affect the overall plausibility illusion.

It appears that the design of object manipulation is torn by two conflicting requirements regarding physics. On one hand, it seems inevitable to break some aspects of physical rules because of technological limitations and the desire to enable interactions beyond reality; on the other hand, conforming to physical rules seems to be an important aspect of the coherence for VR object manipulation that could affect user experience. This dissertation proposes the concepts of *physics coherence* and *physics plausibility* to capture the nuances in such conflict. *Physics coherence* describes the level of compliance to physical laws in a manipulation technique, and *physics plausibility* describes the corresponding psychological response to the physical motion and behavior presented by the system—the perception of whether the manipulation is physically believable or not. Based on these concepts, we investigate how physics coherence could affect user experience, and how to design novel interaction techniques that preserve the core of physics coherence while enabling some desirable capabilities in control. We hope the effort here can serve as a “toolbox” for physically coherent interactions, from which the readers can draw inspirations for the design of future VR object manipulation techniques.

1.2 Definitions

To facilitate the discussion about coherent interaction, related fundamental concepts that describe the VR system, the interface and the user experience are defined as follows:

User Interface (UI): The medium through which the communication between users and computers takes place (Hix and Hartson, 1993; Bowman et al., 2004).

User experience: A person’s perceptions and responses resulting from the use and or anticipated use of a product, system or service (Mirnig et al., 2015).

Virtual Environment (VE): A synthetic, spatial (usually 3D) world seen from a first-person point of view. The view in a VE is under the real-time control of the user (Bowman et al., 2004).

Virtual reality (VR): Synonymous with VE (Bowman et al., 2004).

3D interaction: Human-computer interaction in which the user’s tasks are performed directly in a 3D spatial context (Bowman et al., 2004).

3D user interface: a UI that involves 3D interaction.

Navigation: the fundamental human task of movement in and around an environment. Navigation includes both Travel and Wayfinding (Bowman et al., 2004).

Travel: the motor component of navigation, namely the low-level actions that control the position and orientation of the viewpoint, which moves us from our current location to a new target location or in the desired direction (Bowman et al., 2004).

Wayfinding: the cognitive component of navigation, namely the high-level thinking, planning, and decision making related to user movement (Bowman et al., 2004).

Object Manipulation: The act of handling an object that maps the user's input into the desired actions of the object (Bowman et al., 2004).

Place illusion (PI): "The sense of being there", the illusion that you are placed in the location depicted by the VE in spite the sure knowledge that you are not (Slater, 2009).

Immersion: A property of the VR system that concerns the set of valid actions that are possible within the system. Immersion provides the boundaries within which place illusion can occur (Slater, 2009).

Plausibility illusion (Psi): Psi is the illusion that what is apparently happening is really happing (even though you know for sure that it is not) (Slater, 2009).

Coherence: Parallel to immersion, coherence of a virtual scenario is defined as the set of reasonable circumstances that can be demonstrated by the scenario without introducing unreasonable circumstances. "Reasonable circumstance" is defined as a state of affairs in a virtual scenario that is self-evident given prior knowledge. "Unreasonable circumstance" is the opposite, states of affairs that are inconsistent with prior knowledge" (Skarbez, 2016). Coherence is treated as the system property that supports plausibility illusion.

Embodiment: Embodiment refers to the representation of a user (also known as an avatar) within a mediated or virtual environment (Skarbez et al., 2017b).

Interaction fidelity: The objective degree of exactness with which a system reproduces real world interactions (action realism) (McMahan, 2011).

Pragmatic and hedonic attributes: Two distinct groups of attributes that describe how users perceive an interactive product, each one is defined as follows (Hassenzahl, 2003; 2004):

Pragmatic attributes describe relevant functionality (i.e., utility) and ways to access this functionality (i.e., usability) in order to manipulate the environment. They are connected to the users' need to achieve behavioral goals. Typical pragmatic attributes of software products are "clear", "supporting", "useful" and "controllable". A product that allows for effective and efficient goal-achievement is perceived as pragmatic (Hassenzahl, 2003; 2004).

All other attributes are categorized as hedonic attributes. They are primarily related to the users' self, in that they provide stimulation, communicate identity, and provoke valued memories. Typical hedonic attributes of an interactive product are "outstanding", "impressive", "exciting" and "interesting". A product that can be perceived as hedonic because it provides stimulation by its challenging and novel character or identification by communicating important personal values to relevant others (Hassenzahl, 2003; 2004).

1.3 Overview

1.3.1 Designing Coherent Redirected Walking

1.3.1.1 Problem Statement

Redirected walking for room-scale VR often relies on discrete reset events to confine the user's movement. Introducing these events without a reason that's appropriate for the scenario could break the coherence of the VE. In this dissertation, we look for creative solutions that could seamlessly reset the user's walking path without breaking the narrative of the VR experience.

1.3.1.2 Research Questions (RQs)

RQ1: How can we design reset-based RDW techniques that maintain the coherence of the virtual scenario?

RQ2: Do these novel, coherent RDW techniques provide measurable benefit for user experience?

1.3.1.3 Approach

We introduce a family of techniques called "narrative driven cell-based redirection (CBR)", which attempts to turn the redirection events into built-in features of the VE and narrative. Using this approach, the manipulation of redirection becomes an inherent mechanism of the VR experience itself, so the overall coherence is maintained.

We'll use two concrete examples to introduce this approach. The first one is called "bookshelf and bird", which uses overt, user-initiated redirection events but present it as internal functionalities of the VE. To address RQ2, an experiment was conducted to measure the user's spatial understanding when using these techniques in comparison to their less-coherent counterparts. The second example is called "lighting-driven redirection", which attempts to make the redirection less noticeable by hiding it behind the features of the scenario.

(1) Bookshelf and Bird (Section 3.1) (Yu et al., 2017)

The techniques, called Bookshelf and Bird, provide narrative-consistent redirection to keep the user inside the physical space, and require the user to walk to explore the VE. Both of these are explicit redirection methods that maintain a consistent mapping between physical and virtual spaces after the redirection.

The virtual world is first divided into discrete cells that have the same size as the physical tracking space. The techniques then can redirect the users without altering the relationship between the user, the physical space, and the virtual cell. Bookshelf allows the user to go from one cell to the next by virtually rotating her on a spinning bookcase. After the virtual rotation, the destination room is on the same side as the physically available space. The Bird technique translates the user to another cell by lifting him over ground obstacles in the virtual world. It automatically drops him in a virtual location corresponding to his location in the physical space. These two techniques guarantee that the current cell is always mapped exactly to the physical workspace, keeping the entire cell accessible via real walking.

(2) Lighting-Driven Redirection (Section 3.2) (Yu et al., 2018)

We built an educational virtual reality experience in which users physically walk through virtual tunnels representative of the World War I battle of Vauquois. Walking in only a 15- by 5-foot tracked space, users are redirected through subtle, narrative-driven resets to walk through a tunnel nearly 50 feet in length. The reset events use mechanisms that are naturally embedded in the VE and narrative, so that they not only allow the user to continue walking in the virtual environment but also enhance the experience itself. This work contributes approaches and lessons that can be used to provide a seamless and natural virtual reality walking experience in highly constrained physical spaces.

1.3.1.4 Contributions

- The concept of "narrative driven cell-based redirection", which is a powerful design tool to facilitate real walking in VR given limited physical space
- Several example techniques that implement this idea, with one of them being successfully applied to a real-world educational VR experience
- Evaluation of how these coherent RDW techniques could potentially influence spatial understanding, suggesting promising future work

1.3.2 Designing Physically Coherent Object Manipulation

1.3.2.1 Problem Statement

Existing VR manipulation techniques usually break coherence with respect to physical laws. It is unclear how such deviation from physical principles affects user experience, and there is a lack of knowledge about how to design techniques that maintain this aspect of coherence. We intend to provide theories, empirical evidence, and design knowledge to address these questions.

1.3.2.2 Research Questions (RQs)

RQ3: How can we design physically coherent interactions without replicating the physical contact between the hand and the object?

RQ4: How does physics coherence in object manipulation affect user experience?

1.3.2.3 Approach

(1) The Theoretical Framework (Section 4.1)

We start by introducing the concepts of physics coherence and physics plausibility. An operational model for constructing these properties is also presented, emphasizing the display of correct motion and force on individual sensory channels and the inter-channel consistency between them.

(2) Designing and Evaluating Physically-Plausible Techniques (Section 4.2 and 4.3)

We then explore the targeted design space by proposing two physically coherent techniques for remote manipulation. The first one is "physically-coherent virtual hand". It extends the simple virtual hand control mechanism with physics-based behaviors. The second one is called Force Push. It uses a novel metaphor that's inspired by telekinetic abilities in popular culture and also derived from real-world physical phenomena. These two techniques were analyzed with the proposed operational model to show how physics coherence is constructed from several aspects.

Then we carry out two sets of comparative studies. The first focuses on comparing the task performance between the simple virtual hand based control mechanism and the novel input-to-output mapping algorithm of Force Push (Yu and Bowman, 2018a; b). The second set compares the full-scale subjective user experience of the two proposed techniques with a naïve virtual hand based technique that is designed to have little physics plausibility. Through the comparative evaluation of these techniques, we find empirical evidence to support the importance of physics coherence and useful clues about how these techniques should be designed to exhibit desired qualities. We summarize what we learned in the experiment into a set of design guidelines, focusing on physics-driven motion, physically based correlation between visual movement and kinesthetic cues, and the application of relatable metaphors. On top of this, we propose the design of two novel techniques based on the guidelines.

From there, we continue to explore the design space but focus on introducing physically coherent features to the simple virtual hand based control mechanism. We propose several pseudo-haptic techniques to add physics coherence to rotational motion and demonstrate how these techniques could successfully induce a sensation of mass, which is a direct implication of perceived physics plausibility.

1.3.2.4 Contributions

- The concepts of physics coherence and physics plausibility, along with an operational model that describes the requirements for achieving physics coherence through multi-sensory display
- Empirical evidence that physics coherence has a significant positive influence on hedonic user experience
- Design guidelines for creating physically coherent remote manipulation techniques, which focus on leveraging physics-driven motion, kinesthetic cues and relatable metaphors
- Empirical evidence that physically coherent rotation techniques can produce effective illusions of mass and mass distribution
- The design of several novel physically coherent manipulation techniques
    - Physically coherent virtual hand adds physics coherence to simple virtual hand based control without breaking the superior controllability of zero-order mapping
    - Force Push uses rich features of dynamic gestures to allow users to naturally control complex object movements, which leads to a novel, controllable gesture-to-force mapping algorithm for remote object manipulation
    - Tennis controller and Slingshot controller execute the design guidelines we summarized from the experiment, which could potentially lead to high hedonic values

- Pseudo-haptic rotational control techniques

2. RELATED WORK

2.1 Presence, Plausibility Illusion and Coherence

Skarbez et al. recently provided a thorough survey of presence and related concepts, and this section follows this paper to review the related literature (Skarbez et al., 2017b).

2.1.1 Defining Presence

According to Steuer, the term “presence” was originally defined by psychologist James Gibson as follows (Steuer, 1992):

“Presence can be thought of as the experience of one’s physical environment; it refers not to one’s surroundings as they exist in the physical world, but to the perception of those surroundings as mediated by both automatic and controlled mental processes.” (Whitehead, 1981)

This original appearance defined presence as a subjective feeling in the real world. With the advancement of virtual reality systems, this term has been extensively used in this specific context to describe the key character of a VR experience. Many have defined presence in their own terms and here we don’t intend to exhaustively review them all. We follow the taxonomy by Skarbez et al. and introduce some of the representative definitions (Skarbez et al., 2017b).

Skarbez et al. categorized the various definitions of presence into three groups, namely “being there”, “non-mediation” and “other” (Skarbez et al., 2017b).

2.1.1.1 “Being There”

This group of definitions are deeply rooted in the concept of “telepresence”. Steuer first applied Gibson’s original definition of presence into the context of human-computer interaction. He defined telepresence as “the experience of presence (in the sense of Gibson) in an environment by means of communication medium” (Steuer, 1992). Telepresence was also explained by Akin et al. in the context of remote manipulation, as “at the worksite, the manipulators have the dexterity to allow the operator to perform normal human functions. At the control station, the operator receives sufficient quantity and quality of sensory feedback to provide a feeling of actual presence at the worksite.” (Akin et al., 1983).

Many definitions emphasize on the human actions that can be afforded during such experiences. For example, Zahorik and Jenison described presence as ““tantamount to successfully supported action in the environment” (Zahorik and Jenison, 1998). Similarly, Carassa et al. defined presence using situated cognition theory, as “presence depends on the proper integration of aspects relevant to an agent’s movement and perception, to her actions, and to her conception of the overall situation in which she finds herself, as well as on how these aspects mesh with the possibilities for action afforded in the interaction with the virtual environment.” (Carassa et al., 2005) We believe that the “immersion – place illusion” model from Slater et al. largely captures the same idea, as the “sense of being there” (place illusion) is supported by the valid actions that the system could afford (immersion) (Slater, 2009).

Some other definitions describe presence more passively without specifically addressing the user's actions. Witmer and Singer defined presence as a psychological state of "being there mediated by an environment that engages our senses, captures our attention, and fosters our active involvement." (Witmer and Singer, 1998) Spagnolli and Gamberini stated that "Whenever a person is qualified as 'present'...her location is the salient, characterizing feature." (Spagnolli and Gamberini, 2004)

In this proposal, we rely on the model from Slater et al. and treat the "sense of being there" as place illusion, which is one component that constructs presence instead of the overall definition of presence (Slater, 2009). This model is explained in section 1.1.

2.1.1.2 "Non-mediation"

Lombard et al. defined presence as the "the perceptual illusion of nonmediation" (Lombard et al., 2000). It specifically emphasizes on the illusion that the existence of media disappears to the user in an experience that's actually mediated. We believe that this definition is in accordance with the seamless quality of interaction that we are pursuing, and this is the definition we adopt throughout this document.

Related to this definition, Slater et al. interpreted presence as a "suspension of disbelief that one is in a world other than where one's body is located" (Slater and Usoh, 1993). This explanation describes a similar phenomenon—the user is fully aware that she is in a mediated environment, but somehow the illusion of presence overcomes that awareness and creates a belief (or suspension of disbelief) that she is in a non-mediated "real" world.

2.1.1.3 "Other"

In Skarbez et al.'s survey, he also briefly reviewed some alternative definitions of presence and categorized them as "other" (Skarbez et al., 2017b). Here we introduce two such interpretations that are strongly connected to our vision of seamless interaction. Kwan Min Lee puts an emphasis on the objects in the world and describes presence as "a psychological state in which virtual (para-authentic or artificial) objects are experienced as actual objects in either sensory or nonsensory ways" (Lee, 2004). We believe this definition echoes with the seamlessness illusion in object manipulation, as the user would control the objects in a natural and unobtrusive way. Parola et al. defined presence as "the sense of feeling real" and specifically emphasizes that the stimuli should meet the user's expectation and stay consistent, which is in accordance with the concept of plausibility illusion and coherence of the virtual world (Slater, 2009; Parola et al., 2016).

2.1.2 Measuring Presence

Methods of measuring presence can be categorized into three groups, namely "self-report, behavioral, and physiological" (Skarbez et al., 2017b). Here we review some representative examples for each category.

2.1.2.1 Self-Report

This type of measurements asks the user to actively report their experience. The most common approach is to use post-experiment questionnaires and here we briefly introduce the five "canonical" ones (Rosakranse and Oh, 2014):

- Slater-Usoh-Steed (SUS) questionnaire (Usoh et al., 2000): This is a commonly-used questionnaire and we also adopted it in one of our completed experiments. It presents six statements describing the VR illusion according to different criteria, and asks the user to what degree he would agree with each statement on a 7-point Likert scale. It doesn't explicitly categorize the six questions into sub-groups.
- Witmer and Singer Presence Questionnaire (Witmer and Singer, 1998): This questionnaire has 19 questions and tries to understand the user's perception on four major factors that constructs presence, namely control, sensory, distraction and realism.
- Lombard et al. questionnaire (Lombard et al., 2000): This questionnaire addresses presence based on six dimensions, namely social richness, realism, transportation, immersion, social actor within a medium and medium as a social actor. They are derived from the authors' theoretical model of the components that construct presence.
- Igroup Presence questionnaire (Schubert et al., 2001): This one evaluates the experience on four subscales, namely presence (in its own term), spatial presence, involvement, realness. It emphasizes the connection between presence and supported actions.
- ITC-Sense of Presence Inventory (Lessiter et al., 2001): This questionnaire has 44 items, with sub-scales including sense of physical space, engagement, naturalness and negative effects. It's meant to be applicable for any mediated experience including non-immersive or even non-interactive products such as television.

Besides post-experiment questionnaire, there exist other evaluation methods that rely on self-report. One example is to ask the user to report moments of "breaks" that they don't feel present, and the count of such events would give an indication of the overall presence of the experience (Slater et al., 2003). Another case is to let the user compare their sensory perception from reality and from VE, such as comparing visual illusions in different conditions (Riener and Proffitt, 2002).

2.1.2.2 Behavioral Measures

Sheridan proposed to measure presence based on the user's reaction to threatening stimuli in the VE (Sheridan, 1992). For example, the action of dodging a virtual ball generally indicates a strong sense of presence. Adopting the similar approach of behavioral measurement, Slater et al. asked the user to pointing at an object that exists both in reality and in VE. It was assumed that the tendency of carrying out this same action in VE rather than in reality indicates greater presence (Slater et al., 1995). Thie et al. measured the user's willingness of re-entering the VE after the experiment is finished and used this to interpret how strong the sense of presence was (Thie and Van Wijk, 1998).
Freeman et al. tried to measure the sense of presence by observing the user's magnitude of postural response in a seated viewing experience, but no significant relationship was found between this behavioral response and self-reported presence (Freeman et al., 2000).

2.1.2.3 Physiological Measures

In the famous “visual cliff” experiment, Meehan measured the participant’s physiological responses such as heart beat rate and skin temperature when facing stressful stimuli (Meehan et al., 2002). They found that the increase in heart rate was significantly correlated with self-reported presence using the SUS questionnaire, which indicates that physiological measurements could work as a tool to objectively gauge the sense of presence. Dillon et al. used similar physiological measurement coupled with the “ITC-Sense of Presence Inventory” presence questionnaire in a video viewing experience, but couldn’t find strong correlation between the two (Dillon et al., 2002). Baumgartner et al. used fMRI to identify the brain regions that are most related to feeling presence (Baumgartner et al., 2008). Slater et al. also used physiological measurement to examine the user’s response when “breaks” from presence happen (Slater et al., 2003).

2.1.3 Plausibility Illusion and Coherence

Skarbez proposes the concept of coherence as a parallel construct to immersion. It serves as objective property of a VE that leads to plausibility illusion, just like immersion to place illusion. Formal definitions of these concepts can be found in Section 1.1 and 1.2.

Notice that coherence is not equivalent to “realism.” Skarbez specifically argues that coherence is “a superset of realism or fidelity.” It emphasizes the validity and consistency of the VE’s internal logic. As long as the scenario has reasonable logic, the VE could be coherent even though it may not be realistic. An example is that, if the user is primed to believe he is experiencing a fantasy world, having characters flying around could still provide a coherent scenario.

This gives rise to an important question about the nature of coherence: if coherence depends on the prior belief in the user’s mind, how can it still be objective? To answer this question, Skarbez points out that since every VR experience assumes some prior knowledge, coherence cannot be purely objective. However, it is still useful to treat coherence as an objective characteristic, since a VE designer does have control over whether the events in the VE are internally consistent and he can minimize the dependence on prior knowledge by properly priming the user with certain expectations.

Previous research has tried to understand how to increase coherence of a VE in order to support Psi. In a study investigating both PI and Psi, Slater et al. found that correlation between self-actions and events, along with having a realistic illumination model seemed to be important for Psi, while having a virtual body contributed to both criteria (Skarbez et al., 2017c). Skarbez et al. specifically investigated Psi and coherence. They found that “having an accurate and well-behaved representation of oneself in the virtual environment is the most important contributing factor to Psi”. Llobera et al. considered narrative as an important component for building coherence (Llobera et al., 2013). Wang et al. investigated the specific case of interacting with a virtual character through hand shaking, and found that improving haptic quality increased plausibility (Wang et al., 2010). Biocca et al. presented empirical evidence that multi-modal integration through cross-modal transfer might contribute to the sense of presence, which is related to our efforts in adding physics coherence through pseudo-haptic feedback (Biocca et al., 2001).

2.2 Walking Interfaces

2.2.1 Benefits of Real Walking

Travel interfaces can be categorized by their *interaction fidelity*, which is defined as “The objective degree of exactness with which a system reproduces real world interactions (action realism)” (McMahan, 2011). A typical low-fidelity travel interface uses a gamepad by mapping the two joysticks to translation and rotation control accordingly. A high fidelity interface preserves real, natural walking as it’s how we move around in reality.

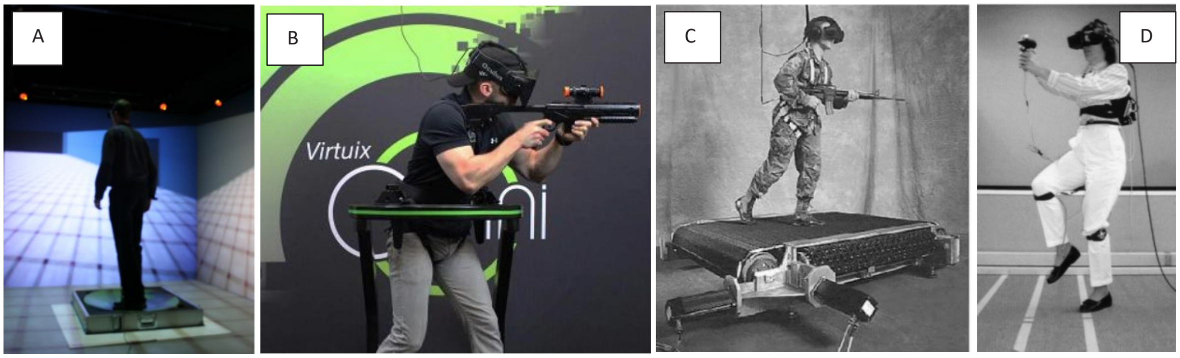


Figure 2-1: Walking interfaces with mid-level fidelity. A: the Wizdish by Swapp et al.(Swapp et al., 2010); B: the Virtuix Omni by Goetgeluk et al. (Goetgeluk et al., 2015); C: The omni-directional treadmill by Darken et al. (Darken et al., 1997); D: The “walk-in-place” interface by Templeman et al. (Templeman et al., 1999)

There exist various locomotion devices and techniques that lie between these two extremes. They try to simulate the locomotion of walking on some level with specialized input-to-output mapping algorithms. Figure 2-1 shows some examples of these mid-level fidelity interfaces. Figure 2-1A shows the device called *Wizdish,* with which the user wears low-friction shoes and “slides” his feet back and forth to simulate the action of walking (Swapp et al., 2010). More recent devices often provide harnesses to help the user keep balance, such as the Virtuix Omni shown in Figure 2-1B (Goetgeluk et al., 2015). With this added safety measure, users are allowed to carry out more complicated actions, such as dodging while running or even jumping. People have also tried to turn treadmills into such travel devices. The “omni-directional treadmill (ODT)” shown in Figure 2-1C lets the user stand on a 1.3m by 1.3m surface and supports turning while walking using a two-orthogonal belt arrangement (Darken et al., 1997). Other approaches let the user “walk in place” (WIP)—they require one to carry out the body movement of walking without actually moving forward, and map that movement to the translation of viewpoint. Figure 2-1D shows an example of WIP called “Gaiter”. It uses pressure sensors placed inside the user’s shoes and 6-DOF tracker placed on the knees to estimate the intended walking direction and speed (Templeman et al., 1999).

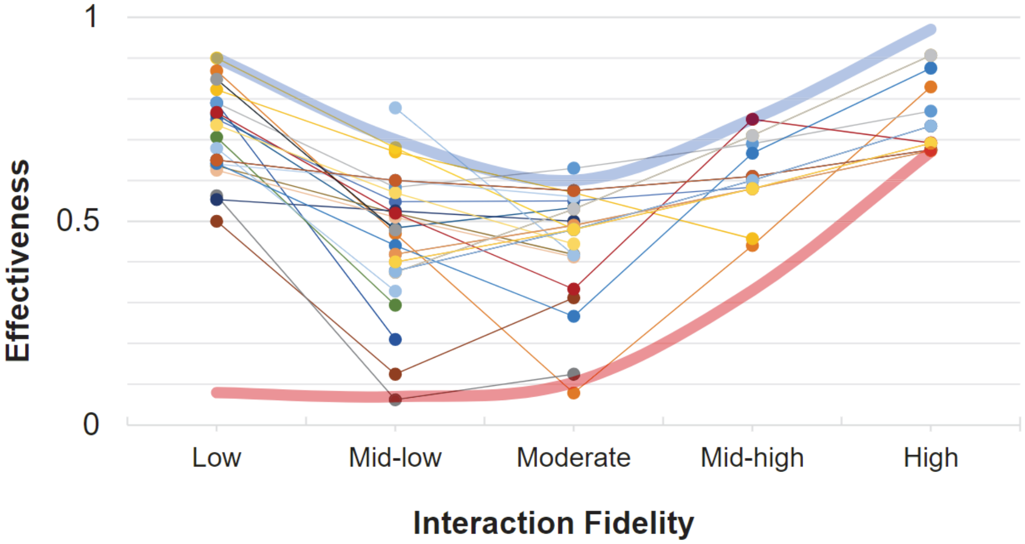


Figure 2-2: The "uncanny valley" effect of how interaction fidelity affects the effectiveness of locomotion interfaces, by Nabiyouni et al. (Nabiyouni, 2016)

Nabiyouni et al. presented a meta-analysis of how interaction fidelity affects the effectiveness of travel interfaces (Nabiyouni, 2016). They found empirical evidences to support an "uncanny valley" effect, which is shown in Figure 2-2. Moderate levels of interaction fidelity, such as the techniques in Figure 2-1, often result in the worst effectiveness compared to low-fidelity and high-fidelity approaches (Nabiyouni, 2016). In fact, many studies have shown that real, natural walking is the travel interface with the best performance and user experience (Steinicke et al., 2013). Peck et al. compared real walking in a large VE (facilitated by their redirection technique) with walk-in-place and joystick on cognitive performance for tasks of navigation and wayfinding. Real walking performed significantly better than the other two, as the user could point to targets more accurately, travel shorter distance and make fewer mistakes in turning (Peck et al., 2012). Similar results that support real walking as the most effective interface can be found in the evaluation by Feasel et al. They developed a novel walk-in-place technique called LLCM-WIP, which was meant to reduce the system latency and generate more natural, continuous viewpoint movement. They evaluated the proposed technique with gamepad and real walking, and real walking performed significantly better than the other two (Feasel et al., 2008). Marsh et al. compared the ability to keep spatial memory during a locomotion task between three techniques, namely a low-fidelity controller, a mid-fidelity interface that maps positional input to velocity and high-fidelity real walking. They found that real walking required the least cognitive load compared to the others (Marsh et al., 2012). The study by Usoh et al. also demonstrates that walking could result in the strongest sense of presence compared to walk-in-place or the low-fidelity technique of flying (Usoh et al., 1999). Another experiment by Chance et al. shows that real walking results in the least motion-sickness and best performance in a task that requires for spatial reference (Chance et al., 1998).

2.2.2 Redirected Walking (RDW)

To preserve real walking in a VE larger than the motion-tracked space, the concept of redirected walking has been extensively studied during recent year.

Razzaque et al. first introduced the concept of RDW (Razzaque et al., 2001; Razzaque, 2005). It reorients and/or repositions the user to confine his movement to the physical space, while preserving the illusion that he is walking through a much larger VE. The original RDW algorithm is shown in Figure 2-3. It lets the user walk in zig-zag pattern in the VE and reorients him at the turning points on the path, which makes him walk back-and-forth in reality. The reorientation algorithm used a rotational gain that was the maximum of the following three components:

(a) A small, fixed amount of baseline rotation that was constantly applied at the turning points;

(b) A component of rotation based on the user's walking speed;

(c) A rotation that was proportional to the user's angular velocity when turning her head.

A key element of RDW is to make such manipulation unnoticeable, so that users perceive themselves as walking normally. Follow-up studies have shown that a fairly large physical space is needed to effectively hide this redirection from the user, although the exact threshold value of this size might depend on the estimation method and the specific design of the redirection algorithm (Steinicke et al., 2010; Grechkin et al., 2016; Langbehn et al., 2017). Recently, Langbehn et al. presented an undetectable RDW technique for room-scale VR using bent paths, but the approach was limited to strictly pre-defined paths (Langbehn et al., 2017). Suma et al. demonstrated that infinite walking could potentially be supported using a 6m by 6m physical space with intelligent algorithms, but more evidence is needed to validate the effectiveness of this approach (Suma et al., 2015).

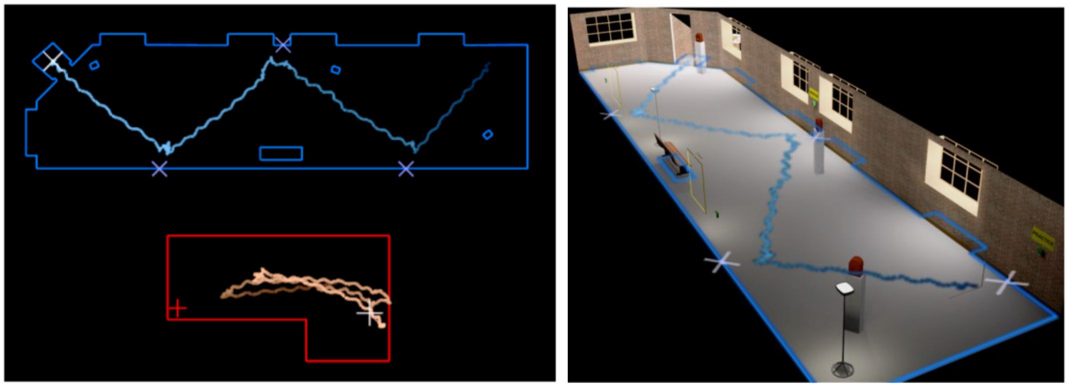


Figure 2-3: The original RDW by Razzaque et al. Left: Overhead views of the path taken by the user in the virtual environment (above in blue) and the laboratory (below in red). The user walked in a zigzag pattern through the VE while walking back and forth within the tracker space. Right: The user's path superimposed onto the virtual environment.

Since the original RDW work, rich variations of this method have been proposed. Hodgson et al. compared different steering strategies of RDW algorithms (Hodgson and Bachmann, 2013). Suma et al. proposed a taxonomy for all RDW methods based on three criteria: whether the redirection is based on reorientation or repositioning, whether the redirection is presented as subtle change or overt alteration, and whether the redirection is actuated continuously or as a discrete event (Suma et al., 2012).

2.2.3 Coherence and RDW

Williams et al. tested several straight-forward realizations of redirected walking based on resetting the user's position or orientation: When the user approached the boundaries of the physical space, the viewpoint of the virtual environment was frozen and she was explicitly asked to step back or turn around. They also presented a method called "2:1 turn," in which the user was asked to turn around by 180 degrees while the viewpoint in the virtual world was rotated by 360 degrees. After one of these resetting actions, she could walk further in the virtual world while in reality she was moving back to the physical space (Williams et al., 2007). All these approaches employ overt redirection that explicitly resets the orientation or position. When these reset events happen, the experience is interrupted, and the user is directly instructed to comply without a valid reason that's an inherent part of the experience. We suggest that these brute-force measures fundamentally break the coherence of the interaction.

Previous research explored ways to make the redirection less intrusive. Even though they don't explicitly relate these efforts to the concept of coherence, we identify them as in the same line with our motivation, as less disruptive redirection could improve coherence compared to brute-force reset. Suma et al. proposed to alter the virtual environment by leveraging "change blindness," a phenomenon that people might fail to notice the change of virtual scene when it happens outside one's field of view (Suma et al., 2011). This approach relied on modifying the virtual scene itself instead of just changing the position and/or orientation of the viewpoint, which might not be desirable in certain situations. Peck et al. used a visual distractor to draw the user's attention while manipulating the viewpoint's rotation (Peck et al., 2009). Similarly, Neth et al. proposed to use virtual characters in the scene to affect the user's walking path (Neth et al., 2012). However, the distractors or virtual characters in these techniques are added to the virtual environment and might be incompatible with the virtual world's context (thus, Suma et al. categorized visual distractor as overt redirection instead of subtle) (Suma et al., 2012). On the other hand, Grechkin et al. emphasized that redirection techniques could be integrated into the virtual world itself, which is similar to our design goal (Grechkin et al., 2015). They proposed the concept of adding a secondary target that was consistent with the narrative, in order to seamlessly manipulate the user's travel path.

One of the experiments we conducted focused on the user's spatial understanding as a potential benefit of coherent RDW. People have investigated how virtual reality and different interaction techniques could affect this specific aspect of user experience, and here we briefly review representative existing work. Cohen provides a thorough overview of the theoretical background about spatial cognition (Cohen, 2013). Schnabel et al. investigated the user's ability to understand spatial volumes when using 2D plans, screen-based 3D VE representation and motion-tracked VE (Schnabel and Kvan, 2003). They found that the VE setting enhanced the design's understanding of complex volumes and their spatial relationships. Suma et al. compared real walking with virtual travel techniques (e.g., pressing a trigger to move) on an information gathering task in structurally complicated environments. The result suggested that virtual travel is a reasonable substitute for walking, but walking is still beneficial for applications that require fast navigation or actions that are close to real-world behaviors (Suma et al., 2009). Similarly, Ruddle et al. found that walking could improve the cognitive map of the VE when the environment is large in both scale and extent (Ruddle et al., 2011). Larrue et al. found that embodied interactions (e.g., turn by physically rotate one's head or body) provide body-centered information that help transfer one's spatial understanding in navigation-related tasks (Larrue et al., 2014). Kim et al. proposed the technique of "finger walking in place (FWIP)" (Kim et al., 2008) that transfers the action of walking onto the hand, and found evidences that this "action-transferred design approach" can also improve spatial knowledge acquisition and usability in navigation-related tasks (Kim et al., 2015a).

2.2.4 Using RDW in Real-World Applications

Previous research has shown successful examples of applying RDW in a domain-specific application. The "Arch-explore" system was one such example, in which redirection was applied when transitioning between rooms in an architectural walk-through experience (Bruder et al., 2009). Closer to one of our techniques that uses a historical site, the "Bema" system created a VR experience of "ancient political assemblies at the hill of the Pnyx", in which architectural modeling and crowd simulation of high fidelity were used to support experiential analysis of mass gatherings at that specific historical phase (Kim et al., 2015b).

2.3 Object Manipulation Techniques

2.3.1 Virtual Hand Techniques

We instinctively manipulate objects using our hands in real life, so it's not a surprise that the majority of the object manipulation techniques are derived from this metaphor of direct manipulation.

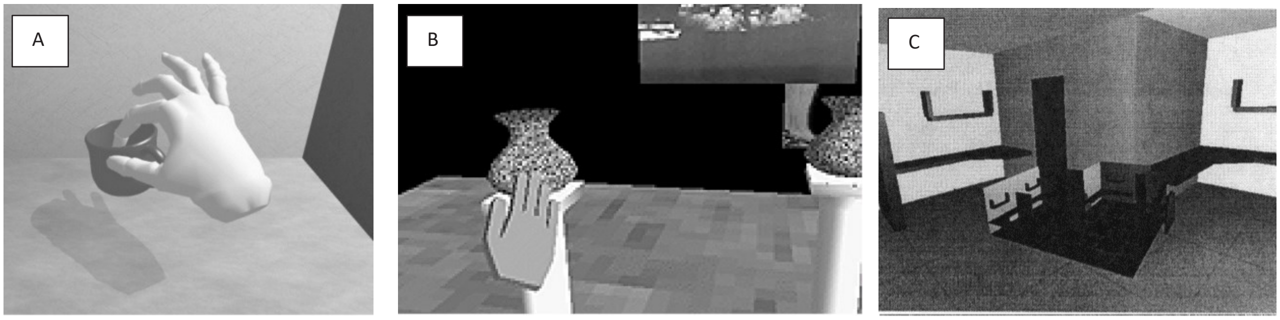


Figure 2-4: Some representative object manipulation techniques. A: Simulating the physics-based process of real-world interaction (Borst and Indugula, 2005); B: Naïve implementation of simple virtual hand (Poupyrev et al., 1996); C: World in Miniature technique (WIM) (Stoakley et al., 1995)

With the advancement of motion tracking and real-time simulation technologies, many have tried to re-create the real world physics-based manipulation in VR. It would usually rely on proprietary algorithms to simulate the friction and hand deformation using complicated computational models (Borst and Indugula, 2005; Hilliges et al., 2012; Talvas et al., 2015; Höll et al., 2018). Borst et al. developed such a system that "couples tracked hand configuration to a simulation-controlled articulated hand model using a system of linear and torsional spring-dampers" (Figure 2-4A). They also tried to address the issues of visual interpenetration between the hand and the object while providing haptic rendering using force feedback gloves. To improve the applicability of such methods, Holl et al. recently proposed to use Coulomb friction model to decrease the computational complexity, and their method can be incorporated into a commercial game engine (Höll et al., 2018). Overall, these approaches try to create high-fidelity interaction that closely mimics reality, which are not directly applicable for remote object manipulation. Without authentic haptic feedback, interpenetration could always break the plausibility illusion.

The more common method uses a (certain variation of) simple virtual hand metaphor. It treats the entire hand as a rigid body without simulating the physics-based contact between the skin and the object. A naïve implantation renders the user's hand (or hand-held device) as a virtual hand at the exact same location (Figure 2-4B). To select and object, the user simply intersects this virtual hand through the object and uses a gesture (or press a button) to "grab" it, then the position and orientation of the object would start following the hand. This interaction technique is very intuitive and provides good usability. The problem is that the user cannot control objects that are out of reach (Poupyrev et al., 1996).

To enable remote control, simple virtual hand has been extended by breaking the isomorphic mapping. The "go-go" technique "grows" the user's virtual arm to reach far-away objects once the hand is stretched out beyond a certain distance from the body. Once this happens, it uses a non-linear mapping between how far the real hand has reached out and how far the virtual hand is placed, so the user could reach locations that are at great distances. The downside of this approach is that the nonlinear mapping makes the control precision decrease dramatically once the hand is far away (Poupyrev et al., 1996). The HOMER (hand-centered object manipulation extending ray-casting) technique uses ray-casting to select distant objects. Instead of attaching the object to the ray during the manipulation, it uses a metaphor that the virtual hand instantly travels to the object and attaches to it, so the user could directly manipulate it remotely (Bowman and Hodges, 1997; Bowman et al., 2004). These approaches focus solely on task performance and usability, while ignoring the physics-related perceptions of the interaction.

2.3.2 Indirect Techniques

Stoakley et al. proposed the World-in-Miniature (WIM) technique, shown in Figure 2-4C. Instead of extending the user's arm, it "shrinks" the entire virtual world to a smaller scale. The user could manipulate objects in the VE by controlling its counterpart in this miniature. This effectively makes the entire virtual world within reach. The limitation is that scaling a large VE will make each object very small, which could add difficulties to the selection and manipulation actions (Stoakley et al., 1995). The scale-world-grab technique combines this idea with the HOMER technique, so the world is scaled to facilitate control instead of the arm (Mine et al., 1997).

Pierce et al. proposed the "Voodoo Dolls" technique. It allows the user to control virtual object using temporary, miniature, handheld copies of them (which are called "dolls"). Because such proxies are only created for the objects that we want to control, it doesn't need to scale the entire world, which solves some problems of the WIM technique; The method also allows the user to explicitly and interactively define a frame of reference for the manipulation using the other hand, which makes the interaction easier (Pierce et al., 1999).

2.3.3 Transfer Functions for Object Manipulation

Card et al. presented a framework for analyzing input devices, in which "expressiveness" was used to measure the capability of input-to-output transfer function (Card et al., 1990). A more expressive mapping function could enable larger range, more precise or more valid actions in control. For transfer functions in object manipulation, the general consensus is that lower-order mappings between hand and object provide better user experience and performance compared to indirect mappings such as rate control or acceleration control (Zhai, 1998; Hinckley et al., 2004). Although many techniques have been proposed to break the isomorphic mapping (e.g., (Poupyrev et al., 1996; Poupyrev et al., 2000)), the nature of the transfer function is still usually a zero-order mapping.

Some proposed interfaces use transfer functions that are different from the naïve 1:1 mapping. In the bimanual interface presented by Levesque et al., the object was attached to a ray cast from the user's right hand, and "the distance from the hand to the object remains constant throughout the whole interaction" (Lévesque et al., 2011). Benko et al. designed a "DepthTouch" interface that was placed above an interactive surface, where translations of the hand along certain dimensions were sometimes mapped to the rotational angle of the object (Benko and Wilson, 2009). Similarly, Song et al. used a "handle bar" metaphor, where circular movements of the hand could be mapped to the angle of rotation of the object (Song et al., 2012). Notice that even though these cases differ from the naïve direct position-to-position mapping, the transfer function was still zero-order.

To overcome the anatomical limits of the human hand such as hand instability and limited range of movement, enhancements of direct mappings have been proposed. Kopper et al. let the user toggle between two separate Control/Display ratios (C/D ratios) in order to achieve both ballistic and precise control, which is very similar to the direct control interface employed in one of our studies (Kopper, 2011). Frees et al. proposed a solution called PRISM that dynamically adjusted the C/D ratio by interpreting the user's intention from the hand's movement speed (Frees et al., 2007). When the hand is traveling at a high speed, it would decrease the C/D ratio to enable faster movement for a long distance; when the hand is moving slowly, it would increase the C/D ratio to improve control precision. Dominjin et al. also found that changing C/D ratio in a direct manipulation interface could create pseudo-haptic illusion of weight variation. Specifically, increasing C/D ratio would make the object feel heavier and decrease the ratio would make it feel lighter.

Previous research in 2D UI has put particular focus on using physics-based actions to enhance usability. Specifically, the interaction of using a flick gesture to "throw" an object to far-away locations has been widely adopted to move an icon beyond arm's reach (Geißler, 1998). Moyle et al. analyzed the physical properties of the flick gesture itself when using mouse and pen input (Moyle and Cockburn, 2002). Aliakseyeu et al. evaluated the effectiveness of the flick interaction for scrolling through documents with pen interfaces (Aliakseyeu et al., 2008). Reetz et al. proposed the "Super Flick" technique that adds an optional correction phase to the simple flick interaction to improve its accuracy (Reetz et al., 2006). Similarly, techniques such as "drag-and-throw" and "push-and-throw" were introduced to help the user accurately define the trajectory of throwing (Collomb et al., 2005; Hascoët, 2008).

2.3.4 Gesture-Based Interaction

Gestural commands can be classified as either postures or dynamic gestures. A posture is a "static configuration of the hand", while a dynamic gesture is a "dynamic movement" of the hand. An example of posture is the "thumbs up" sign, and an example of dynamic gesture is hand waving (Bowman et al., 2004; LaViola Jr et al., 2017). Gesture-based interaction has drawn increasing attention in recent years, as it is an essential part of the emerging "natural user interface" paradigm (Bowman et al., 2004; Wigdor and Wixon, 2011; LaViola Jr et al., 2017). Many have tried to find the best set of gestures for a certain task by understanding human preference through user studies. One such example is the classification system for mid-air gestures by Aigner et al., in which they analyzed human-to-human gestures in terms of hand usage and gesture type (Aigner et al., 2012). Norton et al. investigated design strategies for full body gestures in video games by studying human preference when given complete freedom of choosing gestures (Norton et al., 2010). In fact, there is a whole family of user studies called "gesture elicitation" that tries to derive a suitable gesture set from the users themselves (Ortega et al., 2017). Overall, these efforts usually focus on finding the best mapping from a group of gestures to a group of actions required by the task.

Beurden et al. compared the pragmatic and hedonic qualities between gesture-based and traditional interfaces (van Beurden et al., 2011). They found that "more embodied interaction reveals higher scores in terms of hedonic quality and fun than mouse-based interaction." On the other hand, more embodied interaction could result in more body fatigue. This finding implies that gesture-based interaction may have some inherent qualities that lead to superior user experience on aspects other than performance alone.

2.3.5 Physics Coherence in Object Manipulation

Research in animation has proposed the concept of "plausible motion," which describes "motion that could happen, given what is (un)known about the system" (Barzel et al., 1996). We find this concept to be strongly related to the definitions of physics coherence and physics plausibility, which also emphasize making physics-related behaviors appear "reasonable" for a specific scenario. The difference is that animation is concerned with the visual display alone, and the user does not directly influence the motion with his body. Related research often focuses on exploiting the space between accurate simulation and plausible motion, in order to generate high-performance algorithms that sacrifice some accuracy yet remain plausible (e.g., (Chenney and Forsyth, 2000; O'Sullivan and Dingliana, 2001; O'Sullivan et al., 2003a; O'Sullivan et al., 2003b)). Our research takes the perspective of 3D UI design and emphasizes multi-sensory consistency in the display of motion when the human body is actively participating in the interaction. O'Sullivan et al. (O'Sullivan et al., 2004) present a thorough review of perceptually-adaptive computer graphics.

When the user's body becomes a part of the interaction, physics-related perceptions now include not only visual information of motion but also haptic information of force. Thus, the discussion about physics plausibility has to include haptic sensation. While there has been a tremendous amount of effort in developing more capable haptic display devices (Laycock and Day, 2007; Lin and Otaduy, 2008), we view them as efforts that focus more on improving the immersion of the system but less on coherence. Different from introducing new hardware, our work on physics coherence focuses on the design of novel algorithms for input-to-output mapping given the limited immersion of current commercial VR systems. The research area that relates closely to this direction is *pseudo-haptic* display, which also tries to generate physics-related perception (perception of motion and force) by altering the input-to-output mapping without changing immersion. Here we briefly review the representative work in this area.

Pseudo-haptic display manipulates the interplay between multiple sensory channels to generate haptic sensations. The history of how visual stimuli could affect perception of force-related properties traces back to the Carpentier's size-weight illusion (Murray et al., 1999). Previous research in VR has identified various ways to alter the user's haptic perception using pseudo-haptics. Pusch et al. found that by displacing the position of the user's avatar hand, an illusion that a "force field" exists in the air could be effectively generated. The user could feel that a certain amount of force is pushing his hand when his proprioceptive perception does not match the visual feedback (Pusch et al., 2008). Dominjon et al. demonstrated how to induce a sense of weight variation between different objects by altering the control/display ratio in zero-order mapping (Dominjon et al., 2005). Similar approaches of generating weight perception using displaced movement was demonstrated by Rietzler et al. (Rietzler et al., 2018b) and Samad et al. (Samad et al., 2019). Ban et al. found that the user' perception of an object' shape could be strongly influenced by displacing the visual display of the hand that's touching the object (Ban et al., 2012). Lecuyer et al. induced a sensation of friction by changing the object's movement speed when it's dragged on a surface, and the result suggests isometric input devices are applicable to these pseudo-haptic methods (Lecuyer et al., 2000). Gurari et al. used similar tricks on manipulating proprioceptive cues relative to visual cues to influence the perception of stiffness (Gurari et al., 2009). Sense of softness was also simulated using pseudo-haptic techniques in AR by Punpongsanon et al. (Punpongsanon et al., 2015).

In proposing the novel pseudo-haptic techniques in this dissertation, we specifically focused on the perception of mass distribution. Lukos et al. finds that people would change how they grab an object if they can visually predict its center-of-mass (CoM). This study was conducted in real-world interaction instead of VR. Hashiguchi et al. used diminished reality (DR) to hide a part of a rod to influence the perception of its mass and mass distribution (Hashiguchi et al., 2018). Zenner et al. developed a "weight-shifting dynamic haptic proxy," which could change its mass distribution on the fly (Zenner and Krüger, 2017). This is an effort that directly addresses the challenge of displaying mass distribution by introducing a novel hardware that combines active and passive haptic display. Zenner's thesis also provides a thorough overview of materials related to this issue (Zenner, 2016). Similar to this, Cheng et al. proposed a liquid-based haptic proxy that can dynamically change its mass property (Cheng et al., 2018).

3. DESIGNING COHERENT REDIRECTED WALKING

We propose a family of techniques called “narrative driven cell-based redirection (CBR)” that naturally embeds the redirection events into the narrative and the VE. Using specially designed mapping algorithms, these techniques not only allow the user to continue walking seamlessly in a coherent experience but also create opportunities to enhance the interactivity of the experience itself. Two examples are provided to illustrate the idea. The first one, “Bookshelf and Bird”, uses plausible metaphors to disguise the redirection process as a cohesive part of the overall narrative (Yu et al., 2017). We’ll use this first example to introduce the concept of CBR. The second one, “lighting driven redirection”, uses the narrative to guide the user’s actions so they conform to what’s ideal for completing the redirection (Yu et al., 2018). We’ll use this project to show how CBR could be applied to a real-world application.

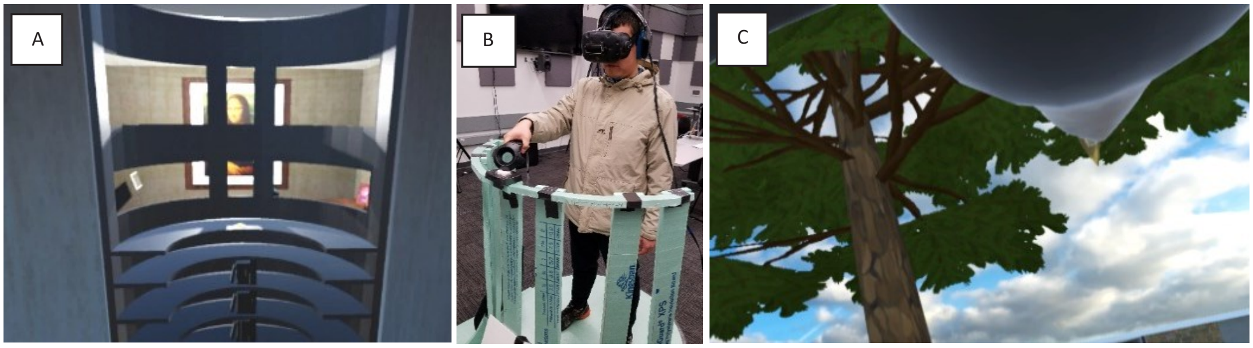


Figure 3-1: The bookshelf and bird techniques. A) The bookshelf uses a virtual rotation to allow the user to proceed to an adjacent room; B) The user standing on the physical prop and touching the button that activates the Bookshelf redirection; C) The bird translates the user to another cell while maintaining the correct relationship between the user, the physical workspace, and the virtual cell.

3.1 Bookshelf and Bird Techniques

This section (except Section 3.1.5) originally appeared in ***Yu, R.****, Lages, W.S., Nabiyouni, M., Ray, B., Kondur, N., Chandrashekar, V., et al. (2017). "Bookshelf and Bird: Enabling Real Walking in Large VR Spaces through Cell-Based Redirection", in IEEE Symposium on 3D User Interfaces. 116-119.*

The virtual environment is first divided into discrete “cells” with the same size and shape as the tracked area. In this way, the cell that the user is currently placed in is entirely accessible by real walking, since it’s perfectly mapped to the real world space. The two techniques help move the users between cells in a coherent manner.

3.1.1 The Techniques

3.1.1.1 The Bookshelf Technique

Movies and games sometimes feature a false bookshelf or fireplace that, when activated, can spin around its vertical axis, taking a person standing next to it into an adjacent, secret room. The Bookshelf technique uses this metaphor (Figure 3-1A).

When the user intends to travel to the next cell connected by a wall, he can step onto the platform attached to the bookshelf and push a yellow button. Once activated, the bookshelf will rotate itself (along with the user) by 180 degrees in the virtual scene, placing the user in the virtual room on the other side of the wall. Since no such rotation happens in reality, the user is actually still standing in the same lab space facing the original side of the real wall (Figure 3-2 A-C). The destination virtual room is now reoriented to lie on the same side as the physically available lab space (Figure 3-2C). The user can then simply turn around and walk to traverse the virtual room, which is perfectly mapped to the physical motion-tracked space (Figure 3-2D). Note that this requires the bookshelf to be placed in the middle of the wall. Repeating this redirection between each pair of adjacent cells enables the user to traverse a much larger area than the physically available space without breaking the narrative of the game.

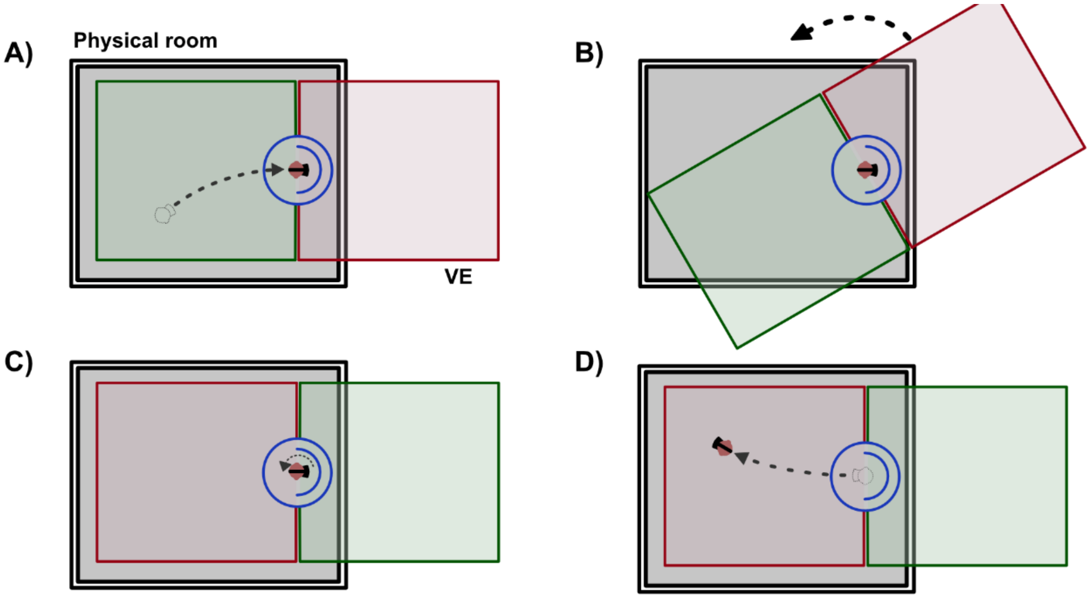


Figure 3-2: Redirection in the Bookshelf technique. The black rectangle indicates the tracked physical room, while the red and green ones indicate two virtual rooms. A: The user starts from the green room, steps onto the bookshelf, and activates redirection; B: The bookshelf rotates with the user in the virtual world while he stands still in reality; C: The physical space now lies on the same side as the red virtual room, and the user turns around to face it; D: The user steps off the bookshelf to walk through the red virtual room.

To be effective, the Bookshelf needs to convince the user that he is indeed rotating. We help create this illusion in several ways. First, we think the metaphor itself is accessible and familiar from popular media. For example, a very similar design of a secret door disguised as a rotating bookshelf appeared in the 1974 movie “Young Frankenstein”, in which a candle is used as the trigger to activate its rotation. Second, part of the virtual bookshelf is made of translucent glass (Figure 3-1A), so the user can see the next room during the reorientation. This provides visual feedback that makes the perceived rotation more convincing. Third, we provide spatial sound in both cells, and use a subwoofer to create vibration-based haptic feedback as the bookshelf turns.

In order to re-map the physical space to the new cell, the bookshelf needs to rotate continuously for 180 degrees. If the user steps off the bookshelf during the rotation, he would perceive himself as being rotated for no reason and perhaps going through virtual walls, which breaks the VR experience. To discourage this from happening, we dim the display while the bookshelf is rotating to create a “cut scene” effect, which makes the user feel like an observer. We also built a physical prop to give the user a tangible platform to stand on, which is expected to make him reluctant to step off it during rotation. Finally, the button that activates redirection is placed so that the user is required to stand on the bookshelf platform before activation (Figure 3-1B).

A potential issue with the Bookshelf is cybersickness created by visual-vestibular mismatch during the virtual rotation. To minimize this effect, we designed a semi-circular bookshelf that places the user close to the pivot point of the rotation, resulting in reduced virtual translation during the bookshelf movement. The dimming of the view is also designed to reduce perceived optical flow, and we anecdotally believe that this could potentially reduce cybersickness.

3.1.1.2 The Bird Technique

The Bird technique (Figure 3-1C) is designed for outdoor scenes, where cells are separated using low obstacles. Instead of relying on redirection through rotation, the Bird translates the user to a new cell.

The player first selects a target by looking at the ground plane of the desired destination cell. We cast a ray from the user’s eye point in the direction his head is facing, and the system selects the cell whose ground plane is intersected. A semitransparent sphere appears on the ground in the selected cell to provide visual feedback. Once satisfied with the choice, the user calls the bird by pushing a button on a hand-held controller. The bird approaches the player, grabs and lifts him, flies to the new cell and then descends to drop him there. Once the user perceives no more movement, he is free to walk around in the destination cell.

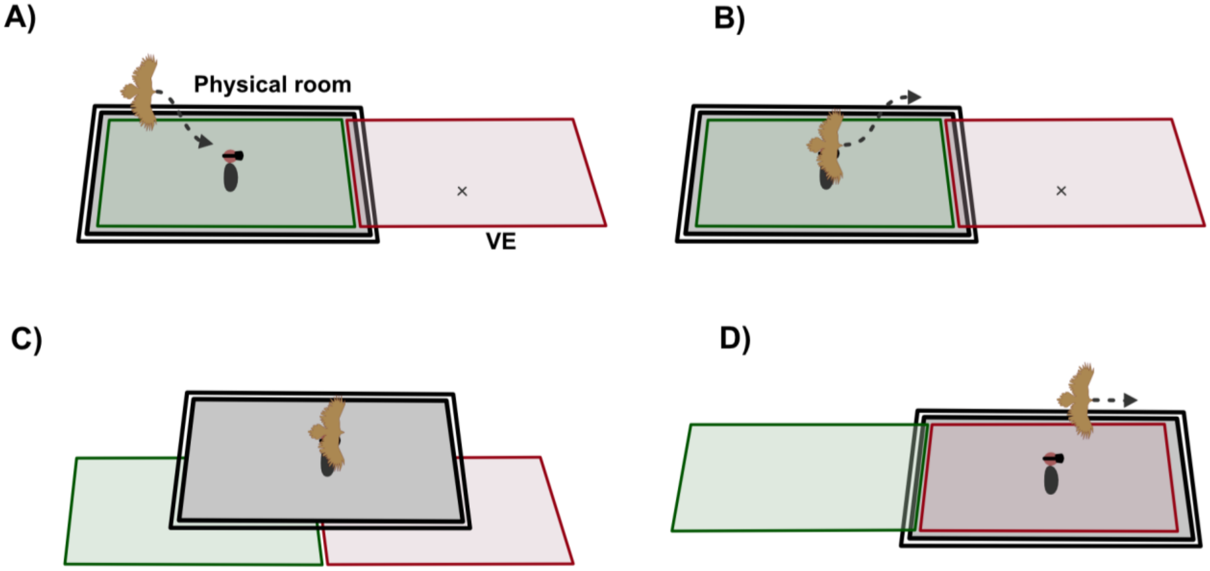


Figure 3-3: Redirection in the Bird technique. The black rectangle indicates the tracked physical room, while the red and green ones indicate two virtual cells. A: The user starts from the green cell and selects the red cell as the target. B: The bird approaches and picks him up; C: The bird lifts him and flies to the destination cell; D: The user is dropped at the corresponding position in the destination cell as his previous location in the original cell.

Although this sounds similar to teleportation, it differs in two key respects. First, the Bird does not drop the user at an arbitrary point in the destination cell, but rather translates the user from her position in the current cell to the corresponding position in the destination cell (Figure 3-3). This means that the physical boundaries are still aligned with the virtual edges of the cell, and the user can thus walk to any location in the cell. Second, the Bird uses a rapid but continuous virtual movement, rather than an instantaneous repositioning as in teleportation or a portal-based technique. This not only provides a physically plausible metaphor, but also may help the user maintain spatial orientation (Bowman et al., 1997).

To maintain consistency with the overall narrative, the Bird technique is applied in outdoor parts of the scene. The boundaries between cells are low obstacles (e.g., plants, rocks, or streams) instead of walls, so the user can see the destination region and will not pass through objects when carried by the Bird.

3.1.2 Demonstration

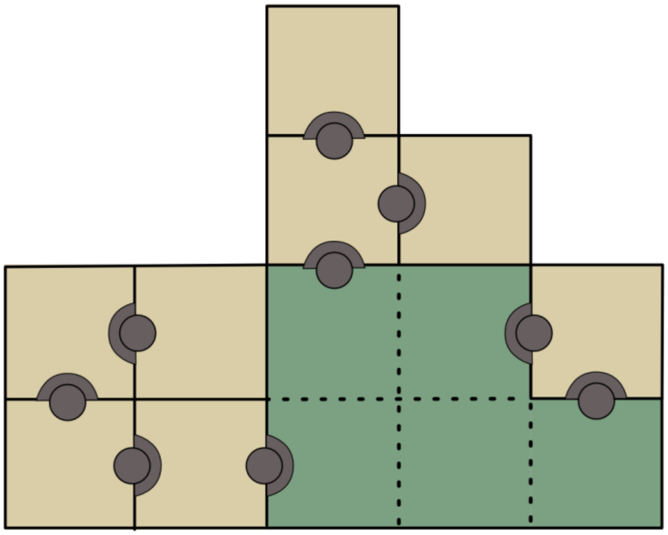


Figure 3-4: Layout of the “treasure hunt” VE. It has 8 rooms (yellow cells) and a yard (green cells), with 9 rotating bookshelves.

We built a simple “treasure hunt” game using these two redirection techniques. The game was implemented in Unity3D and is displayed in an HTC Vive head-mounted display. The display and the Vive controller are tracked by the Vive’s Lighthouse tracker in an area of about 5m x 5m. The user’s goal is to find words that compose a sentence that reveals the location of a treasure. The scene is a house with 8 rooms, some of which connect to an indoor yard with 5 regions. Each room and each region in the yard is a cell with the same dimensions as the physical workspace. The entire virtual scene spans over 325 square meters. Some rooms are connected by bookshelves and the yard is segmented by a small water channel and bushes (Figure 3-4).

The player can traverse the environment using a combination of walking and the Bookshelf/Bird techniques. Each cell contains a clipboard that reveals the next word in the secret sentence when the user walks close to it. The full sentence directs the user to return to a particular room, where the treasure appears.

We first demonstrated both the whole game and the individual techniques to a number of users and collected informal user feedback.

In our experience with the game demo, no users have had any difficulty in learning how to use the Bookshelf or Bird. Players naturally use real walking to move around within each cell. During the session, they seemed to easily recall the locations of previously visited landmarks and which cells have not yet been traversed. Many users felt surprised that they could explore such a large space. Users seemed convinced that they were walking through a virtual scene much larger than the motion-tracked area.

The Bookshelf and Bird metaphors were believable and effective in our experience. One of the users asked, “Wait… was I really being rotated just now?” after she experienced the bookshelf interface and took off the head-mounted display. When carried by the Bird from one cell to another, most users remain still during the flight. Some people express fear the bird will drop them before landing, which also indicates strong immersion.

	CBR	Teleportation	Traditional RDW	2:1-Turn	Arch-Explore
Requires physical walking to traverse the scene	✓		✓	✓	✓
Does not use translation or rotation gains	✓	✓			
Practical for room-scale physical spaces	✓	✓		✓	✓
Currently accessible virtual space fit in the physical space	✓			✓	✓
Redirection can be fit into the narrative of the VE	✓	✓			
Applicable to any VE layout		✓	✓	✓	✓

Table 3-1: Comparison of characteristics of CBR and related previous techniques

Cybersickness is still an issue for some users with one or both techniques. The most severe sickness is reported when someone tries to step off the bookshelf while it is rotating, which only happens when physical prop is not used.

3.1.3 Narrative Driven Cell-Based Redirection

On the surface, Bookshelf and Bird seem to have limited applicability, since they use highly specific metaphors. However, they are actually instances of a family of techniques we call “narrative driven cell-based redirection” (CBR). CBR techniques divide the VE into discrete cells, with each one being able to fit into the tracked space. Thus, each cell is fully accessible by real walking.

Users move between adjacent cells by invoking redirection through reorientation (e.g., Bookshelf) or repositioning (e.g., Bird). The key characteristics of CBR are that natural 1:1 walking is used within each cell, and that redirection is consistent with the narrative of the VE.

Other metaphors can be used within the CBR family to support different narratives. For example, a revolving door could be used in the place of the Bookshelf to rotate the user into an adjacent room. Similar to the Bird, a “sky hook” system could be presented where the player is carried to other rooms by grabbing onto a hook that’s attached to a rail above. Elevators, escalators, or any other vehicle-based metaphor could be applied to carry the user between cells in such interfaces, as long as the new cell is re-aligned with the real-world workspace after the transition.

Table 3-1 provides a comparison between CBR and some existing techniques. Unlike prior approaches to redirection, CBR combines several desirable qualities: it requires the user to walk physically, never applies a scale factor to physical walking or turning movements, is practical for room-scale spaces, maintains the fit between the physical tracking space and the virtual space, and provides metaphors that can be consistent with the VR narrative. (For Traditional RDW, 2:1-Turn and Arch-Explore, see (Razzaque et al., 2001; Williams et al., 2007; Bruder et al., 2009))

3.1.4 Limitations

Obviously, the most important downside of CBR is that it does not apply to arbitrary VE layouts since the VE must be divided into room-sized cells. To seamlessly embed the redirection into the experience, one has to find metaphors that are appropriate for the environment and consistent with the narrative, which may be difficult depending on the application. For example, placing objects like bookshelves in a VE for architectural planning may not be acceptable since it alters the building’s interior appearance. Overall, CBR is more applicable if one has freedom in designing the environment and story, such as when building a VR game.

Redirection in CBR must be designed carefully to realize a good user experience. For example, we need to encourage the user to remain on the Bookshelf platform when it is rotating and stand still when he is being carried by the Bird. Furthermore, the player will experience the redirection every time he travels from one cell to another, which may become too repetitive. The designer may need to vary the redirection metaphors, increase the speed of redirection, provide interesting visuals, or present tasks that are performed during redirection to keep the user engaged. Finally, the visual-vestibular mismatch introduced by CBR techniques can cause cybersickness. One needs to carefully implement the visual and haptic display to limit such discomfort.

Compared to teleportation and portals, the user is given less control over the redirection action with CBR techniques, but this constraint guarantees that the system can maintain a perfect mapping between the walkable cell and physical space.

CBR is not meant to be a general-purpose locomotion solution. But given metaphors that are consistent with the application’s narrative, it is a powerful design approach that enables walking in VR far beyond the physical boundaries of the tracked space.

3.1.5 Evaluation

To address RQ2, we hypothesized that one measurable benefit from coherent RDW is better spatial understanding. The rationale is that these coherent techniques redirect the user without breaking the continuity of the environment and the narrative, so the user is traversing a seamless world throughout the experience. We think this should help the user keep her bearing compared to brute-force resets that take her out of the experience. We designed a user study to test this hypothesis by comparing the user's spatial understanding when using Bookshelf and Bird techniques and their less-coherent counterparts.

3.1.5.1 Experimental Design

The experiment asked the participant to traverse a large virtual environment composed of cell spaces using a designated RDW technique and measured his spatial understanding when using this technique. The Bookshelf was compared with the 2:1 turn technique (Williams et al., 2007). They both rely on reorientation while inducing a 180-degree rotational gain, so 2:1 turn served as the less-coherent counterpart of the Bookshelf. The Bird was compared with a standard teleportation technique, as both rely on relocation to complete the redirection. Teleportation is very popular with commercial VR games and it was used as the less-coherent comparable technique for the Bird.

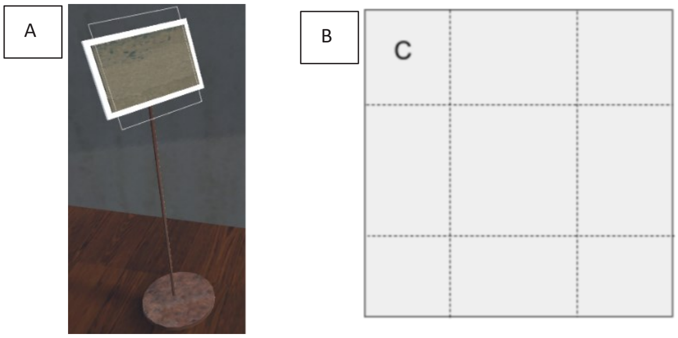


Figure 3-5: A: The post sign to collect clues from; B: An example of marking a sub-cell in answering the first two questions of the post-experiment evaluation

To make the participant traverse each cell space in the entire VE using the chosen RDW technique, the experimental process was implemented as a puzzle game. The user was tasked with collecting "clues" that were spread out in the environment to solve the puzzle. Each cell had one such clue, which was constructed as a sign post as shown by Figure 3-5A. When the user approached a new post that had not been collected before, the sign revealed a new word. The user had to visit every cell space in the VE to collect all these words, and they composed a complete sentence that in turn revealed the location of a final "treasure". The user's last step in completing the experiment was to go to the cell space that the treasure was placed in to find it. Note that the path that each individual took to traverse the VE was up to the user, so the order of how these clues were collected differed between individuals. The system only ensured that each cell had to be visited at least once using the chosen RDW technique before the final treasure was revealed.

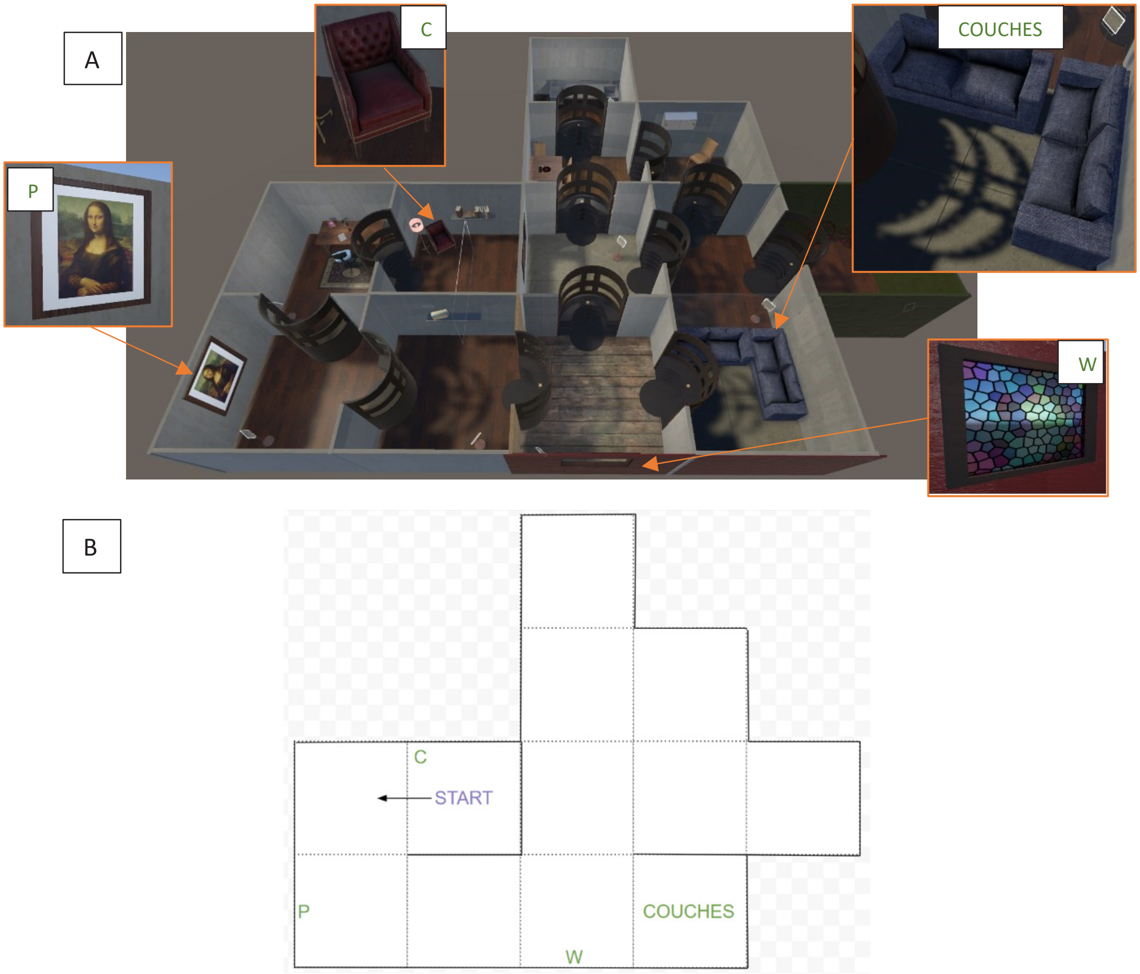


Figure 3-6: The layout of the VE used for Bookshelf and 2:1 turn. A: the 3D overview; B: the 2D layout. The landmarks used for the pointing task and post-experiment evaluation are marked by letters. "P" = the painting of Mona Lisa; "C" = the red chair; "W" = the stained windows; "COUCHES" = the blue couches

To measure the user's spatial understanding during the traversal process, the virtual environment was filled with "landmarks" that we would explicitly ask the participant to remember. After the user had encountered one of the landmarks and moved on to a new specific cell-space, the experimenter would ask the user to stop walking and point to the location of this landmark (using the hand-held controller) based on his current understanding and memory of the space. We used the accuracy of the user's pointing direction as an indication of his ability to grasp the structure of the VE. Specifically, we recorded the angle between the vector given by the user and the vector between the landmark's location and the user's current position. This angle was used as the quantitative measurement of his level of spatial understanding—the smaller this angle was, the more accurate the user was, which in turn indicated a better understanding. Figure 3-6 shows the layout of the environment used for Bookshelf and 2:1 turn. In this scene, the three target landmarks were the red chair, the painting of Mona Lisa and the stained glass window. Figure 3-7 shows the scene in which Bird and Teleportation are applied, and the targets included the sculpture, the tall tree and the well.

Figure 3-7: The layout of the VE used for Bird and Teleportation. A: the 3D overview; B: the 2D layout. The landmarks used for the pointing task and post-experiment evaluation are marked by letters. “F” = the fountain; “W” = the well; “T” = the tall tree; “S” = the sculpture; “FLOWING WATER” = the flowing water.

To further measure the user’s spatial understanding after the walking experience, a post-experiment questionnaire was conducted. The questionnaire showed a 2D top-down layout of the scene that the user just traversed and asked him to mark the locations of four landmarks on the map. Some of these landmarks were the same as the ones used in the pointing task and some were different. For the first two questions on each evaluation, each cell space was further divided into nine sub-cells. Only one of these sub-cells within the correct cell space corresponded to the accurate location of where the landmark was placed in, as shown in Figure 3-5B. The user’s answer was graded by the following rule: if he chose the right cell space, he could get two points; if he further picked the correct sub-cell within that cell space, he could get an extra point. For Bookshelf and 2:1 turn, these two questions asked about the locations of the Mona Lisa painting and the stained glass window; for Bird and Teleportation, the first two questions targeted at the tall tree and the fountain.

The third questions differed between the two evaluations. The re-orientation one required labeling the room that contained the blue couches. This only asked for the proper cell without differentiating between sub-cells, and it was worth two points without any extra ones. On the relocation evaluation, it required identifying the dotted cell borders which contained flowing water. This was again two points for proper marking, with no partial credit given. The fourth question was to mark which cell they collected the final clue (the cell where they completed the clue sentence which directed them towards the final treasure). Like the third question, this was out of two points only for marking the correct cell.

The experiment used a between-subject design. Each participant would experience one of the reorientation-based technique (either Bookshelf or 2:1 turn) and one of the relocation-based technique (either Bird or teleportation). The order of how these techniques were presented was counter-balanced to overcome potential learning effect. For example, the first participant might use Bookshelf first and then Bird. The next subject would use 2:1 turn then Teleportation. Then, Bookshelf and Teleportation, 2:1 turn and Bird, and finally a reversal of the order so the next would use Bird and *then* bookshelf. This rotation continued throughout the experiment, until the last few subjects, where techniques and orders were selected to help account for previous incomplete trials on a technique.

3.1.5.2 Hypothesis

The main hypothesis is that the coherent RDW techniques will lead to superior performance on these spatial understanding related tasks. Specifically, Bookshelf and Bird are expected to provide more accurate results for both the pointing task and the post-experiment marking task compared to 2:1 turn and teleportation accordingly.

3.1.5.3 Apparatus

We used a first-generation HTC VIVE system. The head-mounted display has a refresh rate of 90 Hz and 110 degree of field of view. Each eye has a display resolution of 1080*1200, forming a combined resolution of 2160*1200. The user holds one of the VIVE controllers in hand to and uses the index finger trigger to complete the pointing task.

The VE and interaction was developed using Unity 5.5, and the experiment was carried out on a desktop computer with a CPU of Intel i7-8700k, 16GB of memory and a GPU of Nvidia GTX 1070.

3.1.5.4 Participants and Procedure

A total of 16 participants were recruited. Since this experiment used between-subject design, each technique had 8 samples.

Each participant began by reading and signing the consent form. A background questionnaire was also presented to gather basic demographic data. Once the paperwork was out of the way, each participant was given an introduction to the experiment, the devices and VR in general, so that each of them could start the procedure with clear understanding of the task and familiarity with the devices. The introduction also highlighted important features such as the chaperone, which was important for minimizing the number of breaks in presence and reducing chances for collision with real objects.

After this, the subject was placed into the practice environment to try the techniques assigned to him. Each one ran through the practice using one redirection (Bookshelf or 2:1 turn) and one relocation (Teleportation or Bird) technique. The practice environments consisted of 4 or 5 cells so that the subject could familiarize themselves with both the specific technique used for moving around the environment and the tasks required between all trials. That is, each practice environment contained a number of the posts used for collecting the clues so that the participant could see these and know what to look for as they moved around the study environments. They also contained an (ungraded) pointing task, which let the user become familiar with what they need to do for completing these tasks for measuring egocentric spatial awareness.

Once the subject affirmed they felt comfortable with the technique and tasks, they were placed into the study environment. Each subject was given a brief introduction, reminding them of their tasks, pointing out an example of a landmark they need to locate (visible from the starting position), and mentioning the timed element of the study. It was important to remind them that they were being timed, but to prioritize accuracy and precision over speed. This was simply so that users do not spend excess time on unnecessary details, which could impact their overall grasp of the space.

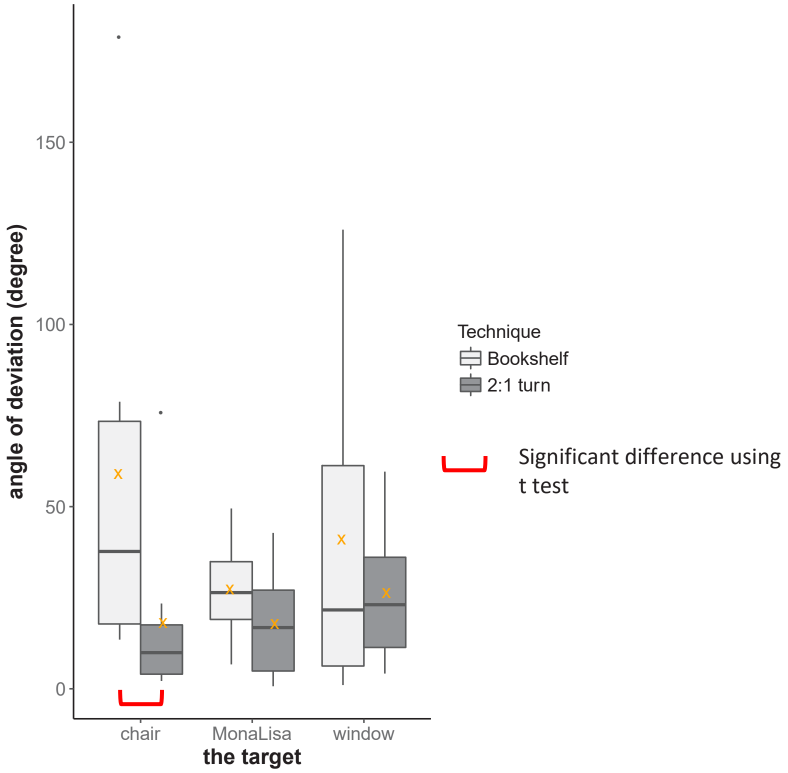


Figure 3-8: The result of the pointing task for Bookshelf and 2:1 turn

Once the user completed the final task, the experimenter helped her take the VR gear off and began proctoring the post-session evaluation. Again, the process from practice environment to the post-session evaluation would be completed *twice* for each participant, once each for two different techniques.

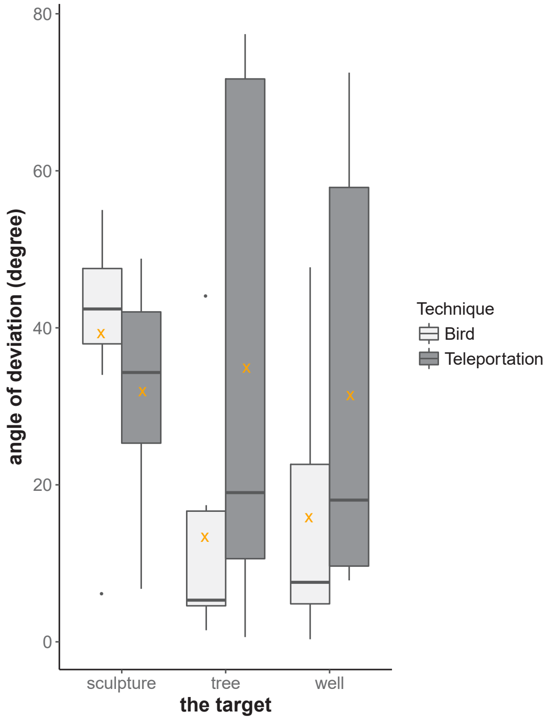


Figure 3-9: The result of the pointing task for Bird and Teleportation

3.1.5.5 Result

Figure 3-8 shows the result of the pointing task for Bookshelf and 2:1 turn. On average, 2:1 turn performed better than Bookshelf (the smaller this angle the more accurate the user was). Due to the limited sample size, we treat this data as non-parametric and ran a Wilcoxon rank sum test between the two techniques for each target. The result shows that 2:1 turn performed significantly better than Bookshelf for the “chair” target (P-value = 0.0289), but not the other two.

The result of the pointing task for Bird and Teleportation is shown in Figure 3-9. Wilcoxon rank sum tests do not show any significant difference between the two techniques for any of the three targets. Judging by the graph, it seems that the distribution of Teleportation has larger spread than Bird, which could indicate larger variation. We ran a Levene’s test to evaluate the equality of variances between the two techniques. The result does not show any significance.

Figure 3-10 shows the result of the post-experiment evaluation for Bookshelf and 2:1 turn. Contrary to the pointing task, Bookshelf performed better on average for all the questions. No significance was shown when we performed a Wilcoxon rank sum test for each question.

The result of the post-experiment evaluation for Bird and Teleportation is shown in Figure 3-11. On average, Bird seemed to perform better than Teleportation except for the case of the "fountain". Again, Wilcoxon rank sum tests don't show any statistical significance.

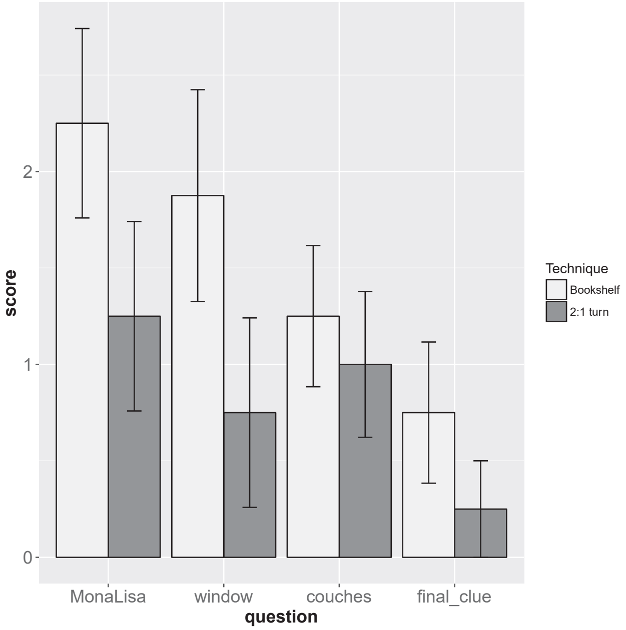


Figure 3-10: The result of post-experiment evaluation for Bookshelf and 2:1 turn

3.1.5.6 Discussion

Due to the very limited sample size, it's difficult to find much statistical significance in the result. However, we find some hints in the experimental results to inspire future work.

For the comparison between Bookshelf and 2:1 turn, Bookshelf performed worse on average for the pointing task, but seemed to be superior for the post-experiment evaluation. This result is surprising, as we hypothesized the coherent redirection of Bookshelf should help the user get a better understanding of the space. One potential benefit of 2:1 turn compared to Bookshelf is that the user has control over how fast the re-orientation is completed—to get redirected faster, the user just has to turn quicker. Another difference between the two techniques is that there is less visual occlusion with 2:1 turn, as there is no Bookshelf placed between the two rooms. We think these are potential factors that may have helped 2:1 turn in spatial understanding. More specific study is needed to investigate these hypotheses.

As for the comparison between Bird and Teleportation, visual inspection of the data from the pointing task hints that Bird seemed to be more "stable" than Teleportation, as the distribution is less spread out. This could be due to the constraints of Bird compared to Teleportation, as Bird always maintains the perfect mapping between the current virtual cell space and the physical play area, while the user could teleport in a lot of different ways. We believe that this unique characteristic of CBR is indeed an advantage in constraining the possible relationship between the virtual and real space, and we expect that future work could provide concrete statistical evidence supporting that Bird results in a more stable spatial understanding compared to Teleportation.

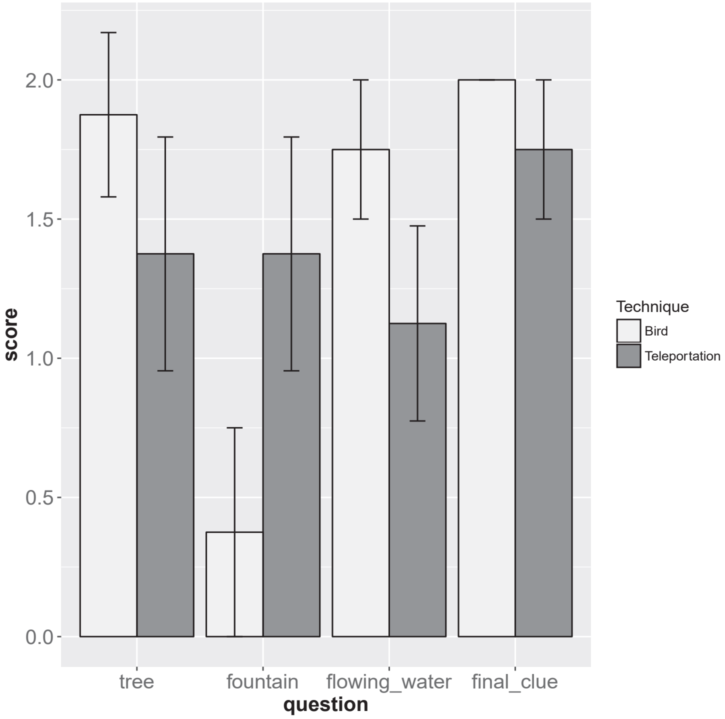


Figure 3-11: The result of post-experiment evaluation for Bird and Teleportation

Overall, this experiment is severely limited by the sample size, so it's not a surprise that little statistical significance was found. On one hand, we cannot conclude that the proposed hypothesis is supported by the experimental result; on the other hand, we hope the discussion here has inspired meaningful future work in finding statistically measurable benefits of coherent RDW.

3.1.6 Summary

This work proposes the Bookshelf and Bird techniques as examples of narrative driven cell-based redirection in VR locomotion. By mapping discrete cells in the virtual world to the physical workspace, the user can always access the entire cell with real walking while traveling between cells using seamless and believable redirection. Users can walk without worrying about the physical boundaries and can access a virtual space much larger than the physical one. We hypothesized that these coherent RDW techniques could lead to superior spatial understanding compared to their non-coherent counterpart and an evaluation was conducted to test this hypothesis. Due to the limited sample size, not much statistical evidence could be found to validate the hypothesis. We summarized the clues we found in the experimental results, hinting on promising future work.

3.2 Lighting-Driven Redirection

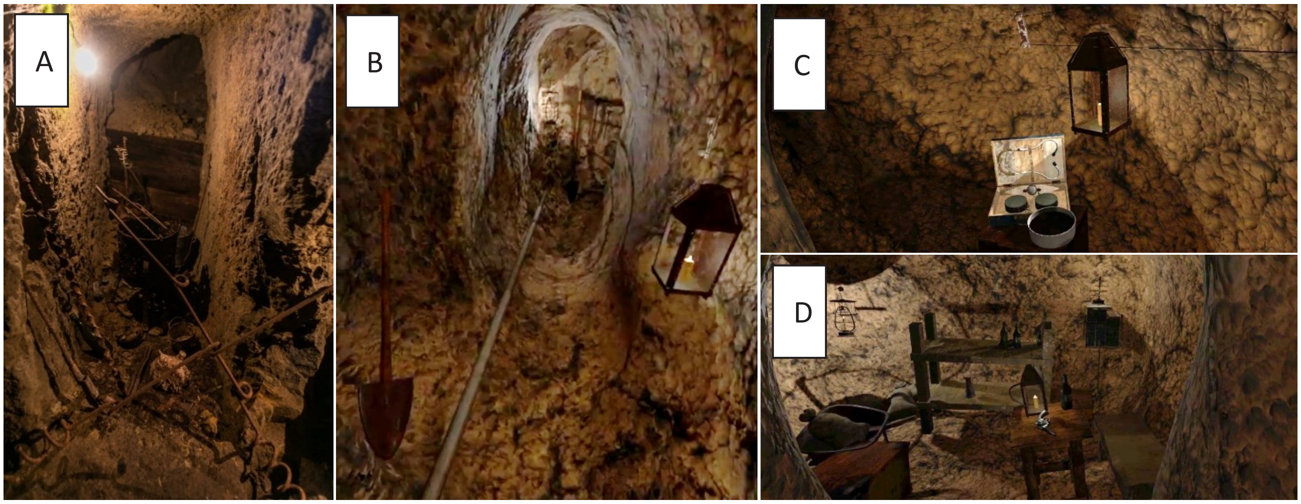


Figure 3-12: A. Photograph of actual Vauquois tunnel; B. VR re-creation of the tunnel; C. A corner that connected two adjacent segments of the tunnel, which served as the “redirection zone” where re-orientation was applied; D. A room in the tunnel where soldiers used to sleep

This section originally appeared in ***Yu, R.****, Duer, Z., Ogle, T., Bowman, D.A., Tucker, T., Hicks, D., et al. (2018). "Experiencing an Invisible World War I Battlefield Through Narrative-Driven Redirected Walking in Virtual Reality", in 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). 313-319.*

In this work, we explore how the idea of narrative driven cell-based redirection could be applied to a real-world application. We aim to embed the reset events into the narrative of the VR experience, so that they not only allow the user to continue walking in the virtual environment but also enhance the experience itself. We present the design of an educational VR application with the goal of letting users experience some of the horrific conditions faced by soldiers in World War I. In particular, the experience puts users in a tunnel representative of the Battle of Vauquois. Our system design allows users to physically walk through a tunnel nearly 50 feet in length, but only requires an exhibition space of approximately 15×5 feet. We discuss the lessons learned from this narrative-driven redirection approach, and how it can be applied to other settings for a coherent and natural walking experience in VR.

3.2.1 Recreating a World War I site in VR

3.2.1.1 The Hill of Vauquois

The rural village of Vauquois sat atop a gentle hill in rural France, 155 miles to the northeast of Paris. The village dates to the early medieval period, and prior to World War I was a small, quiet village of fewer than 200 residents. Because of its location and topography, Vauquois became an important strategic location for both the German and French armies. The hill of Vauquois provided an ideal position for the German Army to observe the French fortress-city of Verdun and control vital transportation routes in the area.

The four years of fighting from 1914 to 1918 at Vauquois represent the full spectrum of combat in World War I, from street fighting, to the development and use of trenches and underground shelters, culminating in a war conducted entirely underground through the extensive use of mining and tunnel warfare. Other sites along the Western Front saw mining used for attack, but Vauquois is unique in the extent of the mining and the fact that both armies attacked and counterattacked one another underground in one location for over three straight years. Vauquois was essentially two underground cities connected through subterranean combat.

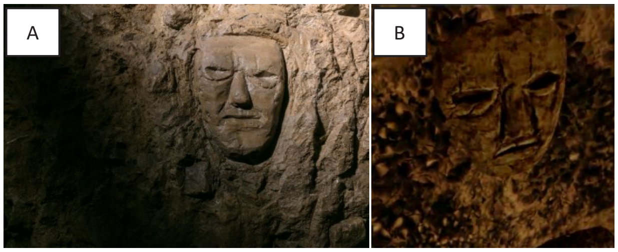


Figure 3-13: A. Carvings on the wall in the real Vauquois tunnel; B. The re-creation in the VR scene.

Today, there is no hilltop village of Vauquois. What remains are two hills instead of one, with massive craters, that 100 years later are still over 60 meters deep in places. Below the surface, tens of kilometers of tunnels remain, including barracks, command posts, power generation rooms, attack tunnels, hospitals, chapels and more. A photo taken in one of the tunnels can be seen in Figure 3-12A. Today, mining tools, weapons (such as rifles, pistols, and hand grenades), heaters, ventilation equipment, rail lines, and personal items (such as bedding, tables, and eating utensils), still remain in the tunnels. Most human of the items left behind are the carvings made by the soldiers to honor friends and loved ones, or mock feared generals (Figure 3-13).

Visiting the site itself presents challenges in terms of not only its geographic isolation but in its challenging terrain. Visiting the underground aspects of the site is only possible for those who are not only able-bodied but also adventuresome and comfortable in dark, confined spaces. But Vauquois is a well-preserved representation of the futility of war and the deadlock that typified combat in World War I, while also offering a unique view on types of combat seldom thought of by anyone not intimately knowledgeable of World War I history. The team wanted to create a high-quality representation of Vauquois that could be experienced by anyone, anywhere.

3.2.1.2 Data Capture

Our field team employed a terrestrial LIDAR (Light Detection And Ranging) unit (FARO scanner) as well as photogrammetry both above and below ground to acquire the data needed to create the virtual environment representation. Over a six-day period, the team scanned three different tunnels with multiple entrances from both the German and French sides including above ground entrances and side trails. Over 381 LIDAR scans were taken. They also utilized a photogrammetry process based on thousands of photos taken with an SLR camera for difficult tunnels where the scanner could not reach because of small crawl spaces or difficult ground surfaces.

Both photogrammetry and laser scanning generate a 3D point cloud. Multiple point clouds from both techniques were combined (by matching at least three points from adjacent scans) to create a detailed and accurate structure of the tunnels and trenches. It took over four months to create the combined file, which had a size of 210 GB and over 3.5 billion points.

As part of data collection at the site, the field team also took hundreds of photos of items we wanted to create 3D models of in the future; including carvings, listening devices, lanterns, electrical switches, bedding, tables, and cutlery. Finally, multiple audio recordings of not only ambient sound in the tunnels, but also oral histories by local historians were collected.

3.2.1.3 3D Modeling

Knowing the project was going to be used in VR, the modeling team had to balance realism and real-time rendering requirements. We used game modeling techniques to reduce the models' size and weight. We created low-polygon geometry and placed normal, specular and diffuse maps on each object to create a realistic illusion of the space without slowing down the frame rate.

For the VR tunnel experience, we had to find a part of the actual tunnel that was relatively straight. Then we processed the point cloud data of that section into a dense mesh. The team then cleaned up the surface of this heavy mesh and retopologized it into a low-polygon structure. Using ZBrush software, we baked the dense mesh onto the low-poly mesh, creating a normal map. This produced 4K maps to use in the game engine. The scanner yielded poor RGB data for the wall surface, but we were able to use photos taken by team members to create the texture of the tunnel interiors. Using Zbrush and UV software, we were able to create a 4K texture map without seams to be used in the game engine.

We also modeled over 40 props to be used in the VR space. Many of these were hand modeled in Maya and ZBrush and textured in Photoshop. Some objects took up to 12 hours each to hand model.

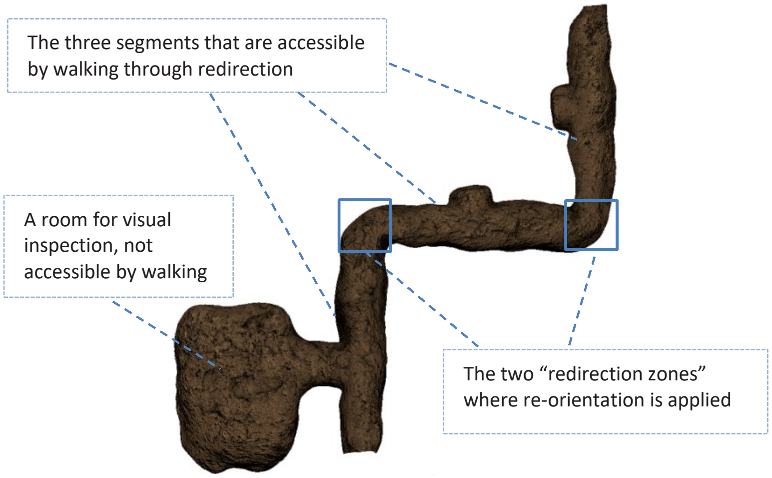


Figure 3-14: Overview of the virtual tunnel layout. The tunnel comprised three straight segments connected by 90-degree turns.

3.2.2 Narrative-Driven Redirected Walking

3.2.2.1 Experience Goals and Design Constraints

The overall goal of this VR experience was to create a convincing illusion of walking through the tunnels of Vauquois and ultimately to gain a sense of what life was like for the people who lived and fought there. The hope was to provide this feeling through as natural an experience as possible, with the narrative elements and objects telling the story, and the physical movement conveying some sense of the confining space and elaborate network of mines navigated daily by the soldiers who lived there. Educational information should, in our view, be naturally conveyed through the realistic feeling of being placed in an authentic site of history.

The experience was designed to open up opportunities to develop a sense of place and perspective regarding the day to day experiences of soldiers (and civilians) during World War I, and to provide participants, who had minimal knowledge of World War I, insights into its impact and effects on French citizens and the practice of warfare.

Our redirected walking design needed to serve these goals by meeting two requirements. First, introducing any overt disruption that might cause a break in presence was likely to be undesirable. The redirection needed to be seamlessly embedded into a continuous narrative and the virtual environment itself. Second, objects in the scene should not deviate too much from historical facts, so we had little freedom in adding metaphors that might facilitate redirection, such as vehicles or spinning bookshelves.

More design constraints came from the real-world setting of the exhibition. We only had a small and narrow physical space to create the entire walk-through experience (15 feet by 5 feet). The width of this real world space roughly matched the width of the tunnel itself, but the length was only enough to cover a very small segment of it. Meanwhile, we were expecting dozens of visitors to go through this demo in one day, so completion time needed to be fairly short for each individual.

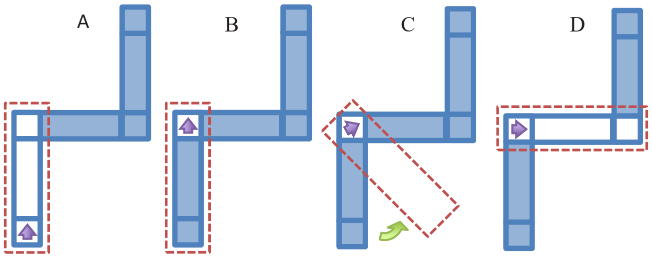


Figure 3-15: The reorientation process. The dotted red rectangle represents the physical space, while the shading of each tunnel region marks whether it is lighted. A. The user is facing a new segment that is lighted up and accessible by walking; B. After she enters the redirection zone, all other regions' lights are turned off, so she needs to inspect the surrounding environment and search for the light switch; C. As she turns her head, the entire virtual world is rotating around the center of redirection zone relative to the real world; D. After the redirection is done, the next segment is aligned with the physical space and lights are turned on for this new area.

3.2.2.2 Reorientation Technique

Redirected walking techniques create walkable space beyond physical constraints by manipulating the mapping between the physical and virtual worlds. In our case, we were facing a severe spatial constraint—the width of the narrow workspace, which was only around 5 feet. Since the virtual tunnel was also thin and had a similar width as the physical space, aligning one small segment of the virtual tunnel with this physical corridor was the natural choice of mapping. Following this idea, we first broke the tunnel into multiple segments, each having the same size and shape as the tracking space. This facilitated a naïve 1:1 mapping between one such segment and the physical space, so that the user could naturally walk through this part of the tunnel using real walking without any redirection.

Now it was the redirection technique's job to enable walking beyond one such segment. The goal was essentially to make the user walk back-and-forth within the tracked area while creating the illusion that he was traversing multiple segments that composed the entire tunnel. Note that this is very similar to the fire-drill task in the original RDW paper, in which the user was to walk in a zig-zag pattern through several line segments connected by waypoints in the virtual world while walking back-and-forth in the smaller physical space (Razzaque et al., 2001). However, we had a more severe design constraint because of the five-foot width of the workspace—this space was so narrow that the user could easily walk out of it with one or two steps. Any mapping that misaligned the physical space with the virtual tunnel segment, even for a short period of time, would risk the user walking into the boundary. Thus, Razzaque's algorithm (Razzaque et al., 2001) was incompatible with our setup, as their technique redirected the user continuously and dynamically during walking. It relied on the fact that even though the user might deviate from the ideal path, the algorithm would redirect him back in the next few steps. In fact, Razzaque et, al. explicitly mentioned that they had to increase the width of the tracking area from 3m to 4m to prevent the algorithm from failing (Razzaque et al., 2001).

Our solution was to separate walking from redirection: a reset event that re-oriented the user was created at each joint that connected the current and next segments. The goal was to make the user stop at this joint and physically turn back (by 180 degrees) to the tracking space while convincing him that he was about to walk further into the next segment of the virtual tunnel. After each redirection event, the next segment was required to be aligned with the physical tracking space, so the entire new segment became walkable. By repeating this walk-reset-walk pattern, the user was always prepared to enter the new segment; thus, a tunnel that was much longer than the physical workspace could be traversed by walking.

Because we intended to make this reset event subtle and unnoticeable, it was desirable to keep the rotational gain small. With this in mind, the segments of tunnel were connected with 90-degree turns rather than as a straight line (Figure 3-14 and 3-15). Using this design, only 90 degrees of added rotation was required to make the user face the next segment in the virtual tunnel (a 90-degree virtual turn) while turning back to tracking space in reality (a 180-degree physical turn), instead of 180 degrees of added rotation if the tunnel was entirely straight.

A square shaped "redirection zone" was created at each joint corner connecting two segments, as shown in Figures 3-14 and 3-15. Once the user entered this zone, the entire virtual environment needed to rotate around the center of the redirection zone. After 90 degrees of such rotation, the new segment would be aligned with the physical space, marking the completion of the reset event (Figure 3-15). Note that since the redirection zone was pretty small, the user was always standing roughly at the pivot point of rotation during the reset event.

The key issue now was how this extra rotation should be applied over time. To determine this function, we drew inspiration from the original RDW paper, in which the maximum of the following three components was used as the basis of rotational gain (Razzaque et al., 2001):

(a) A small, fixed amount of baseline rotation that was constantly applied

(b) A component of rotation based on the user's walking speed;

(c) A rotation that was proportional to the user's angular velocity when turning her head.

Among these three, (b) was not applicable in our case, since our intention was to separate re-orientation and walking. We tested (a) and (c), both combined and individually, with lab visitors and among ourselves, and collected informal feedback. What we found was that (a) always made the re-orientation very noticeable, as most users reported that the rotational distortion became especially obvious when their heads stayed still while the world was slowly rotating. Using (c) alone, on the other hand, didn't suffer from this issue and was perceived as much less noticeable. Based on these observations, we decided to use a rotation that was applied only while the user's head was turning in the redirection zone, with a rotational speed proportional to the speed of the head's yaw rotation.

Next, we needed to determine the actual value of this turning speed. Making this rotation too fast would, obviously, make the reorientation very noticeable, while making it too slow would require too much head turning to complete one reset event. Based on empirical tests with a range of values, we found that Equation (3-1) provided a good balance between keeping the manipulation subtle while demanding a reasonable amount of head rotation.

$$r = 0.1 \times |y| \tag{3-1}$$

In this equation, r was the amount of rotation applied for reorientation during every frame update, and y was how much the head rotated on the yaw dimension in that same time step (the relative rotation between the previous frame and the current frame). Note that r was always in the direction the world needed to rotate to align the next segment of the tunnel with the tracking space, while y could be in either direction. Thus, when the user turned her head in one direction, the rotation would appear to be amplified by 10%, while in the other direction the rotation would appear to be scaled down by 10%.

3.2.2.3 Embedding Reorientation into the Narrative

Figure 3-16: A. View as the user approaches the first redirection zone. The virtual tunnel continues around the corner to the right, but that next segment is dark until the reset has been completed; B. User reaching with a virtual hand to turn on the lights in the next tunnel segment.

Two requirements needed to be fulfilled during the reset event before moving on to the next segment:

(a) Once the user has entered the redirection zone, he must not exit it before the reorientation is completed.

(b) 90 degrees of added rotation needs to be applied in order to align the next segment with the tracking space. This requires the user's head to accumulate 900 degrees of yaw according to (3-1).

A naïve solution to (a) would be adding barriers (e.g. walls, obstacles) in the scene to stop the user from exiting the zone. However, these barriers would need to be dynamic (e.g., the obstacle between the redirection zone and the next tunnel segment should be removed once the reset was done). Such moving barriers would not be historically accurate. We also considered adding explicit visual or auditory instructions to regulate the user's movement during the walk-reset-walk process. This was not desirable either, as it could introduce disruptions to the otherwise seamless VR experience.

To fulfill requirement (a) without breaking the environment or the flow, we decided to constrain the user's movement by manipulating the lighting—to stop the user from entering a certain region, we took advantage of people's natural impulse to not walk into a dark, unknown space by turning off all the lights in that area. Compared to the previously mentioned two solutions, turning the lights on and off was more consistent with the historical context, as the tunnels were indeed illuminated by sometimes unreliable man-made lighting systems.

The manipulation of lighting was designed in the following way: the new tunnel segment beyond the corner was initially unlit and almost completely dark, and the previous segment immediately grew dark once the user entered the corner for redirection (see Figure 3-15, A and B and Figure 3-16A). This effectively stopped the user from stepping into the unwanted areas. Only after the needed rotation had been accumulated would it be possible to light up the new area, while the previous segment would always stay dark. This naturally guided the user to walk into the lit, new segment that was by then mapped to the tracking space, as shown in Figure 3-15, C and D.

In order to accumulate 900 degrees of rotation (requirement b), we had to motivate the user to keep turning her head. The first solution we considered was adding an animated item that drew the user's attention, like a visual distractor (Peck et al., 2009). Again, it was difficult to design this type of distraction without breaking the authenticity of the scene or the intended narrative.

Another way to encourage someone to turn her head was to give her something to search for, as one would have to look left and right in order to locate an object that was difficult to spot. We found this idea particularly appealing: by asking the user to search for a hidden light switch, we could integrate this motivation for head turning into the lighting mechanism mentioned above.

First, the user was verbally informed beforehand that there would be dark hallways and that she would need to find a light switch at each corner in order to continue. When she entered the redirection zone, the lights in the previous segment were turned off automatically, while the lights in the next segment would need to be turned on manually with a light switch. This switch was initially invisible, encouraging users to turn their heads repeatedly searching for it. Once the user had accrued enough added rotation, it was made visible, but was deliberately placed in a location that would be difficult to spot initially (e.g., a dark corner much lower than eye-level). In this way, the user would be less surprised to see it. Instead, she would be likely to believe that the switch had always been there and that she couldn't find it because she didn't look carefully enough in the first place. The user could then reach out with a tracked handheld controller, represented as a virtual disembodied hand, to touch the light switch and turn on the lights of the next tunnel segment (Figure 3-16B). In case the user couldn't find the light switch even after the redirection had been completed, the lights in the next hallway were automatically turned on after five seconds as a fail-safe solution to let the user continue with the experience.

To further help the user accumulate enough rotation naturally, two additional narrative-driven tools were implemented:

- The redirection zone was filled with artifacts of historical value as shown in Figure 3-12C and Figure 3-16A, specifically designed to encourage users to stop and look around, causing them to stay within the redirection zone and rotate their head repeatedly.
- An audio clip of narration was played to direct the user's interest to those objects, again, to encourage her to stay there and look from various angles.

Note that the light switches were not actually present in the historical site. Instead, they were tools to facilitate the redirection technique. However, the entire lighting mechanism was seamlessly embedded as part of the narrative, and it didn't break the fundamental authenticity of the virtual environment. Furthermore, the light switch task not only stopped users and encouraged them to turn their heads, but also made the VR experience more interactive. Users were not simply viewing the virtual tunnel passively; they took on a more active role in the exploration. We expected this narrative element to further increase user engagement and plausibility of the virtual environment.

3.2.3 Implementation

The system hardware was built around the HTC Vive head-mounted display. The Vive was used because it provided the ability to track the user in the entirety of the available 15’x5’ exhibition space. Two Vive lighthouses were placed on either end of the tracking area, approximately 15’ apart and 9’ above the floor. Users wore an MSI VR One backpack computer which drove the simulation and connected to the Vive HMD. All cables were connected directly to the backpack, allowing the user to walk untethered through the tracking area (and potentially out of it). A docent walked next to each user to ensure they did not leave the tracking area, and to answer questions if necessary.

The virtual experience was rendered in the Unity game engine. The textured models of the tunnel and interior props were each imported into Unity separately, as well as audio files with narration and sound effects, and an equirectangular image captured outside the exit of the actual tunnel. In Unity, the assets were arranged inside the tunnel to create a scene reminiscent of what the tunnel might have looked like during active use in the war, and to facilitate narrative-driven reorientation. Reorientation was implemented in Unity scripting to control the virtual camera assets and interface with the HTC Vive system software.

At the midpoint of the first hallway, there was a connected side room approximately 6’x10’ which soldiers used as a sleeping quarters (Figure 3-12D). In the virtual re-creation of the room, it was filled with a bunk bed, food sacks, lanterns, a gun, and several other props. It was included to give the sense that soldiers lived in the tunnels, rather than treating it as a transient space. However, the room was outside the usable tracking area. To discourage users from exploring the room and walking out of the tracking area, the virtual doorway was blocked with pipes, crates, and shovels.

Atmospheric lighting and sound effects were added to increase immersion. Virtual lights were added inside props of candles and lanterns. The lights flickered and occasionally dimmed in order to create the sense of real light sources in the tunnel. Field recordings of ambient sound recorded at Vauquois could be heard, layered with foley sounds of distant explosions and dripping water. A recording of footstep sounds on rock was played when the user’s movement delta was high enough to indicate probable walking. In two areas of the tunnel, the user’s presence triggered an explosion event designed to represent a mortar shell exploding directly above the tunnel, with a loud boom, shaking of the virtual camera, and a dust and rock particle effect coming from the ceiling directly above the user’s perspective. Short voiceover clips were triggered as the user walked the tunnel, describing what he was seeing and how soldiers described their experience in the tunnels during the war.

3.2.4 Observations

The project was exhibited at the Institute for Creativity, Arts and Technology (ICAT) of Virginia Tech during the “ICAT Day” event in Spring 2017, in which collaborative interdisciplinary research projects were demonstrated to the public. The VR experience was a component of a larger exhibit that included a large physical diorama of Vauquois Hill upon which animation and maps of the site were projected, surrounded by a 32-foot cylindrical screen displaying 360-degree video of Vauquois along with period photographs and documents as well as photos taken during the team's field work. The exhibition of this VR project ran for 2 hours and 45 minutes in total, and dozens of visitors experienced it in turn. It took each individual around 5-10 minutes to walk through the entire tunnel.

Our observations indicated that the VR experience as a whole was very effective. Users' behavior was indicative of a strong sense of presence. For example, many users ducked their heads to avoid the low ceilings of the virtual tunnel, and they often stopped and exclaimed verbally when the simulated mortar shell exploded, dropping rock and dust. We asked several visitors about their impressions of the experience and about what they learned. Responses included:

- *I think the VR exhibit kind of gave me a better sense of what ... the underground warfare was like, how terrible and cramped and miserable that was.*
- *I thought the use of the virtual reality was really unique and really inventive and really made it a very immersive experience and made the history really come alive.*
- *I think in the history classes I have taken ... you don't really think as much about the people that were there and you don't really think about the human aspect of it, because it is not presented to you in a human way, it is presented to you as facts and figures and dates … I loved that I walked in and I felt like I could have kept walking. I am definitely somebody who ... was very bored by history, because it was so unreachable. And so feeling like you could kinda touch it made it all the more interesting.*
- *When you are able to use technology and bring the lived experiences of soldiers into the classroom, to give it context, it makes history much more meaningful.*

As for the redirection technique, it appeared to be effective in these respects:

- The users were effectively stopped from trying to continue through the tunnel or going back to the previous region when they reached the redirection zone. They found it natural to stop and look for the light switches, and no one tried to enter the dark areas.
- Users were generally interested in the objects placed in the redirection zone along with the audio introduction of these items. Most of them were willing to spend a little extra time there inspecting the environment while looking for the light switch, which helped to accumulate enough head rotation in a reasonable amount of time.
- For those who spotted the light switch soon after enough head rotation had been accrued, the experience worked as we intended—they could interact with the light switch and naturally continue to the new regions seamlessly.
- Many expressed amazement that they could have walked so far in virtual reality when they later understood the path they had walked in physical space.

For some of the users, the experience didn't work completely as expected, which indicated the drawbacks of our approach:

- Some couldn’t locate the light switch in time, and the lights in the new region were still turned on as a fail-safe measure (as mentioned in Section 3.1.2.2). This introduced some confusion, as they were expecting to manually turn on those lights. Most of these users explicitly asked whether they should keep walking to this new area even though they had not used the switch.
- The rotational distortion was not unnoticeable for all users. Specifically, some reported that they felt something strange happened in the redirection zones. Among these users, some explicitly told us that they considered that to be a system error from the instability of the equipment (e.g., tracking errors). These cases indicated that it might be necessary to improve the design of the reset event itself and/or the empirically determined function (1) in computing the rotational gain over time.

3.2.5 Summary

Motivated by a large interdisciplinary history education project, we have developed a unique VR application allowing users to experience what it was like to be a soldier in the underground tunnel warfare of World War I. The unique feature of this experience is the use of subtle redirection in a very small physical workspace to enable natural walking through a lengthy virtual tunnel. We have explored how narrative cues and embedding the redirection into the story can make redirection less noticeable and actually enhance the user’s experience. The design of this application contributes ideas and principles for naturally stopping the user at the boundary of the physical workspace, providing a search task consistent with the narrative that requires the user to turn her head back and forth, applying subtle rotation gains during the search task that reorient the user toward the workspace, and employing visual effects that naturally tell the user when she can start walking again. In addition, the application is an exemplar of a high-quality, multi-sensory VR experience that can enhance engagement, presence, and learning about history.

4. DESIGNING PHYSICALLY COHERENT OBJECT MANIPULATION

Much research in computer graphics aims to create physically based motion that is visually plausible—motion that looks correct given our experience with the physical world and understanding of physical laws (Barzel et al., 1996). It tries to present convincing behaviors of individual entities (e.g., trajectory affected by gravity) and believable interactions between different entities (e.g., collision and deformation). Displaying plausible motion in virtual reality (VR), however, is even more challenging. Unlike animation or traditional video games, VR directly places the user's body into the virtual environment (VE) as another entity in the physical system. To maintain the overall physical integrity, the interaction between the human body and other virtual entities should also conform to consistent physical laws. Generating such conformity requires not only correct visual information of motion, but also corresponding stimulation both on and inside the human body through somatosensory information. This chapter proposes the concept of *physics coherence* to describe the level of compliance to physical laws in VR object manipulation.

Because current technologies cannot provide high-fidelity haptic feedback in all situations (Bowman et al., 2004; Laycock and Day, 2007; Lin and Otaduy, 2008; LaViola Jr et al., 2017), existing VR manipulation techniques usually have to make choices about how they're going to compromise physics coherence. In general, there are two approaches. The first tries to simulate real-world object manipulation using accurate hand tracking and physics-based simulation of forces (pressure and friction) between the object and individual fingers (e.g., (Borst and Indugula, 2005)). With this approach, physics coherence decreases when the hand makes contact with the virtual object—without authentic haptic display of the object's shape, weight and material, the interface generates conflicting information between visual display of motion and somatosensory cues of force. This often leads to interpenetration and other physically incorrect behaviors (Borst and Indugula, 2005; 2006).

The second approach creates an artificial mapping from the hand to the object's placement to provide good utility (Zhai, 1998; LaViola Jr et al., 2017). The typical example is to mimic the "grab and drag" action using a "simple virtual hand" metaphor (Poupyrev et al., 1996; Bowman et al., 2004; LaViola Jr et al., 2017). It treats the entire hand as one rigid body, and manipulation is achieved by attaching the object's rigid body to it. This technique can be extended easily to remote control (Bowman and Hodges, 1997). With such a technique, the controlled object is deprived of its physical properties such as weight and inertia. Its motion is directly set by the hand's position and orientation instead of exhibiting acceleration/deceleration according to its weight. Defining objects' behavior after collisions is also problematic when one of the objects is held rigidly in the user's hand.

It is unclear how such deviation from physical laws affects user experience, and there is a lack of knowledge about how to design physically coherent VR manipulation techniques. This chapter provides a theoretical model, empirical evidence, and design knowledge to address these questions. Our approach is to focus on *remote manipulation* techniques, which preserve some important aspects of physics coherence without simulating direct contact between the hand and the object. This allows us to explore the impact of physics coherence without solving the problematic issues of accurate hand tracking and realistic haptic rendering. The techniques and design guidelines in this chapter are applicable to current mainstream commercial VR devices.

4.1 Defining Physics Coherence and Physics Plausibility

4.1.1 Conceptual Definition

We define the *physics coherence* of a 3D manipulation technique as the degree to which the motion and force displayed by the interface conform to internally consistent physical laws that are reasonable for the presented scenario. Physics coherence is a component of the overall coherence of a VR scenario (Slater, 2009; Skarbez et al., 2017b). Like overall coherence, it is an objective characteristic of the VR system. By analogy, *physics plausibility* would be the user's subjective psychological response to the physical motion and behavior presented by the system—the perception of whether the manipulation is physically believable or not. Physics plausibility is supported by physics coherence, just as Psi is supported by coherence (Slater, 2009; Skarbez et al., 2017b). In this section, we do not attempt to measure physics plausibility directly; rather, we analyze what comprises physics coherence and how it affects broad user experience.

Note that just as coherence does not require the VE to faithfully reproduce reality, physics coherence does not necessarily require the design to blindly follow all aspects of real-world physics (Skarbez et al., 2017b). Instead, it complies to fundamental real-world physical concepts, but with some "wiggle room" to account for differences in setting or story. Here we explain this definition in detail.

There are 3 components of "physical rules" and distinguish those which need to correspond to real-world physics and those which could deviate from it:

(1) The relationship between force and motion. This component should follow real-world physical rules (e.g., heavier objects require more energy to move by the same distance). This core principle is what defines "reasonable physics" for macroscopic objects.

(2) Physics-related parameters of the environment, such as the strength of gravity or friction. This component can change according to the story told. For example, if the VE is in space, the absence of gravity is reasonable.

(3) How the user initiates force in the system. This component could also adapt to context.

For example, the simple virtual hand technique uses a zero-order mapping from the hand to drive the object's movement instead of a physics-based simulation that relies on force. This means the object does not exhibit realistic acceleration and deceleration when moving. Instead, it would have "jumpy" movements, as if it has negligible weight and resistance. Visually, this could still have high physics coherence if the object is presented as being small and light, and the technique uses an isomorphic mapping that presents it as being held in the hand, because such movements look appropriate in the scenario where the force exerted by the hand is overwhelmingly large compared to the object's weight. However, the same movement would be less coherent visually if the object appears to be heavy and/or the technique is used remotely since it breaks component (1). A more appropriate presentation or control model might be needed to make it physically coherent in this scenario.

This also means that "magic" techniques that do not mimic real-world interactions can be physically coherent, as long as the core principles that explain motion under the influence of forces are respected in the presented scenario. For instance, the ability of "telekinesis" is often depicted in movies as a superpower to remotely control an object. It typically relies on the idea that the user can remotely apply a certain amount of force to the object—a phenomenon artificially designed by altering component (3). However, even if it is unrealistic, movies usually take great care to depict the motion of the object so it appears to be driven by real forces, so that component (1) is respected. It usually demonstrates appropriate acceleration, momentum and resistance, along with correct behavior when colliding with the environment. In other words, the core algorithm that drives its moment-to-moment action still conforms to physical laws. In this way, the scenario becomes believable even if it goes beyond reality

4.1.2 Operational Model

In order to have a clear idea of how to achieve physics coherence in interaction design, here we propose an operational model that describes the components comprising physics coherence in a manipulation technique.

We limit the discussion to interfaces that are based on hand input, since the hand is the most common agent for object manipulation. While a complete model would include all human sensory channels, we only discuss a subset most relevant to manipulation in this document. We focus on the two sensory channels that are strongly related to force and motion for hand-based interaction, namely the visual and somatosensory channels. Finally, we focus on the manipulation part of the interaction, assuming the object in interest is already selected by separate selection techniques (Bowman et al., 2004; LaViola Jr et al., 2017).

In object manipulation, physical laws are demonstrated through the temporal characteristics of motion and force, along with the interplay between the two. Motion and force information is displayed through multiple sensory channels in the real world. Following this observation, our model consists of two parts. The first is **physical correctness of information on individual sensory channels**, which describes the temporal coherence of the information on each sensory channel based on the physical rules of the scenario. The second is **inter-sensory physical consistency**, which describes how well the cues from multiple sensory channels match each other according to those rules.

4.1.2.1 Physical correctness of information on individual sensory channels

The first part of the model has two components, one for each of the relevant sensory channels.

Temporal coherence of visual cues. This component describes to what extent motion presented by the VE conforms to physical laws, judging by visual cues alone. This is similar to the pursuit of "plausible motion" in animation research (Barzel et al., 1996; O'Sullivan et al., 2004). Achieving high coherence for this component usually means employing physics-based algorithms to drive the moment-to-moment actions of the virtual objects. When developing VR interaction using a game engine, this is typically achieved by using the physics engine's simulation thread to realize force-driven motion instead of directly setting the objects' positions and orientations in the rendering thread.

Temporal coherence of somatosensory cues. Somatosensory information conveys motion through sensation of touch and force (LaViola Jr et al., 2017). Inside the human body, it includes kinesthetic and proprioceptive cues, which involve forces from muscular tension and position and angle of body joints (Dickinson, 1974; Goldstein and Brockmole, 2016). In this document, we refer to all these internal cues as kinesthetic cues (Goldstein and Brockmole, 2016). Outside the body, somatosensory cues include haptic sensations imparted by the external environment, which we call external haptic stimulation (Lin and Otaduy, 2008). Note that these two parts cannot be separated into two independent sensory channels, as the external forces of haptic stimulation always result in corresponding internal forces experienced through kinesthetic cues (Stuart, 2001; LaViola Jr et al., 2017). The temporal consistency of somatosensory cues evaluates if these sensations by themselves conform to the physical laws in the presented scenario. For example, suppose the VE presents a scenario in which a virtual object is placed in the user's bare hand. A system of high somatosensory coherence would display its correct weight, shape, material and resistance through haptic rendering as she (kinesthetically) moves it around.

An important note is that somatosensory cues can be generated not only by adding active haptic stimulation using external devices, but also by creating kinesthetic cues through altering the user's action (i.e., moving the hand generates kinesthetic information that the user can interpret as force or weight). Since our research does not focus on high-fidelity active haptic rendering, we try to bypass the need for it by using different control metaphors (technique "FP" introduced later), by altering the user's physical input to create kinesthetic cues (technique "PCVH", "FP", "TB" and "SL" introduced later), or by rendering low-fidelity haptic stimulation (e.g., using vibration on a controller to roughly present some haptic events, see technique "TB" and "SL").

4.1.2.2 Inter-sensory physical consistency

The second part of physics coherence evaluates if motions and forces displayed through multiple sensory channels match each other according to physical laws. Since we focus on only two sensory channels, this part only has one component, namely **inter-sensory consistency between visual and somatosensory cues**. In object manipulation, an important consideration is the correlation between the object's (change in) visual motion resulting from the user's manipulation and the corresponding force felt through the human body. In the simple virtual hand technique, no matter how heavy and how large the object looks, the kinesthetic and haptic cues always generate the same "force feedback" when moving an object a certain distance. Haptic rendering of the object's weight and momentum to match its visual motion would improve the consistency, but this is not always feasible (Lin and Otaduy, 2008). In our work, we instead attempt to generate internal kinesthetic cues that correspond to visual motion using the design of the input-to-output mapping (Sections 4.2.1.3, Section 4.2.1.4 and Section 4.3.1).

Note that this model describes the elements of physics coherence inherent in the design of the interaction rather than the subjective perception of the user, since we treat coherence as an objective property of the system (Skarbez et al., 2017b). Thus, stimulation felt through sensory substitution is not considered as actual information on the sensory channel being distorted, since this substitution happens in one's mind rather than in the VR system (LaViola Jr et al., 2017). Instead, we use the inter-sensory consistency between the two channels at play to capture such phenomena. For example, we consider pseudo-haptic feedback (Pusch and Lécuyer, 2011) to increase the inter-sensory consistency between visual and somatosensory cues.

4.2 Physics Coherence: Design and User Experience

Having defined and operationalized the concept of physics coherence, we now turn to the questions of how physics coherence affects user experience and how to design manipulation techniques with this property. We propose the design of two physically coherent interaction techniques and compare their user experience with a non-coherent one. Two experiments were conducted, one focusing on task performance while the other one focusing on subjective experiences.

4.2.1 Technique Design

In order to understand the influence of physics coherence on user experience, we need to compare two techniques that have different levels of physics coherence while being as similar as possible in all other aspects. To achieve this similarity, they should share the same core control mechanism. We have observed that the simple virtual hand metaphor is very popular among commercial VR applications and is more or less considered the "standard" design. Thus, we designed our two techniques to use this metaphor. The first, which we call "simple virtual hand" (SVH) is a naïve implementation of this metaphor with a very low level of physics coherence; the second, which we call "physically coherent virtual hand (PCVH)" includes many features that add to its physics coherence without deviating from the same core control mechanism.

In order to explore different approaches to achieve physics coherence, we also propose the design of a completely novel technique called Force Push (FP). This technique is specifically designed to preserve physics coherence for remote manipulation (Goldstein et al., 2002). It uses a completely different control metaphor to realize physics-driven movement and display of weight. We hope to gain insights on how to design novel physically coherent interactions by comparing it with virtual hand based techniques.

4.2.1.1 Selection

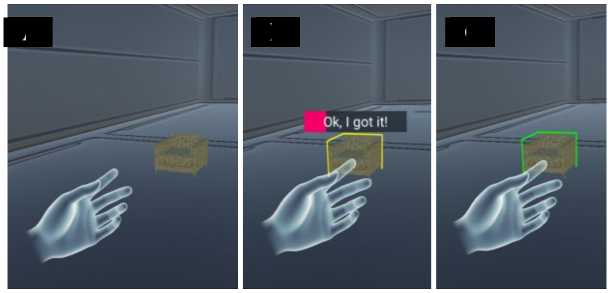


Figure 4-1: The timer-based ray-casting selection process. A: before attempting to select the chair; B: Attempting to select the chair but not yet confirmed; C: Confirmed selection of the chair.

Because we are focusing on the manipulation interaction, we designed a ray-casting based selection mechanism (Figure 4-1) that can be used by all the techniques in the experiment. A ray is cast from the user's dominant eye to the tip of the index finger every frame. When this ray hits an object that is not the currently selected one and the user is not actively manipulating the selected object, the system attempts to select this new one and a progress bar is displayed. If the hand hovers there until it is filled (which takes two seconds), this item is selected and all subsequent manipulation input will be directed to it. Such a selection technique effectively separates the selection process from manipulation.

4.2.1.2 Simple Virtual Hand (SVH)

The first manipulation technique, used as the control condition, is designed to have low physics coherence through a naïve implementation of the "grab and drag" metaphor. The previously described selection technique effectively extends it to be capable of remote control.

The user holds an Oculus Touch controller (displayed as a hand) and uses the index finger trigger to differentiate between two hand postures, namely grab and release. When the trigger is held down, the currently selected object's positional and rotational change follows the user's hand using a zero-order mapping until it switches to a release posture. Note that the action of grabbing and releasing has nothing to do with the action of selecting, since the currently selected object is always the one being moved around no matter where the hand is. Hence, the user can start the manipulation by making the grabbing posture anywhere within the hardware's tracking range. It maps the hand's relative change in position and orientation to the object when the hand is forming the grabbing posture, as if a remote virtual hand appears at the object's location, grabs it and follows the real hand's movement.

The algorithm overwrites the object's position and orientation each frame in the rendering thread that is completely detached from the physics simulation in the game engine. Thus, objects have no gravity and float in the air once they are dropped. They do not have physical rigid bodies, so they can also pass through other objects. The movement uses a constant control/display (C/D) ratio.

We can use the proposed operational model to evaluate the physics coherence of SVH. Since its visual motion is detached from the physics-based simulation of the virtual world, it is missing many important **temporally coherent visual cues***,* such as proper acceleration/deceleration, collision, and momentum. When it comes to the **temporal coherence of somatosensory cues**, the user is always holding the same controller, so the haptic information contradicts the physical nature of the virtual scenario. As for **inter-sensory consistency between visual and somatosensory cues**, the interface always provides the same somatosensory information when moving an object by the same distance visually, no matter how heavy the object is. Overall, SVH generates physically incoherent cues for all of the components of the model.

4.2.1.3 Physically Coherent Virtual Hand (PCVH)

PCVH adds many physically coherent features to SVH without altering the fundamental control mechanism. First, to improve the **temporal coherence of visual cues**, objects have rigid bodies with appropriate weight, and their movements are now controlled by the physics engine instead of the rendering engine. Control is not achieved by directly overwriting the object's position and orientation; instead, the physics engine "drags" the selected body to the location specified by hand input. In this way, manipulation becomes a part of the physics-based simulation that is consistent with the rest of the VE—the controlled object can react to other forces in the environment that might influence its status, such as collision with other objects.

With the Unity physics engine, we found that using a force to make the object follow the hand results in too much lag behind the hand movement, such that the technique did not feel like direct manipulation anymore. After trying various alternative approaches, we discovered that setting the rigid body's velocity to point towards the intended position and orientation results in the desired behavior—the object closely follows the hand input as if using a zero-order mapping, and can still behave properly when colliding with other objects according to their weight and momentum. For example, the object can be stopped if a much heavier object is in its path, or it can knock over a lighter object in a realistic manner. Granted, this is not true physics-based simulation that generates motion from forces, and the resulting movement does not conform to physical laws in every aspect (e.g., the acceleration/deceleration is technically inaccurate according to the object's weight). However, we felt this was a reasonable compromise to add many physically coherent features to the motion without breaking the input-to-output mapping of the control metaphor.

Besides collision, gravity is enabled for all rigid bodies when they are not directly grabbed, so they drop to the ground when released. We also give them momentum; if an object's speed is larger than a certain threshold at release, we maintain that speed and let the physics engine simulate its movement. This enables the action of "throwing" the objects around.

PCVH also improves the **inter-sensory consistency between visual and somatosensory cues**. Here, we adopt an established method from pseudo-haptics research. Specifically, we apply the technique of changing C/D ratios based on object weight (Dominjon et al., 2005) to form a relationship between kinesthetic cues and visual display that reflects physical laws. With this method, heavier objects have larger C/D ratios, while lighter ones have smaller C/D ratios. To move a heavier object by the same distance visually, the user would need to move his hand by a longer distance because of its larger C/D ratio, which scales up the force sensations the user experiences as internal kinesthetic cues. This display of weight variation effectively creates a physically based connection between visual information and somatosensory information. All the enhancements applied to PCVH to increase its physics coherence are summarized in Table 4-1.

	physical correctness of information on individual sensory channels		**inter-sensory physical consistency**
	visual cues alone	somatosensory cues alone	visual – somatosensory consistency
SVH	baseline	baseline	baseline
PCVH	the object has a rigid body in the physic engine, which enables collision, momentum and gravity	baseline	changing C/D ratio according to weight results in a physics-based correlation between kinesthetic cues and the object's visual motion
FP	native physics-based motion driven by force	bypassed by the metaphor	mapping the amplitude and speed of gestural input to force generates a physics-based correlation between kinesthetic cues and object's visual motion
TB	native physics-based motion driven by force	vibration played based on how hard the user hits the tennis ball	mapping the direction and how hard the hands "hit" each other to force creates a physics-based correlation between kinesthetic cues and the object's visual motion
SL	native physics-based motion driven by force	vibration played based on the tension of the rubber bands	mapping the relative positions and distance between two hands to force on the object creates a physics-based correlation between kinesthetic cues and object's visual motion

Table 4-1. Physics coherence of the techniques based on the operational model

4.2.1.4 Force Push (FP)

Section 4.2.1.4 and Section 4.2.2 originally appeared in the poster, ***Yu, R.**, and Bowman, D.A. (2018). "Force Push: Exploring Expressive Gesture-to-Force Mappings for Indirect 3D Object Manipulation", in 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). 733-734.* It also appeared in the journal paper ***Yu, R.**, and Bowman, D.A. (2018). "Force Push: Exploring Expressive Gesture-to-Force Mappings for Remote Object Manipulation in Virtual Reality". Frontiers in ICT 5, 25.*

Force Push is a gesture-based technique designed specifically to preserve physics coherence for remote manipulation (Yu and Bowman, 2018b; a). At a high level, it is based on a control metaphor inspired by "telekinesis" hand gestures that frequently appear in popular media. The interface reads gestural input and maps it to force-driven movement of the object: pushing or pulling gestures result in translational force in the corresponding direction while drawing circles in the air results in corresponding torque. The quantity of force or torque applied is determined by the amplitude and speed of the gestural input. Unlike the previous two techniques, Force Push doesn't use an Oculus Touch controller. Instead, it uses a Leap Motion sensor to recognize hand gestures. Although this introduces a confound in our experiment, we chose to use the Leap Motion as a more natural and appropriate device for a technique based on fluid hand gestures. Here we introduce the design of this technique in detail.

(1) The Metaphor

Inspired by the magic power of telekinesis in popular culture, such as the *Star Wars* Force Push, we set out to design a remote manipulation technique in VR that was based on hand gestures and a gesture-to-force mapping. In designing the details of the mapping, however, we found additional metaphors with roots in reality.

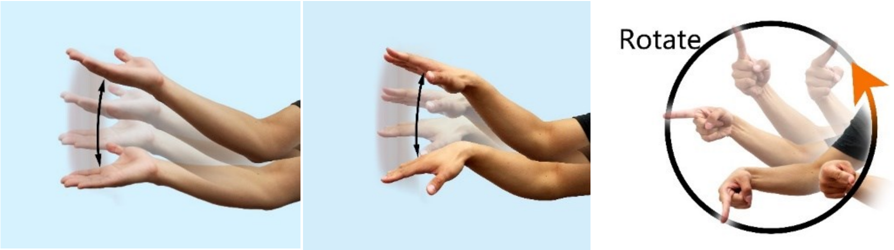


Figure 4-2: The Force Push gestures. A: Translation that moves the object up; B: Translation that moves the object down; C: Rotation

People often use a hand waving gesture to indicate desired movement to others at a distance. The path of the hand motion along with the orientation of the palm typically indicates the intended movement direction, which creates easy-to-understand commands such as "come closer," "go further," or "move more to the left.". Think about the case of communicating with a person parking a vehicle: people naturally wave their hands to express "straight back quickly," "turn slightly this direction," or "slow down." This set of simple gestures seen in human-human interaction are not based on direct movement specification using a positional mapping. At the same time, they are commonly seen and widely accepted, which makes them natural and easy to understand.

Another metaphor based on physical phenomena may be at play in such gestures. The user may think of this as pushing the air with hand motions, so that the airflow moves the target object. One may also think of these gestures as "nudging" actions applied remotely to the target object, similar to the motions used to tweak the position or orientation of an object sitting on a table or floor—rather than picking up the object and placing it down again, the user simply pushes the object with a finger or hand. In both scenarios, the user's movements are being translated into force, rather than positions.

The high-level design of Force Push follows these inspirations and metaphors: with the hand in an open posture (all fingers stretched out), a quick burst of hand movement along the direction that is perpendicular to the palm would translate the object in the same direction. Figure 4-2 (A and B) shows this hand gesture when it's moving the object on the vertical dimension.

Following the same idea of utilizing intuitive gestures, we also implemented a gesture for rotation: a circle drawn by one finger around a certain axis would rotate the object in that direction by applying torque, as shown in Figure 4-2C.

(2) Gesture Detection

In order to formalize the Force Push technique, several important questions needed to be answered:

- How to detect the gesture
- How to determine the exact direction of translation
- How to determine the quantity of translation (distance, speed)

In answering the first two questions, a naïve design would allow “pushing” at an arbitrary direction: Any time a hand moves along its palm’s normal, the gesture would be activated and start translating the object along that direction. This naïve interpretation would create too many false positive cases for the gesture recognition algorithm, as we often accidentally move a hand along the direction perpendicular to the palm without the intention of activating the gesture. Moreover, we found that people tended to be inaccurate when defining a direction through either a trajectory of hand movement or the palm’s orientation in our early tests. To overcome these issues, we added some constraints to the possible translation directions that made it easier to control:

- Limit the gesture’s moving directions to be along the left/right and forward/back axes of the camera’s coordinate system, plus the up/down axis of the world coordinate system. From the user’s perspective, this limited the choice of directions in translating the object to six directions that were aligned with his egocentric point of view.

- Among these six options, the one that was closest to the direction that the palm was facing (normal of the palm) was chosen. This effectively “snapped” the hand gesture direction to the closest option that was aligned with the egocentric view.

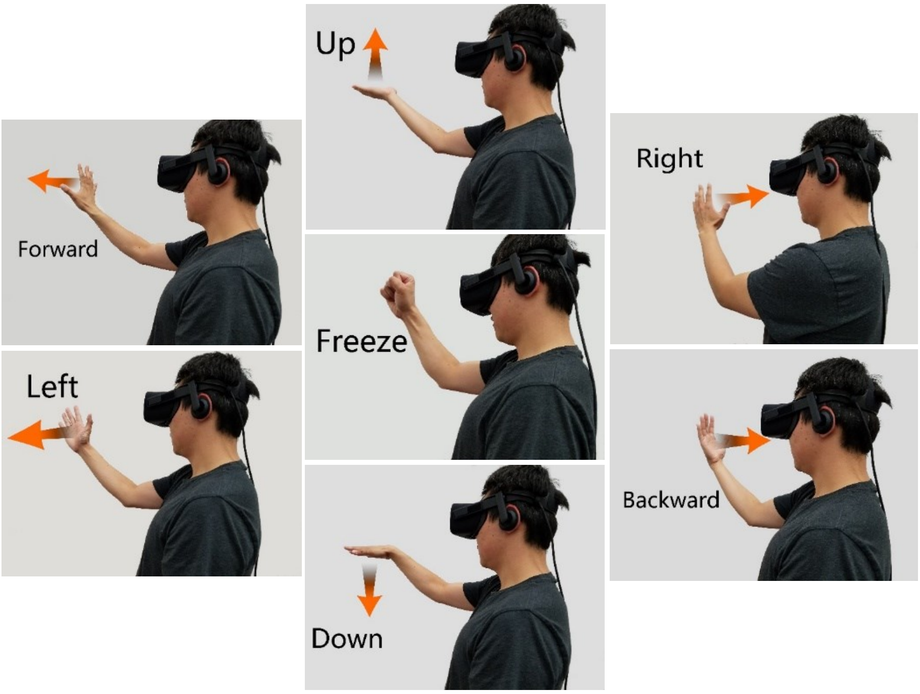


Figure 4-3: The direction the palm is facing defines the translation direction of the manipulation. The “freeze” gesture stops the object from moving. (The depicted individual is the author of this document and these images are published with his consent)

Figure 4-3 shows how the user controls the translation direction of the object using hand orientation and hand choice (the right hand is used to make the left translation gesture, and vice-versa). With the constraints added to gestural movement directions, a simple gesture recognition algorithm was implemented: a ring buffer was built to record the most recent 10 frames of hand motion; it pushes in a new frame’s data every 0.1 second (a sliding window of 1 second). The gesture detection algorithm continuously analyzes the recorded hand motion and determines whether to activate the gestural interaction based on two criteria:

- We first examine whether the palm is generally facing one of the six directions aligned with the egocentric view. This is done by examining the similarity between the palm’s normal and these directions. In every frame, we calculate the Euclidean distance between the value of the palm’s normal and the value of the normalized vector pointing to each of these six directions. We sum this distance over the past 10 frames and see if it’s smaller than a certain threshold. This threshold was empirically determined to be 8.0 in the Unity engine.
- If the above criterion is met, we inspect whether the fingers’ motion exhibits a burst of movement along that direction. To do this, we examine the similarity between a long vector pointing to this direction and the fingers’ relative translation. Specifically, we calculate the Euclidean distance between the value of a normalized vector (with a length of 1 meter in Unity) pointing in that direction and the value of each finger tip’s translation in every frame.

We sum this distance over all five fingers and over the past 10 frames to see if it's larger than a threshold. This threshold was empirically determined to be 49.5 in the Unity engine.

(3) Physics-Driven Control

Having determined the gesture detection algorithm and the direction of object translation, next we needed to decide how to map hand movement to the movement of the object. In an earlier gesture-based manipulation prototype, we actually started with a position-to-position mapping: the travel distance of the hand along the palm's normal direction was applied to move the object by the same distance. However, we found that this naïve mapping created a jerky movement for the object. As the hand traveled forward, for example, the object would be translating forward at the same pace, while as the hand travelled backward to reset its position, the object would stop moving. This was considered undesirable and hard to control.

To improve the interaction experience, we drew upon the force-based metaphors described before and applied a constant amount of force to the object, driving its movement using a physics-based simulation every time the gesture was activated. This seemingly simple improvement was actually a fundamental change to the interaction design: the interface was changed from position-based direct mapping to acceleration-based indirect control.

Since this manipulation interface was driven by physics, inertia would keep the object moving for a while even when the gesture stopped. This made it hard to control since one could not stop the movement when the object reached the target location. Here we applied two improvements. The first one was adding resistance. In the Unity3D game engine, this was done by setting a positive value to the "drag" parameter of the rigid body. With resistance, the object would gradually slow down and stop its movement after the gesture stopped. We also implemented another static posture called "freeze" that would stop the object immediately from moving further: whenever the user closed his hand to make a fist (as shown in the center of Figure 4-3), the object would instantly stop. We found that the addition of the freeze gesture greatly enhanced the controllability since the user could stop the object at a desired location any time.

We let a number of users try out this prototype and collected informal feedback. No one had any difficulty in understanding the interface and control. Most people felt the interaction was fun and magical, as it gave them a "super power" that was similar to what they see in popular media. Some users also told us the physics-based smooth movement added to the user experience as it made the object feel more realistic. All this early informal feedback indicated that Force Push might provide desirable hedonic qualities compared to traditional direct control techniques. However, we observed that users still had trouble moving objects accurately to the desired location.

(4) Expressive Gesture-to-Force Mapping

Previous research told us that acceleration-based indirect manipulation provides significantly worse controllability than position-based direct manipulation (Zhai, 1998), and Force Push was no exception. To make things worse, the technique could only apply a constant amount of force once activated. Without means to control how fast the object moved, one had to rely heavily on the freeze gesture to try to stop it at the right time. In order to provide a good balance between hedonic qualities and task performance, we needed a more effective method to define the quantity of translation.

Real-world object manipulation tasks may pose different requirements. The two main factors here are the initial distance between the object and the target location, and the required accuracy (i.e., how close does the object need to be to the target for the task to be considered completed). According to Fitts's law, the difficulty of an aimed movement task is determined by the amplitude (distance) and the target width (required accuracy). In certain difficult situations, even a direct, position-based manipulation can become cumbersome. If the target is far away, the task demands a small C/D ratio (i.e., a small movement of hand mapped to a large movement of the object). Otherwise, the user will have to repetitively clutch (grasp, move, and release) the object, which is time consuming and fatiguing. On the other hand, when the required accuracy is high, one needs a large C/D ratio, since hand instability makes it hard to align the object with the target accurately. In fact, these two factors pose requirements that contradict each other. Thus, it is difficult to design an efficient, direct control interface to align an object to a far-away target with very high accuracy. Realizing this, many researchers have designed techniques that attempt to sidestep the requirements of rapid movements across distance, precise fine-grained control, or both (e.g. (Wingrave et al., 2002; Frees et al., 2007)).

We observed that there are some inherent qualities of acceleration-driven control that may help in these difficult scenarios. When the target is far away and coarse-grained movement is needed, inertia may enhance our ability to quickly move the object by a large distance with less user input. With the help of physics and inertia, one can "throw" the object towards the target by applying a short burst of force that results in fast, lasting movement without further user interference. When placing the object with a high accuracy requirement, mapping hand movement to acceleration instead of the positional change of the object serves as a low-pass filter—hand tremor does not directly result in shaky movement of the object (Zhai, 1998). These qualities of a force-based mapping could help us improve the performance of Force Push.

In order to meet both the requirements for coarse-grained movement and fine-grained alignment, the object's physics-based movement needed to be adaptive to these two ends of the spectrum. Moreover, the adaptation must happen based on the user's intention. Fortunately, we found that users tended to naturally express their intention into the dynamic gesture of Force Push itself. When they wanted to move the object by a large distance, they tended to extend the range of their arm/hand movement and increase the speed of the repetitive push gesture. When they hoped to manipulate the object precisely, they often resorted to small-ranged finger movement, with a slower pace of gestural input.

This observation told us that a gestural input carried much more information than the simple binary message of activating the action or not. Instead, it contained quantitative information that directly reflected the user's intention. Thus, we could alter the parameters of the interaction in order to meet the user's need, and this adaptation could happen seamlessly during the interaction. Specifically, we could alter the properties of the physics simulation to help the object move faster when we detected that the user preferred coarse-grained, ballistic movement, or we could make the object move slower and more accurately when we detected that the user needed precise control instead.

Conceptually, this is very similar to the PRISM technique, where the C/D ratio of the direct positional mapping is changed on the fly based on the user's hand movement speed (Frees et al., 2007). The difference here lies in the mapping between the input's feature and the output's property. In PRISM, the positional information of input is already fully dedicated to the direct control itself and cannot carry more information. To interpret the user's intention, the technique needs to search in the first derivative space, namely the speed. In the Force Push gesture, however, because we are using a gesture-to-force mapping instead of a position-to-position mapping, the positional information of the hand becomes available to carry the user's preference for the interaction instead of the control itself. Furthermore, since this dynamic gesture is a sequence of movements, its temporal characteristics expand the expressiveness of the feature. In a way, our gesture-to-force design facilitates a richer, more expressive mapping from the user's intention to the properties of interaction.

Assuming the mass of the object to be constant, there were two parameters we could change in the physics engine of Unity3D that would alter how the object moved: the amount of force applied and the physical resistance (drag) of the object. The general idea was simple:

- When the algorithm interpreted that the user wanted to move the object quickly (large range of hand movement with high travel speed), we would increase the amount of force and decrease the drag. In this way, a large amount of acceleration would drive the object to quickly travel a long distance without much resistance.
- When it appeared that the user wanted fine-grained control (small range of hand movement with low travel speed), we would decrease the amount of force and increase the drag. This way the object would move slowly with a small amount of acceleration and the large resistance would stop it quickly when the gesture stopped.

Because of the purely artificial nature of this gesture-to-force mapping, the actual mathematical function to dynamically change these parameters was not based on a real-world model. This implied that we had complete freedom to design it arbitrarily. We developed the equation by trying a variety of different functions and empirically choosing the best one.

We designed the following function to change the amount of force applied in real time when the "pushing" gesture has been recognized. The purpose of this design is to scale the force using both the amplitude of the hand gesture and the hand's travel speed.

$$F = \alpha \times R^2 / (N + 1) \tag{4-1}$$

F (in newtons) is the resulting amount of force; α is an empirically derived parameter to scale the value; *R* (in meters) is the total distance that the hand has travelled along the axis of object translation (the axis along which the "pushing" gesture is performed) within a constant amount of time, as an indication of how much the hand has moved in that direction. To calculate this value, we sum up all five fingertips' forward and backward translation distances along that axis in each 0.1 second interval (1 frame in the gesture recognition algorithm), and accumulate this value over the last one second (10 frames, same sliding time window in gesture recognition). *N* is how many times the fingertips have switched travel direction along the dimension of object translation during the same time period. To calculate this value, we look at the travel directions of each fingertip during the latest frame and the previous frame (again, each frame lasts for 0.1

second). If these directions are opposite to each other along that axis of object translation, the finger has switched travel direction once. We sum up how many times this kind of switch happens for all five fingertips in the last one second (10 frames) as the value of N. To avoid division by zero, we add one to N to serve as the divisor.

This function takes both the travel speed and the amplitude of the "pushing" gesture into consideration. For example, when the hand travels at the same speed but increases the range of movement, R will remain the same while N becomes smaller in the fixed amount of time, which leads to an increased amount of force. On the other hand, if the gesture maintains the same amplitude but the hand travels at a faster speed, R and N will increase at the same rate, which also leads to a stronger force since R is squared.

The implementation directly uses the tracking data given by the Leap Motion API without performing any filtering to account for hand tremor. Although filtering might be more ideal, we found that the lack of filtering does not affect performance much since hand tremor typically results in small travel distances, which will always result in small values using (4-1).

For the drag parameter, we applied a linear function that scaled the value based on the result of (4-1). When the amount of force increased, it would decrease the drag value; when the force was weak, it scaled up the drag. This function is shown in Equation (4-2), in which *D* is the value for the drag parameter; *F* is the real-time result of (4-1); and *a* and *b* are two empirically derived constants that linearly scale the amount of drag based on F.

$$D = a \times F + b \tag{4-2}$$

Both (4-1) and (4-2) are empirically designed functions to achieve our goals, and there are many other possible ways to define a mapping from gesture to force and to set the parameters of the physics simulation. Further research is needed here to gain a better understanding of this design space. Through empirical testing, we found that the mapping and parameters described here provide a good balance of expressiveness, control, comfort, and naturalness.

Similar to translation, rotation is scaled by the size of the circle one draws (in performing the "drawing circle" gesture) and the hand's travel speed. We use the following function to determine the amount of torque applied to the object:

$$T = \beta \times P^2/A \tag{4-3}$$

In this function, T is the amount of torque applied; β is an empirically determined scale parameter; P is the distance that the hand had traveled in the past constant period of time; A is the accumulated turning angle during that same period of time. If the hand travels at the same speed but draws a larger circle, A will decrease while d remains the same, which leads to a larger torque; if one draws the circle with the same size but increases the hand travel speed, A and P will increase at the same rate, which also leads to a larger T since P is multiplied twice.

There is no need to alter the angular drag for rotation, since rotation tasks never require for a large moving distance. The task amplitude is always smaller than 360 degrees, and we believe function (4-3) provides enough expressiveness for such interaction.

(5) Analysis

The most important characteristic of this technique is that it uses a completely different metaphor that does not rely on making contact with the object. Because of this, we bypass the requirements related to haptic rendering for maintaining physics coherence (i.e., the user does not expect to feel forces from hand-object contact). For example, **temporal coherence of somatosensory cues** is naturally kept intact since the scenario assumes the hand doesn't touch anything throughout the interaction. Also, because hand gestures are mapped to force and torque, the movement of the object is directly controlled by the physics engine. By using algorithms specifically based on physical laws to drive motion, a high level of **temporal coherence of visual cues** is naturally realized.

Just like PCVH, this technique also forms a physics-based relationship between visual and kinesthetic cues in order to enhance **inter-sensory consistency between visual and somatosensory cues**. This is because the amount of force (or torque) applied is determined by the amplitude and speed of gestural input—to move a heavier object by the same distance, the user has to output more force kinesthetically by performing hand gestures with a larger amplitude and faster speed (e.g., pushing harder to translate a heavier object). From a user perception point of view, this can be seen as an attempt to realize pseudo-haptic feedback under the specific control metaphor of Force Push. Table 4-1 shows how Force Push implements each component of physics coherence.

Judging by each component in our model, Force Push should have the overall highest level of physics coherence among the three. On the other hand, its force-based control using separated degrees of freedom (users can only translate/rotate objects along one direction at a time) and the learning curve for a novel control scheme is expected to result in lower usability compared to the two zero-order mapping techniques.

4.2.2 Evaluation: Usability of Force Push

4.2.2.1 Goal and Hypothesis

To gain an initial understanding of the user experience of the Force Push approach, we designed an experiment that compared it with conventional direct control manipulation using the simple virtual hand metaphor.

The first and foremost focus of this first study was on task performance, since we were interested in whether this method was usable at all as the gesture-to-force mapping breaks the common zero-order mapping. Furthermore, we were also curious about whether the desirable qualities of the dynamic force mapping might have performance benefits in some cases, as stated in (4) of Section 4.2.1.4. On top of task performance, we also wanted to gather some initial data on subjective aspects of user experience. Thus, the evaluation included both performance and preliminary subjective experience measures.

We proposed these hypotheses:

- As long as the direct control was equipped with an appropriate choice of C/D ratios, it would typically outperform the Force Push interface.

- With Force Push, we expected there to be a few cases where participants could not finish the task in a reasonable amount of time due to the difficulty of learning and controlling the acceleration-based mapping.
- In cases where even two C/D ratios were not enough (e.g., requiring high accuracy while the target was far-away), Force Push might provide performance superior to direct control.
- Beyond performance, we expected that users would perceive Force Push as more natural and fun than direct control.

In the rest of the document, the phrase “direct control”, “direct mapping” or “simple virtual hand” is used to describe an interaction technique with a position-to-position mapping between the movement of the hand and the object using one or more fixed C/D ratios. The name “Force Push”, “FP” or “gesture-to-force mapping” is used to describe our novel technique that moves the object using force-driven physics simulation while adaptively changing its property using features extracted from the dynamic gesture.

4.2.2.2 Experimental Design

(1) Tasks

Note that there are two factors that differentiate the control mechanisms of direct control and Force Push:

a. Integrated degrees of freedom (DOF) vs. separated DOF: Direct control integrates all six degrees of freedom (DOF) together so that the user can translate/rotate the object in an arbitrary direction at any time. On the contrary, Force Push separates the six DOFs of manipulation, so that users can only translate or rotate the object along one of the limited, pre-defined directions at a time, as mentioned in Section 4.2.1.4.

b. Two different mapping functions that determine the quantity of the object’s movement: Direct control uses a zero-order mapping that maps the distance the hand travels (or the angle it rotates) to the distance the object moves (or the angle it rotates). Force Push uses the novel gesture-to-force mapping function that maps the feature of input hand gestures to the amount of force applied to the object, which in turn determines how much it moves at a given time.

As our first step in understanding the usability and user experience of Force Push, we intended to investigate only factor b while eliminating the influence from factor a. In this study, therefore, we did not require the user to define the movement direction and instead focused only on evaluating the novel, gesture-to-force mapping algorithm when that direction is already set. Thus, we limited the task in the experiment to one-dimensional translation.

The user was required to move an object to a target location along a pre-defined dimension (in this case, vertical). Participants were asked to complete the task with a required accuracy as quickly as possible, and we evaluated performance by measuring completion time. A post-experiment questionnaire was used to gather preliminary data of subjective user experience.

Initial Distances	10	100	1000	10000
Required Accuracy	0.001 (i.e., 1 mm)			

Table 4-2: Task Parameters (meters)

Smaller C/D ratio	1:500
Larger C/D ratio	1:1

Table 4-3: Selected C/D Ratios for the Direct Control Interface

In order to cover a wide range of difficulties, we created several conditions by altering the initial distance between the object and the target location (the amplitude in Fitts's Law). The required accuracy—how close the object needed to be to the target for the task to be completed (the target width in Fitts's law)—was kept constant. To test the robustness of the techniques, all task conditions were designed to be difficult with (relatively) large initial distances and a high accuracy requirement. Table 4-2 shows the parameters for the four conditions.

To be fair in comparing the two techniques, they should be equipped with near-optimal parameters. Since we were covering a range of difficulties, the pre-defined parameters we chose wouldn't be perfect for each condition. But we ought to choose them to be as adaptive as possible for all the testing conditions.

Due to the high difficulty of these tasks, the direct mapping interface was equipped with two C/D ratios for the user to switch between, with a smaller C/D ratio for coarse, ballistic movement and a larger C/D ratio for fine-grained accurate alignment to reach the target. The values we chose are shown in Table 4-3. In choosing these values, we took into consideration both the human arm's limited range and hand tremor, so an average person could complete the tasks with a reasonable amount of clutches while hand instability would not keep them from achieving the required accuracy.

An alternative approach here would be to use the PRISM technique to adaptively change the C/D ratio based on the hand's movement speed (Frees et al., 2007). However, the tasks in this experiment were extremely difficult, with requirements for both high accuracy and long travel distance. This means the discrepancy between the required largest C/D ratio and smallest C/D ratio was huge. This considerable gap is reflected by the chosen two C/D ratios in Table 4-3, which are the values that we found could cover both ends of the requirement while keeping this gap as small as possible. If PRISM were to be used, it would map the variation in hand movement speed (which provides a quite limited range) to this huge change of C/D ratio. We doubted that the change in hand speed alone could provide enough expressiveness to precisely reach an intended C/D ratio between the two extremes, and we surmised that it might feel similar to switching between the two extremes since the C/D ratio would change very quickly with even a subtle change of hand movement speed. Thus, we chose to provide the two C/D ratios instead. We will consider a comparison between Force Push and an adaptive technique like PRISM as future work.

For Force Push, we fixed the object's mass at 1.0 units in Unity3D. Equations (4-1) and (4-2) in Section 4.2.1.4 explained the algorithm to determine the amount of force and the value of the drag parameter. By trying a range of options for the empirical constants in the equations, we found the values shown in (4-4) to be a reasonable choice, with which the gesture appeared to be expressive enough to achieve both ballistic movements and precise control.

$$\alpha = 1.0, a = -0.17, b = 87.40 \tag{4-4}$$

(2) Environment

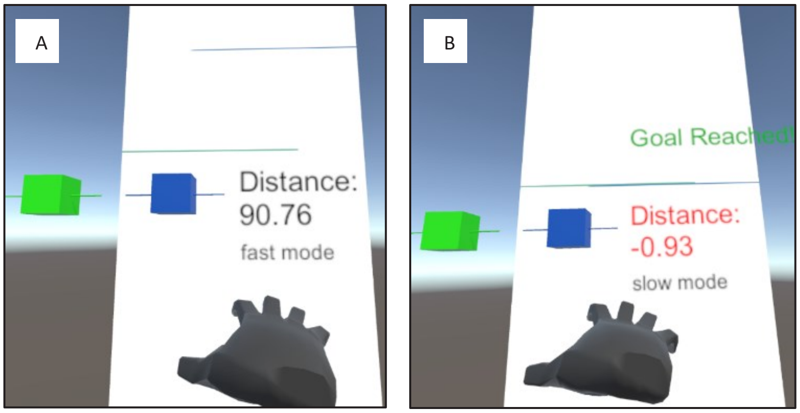


Figure 4-4: The virtual environment. A: The object is above and close to the target; B: The task is complete.

To visualize the task, we designed the simple virtual environment shown in Figure 4-4. The green cube marked the target location and it remained static throughout the experiment, while the blue cube symbolized the object being manipulated. Since we were only examining 1-D translation on the vertical dimension, the blue cube would only move vertically. The task was to move the blue cube up and down until it had exactly the same height as the green one.

Because some of the difficult conditions would place the object too far-away, (e.g. 10000 meters), the displayed distance between the green cube and blue cube was scaled down by a factor of 1000. This meant the blue cube would appear at a location that was 10 meters away from the green one when their (conceptual) distance was meant to be 10000 meters according to the task. To maintain accurate information of the task, the original distance was displayed as a number using a unit of 1 millimeter. The sign and color of the number indicated whether the object was above or below the target, as seen in Figure 4-4. The user could always refer to it to know exactly how far away the object was from the target and which direction to move further. When the magnitude of this number was smaller than 1.0 and the object was released from interaction, the task was accomplished and the timer would stop (Figure 4-4B).

To help the user with accurate alignment when the object got very close to the target, the white board behind the object showed a zoomed in visualization of the two lines attached to the two cubes. The distance between the two lines on the whiteboard was a scaled up version from the two lines on the two cubes. One simply needed to make the two lines overlap on the white board in order to complete the task. Figure 4-4B shows the display when the task was completed.

The text “fast mode” and “slow mode” seen in Figure 4-4 indicated which C/D ratio was currently being used, and was only displayed in the direct control condition. When using Force Push, this text read “using gesture.” As soon as the task was accomplished, “Goal Reached!” would be printed on the screen (Figure 4-4B). A virtual hand was always rendered on the screen to indicate whether the hand was being tracked at this moment.

During the experiment, we spent time explaining this visualization as part of the training session. We found that no one had difficulty in understanding it and could all finish the tasks in a reasonable amount of time by relying on them.

(3) Questionnaire

The questionnaire asked about the participant’s subjective impressions of the two techniques. The questions were designed to gather preliminary information about seven aspects of user experience:

- Accuracy: Which technique did you feel was more accurate?
- Ease of use: Which technique was easier to use?
- Speed: Which technique did you feel was faster in completing the task?
- Controllability: Which technique did you feel more in control?
- Naturalness: Which technique felt more natural?
- Fun: Which one was more fun to use?

The answers to these questions were collected after the experiment was finished. Participants were also welcomed to provide other comments about the techniques they used.

(4) Apparatus

An Oculus Rift CV1 was used for display and a Leap Motion was applied for hand tracking. The virtual environment was developed in Unity 5.5 and the native physics engine of Unity was used to drive the physics-based simulation of the Force Push interface.

With the direct manipulation technique, the user needed to be able to activate the “grab” and “release” actions and switch between two C/D ratios. A naïve implementation would use the grabbing and releasing postures of the bare hand to activate and deactivate control, but we found that Leap Motion would experience significant tracking loss when the hand is closed (in the grabbing posture). Hence, we let the user hold an Oculus Touch controller in her non-dominant hand and use its trigger and one button to realize these functions, while the translation of the object was mapped to the open, dominant hand motion tracked by the Leap Motion. This separation of function might seem counter-intuitive, but it maintains the same tracking accuracy between the two techniques—the actual movement of the object was always driven by the dominant hand with an open posture tracked by the Leap Motion device.

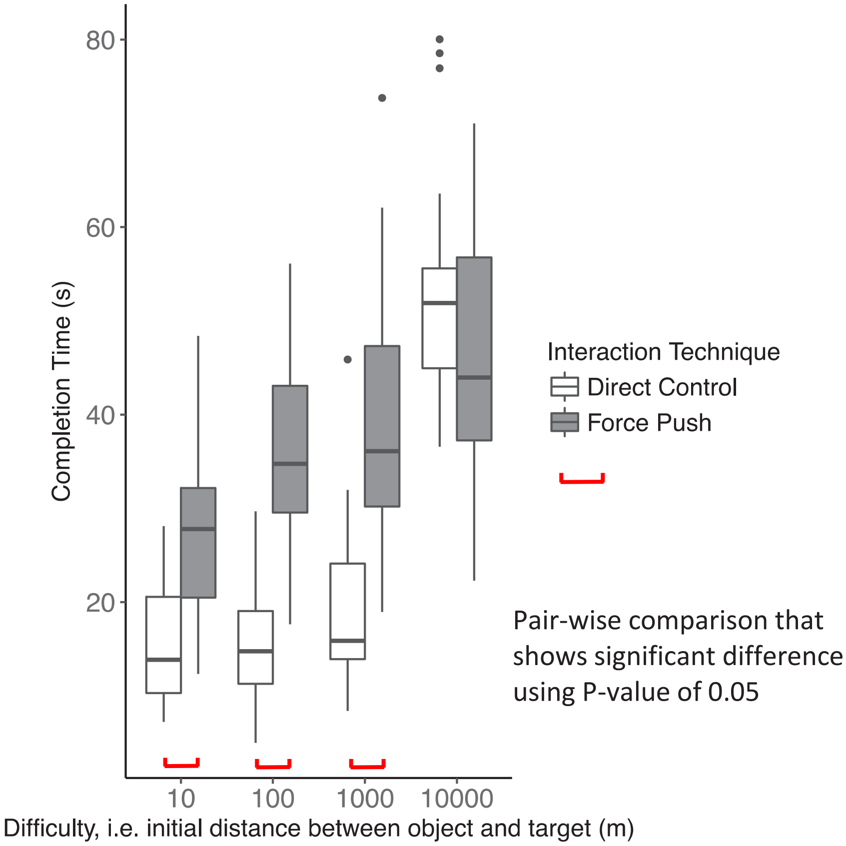
Activating direct control was achieved through holding down the index finger trigger on the Oculus Touch controller. Once activated, the relative positional change of the hand on the vertical dimension was directly mapped to the translation of the object, using the chosen C/D ratio.

Toggling between the two C/D ratios is easily done by pressing the “X” button on the Oculus Touch using the thumb if it’s the left hand controller (or the “A” button at the same position if it’s a right hand controller).

(5) Participants and Procedure

A total of 20 people were recruited on campus as participants. We required them to have normal range of hand and arm movement. Three of them served as pilot participants in order to test and optimize the procedure. The data from the other 17 participants were collected as the experimental data. Among these 17 participants, ten were male and seven were female. Two participants were above thirty years of age, while the others were all between 20 and 30. Only one participant was left-handed.

Using a within-subject design, each participant tested both techniques. There were four difficulties (initial distance between object and target) for each technique and each difficulty had two trials. Hence, each participant would perform 16 trials (2 techniques × 4 levels of difficulty × 2 trials for each difficulty).



Pair-wise comparison that shows significant difference using P-value of 0.05

Figure 4-5: Box plot of the completion time

The experiment always started with one technique, completed all its trials and then switched to the other one. The choice of which technique to start with was counterbalanced—nine participants began with direct control and the eight with Force Push. To avoid bias induced by the order of the eight trials, we generated two random orderings of eight trials beforehand. The first random ordering was always used for the technique presented first, and the second random ordering was used for the second technique. The participants were asked to go through a training session with five trials for each technique before the experiment started. Everyone expressed confidence in understanding the technique and controlling the object after the training session.

4.2.2.3 Results

(1) Task Performance

Figure 4-5 shows the box plot of the completion time for each technique with the four difficulties.

We ran a two-way repeated-measures ANOVA for the task completion time metric. A significant interaction between the two independent variables was found ($F(3,14) = 29.72$, $p < 0.0001$). This interaction can also be easily observed from the box plot.

We conducted post-hoc Bonferroni-corrected t-tests to compare the two techniques at each level of difficulty, and found that the direct control technique was significantly faster than Force Push at the three lower levels of difficulty ($p < 0.001$). However, we found no significant differences between the two techniques at the highest level of difficulty, even though Force Push had better performance on average ($p = 0.32$).

(2) Subjective User Experience

Figure 4-6 shows the distribution of answers to the questionnaire. Most participants perceived the direct control technique to be more controllable and accurate, but perceived the Force Push technique to be more fun and natural. Slightly more participants perceived Force Push to be more comfortable than direct control. Preference for the two techniques was approximately equal for the criteria of speed and ease of use.

4.2.2.4 Discussion

A direct observation from the performance result was that all participants could complete all the tasks in a reasonable amount of time. This was not surprising for the direct control interface since it was equipped with two C/D ratios optimized for long-distance and precise movements respectively. However, this was very encouraging in the case of Force Push, as it used a second-order mapping, which was inherently more difficult to control, and some of the tasks were extremely difficult. Even though Force Push did not outperform direct control overall, it provided reasonable usability in most cases. Observing this result, we tentatively state that, by dynamically mapping expressive features of gestures to properties of the physics simulation, we created a controllable gesture-to-force mapping for remote object manipulation that was usable even in extremely difficult cases. However, as this experiment was limited to a 1D manipulation task, future work is needed to verify that the Force Push technique remains controllable in the full 3D case.

The fact that direct control had much shorter completion time in most conditions was consistent with the general consensus that a zero-order mapping is easier to control than an acceleration-based mapping. What was more interesting was how each technique's performance changed with increasing difficulty. The time used in Force Push seemed to grow approximately linearly with the log of the initial distance, while the time for direct control seemed to grow at an exponential rate. Granted, the number of difficulty settings and the number of trials was too small here to model these relationships precisely, but we have some preliminary evidence that the gesture-to-force mapping was more robust against the increasing difficulty. One way to interpret this is that the gesture-to-force mapping seems to be more adaptable to extremely difficult situations where the object was placed far away and direct mapping struggled to cover the two ends of the spectrum even with two distinct C/D ratios. We speculate that if the task becomes even harder than the conditions presented in this study, the performance advantage of gesture-to-force mapping will become more apparent.

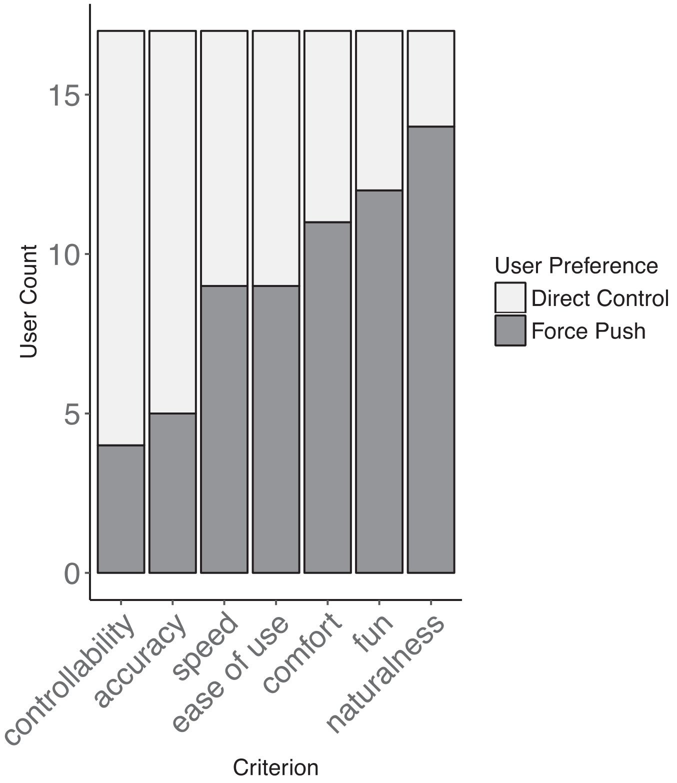


Figure 4-6: Bar chart for the distribution of user preference based on seven criteria

The participants' subjective evaluation also showed some interesting results. Even though Force Push was less efficient than direct control in most cases, the criteria of speed and ease of use showed equal preference. More surprisingly, the answers to "Which one feels more comfortable to use?" actually favored the gesture-to-force method. We speculate that this result is due to the significant muscle tension required in the direct control interface during the fine tuning stage of the manipulation task. With Force Push, although the hand is held in mid-air, the position of the hand does not have to be controlled precisely.

The comparison on the criteria of fun and naturalness supported our hypothesis, as the majority of the participants favored the Force Push interface over traditional direct manipulation.

We surmise there are three reasons behind this result. First, the hyper-natural property of Force Push may have made it feel magical, and this kind of superpower frequently appears in popular movies like *Star Wars*. It is not realistic, yet it feels familiar, which makes the technique fun and natural. Second, we suggest that the force-driven physics-based simulation may have created a richer perception of the virtual world's plausibility (Slater, 2009). For example, the user could have a much stronger sense of the weight of the object using Force Push as compared to the direct control interface, in which the object instantly became "weightless" when it was attached to hand movement. Virtual environments with consistent and explainable physics models may be more plausible and relatable than those without, leading to a deeper sense of presence (Skarbez et al., 2017a). Third, the interaction of Force Push might appear to be more novel, as the majority of users have already used position-to-position mappings in human-computer interaction on a daily basis (e.g., mouse).

Despite these preliminary findings about potential user experience benefits of the Force Push approach, more work is necessary to validate our hypotheses regarding the plausibility, fun, and naturalness of our technique as compared to traditional virtual hand techniques in realistic object manipulation tasks.

The results can be used to make preliminary design recommendations for certain applications. For example, if the target application is entertainment (e.g., VR games), it may be beneficial to choose a gesture-to-force mapping over a traditional position-to-position mapping for object manipulation, even though it may sacrifice a certain degree of user performance. On the other hand, if the target application is more engineering-oriented and requires both speed and accuracy (e.g., VR modeling), a direct control interface may still be more appropriate in most cases.

It is important to realize that our work to date with gesture-to-force mappings is only a small step in exploring the design space of mapping gesture features to interaction properties. One purpose of this work is to inspire more effort in thinking deeply about how to leverage the expressive information that may be contained in an individual gesture instead of only searching for a wider range of different gestures. An important take-away from this effort is that the rich interaction realized by creatively using one gesture is at least as interesting as a number of simple actions accomplished by many gestures.

4.2.3 Evaluation: Broad User Experience of Physically Coherent Techniques

4.2.3.1 Goals and Hypotheses

In this experiment, the three techniques described above were evaluated on a variety of user experience measures. These measures can be categorized into three groups. the sense of presence (Usoh et al., 2000); pragmatic attributes that describe functionality and usability; and hedonic attributes that provide stimulation, communicate identity and provoke valued responses, such as "exciting" and "interesting" (Hassenzahl, 2003; 2004). We applied both quantitative and qualitative methods.

We had these hypotheses:

H1: The two physically coherent techniques (PCVH and FP) would result in a significantly higher sense of presence compared to SVH. This is based on our definition of physics coherence and plausibility, and that plausibility illusion is one of the two components that contribute to presence.

H2: The two virtual hand techniques (SVH and PCVH) would show very similar results for pragmatic attributes, since they are designed to comply to the same control mechanism.

H3: The two physically coherent techniques (PCVH and FP) would result in superior user experience for hedonic attributes. This hypothesis is based on our intuition that physics coherence makes the world more believable (e.g., sensing weight) with richer interactions (e.g., throw and collide), and the user is likely to enjoy such an experience more.

H4: The comparison between the two physically coherent techniques (PCVH and FP) would show trade-offs between pragmatic and hedonic attributes. Specifically, we hypothesize that PCVH would result in superior performance on pragmatic qualities, since a zero-order mapping should be easier to control than the novel gesture-to-force mapping in FP (Goldstein et al., 2002). On the other hand, we hypothesized that FP would show advantages on hedonic qualities because of its overall higher level of physics coherence.

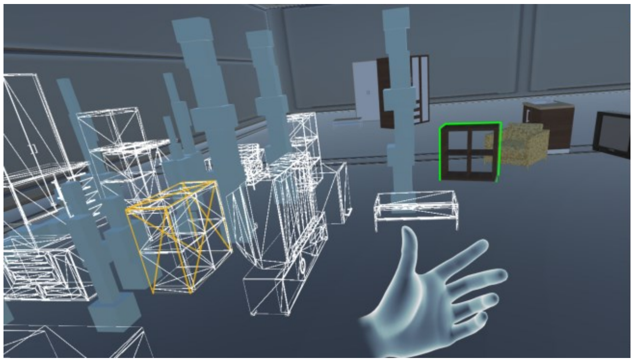


Figure 4-7: The virtual environment.

4.2.3.2 Experimental Design

(1) Environment and Task

We built a large virtual room with walls, a floor and a ceiling, as shown in Figure 4-7. Various kinds of furniture were scattered on the floor, and the user could select and manipulate any of them. Each piece of furniture could be translated with three degrees of freedom and rotated around the world's vertical axis. Since collision between objects is an important phenomenon that shows the physics coherence of some techniques, we added stacks of cube-shaped "blocks" into the scene to provide opportunities for collision. These blocks could not be selected or directly controlled, but they had physical rigid bodies and could collide with the furniture controlled by the user.

The task was divided into two stages. The first was exploration. The user was asked to actively play around by selecting and manipulating various pieces of furniture for four minutes. Its purpose was to let the user feel the interaction as much as possible without the pressure of having a certain goal. The second stage was goal-oriented—each item in the scene had a target location rendered as wire-frame. The objective was to move the furniture to these target locations. Once an item's position and orientation matched its target (using a threshold of 0.2m for translation and 15 degrees for rotation), visual and audio effects were played to indicate that the objective was achieved. Note that even in this second stage, the experiment was still about subjective user experience instead of objective task performance. Thus, we did not ask the user to optimize speed of task performance, nor did we measure task completion time. Instead, we just asked participants to keep actively interacting with the objects for six minutes.

(2) Measures

Quantitatively, we used the classic Slater-Usoh-Steed (SUS) Presence Questionnaire (Slater et al., 1998; Usoh et al., 1999; Usoh et al., 2000) to measure the sense of presence. It presents six statements to describe presence-related impression and asks participants to indicate their level of agreement on a seven-point Likert scale. The User Experience Questionnaire (UEQ) (Laugwitz et al., 2008) and Game Experience Questionnaire (GEQ) (IJsselsteijn et al., 2007; Norman, 2013) were used to provide a comprehensive evaluation of user experience. UEQ measures pragmatic values ("efficiency," "perspicuity," and "dependability") along with hedonic attributes ("stimulation" and "novelty"). It also has a pure valence component of "attractiveness." It asks the participant to respond to 26 items that describe these components on a seven-point scale. GEQ measures user experience with seven components, among which "competence" and "challenge" are strongly related to pragmatic values, while "sensory and imaginative immersion," "flow," "positive affect," "negative affect," and "tension/annoyance" relate more to hedonic values. Similarly, it asks how much a participant agrees with 33 statements that describe these components on a five-point Likert scale.

Besides questionnaires, we also employed a think-aloud protocol and post-experiment interview to capture users' thoughts on these techniques (Shneiderman et al., 2016). From this feedback we were particularly interested in understanding what features in each component of physics coherence affected the perceived user experience.

(3) Apparatus

Unity (version 2017.2) was used to develop the VE and interaction techniques. All physics-based simulation used the default Unity physics engine. Behaviors that were meant to be running in the physics simulation thread were programmed into the FixedUpdate() function, while algorithms that were meant for the rendering thread were built into the Update() function. An Oculus Rift CV1 was used as the head-worn display (HWD). Oculus Touch controllers were used for SVH and PCVH, while a Leap Motion mounted on the HWD was used for FP to recognize hand gestures.

(4) Participants and Procedure

The experiment was approved by our university's Institutional Review Board. Data from 18 participants were collected as valid results. Among them, ten were female and eight were male.

Five participants were in the age between 30 and 39 and 13 were in their twenties. The experiment used a within-subjects design, so each participant used all three techniques. For each technique, the user first went through a short training session to get familiar with the control and then completed the two stages of interaction. After that, the three questionnaires were presented to her in the order of UEQ, GEQ and then SUS Presence questionnaire. Think-aloud protocol was used throughout the experimental session and a post-experiment interview was conducted at the end. The order that the three techniques were used was counter-balanced between subjects.

4.2.3.3 Results

(1) Quantitative Data

Figure 4-8 presents box plots of the original data from the three questionnaires. We applied Friedman tests for each criterion in the questionnaires. If the test indicated a significant difference, post-hoc pairwise Wilcoxon signed rank tests were carried out with Bonferroni correction. Pairs that were significantly different ($p < 0.05$) are marked in Figure 4-8.

There were no significant differences among techniques on the SUS Presence Questionnaire. On the UEQ, PCVH was judged to be more attractive ($p = 0.0412$) and stimulating ($p = 0.0024$) than SVH; SVH was seen as more dependable ($p = 0.0229$) and perspicuous ($p = 0.0019$) than FP; and PCVH was found to be more dependable ($p = 0.0476$), efficient ($p = 0.0061$) and perspicuous ($p = 0.0017$) than FP. With the GEQ, PCVH had less negative affect ($p = 0.0281$) and greater positive affect ($p = 0.0496$) than SVH; FP exhibited greater challenge than SVH ($p = 0.0032$) and PCVH ($p = 0.0179$); and FP made users feel less competent than SVH ($p = 0.0053$) and PCVH ($p = 0.0023$).

(2) Qualitative Data

We coded the user's comments from the think-aloud process and post-experiment interview. The most common statements made by multiple participants are summarized into four themes here.

The first theme was about the physics-driven motion used in PCVH and FP. All of the participants mentioned that the gravity, momentum and collision effects of PCVH and FP added a lot to the experience in comparison to SVH. They made the interaction feel "immersive," "interesting," and "cool." In particular, many mentioned the force-driven movement of "throwing" in PCVH and "pushing" in FP made the object feel more "realistic," and said they are desirable for long-distance placement since they reduce the necessity of clutching. On the other hand, some also stated that SVH was "simple" and "clean," which they would consider as good qualities for a pure work-related task.

The second theme was about the perception of weight through kinesthetic cues when using PCVH and FP. All of the participants mentioned that they could feel the weight variation between objects using PCVH, which resulted from different C/D ratios. They mentioned that this was "like reality," as heavier objects were "more difficult" to move. All but one participant mentioned that FP generated a strong sense of the object's weight. They all connected FP with the real-world concept that they needed to “push harder” to move a heavier item. Three participants specifically stated that FP generated a stronger sense of weight compared to PCVH.

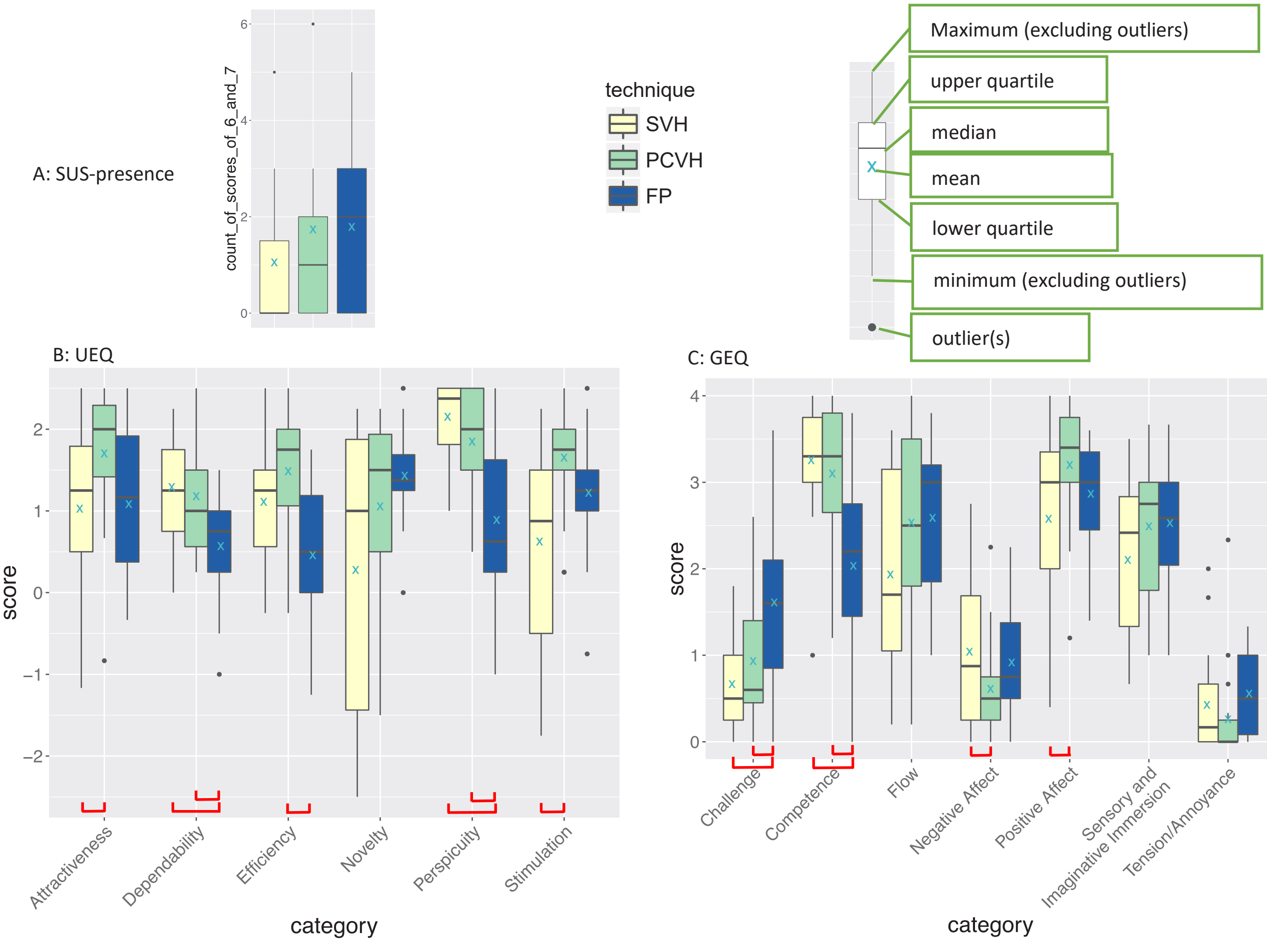


Figure 4-8: Boxplots of the questionnaires’ results. A: SUS-presence questionnaire; B: UEQ; C: GEQ. How to interpret the box-plot is explained on the upper-right. The pairs that show significant difference using Friedman test and post-hoc Wilcoxon signed rank test are linked with a red symbol at the bottom. For SUS-presence questionnaire, count of scores of 6 and 7 is used as the score for each user.

The third theme was about the control mechanism. All of the participants agreed that the two virtual hand techniques (SVH and PCVH) were much easier to control than the gesture-to-force mapping in FP. First, they recognized that FP took a noticeable amount of time to register the hand gesture, which created an annoying delay between input and the object’s action. Second, the separated degrees of freedom in FP overly limited the motion of the object. Lastly, the Leap Motion tracking was not stable at all times, which caused some confusion. Overall, the users appreciated the directness of the virtual hand techniques and preferred them for controllability.

The last theme was about the “telekinesis” metaphor of FP. Many participants mentioned this design was “cool” and “innovative.” Many related it to superpowers from popular culture, and some mentioned that using gesture to give movement commands made the control felt more “intelligent” than using “grab and move” metaphors. Some expressed that if the controllability of FP could be improved, it would become their favorite technique.

4.2.3.4 Discussion

We did not find statistical support for **H1** in the SUS presence questionnaire results. The average scores do favor the two physically coherent techniques, with FP being the highest. Participants' comments about them being "immersive" and "realistic" also hint that increased physics coherence may potentially help improve the overall sense of presence. Other ways of measuring presence might be needed to provide more thorough understanding of its influence on presence (e.g., (Slater et al., 2003)).

The UEQ and GEQ results did not reveal any significant difference between the two virtual hand techniques (SVH and PCVH) on any attribute strongly related to pragmatic values. This includes "efficiency," "perspicuity," and "dependability" in UEQ and "challenge" and "competence" in GEQ. This result is consistent with **H2**, although of course a lack of significant difference does not imply equivalence. As for hedonic values, PCVH had significantly better scores on two of the three hedonic attributes in the UEQ ("attractiveness" and "stimulation"). This difference was also seen in the GEQ results for "positive affect" and "negative affect." Overall, the experiment supports **H3** that physics coherence has a positive influence on hedonic qualities in user experience.

Both the quantitative and qualitative data show that FP suffers from low controllability in comparison with the two zero-order mapping based techniques. In the UEQ results, FP was significantly worse on pragmatic values such as perspicuity, efficiency and dependability compared to one or both of the other techniques. Similarly, in the GEQ, FP was perceived as being significantly more challenging and less competent. The qualitative results also indicated its disadvantage on usability. Overall, the results support the first part of **H4** that FP is inferior to PCVH for pragmatic user experience. On hedonic attributes, we did not identify any significant difference between FP and PCVH, so we cannot confirm the second part of **H4**. The qualitative data shows that both the high level of physics coherence and the "telekinesis" metaphor were appreciated by users, but these positive effects seem to be canceled out by its deficiencies in usability.

We extract three design guidelines for creating effective physically coherent manipulation techniques by inspecting which features of PCVH and FP users appreciated most compared to SVH:

- **Physics-based motion:** all the features dedicated to a more physics-driven display of motion were appreciated by users. To increase physics coherence, physics-based algorithms for driving objects' behaviors should be favored over setting their positions/orientations directly.
- **Physically based correlation between kinesthetic cues and visual motion:** users especially appreciated the display of weight that led to higher consistency between visual and somatosensory cues. In both PCVH and FP, this was achieved by forming a mapping between the object's motion and the internal kinesthetic cues based on physical laws. We consider this to be an effective approach to enhance physics coherence.
- **Relatable metaphors:** although the magical interaction style of FP was very different from reality, users still appreciated it because it was understandable and relatable to their

experiences with popular culture. Thus, it is important that similarly familiar metaphors are used when designing a novel input-to-output mapping.

4.2.3.5 Limitations

The main limitation of this experiment is that there were factors other than physics coherence itself that could have influenced the user experience, yet we could not completely control for these factors. In the comparison between SVH and PCVH, added physical properties such as inertia and rigid bodies enable richer interactions in PCVH such as throwing and colliding, which could have enhanced the perceived user experience. However, we argue that these are inevitable results of higher physics coherence. The novel and "magical" nature of FP could have added to its hedonic experience, but this is often the case with new metaphors and control schemes. On the other hand, Leap Motion provides noticeably lower tracking accuracy compared to Oculus Touch controllers, which could have affected the pragmatic values of FP, yet using such a bare-hand tracker is necessary for FP given its gesture-driven nature. We do not claim that FP is strictly comparable with SVH and PCVH on the dimension of physics coherence. Rather, we included it in our study to explore how to design future physically coherent interactions.

4. 3 Expanding the Design Space of Physically Coherent Manipulation Techniques

4.3.1 Force-Based Novel Techniques

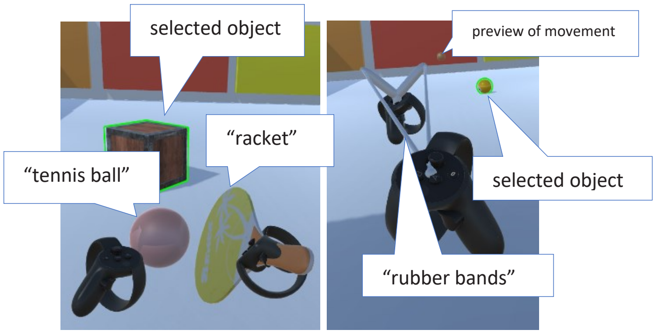


Figure 4-9: A: the tennis ball technique; B: the slingshot technique

Having shown the positive effects on hedonic user experience from physics coherence and extracted corresponding design guidelines, we acknowledge that the two physically coherent techniques we used still have some limitations. For PCVH, the object's visual motion is technically not physically based since it is achieved by customized tweaking in the physics engine instead of the native simulation. Gravity also has to be disabled temporarily when holding the object. PCVH also poses a usability issue in moving heavy objects—because the C/D ratios associated with them could become very large, moving them requires an extensive amount of clutching. As for FP, lag and separated degrees of freedom severely limited its usability. In this section, we explore other physically coherent techniques that illustrate our design guidelines while mitigating some of these issues.

We propose the design of two such techniques: “tennis ball” (TB) and “slingshot” (SL). TB is inspired by the real-world interaction of (table) tennis. The user holds a virtual racket in the dominant hand and a virtual ball in the other. To move the selected object, he hits the ball with the racket—the algorithm applies an impulse force to the object based on the movement direction and speed of the racket at the moment of hitting the ball. SL is inspired by real-world slingshots. With this technique, users form a virtual slingshot between the two hands and use it to apply an impulse force to the selected object. Just like a real slingshot, the force’s direction and strength are determined by the length and orientation of the virtual rubber bands stretched between the hands at the moment they are released. Both techniques are implemented with the Oculus Touch controllers, as shown in Figure 4-9.

These techniques strongly conform to the proposed design guidelines and operational model. First, since both use force-based control through the physics engine, visual motion is completely physics-driven. Second, they also create a physical correlation between kinesthetic cues and the object’s motion. With TB, one can always feel the direction and force of hitting the tennis ball by internal kinesthetic cues from both hands, and the object’s visual movement will conform to it since it is driven by a corresponding impulse force. Similarly, one can always feel the aiming direction and the tension of the rubber bands (determined by the distance between two hands) just by kinesthetic cues, and the object’s motion will comply with such cues once the rubber bands are released. Finally, the metaphors of TB and SL should be familiar to most people. Using these metaphors for force-driven control should feel relatable to one’s experience with the physical world.

To improve the **temporal coherence of somatosensory cues**, we also added some haptic feedback: at the moment of hitting the tennis ball using TB, both controllers display a short burst of vibration based on the calculated force. Similarly, vibration is presented when the slingshot is pulled back, with its amplitude matching the virtual tension of the rubber bands. Features enhancing physics coherence in TB and SL are shown in Table 4-1.

These techniques also help solve one of the usability issues of PCVH. It is reasonable to assume both TB and SL could generate a significantly larger amount of force compared to using bare hands to place or throw the object. Hence, we could use them in combination with PCVH to enhance long-distance manipulation of heavy objects. Our current implementation allows users to freely switch between the techniques. Touching the thumb stick on one controller will turn this hand into the tennis ball and the other into the racket for using TB. When the two hands are very close and no button is touched, a low-frequency vibration is played on both controllers to indicate that the user can use SL. Holding down the index finger trigger on one controller turns this hand into the one holding the rubber bands while the other becomes the slingshot’s base. Releasing the rubber bands is done, naturally, by releasing the trigger. Previews of the object’s movement are also displayed before the release. When the thumb sticks are not touched and the hands are not close to each other, PCVH can be used. With this implementation, one could use TB and SL for ballistic translation, then switch to PCVH to “catch” the object for precise placement and rotation.

These two techniques generate instant movement at the moment the racket hits the ball or the rubber bands are released, which eliminates the lag that plagued FP. Furthermore, they also combine the three degrees of freedom of translation together. Another hybrid interface could use TB and SL to apply strong forces for long-distance placement while using FP to perform "light nudging" for precise movement and rotation (Goldstein et al., 2002).

We have not yet carried out an evaluation of these techniques, and we acknowledge that they have some potential drawbacks (mainly in controllability) that could influence user experience. For example, both playing tennis and using a slingshot require a certain level of dexterity and skill, and it is unclear how fast a novice user could get used to them. This is why we propose to use such techniques in combination with PCVH, instead of replacing it.

We implemented these two techniques using the Oculus Touch controllers, so that they are applicable to commercial VR systems. We can imagine using some customized controllers to provide haptic feedbacks that conform to the metaphors themselves. For example, physical proxies could provide actual collision between two hands when using TB, or we could construct a slingshot with physical rubber bands that provide actual tension for SL. These customized controllers could potentially enhance **temporal coherence of somatosensory cues** and **inter-sensory consistency between visual and somatosensory cues.**

4.3.2 Pseudo-Haptic Display of Mass and Mass Distribution During Object Rotation in Virtual Reality

So far, we've evaluated the task performance and broad user experience of the proposed techniques. However, we've never directly measured physics plausibility. This section makes up this gap by proposing and evaluating novel pseudo-haptic techniques to display mass and mass distribution.

The sensation that a virtual object has appropriate heaviness is a direct indication of perceived physics plausibility. However, providing effective haptic feedback during interactive object manipulation in VR is difficult. In object manipulation, modern commercial VR systems often let the user hold a proxy (called a hand-held controller), and pick up different virtual objects using this same tool (Zenner, 2016). A particular proxy provides the same haptic stimuli for all virtual objects, which contradicts the scenario of holding different objects with distinct mass properties. Thus, an item that looks heavy visually may suddenly feel as light as the proxy once it is picked up (Zenner and Krüger, 2017). Such deviations from the user's expectation may break the physics plausibility of the VE, since fundamental laws of physics are violated by this visual-haptic mismatch, reducing the VE's physics coherence (Skarbez et al., 2017c).

This problem can be addressed through active haptic rendering, which directly displays haptic forces in real-time (e.g., through a robotic arm attached to the user's hand), but such approaches are often cumbersome and expensive (Hummel et al., 2013; Zenner and Krüger, 2017). *Pseudo-haptic techniques*, on the other hand, can produce illusions of haptic sensations such as force without altering the physical proxy manipulated by the user (Lécuyer, 2009). Typically, pseudo-haptic techniques use visual stimuli to influence haptic perception. A convincing example is that varying the control/display ratio (C/D ratio) of an object's translation directly influences the perception of its mass. In other words, changing how the object moves visually relative to how the hand moves kinesthetically affects how heavy it feels (Dominjon et al., 2005). A larger C/D ratio makes objects feel heavier, since the same hand movement results in less visual movement of the object. This technique requires no special-purpose hardware and can be applied to commercial VR systems that use a single physical proxy.

Despite the success of the pseudo-haptic technique of modifying the C/D ratio for translation to influence mass perception, this approach has not yet been applied to object rotation. We hypothesize that scaling an object's rotational motion relative to its mass will also be able to generate accurate perception of relative mass among multiple objects. Our first experiment tests this hypothesis.

Beyond the perception of an object's total mass, however, we recognize that object rotation is also affected strongly by the distribution of mass within an object. For example, swinging a golf club where most of the mass is concentrated in the club head feels very different than swinging a metal rod made of a single homogeneous material, even if both have the same total mass. We hypothesize that pseudo-haptics can be used to create the illusion that different parts of the same object have different masses by manipulating the visual rotational motion of the object.

In this section, we first present a theoretical framework explaining how visual motion combined with kinesthetic cues can affect the perception of mass and mass distribution during object rotation. Based on these principles, we first investigate whether modifying the C/D ratio can effectively produce the illusion of different masses for rotational motion. We present statistical evidence from a psycho-physical experiment to support this hypothesis. We then discuss the design and implementation of two pseudo-haptic techniques to display mass distribution within an object. One of them relies on manipulating the pivot point of rotation, while the other adjusts the C/D ratio on the fly based on the real-time dynamics of the object's movement. Our experiment shows that both techniques can induce effective perceptions of mass distribution, with the second technique being much more effective.

As far as we know, this is the first effort focusing on pseudo-haptic display of mass and mass distribution during rotational motion in VR object manipulation. The contributions of this work include:

- A theoretical framework explaining how to generate pseudo-haptic perception of mass in VR object manipulation, with specific interpretations in the context of rotational motion.
- Empirical evidence that changing the C/D ratio of rotational motion effectively influences the user's perception of an object's total mass.
- Two novel pseudo-haptic techniques for displaying mass distribution during rotation, along with empirical evidence that at least one of these techniques is very effective in creating such illusions.

4.3.2.1 Theoretical Framework

In this section, we present the theoretical basis for the use of pseudo-haptic techniques to create perceptual illusions of mass. These illusions depend on the multi-sensory interaction between visual motion cues and kinesthetic force cues, and the relationship between motion and force according to fundamental physical laws. Essentially, they follow the same theoretical model of physics coherence proposed in Section 4.1, and here we specifically interpret this issue in the context of rotational motion.

The relationship between translational motion and force is given by Newton's second law:

$$F = ma \tag{4-5}$$

In this equation, F is a vector representing the force exerted on the object, m is the object's mass, and a is a vector describing the acceleration of translational motion. In object manipulation, there are generally two forces applied to the object: the force from the hand input, which we write as F_h, and the gravitational force, which we write as mg. Putting them into (4-5) gives us:

$$F_h + mg = ma \tag{4-6}$$

Combining terms that have m together gives us:

$$F_h = m(a - g) \tag{4-7}$$

Based on real-world experience of manipulating objects, humans have an expectation of the general relationship between motion, force and mass that *loosely* conforms to (4-7), and we can leverage this expectation to manipulate the perception of mass. This expectation (*E1*) can be expressed using two statements:

E1a: Putting aside the influence of gravity, objects with larger mass will move more slowly compared to objects with smaller mass when the same amount of force is exerted from the hand input.

E1b: Translating an object against the direction of gravity (lifting) requires more force from hand input compared to translating with the direction of gravity (dropping), and the influence of gravity gets stronger as the object's mass gets larger.

In VR, we perceive motion primarily through visual feedback and perceive exerted force through haptic feedback. Hence, *E1a* and *E1b* become an expectation of the relationship between visual cues, haptic cues and the mass of the object. Thus, scaling visual motion feedback relative to haptic force feedback should influence the perception of mass when the information from these two sensory channels is fused together.

In VR interaction, we can change the haptic cues using two approaches. On the one hand, we can alter the external haptic stimuli. This would require changing the hand-held proxy, which is undesirable in many scenarios, because we want to use a single general-purpose input device as the proxy. On the other hand, one can alter the internal kinesthetic cues by changing the user's input. This can be achieved by manipulating the input-to-output mapping (C/D ratio) of the interaction without changing the hand-held proxy. These pseudo-haptic techniques manipulate this mapping by creating an offset of the visual movement relative to the kinesthetic cues (e.g., (Dominjon et al., 2005; Taima et al., 2014; Rietzler et al., 2018b), influencing the perception of mass.

Note that when we modify the C/D ratio, the movement of the virtual object will not conform exactly to (4-7), because changing C/D ratios relate to changes in velocity, not acceleration. But this approach appears to be "close enough" to match users' expectations and result in effective perception of mass (or at least, relative mass between objects).

In rotational motion, the relationship between force and motion takes on a slightly different form:

$$\tau = I\alpha \tag{4-8}$$

In this equation, τ is the torque, which is the rotational analog of force; I is the *moment of inertia*, which plays the role of mass in rotational motion; and α is the angular acceleration. Again, there are two potential torques in the system in object manipulation, the torque from hand input, which we write as τ_h, and the torque generated by gravity, which we write as τ_g. We re-write (4-8) as:

$$\tau_h + \tau_g = I\alpha \tag{4-9}$$

What's unique about rotation is that both *torque* and *moment of inertia* in (4-9) are only meaningful when tied to a specific rotational axis, and different axes will have different moments of inertia. Torque is defined as a vector generated by the cross product of the radius vector (from the axis of rotation to the point of application of force) and the force vector, and moment of inertia is the sum of the products of each particle's mass in the rigid body with the square of its distance from the rotational axis (for a more thorough explanation of the dynamics of rigid body rotation, please refer to (Goldstein et al., 2002)). By definition, how mass is distributed relative to the axis matters to both of these concepts. Thus, when discussing rotation, we cannot treat the entire object as a point mass, but rather we need to understand its *mass distribution* relative to the rotational axis. Consider again the golf club example: the forces required to rotate the golf club change drastically depending on where it is held and which direction it is rotated.

Similar to the case of translation, people have an intuitive expectation of the relationship between the exerted torque, mass and mass distribution, and rotational motion. In analogy to *E1*, this expectation (*E2*) *loosely* follows (4-9) and has two components:

E2a: Putting aside the influence of gravity, rotation against a larger moment of inertia will result in less movement compared to rotation against a smaller moment of inertia when the same amount of torque is exerted from hand input.

E2b: Rotating against the direction of gravity will require more torque from hand input compared to rotating along that direction (but note that torque due to gravity becomes zero if the rotational axis goes through the CoM of the object).

In analogy to translation, we feel the exerted torque through kinesthetic cues while seeing the rotational motion visually. Because both total mass and mass distribution are now in the equation, scaling visual motion relative to kinesthetic cues give us an opportunity to display both.

It is important to point out that total mass and mass distribution are two different physical properties. On the one hand, two objects can have the same shape and total mass but distinct characteristics in mass distribution; on the other hand, two objects can have the same shape and the same relative mass distribution between different parts of the body, but different total masses (e.g., because one object is made of a different material). This means that we need to examine pseudo-haptic illusions for these two properties separately. In sections 4.3.2.2 and Section 4.3.2.3, we describe techniques for generating these two illusions and evaluate their effectiveness.

4.3.2.2 Pseudo-Haptic Display of Total Mass During Rotation

(1) The Technique

All the interaction techniques in Section 4.3.2 are based on the common “simple virtual hand” metaphor, in which a zero-order mapping maps the hand-held controller’s movement to the object’s movement (translation and rotation) (Poupyrev et al., 1996). To enhance the applicability of these techniques, we use the “remote control” version of this metaphor, which maps the hand’s relative movement to the movement of a remotely placed object. This is similar to how we use a mouse to drag icons in a desktop environment, which does not require the hand and the object to be co-located.

Our technique for producing a pseudo-haptic illusion of total mass during rotation is in direct analogy with the technique of changing C/D ratio in translation. We change the C/D ratio for rotation based on the object’s mass—a heavier object will have a larger C/D ratio (i.e., its rotational motion is scaled down) and vice versa. The C/D ratio is applied to rotational motion using several steps. First, for every frame, the orientation (originally expressed as a quaternion) of the hand-held proxy is extracted. Then the difference in the orientation between the last frame and the current frame is used to calculate an axis and an angle, which represent the proxy’s relative rotation between these two frames. The appropriate C/D ratio is applied to scale the angle while keeping the same axis. Finally, these values are used to rotate the virtual object.

This technique is designed to match the user’s expectations about the relationship between mass, torque, and rotational motion. Assuming the pivot point is at the CoM, the only active torque in the system is the torque from hand input, so *E2b* does not apply. As long as the mass distribution is constant, increasing the total mass will always increase the moment of inertia, which makes the object more difficult to rotate. Increasing the C/D ratio has the effect that the same exerted torque from the hand will result in less visual rotation of the object, which conforms to *E2a*.

(2) Experiment I

We designed a psycho-physical experiment to evaluate the effectiveness of this technique. The experiment tests whether a user can perceive a mass difference between two virtual objects rotated with different C/D ratios.

Figure 4-10: The virtual environment for experiment I.

a. Experimental Design. Figure 4-10 shows the virtual environment used in this experiment. In this scene, two cubes with the same appearance (size, shape, and texture) were placed side by side. The user was asked to rotate them in turn and make a two-alternative forced choice about which of them felt heavier than the other. The rotational control of the two cubes was scaled by difference C/D ratios. If the technique is effective, we expected that users would perceive that objects rotating with larger C/D ratios would feel heavier than the other, and vice versa. In this experiment, the pivot point of rotation was always at the geometric center of the cube.

For each trial, the system randomly picked one of the cubes to assign it a non-1:1 rotational C/D ratio, while the other cube had a ratio of 1:1 (used as the reference). The levels of the first C/D ratio are listed in Table 4-4 (note that the table lists the reciprocal of C/D ratios), with values centered around 1.0. We can think of these values as being linearly scaled by the virtual mass of the items. For example, assuming the reference cube with a 1:1 C/D ratio has a mass of 1kg, another cube with a mass of 5kg would have a C/D ratio of 1:0.2 (calculated by $5kg/1kg$).

rotation scaled down				rotation scaled up			
0.2	0.5	0.7	0.9	1.1	1.3	1.5	1.8

Table 4-4: Scaling values used in experiment I (values represent the reciprocal of C/D ratio)

One thing worth noting is that this experiment only evaluates if different rotational C/D ratios could lead to the illusion of one item being heavier than the other. It does not evaluate the perception of the absolute value of the object's mass or the actual mass ratio between two objects. Using the previous example, even though we set the C/D ratios based on the objects' virtual mass, we don't claim that using a C/D ratio of 1:0.2 could lead to the perception that the item weighs 5kg or that it feels 5 times heavier than the 1kg reference cube. It only tests if the one with a C/D ratio of 1:0.2 feels heavier than the reference with 1:1. We consider the perception of the absolute values of mass and mass difference as future work.

To eliminate any influence from translational motion or the translational C/D ratio (which has already been shown to effectively produce a pseudo-haptic illusion of mass perception (Dominjon et al., 2005)), both cubes were fixed in place and only had three degrees-of-freedom (DOF) of rotation. The user held a physical proxy in her dominant hand, with its rotational motion mapped to the currently selected cube using the assigned C/D ratio. Another hand-held controller in the non-dominant hand was used to select one of the cubes using ray-casting. The user indicated which cube felt heavier by clicking on the button next to the corresponding cube (Figure 4-10) using the same ray.

Figure 4-11: Left: VIVE Pro controller. Right: VIVE tracker.

b. Procedure and Measures. With three repetitions for each of the eight C/D ratio levels shown in Table 4-4, each participant made 24 total mass judgments. The order of the trials was randomized. For each trial, if the cube with a larger C/D ratio was selected as the heavier one, we considered it as a “correct” response. For each participant, we measured the percentage of correct responses for each level of C/D ratio as the primary quantitative measure.

Using correct responses alone potentially puts us at risk that we were only measuring participants’ ability to guess the correct answer instead of actually measuring their perception—users may have guessed that the object that rotates slower “should” be the heavier one, but they may not have been experiencing an actual sensation of mass difference. To alleviate this issue, we also adopted a post-experiment questionnaire on the quality of induced haptic perception, proposed by Rietzler et al. (Rietzler et al., 2018a). Customized to fit the context of this experiment, the questionnaire asked participants how much they agreed with the following six statements on a seven-point Likert scale:

Q1: I could feel that the two cubes have different weight

Q2: The representation of weight difference felt realistic

Q3: I had the feeling of manipulating real objects

Q4: The representation of weight difference is sufficient for VR

Q5: I liked this representation of weight difference

Q6: I felt I have control

Each participant first went through an introduction session, which explained the task and the interface. Next, we provided a training session to practice the two-alternative forced choice task four times. After the training, each participant completed the 24 experimental trials. During both the training and the experiment, no feedback was given after each trial; instead, the system simply played a sound indicating the conclusion of the current trial and immediately loaded the next one. Finally, we gave the participant the post-experiment questionnaire.

c. Apparatus and Participants. Participants sat in a chair to complete the experiment. A wireless HTC VIVE Pro head-mounted display was used to display the virtual environment. This headset has a resolution of 1080 x 1200 pixels per eye (2160 x 1200 pixels combined), a refresh rate of 90 Hz and a field of view of 110 degrees. The user used a VIVE controller (Figure 4-11, left) in her non-dominant hand to select between the two cubes and give answers using ray-casting, while a VIVE tracker (Figure 4-11, right) was used to rotate the virtual cubes. The headset, controller, and tracker were all tracked by a HTC Lighthouse 2.0 inertial-optical tracking system. The interface and environment were developed in Unity 2018.3.2f1. Rendering was performed by a PC with an Intel Core i7-8700k CPU, 16GB of RAM, and an Nvidia GTX 1070 GPU.

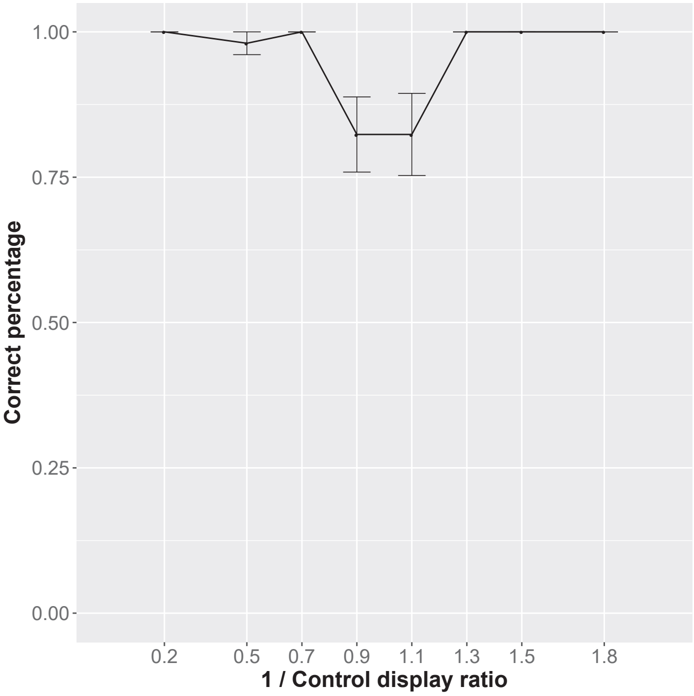


Figure 4-12: Percentage of correct responses at each C/D ratio in experiment I. Error bars represent standard deviation.

A key aspect of the experiment was the choice of the physical proxy. Ideally, it should be generic enough to represent the cube at an arbitrary orientation without biasing the user's perception of mass and mass distribution (note the orientation of the proxy does not necessarily match the orientation of the virtual cube at any time because of the scaled C/D ratio). Hence, the ideal proxy would be a sphere that is light enough to easily hold and rotate. However, it was non-trivial to track a sphere's rotation accurately in our lab. We ended up using the VIVE tracker (Figure 4-11, right) as a compromise. It is lightweight and easy to manipulate with one or two hands, and its form factor is symmetric along two dimensions, helping to ensure that it does not strongly influence perceived mass distribution (with the help of its overall light mass). The participant sat at a table, so he could put the controller down if he wanted to use both hands to manipulate the VIVE tracker. The table could also provide elbow support to mitigate potential fatigue.

A total of 17 participants (4 females, 13 males) were recruited, all aged between 20 to 30. Among them, only two of them identified their left hand as the dominant hand.

d. Results. Figure 4-12 shows the correctness results. As the figure shows, participants achieved almost 100% correctness at all levels of C/D ratio except those closet to 1.0. Even at these levels, the average correctness was greater than 80%.

We treat the correctness percentage as ordinal data. For C/D ratios of 1/0.2, 1/0.7, 1/1.3, 1/1.5, 1/1.8, the answers were 100% correct so no statistical test was needed. For the rest of the cases, one-sample Wilcoxon signed rank tests show that the results are all significantly higher than random chance of 50%, with p-values of 4.103e-14 (C/D = 1/0.5), 0.0001 (C/D = 1/0.9) and 0.0003 (C/D = 1/1.1).

Figure 4-13 shows the results from the questionnaire. One-sample Wilcoxon signed rank tests show that the mean responses are significantly higher than 0 for all questions, with p-values of 0.0004 (Q1), 0.0006 (Q2), 0.0008 (Q3), 0.0002 (Q4), 0.002 (Q5), 0.0001 (Q6).

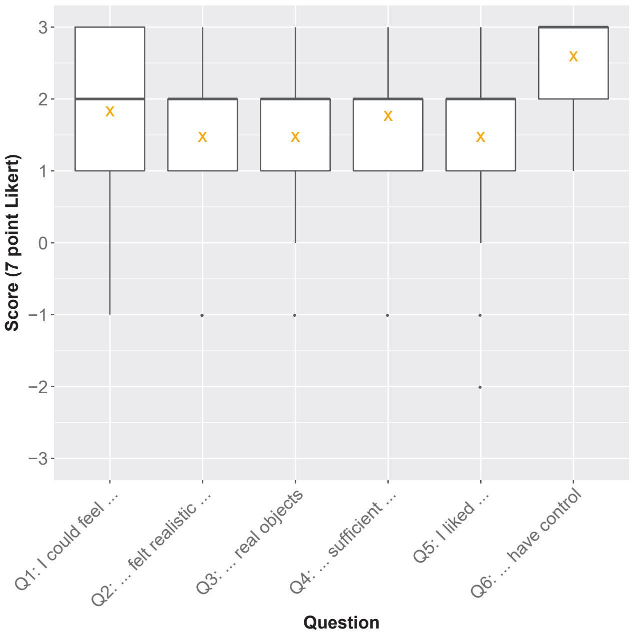


Figure 4-13: Box plot of subjective questionnaire results in experiment I. Mean values are indicated by an ‘x.’

(3) Discussion

The results of experiment I strongly support the effectiveness of the C/D ratio technique for displaying total mass of a virtual object during rotation. Participants could robustly identify the heavier object, even with small mass differences (Figure 4-12). The fact that they could perfectly perform the task (with 100% correctness) in some cases was surprising, and it’s reasonable to doubt that some users might have found the simple rule of always picking the one that rotates slower instead of actually perceiving the pseudo-haptic illusion. The questionnaire was used as another measurement of the users’ actual perception to alleviate such doubt and its result indicated that the illusion was perceived to be strong, realistic, sufficient and desirable (Figure 4-

13). The results from Q6 of the questionnaire also show that adjusting the rotational C/D ratio did not compromise the perceived controllability.

Overall, this technique and experiment complement the previous work by Dominjon et, al. that used translational C/D ratio to generate the sensation of mass (Dominjon et al., 2005). Combining these two techniques to adjust the C/D ratio for both translation and rotation should generate a coherent and convincing illusion of mass variation.

Figure 4-12 shows that users' ability to perceive mass differences does descrease as the C/D ratio gets closer to 1.0. A limitation of this current experiment is that, due to our choice of C/D ratio levels, we were not able to find the threshold of mass difference that will lead to detectability close to random chance. We consider finding that threshold as part of the future work.

4.3.2.3 Pseudo-Haptic Display of Mass Distribution During Rotation

(1) Techniques

Recognizing that mass distribution also plays an important role in rotational motion, we leverage the same interplay between kinesthetic torque cues and visual movement cues to display pseudo-haptic mass illusions at a finer level of granularity. We propose two different techniques, both leveraging the intuitive relationship between torque, rotational motion and mass expressed by *E2a* and *E2b*.

a. Pivot Technique (PV). The first technique, called "the pivot technique," simply sets the pivot point of rotation at the object's CoM while maintaining a constant C/D ratio for rotation around any axis. We expect that users will be able to naturally feel where the CoM is when it coincides with the rotational pivot point. Figure 4-14 illustrates this idea in 2D: imagine that the object is made of the same material but does not have even density across its body. Starting from the upper-left corner, the density gradually becomes greater as we move towards the lower-right corner. With this mass distribution, the CoM will lie close to the lower-right corner instead of at the geometric center. Using *PV*, the object rotates around this CoM, as illustrated by Figure 4-14.

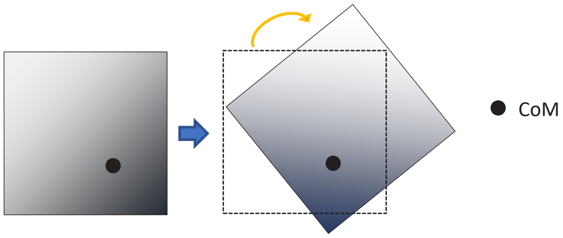


Figure 4-14: Illustration of the *PV* technique. The object rotates around the CoM instead of the geometric center.

In remote object manipulation, we expect the user to feel that he is gripping the object right at the pivot point of rotation. Hence, *PV* creates the illusion that the user is always gripping the object at its CoM. Intuitively, this is how people tend to grab an object in the real world (Lukos et al., 2007), because gravity does not create any torque relative to the CoM, making the object more controllable. *PV* is designed so that users see a constant rotational speed (C/D ratio) around the pivot point, indicating that this point is the CoM, and thus giving the user information about the object's mass distribution.

We can analyze this technique using the proposed theoretical framework. *PV* conforms to *E2b*, which indicates that torque due to gravity should influence the speed (C/D ratio) of rotation unless the pivot coincides with the CoM. According to *E2a*, the C/D ratio should be adjusted according to the direction of rotation since the moment of inertia could change according to this direction. However, for simple objects like cubes (used in the experiment described in this section) that rotate around the CoM, the slight change of moment of inertia would have little influence, so we decided to ignore this factor to simplify the technique and keep a constant C/D ratio.

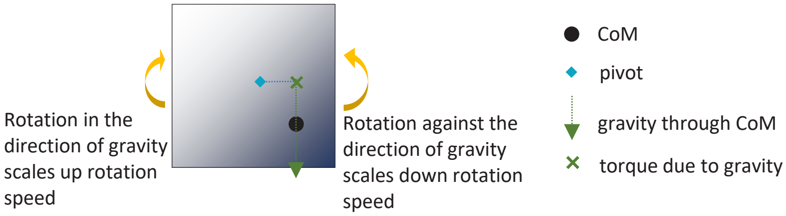


Figure 4-15: Illustration of the *DT* technique. The torque due to gravity is a vector pointing into the 2D plane (marked '×'). It affects the C/D ratio based on the axis and direction of rotation.

b. Dynamic Torque Technique (DT). One limitation of *PV* is that it requires modifying the pivot of rotation. But in practical VR applications, it is preferred that users can grasp objects at any point, motivating the need for a more general-purpose technique. The second technique, called "dynamic torque," achieves by adjusting the rotational C/D ratio using the real-time mechanics of the rotating object given an arbitrary pivot.

The general idea is shown in Figure 4-15. Assume we are using the same square-shaped object in which the CoM is at the bottom right, but the rotational pivot is now at the geometric center rather than at the CoM. This means that gravity will generate a torque relative to this pivot (shown in Figure 4-15 as a vector pointing into the 2D plane). This torque affects how difficult it is to rotate the object based on its relationship with the vector of the rotational axis. If they lie in the same direction (the angle between them is smaller than 90 degrees, as in the case of the clockwise rotation in Figure 4-15), the torque due to gravity will aid the intended rotation. Hence, we scale up the rotational motion using a smaller C/D ratio in this case. Conversely, if the torque due to gravity is opposed to the direction of the rotational axis, we scale down the rotation with a larger C/D ratio, as in the counter-clockwise rotation in Figure 4-15.

DT requires that rotational motion should be scaled according to the direction of rotation at each instant. Thus, we need to adjust the rotational C/D ratio on a frame-to-frame basis. In the following paragraphs, we derive how this ratio can be computed for each frame in a way that is both feasible and intuitive.

For each frame, the so-called rotational C/D ratio is the ratio between the amount of rotation of the physical proxy (A_{ph}) and the amount of rotation of the virtual object (A_v). Given a desired C/D ratio and A_{ph} measured in the current frame, we can compute A_v.

Based on *E2a* and *E2b*, the general relationship we want to achieve between torque, mass distribution and rotation can be expressed as:

$$\tau_{hand} + \tau_{gp} \sim I_v A_v \tag{4-10}$$

In (4-10), τ_{hand} is the torque from hand input, representing how much kinesthetic effort the user is exerting to rotate the object in that frame. The value of this component should be determined by how the hand-held proxy is being rotated.

The other torque component, τ_{gp}, is the torque generated by gravity acting on the virtual object's mass, projected onto the rotational axis. Note that since we are using a zero-order mapping between the physical proxy's movement and the virtual object's movement, the axis of rotation for the object is solely determined by the rotation direction of the proxy in that frame. The torque from gravity should not influence this *direction*; it should only influence the *amount* of rotation. This reduces the problem from 3D rotation to 2D rotation. Thus, torque due to gravity is projected onto the rotational axis.

On the right side of (4-10), I_v is the moment of inertia of the virtual object relative to the rotational axis. Again, A_v is the amount of rotation for the virtual object, which is the value we are looking for.

Next, we turn the general relationship of (4-10) into an equation by giving each component an empirically determined constant (e_0, e_1, e_2):

$$e_0 \tau_{hand} + e_1 \tau_{gp} = e_2 I_v A_v \tag{4-11}$$

We approximate the kinesthetic effort from hand input (τ_{hand}) by the product between the physical proxy's moment of inertia (I_{ph}) and how much it was rotated in the current frame (A_{ph}):

$$e_0 I_{ph} A_{ph} + e_1 \tau_{gp} = e_2 I_v A_v \tag{4-12}$$

To solve for A_v, we re-organize (4-12) into (4-13):

$$A_v = (e_0 I_{ph} A_{ph} + e_1 \tau_{gp}) / e_2 I_v \tag{4-13}$$

In this formula, all the components on the right side can be computed (I_{ph}, τ_{gp}, I_v), empirically set (e_0, e_1, e_2), or measured in real-time from the user's input (A_{ph}).

Directly using (4-13) may cause a problem: the torque due to gravity can counter the hand's intended rotation while having a very large value, which could result in a rotation angle of zero or in the opposite direction of the hand's input. This would break the zero-order mapping of the control. To overcome this issue, we set a minimum scale for the object's rotation to be 10% of the physical proxy's rotation (a maximum C/D ratio of 1/0.1), so the zero-order mapping from hand input to the object's movement is always maintained. This minimum scale was set empirically after testing some different values, and it seemed to be small enough to convey the heaviness in rotating against the torque due to gravity yet large enough to maintain the controllability of zero-order mapping. The final formula to calculate the angle the virtual object should rotate in each frame is shown in (4-14):

$$A_v = \max\left[(e_0 I_{ph} A_{ph} + e_1 \tau_{gp}) / e_2 I_v,\ 0.1 \cdot A_{ph}\right] \tag{4-14}$$

In the next section, we show the values used in the computation of (4-14) based on the specific virtual object used in the experiment.

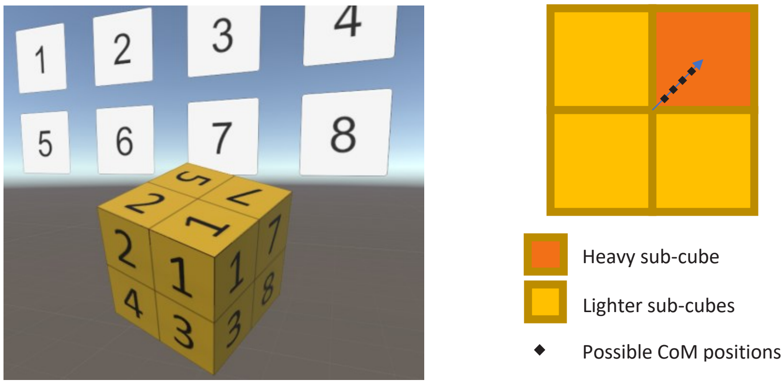


Figure 4-16: Left: VE for experiment II. Right: the four possible CoM positions in experiment II.

(2) Experiment II

To evaluate the effectiveness of the two proposed techniques, we designed an experiment to test whether the mass distribution displayed through these techniques could be correctly perceived.

a. Experimental Design

The experimental task was to manipulate the yellow cube shown in Figure 4-16, left, which is composed of eight sub-cubes at the corners. Among these sub-cubes, one is heavier than the other seven (with these other seven having the same mass). The heavier sub-cube is randomly picked by the system and unknown to the user, since the sub-cubes all have the same appearance. The system uses either *PV* or *DT* to display the mass distribution, and the user must judge which sub-cube is heavier after manipulating the object (an eight-alternative forced choice task).

	Condition 1	Condition 2	Condition 3	Condition 4
Heavy sub-cube mass	30g	47.5g	65g	825g
Light sub-cube mass	10g	7.5g	5g	25g
Heavy/light ratio	3:1	6.33:1	13:1	33:1

Table 4-5: Mass ratio conditions in experiment II.

By adjusting the ratio between the mass of the heavier sub-cube and the lighter ones, the CoM of the entire box will move along the diagonal of the heavy sub-cube (shown in two dimensions in Figure 4-16, right). The larger this ratio is, the closer the CoM is placed to the center of the heavy sub-cube. We set the total mass of the cube at 100g and used four levels of mass ratio to create four conditions, as shown in Table 4-5. These values were chosen so that the positions of the CoM would be evenly spread on the diagonal, as shown on the right side of Figure 4-16. For each condition, *PV* used the corresponding CoM position as the rotational pivot, while

DT used the corresponding values in Table 4-5 to drive its computation of A_v using formula (4-13).

To carry out the actual calculation of (4-14), several values on its right side need to be computed or empirically set. First, we use a constant value to approximate I_{ph} based on the size and mass of the VIVE tracker: $I_{ph} = 0.0004\ kg\ m^2$.

The virtual cube has a size of $0.1m \times 0.1m \times 0.1m$, with each sub-cube having a size of $0.05m \times 0.05m \times 0.05m$. Given the mass properties (Table 4-5) and size of the object, τ_{gp} and I_v can be calculated using their definitions in physics given a rotational axis (Goldstein et al., 2002). Here we briefly introduce how these values are computed.

τ_{gp} is computed by first calculating the torque due to gravity τ_g using the cross product between the object's total gravitational force (marked as $M_v g$ for the big cube) and the radius vector r (the vector from pivot to CoM) :

$$\tau_g = M_v g \times r \tag{4-15}$$

Then this vector of τ_g is projected onto the unit vector of the rotational axis to compute τ_{gp}.

The computation of I_v is a little more complicated. For a composite object such as the big cube we used, I_v can be generated by adding the moment of inertia of each individual component. Hence, I_v is the sum of the moments of inertia from each sub-cube (marked as I_{vi}):

$$I_v = \sum_{i=1}^{8} I_{vi} \tag{4-16}$$

To compute each I_{vi}, the parallel axis theorem is applied. This theorem states that if we know an object's moment of inertia when it is rotating about an axis passing through its CoM (marked as $I_{vi,CoM}$ for each sub-cube), we can further deduce its moment of inertia when it's rotating around any axis that's parallel to the original one. This calculation uses the distance between the two axes (d) and the mass of the object (marked as m_v for each sub-cube):

$$I_{vi} = I_{vi,CoM} + m_v d^2 \tag{4-17}$$

For our cube object, m_v is listed in Table 4-5, d can be computed using the distance from each sub-cube's geometric center (also the CoM for each sub-cube) to the current rotation axis. $I_{vi,CoM}$ is generated using the following formula no matter the direction of rotation, with s being the length of the sub-cube's edge:

$$I_{vi,CoM} = \frac{1}{6} m_v s \tag{4-18}$$

Finally, we need to set the constants e_0, e_1 and e_2. These values determine the influence of each component of (4-14). Enlarging e_0 will increase the influence of the physical proxy's rotation, so the same kinesthetic effort from the hand will result in more rotation of the virtual object. Enlarging e_1 will increase the influence of torque due to gravity, so the resulting C/D ratio will fluctuate more dramatically when the rotational axis changes. Enlarging e_2 will increase the influence of the moment of inertia of the virtual object, so that it is more reluctant to move. After empirical testing, we set the following values: $e_0 = 1$; $e_1 = 0.02$; $e_2 = 1$.

Note that since the display of mass distribution is independent from the display of total mass (explained at the end of Section 4.3.2.1), the overall scale of the C/D ratio for both translation and rotation should not affect the perception of mass distribution. One can uniformly scale the C/D ratio of translation (Dominjon et, al. (Dominjon et al., 2005)) and/or rotation (Section 4.3.2.2) to display the object's total mass while still using either *PV* or *DT* to display the mass distribution. Thus, we do not need to separate translation from rotation in this experiment (unlike experiment I, in which we allowed only rotation, since the C/D ratio of translation would affect the perception of total object mass). To maximize the potential applicability of these techniques, we gave the user full six DOF control in this experiment.

b. Procedure and Measures. This experiment used a within-subject design—each participant used both techniques in turn for the same set of tasks. The order of presentation was counter-balanced. For each technique, there were four repetitions for each of the four conditions (Table 4-5), resulting in a total of 32 trials $(4\,repetitions \times 4\,conditions \times 2\,techniques)$ per participant. For each technique, the order of the 16 trials was randomized. For each trial, the position of the heavy sub-cube was also randomized.

The participant indicated which sub-cube was perceived to be heavier by selecting the corresponding numbered button (Figure 4-16, left) using ray-casting. Each participant's percentage of correct answers was used as the primary quantitative measure. Note that since this is an eight-alternative forced choice, random chance would lead to a 12.5% correctness rate.

We also included the questionnaire (Rietzler et al., 2018a) used in experiment I to evaluate the quality of the haptic sensation. Modified to match the context of this experiment, the questionnaire asks the participant to rate agreement with these statements on a 7-point Likert scale:

Q1: I could feel that one cube is heavier than the other seven

Q2: The representation of mass distribution felt realistic

Q3: I had the feeling of manipulating a real object

Q4: The representation of mass distribution is sufficient for VR

Q5: I liked this representation of mass distribution

Q6: I felt I have control

We also asked participants to indicate which of the two techniques they preferred according to the same criteria. Specifically, we asked the following questions:

Q1: Which technique gave you a stronger sense of mass distribution (one of the sub-cubes is heavier than the others)?

Q2: Which representation of mass distribution felt more realistic?

Q3: Which technique gave you a sensation that's closer to the feeling of manipulating real objects?

Q4: Which representation of mass distribution felt more sufficient for VR applications?

Q5: Which technique do you prefer?

Q6: Which technique felt easier to control?

Participants first went through a tutorial session, in which the concept of mass distribution and the experimental task were explained. Next, participants practiced the task four times with the first technique, and then completed the 16 experimental trials. The system did not give any feedback on correctness during the practice or experimental trials. After using the first technique, participants completed the questionnaire on that technique before repeating the entire process for the second technique. Finally, the preference questionnaire was presented.

c. Apparatus and Participants. The same apparatus from experiment I (Section 4.3.2.2) was used here. The VIVE tracker was used to manipulate the object and the VIVE controller was used to select answers. The participants were the same group of 17 people who completed experiment I. They participated in experiment II immediately following experiment I, in a single session of approximately 40 minutes.

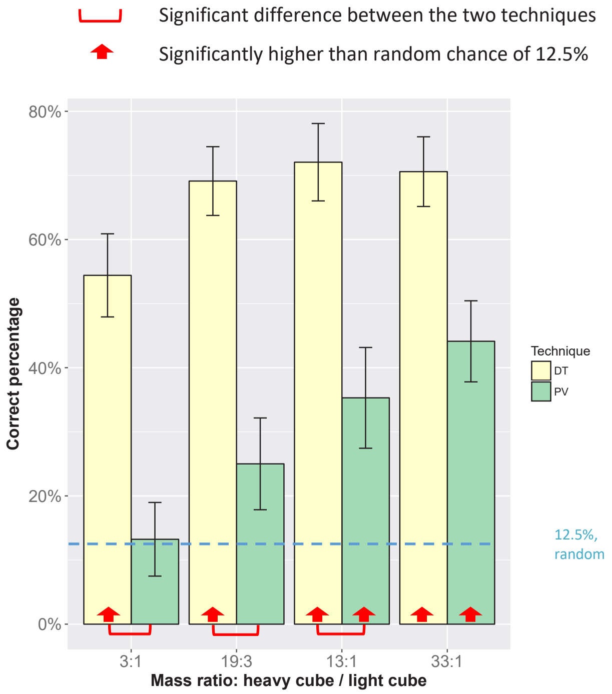


Figure 4-17: Percentage of correct responses by each technique in each mass ratio condition in experiment II. Error bars represent standard deviation.

d. Results. Figure 4-17 presents the correctness results from experiment II. Again, we treated the correctness percentage as ordinal data and performed an ordinal logistic regression, using the technique and mass ratio condition as the two predictors. Both technique ($p < 0.0001$) and mass ratio condition ($p < 0.001$) were significant factors, and there was no significant interaction between these two factors. We conducted a pair-wise comparison between *PV* and *DT* in each condition using Wilcoxon signed rank test with Bonferroni correction. The result confirmed that

DT significantly outperformed *PV* in three out of the four conditions (p = 0.008, 0.011, 0.044, 0.068 respectively for conditions 1-4). We also compared the result of each technique in each condition against random chance (12.5%) using one-sample Wilcoxon signed rank tests with Bonferroni correction. Correctness levels for *DT* were significantly higher than random chance in all conditions (p = 0.002, 0.001, 0.001, 0.001 for conditions 1-4). However, *PV* outperformed random chance significantly only when the mass ratio was 13:1 and 33:1 (p = 2.044, 0.264, 0.010, 0.001 for conditions 1-4). These results are annotated in Figure 4-17.

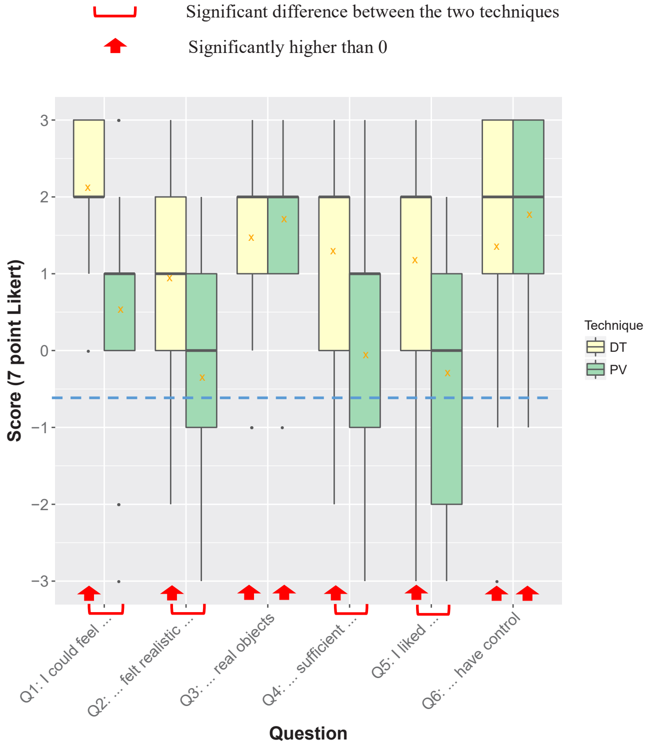


Figure 4-18: Results of the individual technique questionnaire in experiment II. Mean values are indicated by an 'x.'

Figure 4-18 shows the results from the individual technique questionnaire. Wilcoxon signed rank tests showed that *DT* had significantly higher scores than *PV* on Q1, Q2, Q4 and Q5. The p-values were 0.008 for Q1, 0.027 for Q2, 0.386 for Q3, 0.043 for Q4, 0.042 for Q5 and 0.503 for Q6. Responses for *DT* were significantly higher than 0 for all the questions, with p-values of 0.0002, 0.008, 0.001, 0.003, 0.008 and 0.005 for Q1-Q6. In the case of *PV*, the results were only significantly higher than 0 on Q3 and Q6. The p-values are 0.128, 0.862, 0.0002, 0.606, 0.833 and 0.0005 for Q1-Q6. These results from statistical analysis are annotated in the figure.

Figure 4-19 shows the results from the preference questionnaire. Except that slightly more participants believed that *PV* is easier to control than *DT*, more people preferred *DT* on all other criteria.

(3) Discussion

The results of experiment II show that *DT* is a powerful technique in generating the sensation of mass distribution. When using *DT*, participants could usually perceive which part of the object was heavier even when the CoM was only slightly off the geometric center. Evaluation of subjective user experience also shows that the illusion was perceived to be strong, realistic, and sufficient, and that the technique was appreciated. Overall, this empirical evidence shows the high effectiveness of *DT* in producing a pseudo-haptic illusion of mass distribution.

PV, on the other hand, seems to be much weaker in generating this illusion. Even though participants gave correct answers at a significantly higher rate than random in two of the four conditions, these conditions had very high mass ratios (13:1 and 33:1). Participants using *PV* did not find the illusion to be strong, realistic, sufficient, or desirable, and it was generally considered as the sub-optimal choice when compared to *DT*. This could be seen in both the individual evaluation of *PV* (Figure 4-18) and the comparison between the two techniques (Figure 4-19).

Both techniques were considered to be very easy to control (Q6), and slightly more participants thought *PV* was more controllable. This is not a surprise, as *PV* maintains the 1:1 rotational C/D ratio throughout, while *DT* changes it on the fly. When controllability is at high priority, *PV* may still have some value even though its pseudo-haptic illusion seems to be much weaker.

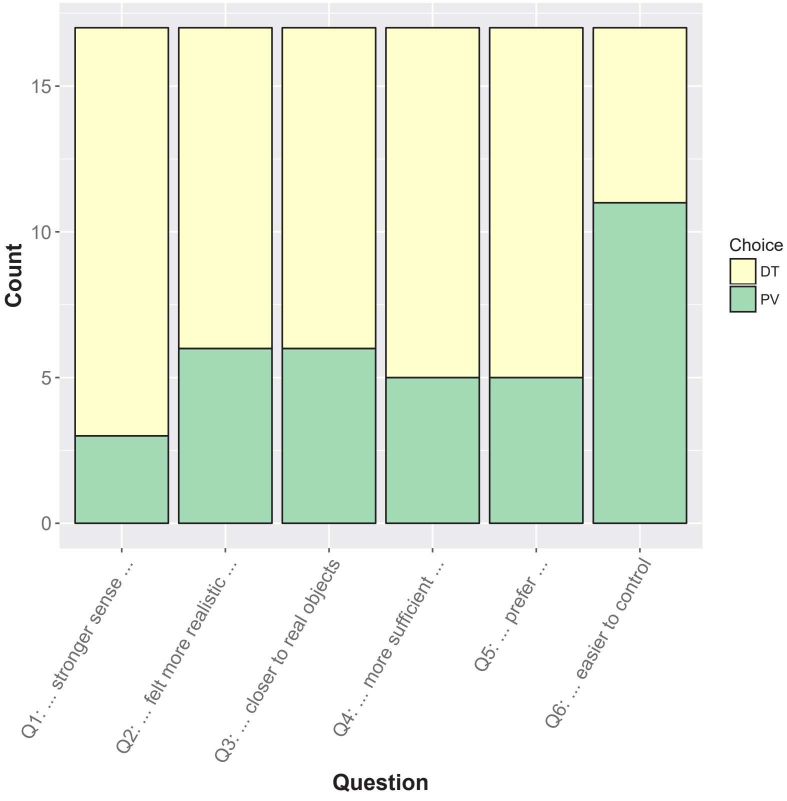


Figure 4-19: Results of the preference questionnaire from experiment II.

A limitation of the *DT* technique is that there are several empirically determined constants in our algorithm. These were specifically tweaked to fit the current experiment, and they may need to be further modified when using *DT* for a new VE. Another limitation is that the influence from the mass distribution of the physical proxy is not explored in this experiment, and we consider it as part of the future work.

4.3.2.4 Summary

This research contributes physics-based theories and several pseudo-haptic techniques to generate mass-related perceptual illusions during rotational motion in VR object manipulation. Experiment I showed that one can induce the illusion of different total masses by scaling objects' rotational C/D ratios. In addition, we also found empirical evidence that by altering the C/D ratio dynamically based on the physical principles that explain the motion of rotation, we can produce a convincing illusion of mass distribution. Overall, these pseudo-haptic techniques combine simple virtual hand technique with perceivable physics plausibility, which should achieve a balance between controllability and physics coherence.

5. CONCLUSIONS AND FUTURE WORK

5.1 Conclusions

This dissertation explores the design of coherent techniques for two fundamental VR interaction tasks, namely walking and object manipulation. While existing studies about coherence usually put the user in a passive role to "view" the behaviors of a VE, our work focuses on designing the user's active interaction with the VE in a coherent manner. In this way, this dissertation contributes in connecting a core concept of VR experience to fundamental 3D UI design.

In this exploration, we proposed several research questions to address two aspects of the issue:

(1) How can we design coherent interactions (RQ1 for walking and RQ3 for object manipulation)?

(2) How does coherence affect user experience (RQ2 for walking and RQ4 for object manipulation)?

In addressing (1), this dissertation provides a "toolbox" for designing coherent interactions that readers could either directly apply or draw inspiration from. For RDW, we propose a family of techniques called "narrative driven cell-based redirection." Several examples were provided to explain the concept and its usage in real-world applications. For object manipulation, we focus on a specific aspect of coherence, namely physics coherence, and provide the theoretical models with several exploratory techniques to inspire the design of physically coherent techniques.

To address (2), we conducted several user studies to evaluate the user experience of proposed coherent techniques. In the case of RDW, a study evaluated the users' spatial understanding when using the novel coherent techniques in comparison with non-coherent counterparts. Even though little statistical significance was found, we identified clues that could inspire promising future work in this direction. As for object manipulation, we found positive influence on hedonic experience from physics coherence, while also identifying potential deficiencies in controllability in some of the novel coherent techniques. On top of that, we directly measured perceived physics plausibility by applying methods from pseudo-haptic research and demonstrated effectiveness of the proposed novel techniques in enhancing physics coherence using rotational motion.

Overall, this dissertation provides theoretical frameworks, design innovations, and knowledge in user experiences in the exploration of connecting coherence to 3D UI design. In the next section, we summarize its limitations into several themes, highlighting potential future work to follow up.

5.2 Future Work

5.2.1 Mechanism of Narrative Driven Cell-Based Redirection

Chapter 3 evolves around the design of narrative driven cell-based redirection. To successfully apply this kind of technique, appropriate metaphors need to be found to fit the context of the VE and the redirection phase needs to be carefully designed accordingly (e.g., the features of dimming brightness and haptic feedback in Bookshelf). Following these requirements, two general directions are worth exploring in the future. First, additional metaphors could be designed to compose a “library” to facilitate coherent RDW. Section 3.1.3 briefly discussed potential ones, such as elevators and revolving doors, and more examples could be found for particular VEs. Second, features to enhance user experience during the redirection phase could be explored. They should aim at providing a better sense of safety, decreasing potential cyber-sickness and increasing the overall coherence. For example, we could install a pressure sensor on the platform of the physical prop for Bookshelf, which could help dynamically adjust some of the visual or haptic feedback depending on the user’s standing posture during the re-orientation. Parameters for the redirection mechanism could be further optimized. For example, the rotation speed of the Bookshelf needs to be carefully set to reduce potential cyber-sickness while maintaining spatial reference for the user.

5.2.2 Potential Benefits of Coherent RDW

One major limitation of the current evaluation (Section 3.1.5) is that little statistical significance could be found to support the benefit of coherent RDW. This is at least partially due to the limited sample size of the study. That being said, we still believe that these coherent techniques could potentially bring differences in spatial understanding. In the future, a similar study with larger sample size could be conducted to collect more data for more powerful analysis. The new experiment could also put more emphasis on the task performance—if the user could complete wayfinding faster using the proposed technique (when he is told to optimize on completion time), it could be an indication of superior spatial understanding.

Apart from understanding of space, there are other aspects of user experience that could benefit from coherent RDW. For example, we could design an experiment that directly measures presence, with the assumption that these coherent techniques should enhance presence since they don’t break the seamlessness of the walking process. Apart from usability and presence, the proposed techniques could also have potential hedonic values compared to the non-coherent ones. For example, users may enjoy the continuous relocation facilitated by the Bird more compared to unrealistic Teleportation; people who use the lighting-driven redirection might find the learning process more engaging than using teleportation to move around.

5.2.3 Potential Trade-off: Physics Coherence vs. Controllability

In Chapter 4, we state that physics-based motion could help enhance physics coherence of object manipulation, especially the component of **temporal coherence of visual cues.** In VR, such motion is usually achieved by simulation algorithms that are embedded in a physics engine. Among the evaluated techniques in this document, Force Push applies this strategy to make the object’s movement appear plausible. However, such physics-based motion is driven by force and acceleration, just like how objects move in the real-world. Controlling force and acceleration appears to be much more difficult than zero-order mapping (Zhai, 1998), and this is shown by the evaluation of Force Push, as its deficiency in controllability severely limits its perceived usability. It seems that there is a potential trade-off between the **temporal coherence of visual cues** (as an important component of overall physics coherence) and the controllability of the interface.

An important question to follow up is whether we use force-driven motion to enhance physics coherence while keeping the technique easy to control?

Our tentative answer to this question is yes, but for most cases this requires force-based control to be combined with zero-order mapping. To create such a combination, two conditions have to be met simultaneously. First, the user must have easy control over how forces exerted onto the object could be configured; second, the object's mass cannot be too large relative to the amount of force generated from hand input, so its movement can be changed quickly in a predictable way without too much influence from its inertia. In this way, force works as a constraint to "lock" the object's movement to the user's desire, making it a force-driven zero-order mapping that's easy to control.

We can use these two conditions to inspect a manipulation technique. First, let's consider how we grab and move a light, small item in reality. In this case, both conditions are met—we can easily maneuver how forces are exerted on the object, as there are multiple contact points with rich haptic feedback when the object is held in hand, and the hand locks the object's movement in a predictable manner because its inertia is insignificant relative to the forces from hand input. Now let's consider how we move a large piece of furniture in real life: the situation becomes much less controllable. First, because of the size of the furniture, we could only leverage one or two contact points to "push" or "drag" it to a certain direction at one time; second, the object's motion will be significantly influenced by the large mass of the object, making it much less predictable. The technique of Force Push is essentially placed in a similar situation, with the added difficulty that the user has to learn how to control the exerted force using hand gestures.

Realizing these reasons behind this trade-off, it seems that it's inevitable to sacrifice either the physics coherence or the controllability when we need to control a large, heavy item in VR. If controllability is prioritized, we can use the simple virtual hand technique that applies a zero-order mapping between the physical proxy's movement and the virtual object's movement. However, the virtual item will move like it's much lighter than it should be, breaking physics coherence and physics plausibility. If we sacrifice controllability, force-driven motion can be used, which makes its movement visually plausible and gives the user an appropriate sense of heaviness, but the usability could be severely damaged compared to zero-order mapping (for most tasks).

That being said, we believe there are some "middle ground" solutions that achieve a reasonable compromise between the two. Some of the pseudo-haptic techniques discussed in this document are good examples in this direction. For instance, by increasing the translational and rotational C/D ratio for heavier objects, the system could generate a perception of relative mass difference between different objects without breaking the desirable zero-order mapping. This is one step closer in achieving physics coherence, since the relationship between motion and force demonstrated by the interface is relatively closer to fundamental physical laws compared to naïve simple virtual hand, yet it does not rely on actual physics-based simulation, which would require force-based motion that breaks the zero-order mapping.

In the future, we'd like to explore this general direction of balancing between physics coherence and controllability. This document could only explore the limited design space that uses a commercial hand-held proxy and bare-hand gestures. It would be particularly interesting when different haptic devices are put in the equation. For instance, how should we approach this act of balancing if we have a mid-fidelity haptic rendering device that provides much more than vibration but not enough to simulate the actual object's mass and shape? What if we use a hand-held proxy that could vary its shape and mass distribution on the fly? Would gesture-based methods, such as Force Push, be improved if richer haptic feedback could be provided? The potential design space is huge and much more exploration is needed to draw a complete picture of the relationship between physics coherence and controllability.

5.2.4 Perception of Absolute Mass

In Chapter 4, we explored how adjusting C/D ratio could induce a sensation of mass difference between two virtual objects and mass distribution within the same item. However, these illusions are both about the *relative* heaviness between two objects or two parts of the same object (i.e. which one is relatively heavier). How the *absolute* quantity of mass is perceived is not discussed.

We can imagine an experiment that asks the participant to directly report the item's weight, similar to the one conducted by Samad et al. (Samad et al., 2019). In their experiment, this report of absolute quantity was generated by comparing the perception of the object's weight to the simulated weight from a haptic device. From there, we could try to train a mathematical model that maps the value of the technique's parameter (e.g., C/D ratio) to the absolute value of the perceived mass. We expect that designers could use this model to accurately induce the intended physics plausibility into the interaction, similar to how Stevens' psychophysical laws are applied (Stevens, 1957). We'd also expect that gesture-based approaches like Force Push might demonstrate unique advantages in this direction. As explained in Section (4) of Chapter 4.2.1.4, gestural input enables richer, more expressive mapping from kinesthetic cues to the object's movement, which could potentially enlarge the range of perceivable mass.

So far, we've only seen one case of this direction being explored in translational control (Samad et al., 2019), and it would be particularly interesting to investigate the perception of actual mass and mass distribution using the novel, rotation-related techniques proposed in this document. For example, in generating the illusion of mass distribution, what's the perceived value of the ratio between the heavier part and the lighter part when using these techniques? How do we tweak the parameters of the techniques to gradually alter this value? These questions require more psycho-physical experiments to address and we consider them as promising future work.

REFERENCES

Aigner, R., Wigdor, D., Benko, H., Haller, M., Lindbauer, D., Ion, A., et al. (2012). Understanding mid-air hand gestures: A study of human preferences in usage of gesture types for hci. *Microsoft Research TechReport MSR-TR-2012-111* 2.

Akin, D.L., Minsky, M.L., Thiel, E., and Kurtzman, C. (1983). Space Applications of Automation, Robotics and Machine Intelligence Systems (ARAMIS), phase 2. Volume 1: Telepresence technology base development.

Ban, Y., Kajinami, T., Narumi, T., Tanikawa, T., and Hirose, M. (2012). "Modifying an identified curved surface shape using pseudo-haptic effect", in: *Haptics Symposium (HAPTICS), 2012 IEEE*. 211-216.

Barzel, R., Hughes, J.R., and Wood, D.N. (1996). "Plausible motion simulation for computer graphics animation," in *Computer Animation and Simulation'96*. Springer, 183-197.

Baumgartner, T., Speck, D., Wettstein, D., Masnari, O., Beeli, G., and Jäncke, L. (2008). Feeling present in arousing virtual reality worlds: prefrontal brain regions differentially orchestrate presence experience in adults and children. *Frontiers in human neuroscience* 2**,** 8.

Benko, H., and Wilson, A.D. (2009). "DepthTouch: using depthsensing camera to enable freehand interactions on and above the interactive surface", in: *In Proceedings of the IEEE Workshop on Tabletops and Interactive Surfaces*.

Biocca, F., Kim, J., and Choi, Y. (2001). Visual touch in virtual environments: An exploratory study of presence, multimodal interfaces, and cross-modal sensory illusions. *Presence: Teleoperators & Virtual Environments* 10(3)**,** 247-265.

Borst, C.W., and Indugula, A.P. (2005). "Realistic virtual grasping", in: *Proceeding of IEEE Virtual Reality 2005 (VR 2005)*. 91-98.

Borst, C.W., and Indugula, A.P. (2006). A spring model for whole-hand virtual grasping. *Presence: Teleoperators & Virtual Environments* 15(1)**,** 47-61.

Bowman, D., Kruijff, E., LaViola Jr, J.J., and Poupyrev, I. (2004). *3D User Interfaces: Theory and Practice.* Addison-Wesley.

Bowman, D.A., and Hodges, L.F. (1997). "An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments", in: *Proceedings of the 1997 symposium on Interactive 3D graphics*. 35-ff.

Bowman, D.A., Koller, D., and Hodges, L.F. (1997). "Travel in immersive virtual environments: An evaluation of viewpoint motion control techniques", in: *Virtual Reality Annual International Symposium, 1997., IEEE 1997*. 45-52, 215.

Bruder, G., Steinicke, F., and Hinrichs, K.H. (2009). "Arch-explore: A natural user interface for immersive architectural walkthroughs", in: *IEEE Symposium on 3D User Interfaces*. 75-82.

Carassa, A., Morganti, F., and Tirassa, M. (2005). A situated cognition perspective on presence.

Card, S.K., Mackinlay, J.D., and Robertson, G.G. (1990). "The design space of input devices", in: *Proceedings of the SIGCHI conference on Human factors in computing systems*. 117-124.

Chance, S.S., Gaunet, F., Beall, A.C., and Loomis, J.M. (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)**,** 168-178.

Cheng, C.-H., Chang, C.-C., Chen, Y.-H., Lin, Y.-L., Huang, J.-Y., Han, P.-H., et al. (2018). "GravityCup: a liquid-based haptics for simulating dynamic weight in virtual reality", in: *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*. 51.

Chenney, S., and Forsyth, D.A. (2000). "Sampling plausible solutions to multi-body constraint problems", in: *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*. 219-228.

Cohen, R. (2013). *The development of spatial cognition.* Psychology Press.

Darken, R.P., Cockayne, W.R., and Carmein, D. (1997). "The omni-directional treadmill: a locomotion device for virtual worlds", in: *Proceedings of the 10th annual ACM symposium on User interface software and technology*. 213-221.

Dickinson, J. (1974). Proprioceptive Control of Human Movement. The Human Movement Series.

Dillon, C., Keogh, E., and Freeman, J. (2002). "It's been emotional": Affect, physiology, and presence", in: *Proceedings of the Fifth Annual International Workshop on Presence. Porto, Portugal*.

Dominjon, L., Lécuyer, A., Burkhardt, J.-M., Richard, P., and Richir, S. (2005). "Influence of control/display ratio on the perception of mass of manipulated objects in virtual environments", in: *Proceeding of IEEE Virtual Reality 2005 (VR 2005)*. 19-25.

Feasel, J., Whitton, M.C., and Wendt, J.D. (2008). "LLCM-WIP: Low-latency, continuous-motion walking-in-place", in: *3D User Interfaces, 2008. 3DUI 2008. IEEE Symposium on*. 97-104.

Freeman, J., Avons, S.E., Meddis, R., Pearson, D.E., and IJsselsteijn, W. (2000). Using behavioral realism to estimate presence: A study of the utility of postural responses to motion stimuli. *Presence: Teleoperators & Virtual Environments* 9(2)**,** 149-164.

Frees, S., Kessler, G.D., and Kay, E. (2007). PRISM interaction for enhancing control in immersive virtual environments. *ACM Transactions on Computer-Human Interaction (TOCHI)* 14(1)**,** 2.

Goetgeluk, J., Soza, R., Bester, D., and Shuffield, J.D. (2015). "Method of generating an input in an omnidirectional locomotion system". Google Patents).

Goldstein, E.B., and Brockmole, J. (2016). *Sensation and perception.* Cengage Learning.

Goldstein, H., Poole, C., and Safko, J. (2002). "Classical mechanics". AAPT).

Grechkin, T., Azmandian, M., Bolas, M., and Suma, E. (2015). "Towards context-sensitive reorientation for real walking in virtual reality", in: *2015 IEEE Virtual Reality (VR)*. 185-186.

Grechkin, T., Thomas, J., Azmandian, M., Bolas, M., and Suma, E. (2016). "Revisiting detection thresholds for redirected walking: combining translation and curvature gains", in: *Proceedings of the ACM Symposium on Applied Perception*. 113-120.

Gurari, N., Kuchenbecker, K.J., and Okamura, A.M. (2009). "Stiffness discrimination with visual and proprioceptive cues", in: *EuroHaptics Conference, 2009 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics 2009. Third Joint*. 121-126.

Hashiguchi, S., Mori, S., Tanaka, M., Shibata, F., and Kimura, A. (2018). "Perceived weight of a rod under augmented and diminished reality visual effects", in: *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*. 12.

Hassenzahl, M. (2003). "The thing and I: understanding the relationship between user and product," in *Funology*. Springer, 31-42.

Hassenzahl, M. (2004). The interplay of beauty, goodness, and usability in interactive products. *Human-computer interaction* 19(4)**,** 319-349.

Hilliges, O., Kim, D., Izadi, S., Weiss, M., and Wilson, A. (2012). "HoloDesk: direct 3d interactions with a situated see-through display", in: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2421-2430.

Hinckley, K., Jacob, R.J., Ware, C., Wobbrock, J.O., and Wigdor, D. (2004). "Input/Output Devices and Interaction Techniques," in *The Computer Science Handbook, Second Edition*. CRC Press, LLC.

Hix, D., and Hartson, H.R. (1993). *Developing user interfaces: ensuring usability through product & process.* John Wiley & Sons, Inc.

Hodgson, E., and Bachmann, E. (2013). Comparing four approaches to generalized redirected walking: Simulation and live user data. *IEEE transactions on visualization and computer graphics* 19(4)**,** 634-643.

Höll, M., Oberweger, M., Arth, C., and Lepetit, V. (2018). "Efficient Physics-Based Implementation for Realistic Hand-Object Interaction in Virtual Reality", in: *2018 IEEE Conference on Virtual Reality and 3D User Interfaces*.

Hummel, J., Dodiya, J., Wolff, R., Gerndt, A., and Kuhlen, T. (2013). "An evaluation of two simple methods for representing heaviness in immersive virtual environments", in: *2013 IEEE Symposium on 3D User Interfaces (3DUI)*. 87-94.

IJsselsteijn, W., De Kort, Y., Poels, K., Jurgelionis, A., and Bellotti, F. (2007). "Characterising and measuring user experiences in digital games", in: *International conference on advances in computer entertainment technology*. 27.

Kim, J.-S., Gračanin, D., Matković, K., and Quek, F. (2008). "Finger walking in place (FWIP): A traveling technique in virtual environments", in: *International Symposium on Smart Graphics*. 58-69.

Kim, J.-S., Gračanin, D., Yang, T., and Quek, F. (2015a). Action-transferred navigation technique design approach supporting human spatial learning. *ACM Transactions on Computer-Human Interaction (TOCHI)* 22(6)**,** 30.

Kim, K., Jackson, B., Karamouzas, I., Adeagbo, M., Guy, S.J., Graff, R., et al. (2015b). "Bema: A multimodal interface for expert experiential analysis of political assemblies at the Pnyx in ancient Greece", in: *IEEE Symposium on 3D User Interfaces*. 19-26.

Kopper, R.A.P. (2011). *Understanding and Improving Distal Pointing Interaction.* Doctoral dissertation, Virginia Tech.

Langbehn, E., Lubos, P., Bruder, G., and Steinicke, F. (2017). Bending the curve: Sensitivity to bending of curved paths and application in room-scale vr. *IEEE transactions on visualization and computer graphics* 23(4)**,** 1389-1398.

Langbehn, E., and Steinicke, F. (2018). Redirected walking in virtual reality. *Encyclopedia of Computer Graphics and Games. Springer International Publishing*.

Larrue, F., Sauzeon, H., Wallet, G., Foloppe, D., Cazalets, J.-R., Gross, C., et al. (2014). Influence of body-centered information on the transfer of spatial learning from a virtual to a real environment. *Journal of Cognitive Psychology* 26(8)**,** 906-918.

Laugwitz, B., Held, T., and Schrepp, M. (2008). "Construction and evaluation of a user experience questionnaire", in: *Symposium of the Austrian HCI and Usability Engineering Group*. 63-76.

LaViola Jr, J.J., Kruijff, E., McMahan, R.P., Bowman, D., and Poupyrev, I.P. (2017). *3D user interfaces: Theory and practice.* Addison-Wesley Professional.

Laycock, S.D., and Day, A. (2007). "A survey of haptic rendering techniques", in: *Computer Graphics Forum*. 50-65.

Lécuyer, A. (2009). Simulating haptic feedback using vision: A survey of research and applications of pseudo-haptic feedback. *Presence: Teleoperators and Virtual Environments* 18(1)**,** 39-53.

Lecuyer, A., Coquillart, S., Kheddar, A., Richard, P., and Coiffet, P. (2000). "Pseudo-haptic feedback: Can isometric input devices simulate force feedback?", in: *Virtual Reality, 2000. Proceedings. IEEE*. 83-90.

Lee, K.M. (2004). Presence, explicated. *Communication theory* 14(1)**,** 27-50.

Lessiter, J., Freeman, J., Keogh, E., and Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators & Virtual Environments* 10(3)**,** 282-297.

Lévesque, J.-C., Laurendeau, D., and Mokhtari, M. (2011). "Bimanual gestural interface for virtual environments", in: *Virtual Reality Conference (VR), 2011 IEEE*. 223-224.

Lin, M.C., and Otaduy, M. (2008). *Haptic rendering: foundations, algorithms, and applications.* AK Peters/CRC Press.

Llobera, J., Blom, K.J., and Slater, M. (2013). Telling stories within immersive virtual environments. *Leonardo* 46(5)**,** 471-476.

Lombard, M., Ditton, T.B., Crane, D., Davis, B., Gil-Egui, G., Horvath, K., et al. (2000). "Measuring presence: A literature-based approach to the development of a standardized paper-and-pencil instrument", in: *Third international workshop on presence, delft, the netherlands*. 2-4.

Lukos, J., Ansuini, C., and Santello, M. (2007). Choice of contact points during multidigit grasping: effect of predictability of object center of mass location. *Journal of Neuroscience* 27(14)**,** 3894-3903.

Marsh, W.E., Putnam, M., Kelly, J.W., Dark, V.J., and Oliver, J.H. (2012). "The cognitive implications of semi-natural virtual locomotion", in: *Virtual Reality Short Papers and Posters (VRW), 2012 IEEE*. 47-50.

McMahan, R.P. (2011). Exploring the effects of higher-fidelity display and interaction for virtual reality games.

Meehan, M., Insko, B., Whitton, M., and Brooks Jr, F.P. (2002). Physiological measures of presence in stressful virtual environments. *ACM Transactions on Graphics (TOG)* 21(3)**,** 645-652.

Mine, M.R., Brooks Jr, F.P., and Sequin, C.H. (1997). "Moving objects in space: exploiting proprioception in virtual-environment interaction", in: *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*. 19-26.

Minsky, M. (1980). Telepresence.

Mirnig, A.G., Meschtscherjakov, A., Wurhofer, D., Meneweger, T., and Tscheligi, M. (2015). "A formal analysis of the ISO 9241-210 definition of user experience", in: *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. 437-450.

Murray, D.J., Ellis, R.R., Bandomir, C.A., and Ross, H.E. (1999). Charpentier (1891) on the size—weight illusion. *Perception & Psychophysics* 61(8)**,** 1681-1685.

Nabiyouni, M. (2016). *How Does Interaction Fidelity Influence User Experience in VR Locomotion?* Doctoral dissertation, Virginia Tech.

Neth, C.T., Souman, J.L., Engel, D., Kloos, U., Bulthoff, H.H., and Mohler, B.J. (2012). Velocity-dependent dynamic curvature gain for redirected walking. *IEEE transactions on visualization and computer graphics* 18(7)**,** 1041-1052.

Norman, K.L. (2013). Geq (game engagement/experience questionnaire): a review of two papers. *Interacting with Computers* 25(4)**,** 278-283.

Norton, J., Wingrave, C.A., and LaViola Jr, J.J. (2010). "Exploring strategies and guidelines for developing full body video game interfaces", in: *Proceedings of the Fifth International Conference on the Foundations of Digital Games*. 155-162.

O'Sullivan, C., and Dingliana, J. (2001). Collisions and perception. *ACM Transactions on Graphics (TOG)* 20(3)**,** 151-168.

O'Sullivan, C., Dingliana, J., Giang, T., and Kaiser, M.K. (2003a). "Evaluating the visual fidelity of physically based animations", in: *ACM Transactions on Graphics (TOG)*. 527-536.

O'Sullivan, C., Dingliana, J., and Howlett, S. (2003b). "Eye-movements and interactive graphics," in *The Mind's Eye*. Elsevier, 555-571.

O'Sullivan, C., Howlett, S., Morvan, Y., McDonnell, R., and O'Conor, K. (2004). Perceptually adaptive graphics. *Eurographics state of the art reports* 4**,** 1-24.

Ortega, F.R., Galvan, A., Tarre, K., Barreto, A., Rishe, N., Bernal, J., et al. (2017). "Gesture elicitation for 3D travel via multi-touch and mid-Air systems for procedurally generated pseudo-universe", in: *3D User Interfaces (3DUI), 2017 IEEE Symposium on*. 144-153.

Parola, M., Johnson, S., and West, R. (2016). Turning presence inside-out: MetaNarratives. *Electronic Imaging* 2016(4)**,** 1-9.

Peck, T.C., Fuchs, H., and Whitton, M.C. (2009). Evaluation of reorientation techniques and distractors for walking in large virtual environments. *IEEE Transactions on Visualization and Computer Graphics* 15(3)**,** 383-394.

Peck, T.C., Fuchs, H., and Whitton, M.C. (2012). The design and evaluation of a large-scale real-walking locomotion interface. *IEEE transactions on visualization and computer graphics* 18(7)**,** 1053-1067.

Pierce, J.S., Stearns, B.C., and Pausch, R. (1999). "Voodoo dolls: seamless interaction at multiple scales in virtual environments", in: *Proceedings of the 1999 symposium on Interactive 3D graphics*. 141-145.

Poupyrev, I., Billinghurst, M., Weghorst, S., and Ichikawa, T. (1996). "The go-go interaction technique: non-linear mapping for direct manipulation in VR", in: *Proceedings of the 9th annual ACM symposium on User interface software and technology*. 79-80.

Poupyrev, I., Weghorst, S., and Fels, S. (2000). "Non-isomorphic 3D rotational techniques", in: *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 540-547.

Punpongsanon, P., Iwai, D., and Sato, K. (2015). Softar: Visually manipulating haptic softness perception in spatial augmented reality. *IEEE transactions on visualization and computer graphics* 21(11)**,** 1279-1288.

Pusch, A., and Lécuyer, A. (2011). "Pseudo-haptics: from the theoretical foundations to practical system design guidelines", in: *Proceedings of the 13th international conference on multimodal interfaces*. 57-64.

Pusch, A., Martin, O., and Coquillart, S. (2008). "Hemp-hand-displacement-based pseudo-haptics: a study of a force field application", in: *3D User Interfaces, 2008. 3DUI 2008. IEEE Symposium on*. 59-66.

Razzaque, S. (2005). *Redirected walking.* Ph.D. dissertation, University of North Carolina at Chapel Hill.

Razzaque, S., Kohn, Z., and Whitton, M.C. (2001). "Redirected walking", in: *Proceedings of EUROGRAPHICS*. 105-106.

Riener, C., and Proffitt, D. (2002). "Quantifying spatial presence", in: *Proceedings of PRESENCE 2002*. 345-352.

Rietzler, M., Geiselhart, F., Frommel, J., and Rukzio, E. (2018a). "Conveying the perception of kinesthetic feedback in virtual reality using state-of-the-art hardware", in: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 460.

Rietzler, M., Geiselhart, F., Gugenheimer, J., and Rukzio, E. (2018b). "Breaking the tracking: Enabling weight perception using perceivable tracking offsets", in: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 128.

Rosakranse, C., and Oh, S.Y. (2014). "Measuring presence: The use trends of five canonical presence questionnaires from 1998–2012", in: *Proceedings of the 15th International Workshop on Presence (ISPR'14)*.

Ruddle, R.A., Volkova, E., and Bülthoff, H.H. (2011). Walking improves your cognitive map in environments that are large-scale and large in extent. *ACM Transactions on Computer-Human Interaction (TOCHI)* 18(2)**,** 10.

Samad, M., Gatti, E., Hermes, A., Benko, H., and Parise, C. (2019). "Pseudo-Haptic Weight: Changing the Perceived Weight of Virtual Objects By Manipulating Control-Display Ratio", in: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 320.

Schnabel, M.A., and Kvan, T. (2003). Spatial understanding in immersive virtual environments. *International Journal of Architectural Computing* 1(4)**,** 435-448.

Schubert, T., Friedmann, F., and Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments* 10(3)**,** 266-281.

Sheridan, T.B. (1992). Musings on telepresence and virtual presence. *Presence: Teleoperators & Virtual Environments* 1(1)**,** 120-126.

Shneiderman, B., Plaisant, C., Cohen, M., Jacobs, S., Elmqvist, N., and Diakopoulos, N. (2016). *Designing the user interface: strategies for effective human-computer interaction.* Pearson.

Skarbez, R., Brooks, F.P., and Whitton, M.C. (2017a). "Immersion and coherence in a visual cliff environment", in: *Proceeding of IEEE Virtual Reality 2017 (VR 2017)*. 397-398.

Skarbez, R., Brooks Jr, F.P., and Whitton, M.C. (2017b). A Survey of Presence and Related Concepts. *ACM Computing Surveys (CSUR)* 50(6)**,** 96.

Skarbez, R., Neyret, S., Brooks, F.P., Slater, M., and Whitton, M.C. (2017c). A psychophysical experiment regarding components of the plausibility illusion. *IEEE transactions on visualization and computer graphics* 23(4)**,** 1369-1378.

Skarbez, R.T. (2016). *Plausibility illusion in virtual environments.* Ph.D. dissertation, The University of North Carolina at Chapel Hill.

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364(1535)**,** 3549-3557.

Slater, M., Brogni, A., and Steed, A. (2003). "Physiological responses to breaks in presence: A pilot study", in: *Presence 2003: The 6th Annual International Workshop on Presence*.

Slater, M., McCarthy, J., and Maringelli, F. (1998). The influence of body movement on subjective presence in virtual environments. *Human Factors* 40(3)**,** 469-477.

Slater, M., and Usoh, M. (1993). Representations systems, perceptual position, and presence in immersive virtual environments. *Presence: Teleoperators & Virtual Environments* 2(3)**,** 221-233.

Slater, M., Usoh, M., and Chrysanthou, Y. (1995). "The influence of dynamic shadows on presence in immersive virtual environments," in *Virtual Environments' 95*. Springer, 8-21.

Song, P., Goh, W.B., Hutama, W., Fu, C.-W., and Liu, X. (2012). "A handle bar metaphor for virtual object manipulation with mid-air interaction", in: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1297-1306.

Spagnolli, A., and Gamberini, L. (2004). The sense of being 'There': a model for the space of presence. *EXPLORING THE SENSE OF PRESENCE***,** 48.

Steinicke, F., Bruder, G., Jerald, J., Frenz, H., and Lappe, M. (2010). Estimation of detection thresholds for redirected walking techniques. *IEEE Transactions on Visualization and Computer Graphics* 16(1)**,** 17-27.

Steinicke, F., Visell, Y., Campos, J., and Lécuyer, A. (2013). *Human walking in virtual environments.* Springer.

Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of communication* 42(4)**,** 73-93.

Stevens, S.S. (1957). On the psychophysical law. *Psychological review* 64(3)**,** 153.

Stoakley, R., Conway, M.J., and Pausch, R. (1995). "Virtual reality on a WIM: interactive worlds in miniature", in: *Proceedings of the SIGCHI conference on Human factors in computing systems*. 265-272.

Stuart, R. (2001). *Design of virtual environments.* Barricade Books, Incorporated.

Suma, E., Finkelstein, S., Reid, M., Babu, S., Ulinski, A., and Hodges, L.F. (2009). Evaluation of the cognitive effects of travel technique in complex real and virtual environments. *IEEE Transactions on Visualization and Computer Graphics* 16(4)**,** 690-702.

Suma, E.A., Azmandian, M., Grechkin, T., Phan, T., and Bolas, M. (2015). "Making small spaces feel large: Infinite walking in virtual reality", in: *ACM SIGGRAPH 2015 Emerging Technologies*. 16.

Suma, E.A., Bruder, G., Steinicke, F., Krum, D.M., and Bolas, M. (2012). "A taxonomy for deploying redirection techniques in immersive virtual environments", in: *IEEE Virtual Reality Workshops (VRW)*. 43-46.

Suma, E.A., Clark, S., Krum, D., Finkelstein, S., Bolas, M., and Warte, Z. (2011). "Leveraging change blindness for redirection in virtual environments", in: *2011 IEEE Virtual Reality Conference (VR)*. 159-166.

Swapp, D., Williams, J., and Steed, A. (2010). "The implementation of a novel walking interface within an immersive display", in: *3D User Interfaces (3DUI), 2010 IEEE Symposium on*. 71-74.

Taima, Y., Ban, Y., Narumi, T., Tanikawa, T., and Hirose, M. (2014). "Controlling fatigue while lifting objects using pseudo-haptics in a mixed reality space", in: *2014 IEEE Haptics Symposium (HAPTICS)*. 175-180.

Talvas, A., Marchal, M., Duriez, C., and Otaduy, M.A. (2015). Aggregate constraints for virtual manipulation with soft fingers. *IEEE transactions on visualization and computer graphics* 21(4)**,** 452-461.

Templeman, J.N., Denbrook, P.S., and Sibert, L.E. (1999). Virtual locomotion: Walking in place through virtual environments. *Presence* 8(6)**,** 598-617.

Thie, S., and Van Wijk, J. (1998). A general theory on presence. *1st Int. Wkshp. on Presence*.

Usoh, M., Arthur, K., Whitton, M.C., Bastos, R., Steed, A., Slater, M., et al. (1999). "Walking> walking-in-place> flying, in virtual environments", in: *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*. 359-364.

Usoh, M., Catena, E., Arman, S., and Slater, M. (2000). Using presence questionnaires in reality. *Presence: Teleoperators & Virtual Environments* 9(5)**,** 497-503.

van Beurden, M.H., Ijsselsteijn, W.A., and de Kort, Y.A. (2011). "User experience of gesture based interfaces: a comparison with traditional interaction methods on pragmatic and hedonic qualities", in: *International Gesture Workshop*. 36-47.

Wang, Z., Lu, J., Peer, A., and Buss, M. (2010). "Influence of vision and haptics on plausibility of social interaction in virtual reality scenarios", in: *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*. 172-177.

Whitehead, B.A. (1981). "James J. Gibson: The ecological approach to visual perception. Boston: Houghton Mifflin, 1979, 332 pp". Wiley Online Library).

Wigdor, D., and Wixon, D. (2011). *Brave NUI world: designing natural user interfaces for touch and gesture.* Elsevier.

Williams, B., Narasimham, G., Rump, B., McNamara, T.P., Carr, T.H., Rieser, J., et al. (2007). "Exploring large virtual environments with an HMD when physical space is limited", in: *Proceedings of the 4th symposium on Applied perception in graphics and visualization*. 41-48.

Wingrave, C.A., Bowman, D.A., and Ramakrishnan, N. (2002). "Towards preferences in virtual environment interfaces", in: *EGVE*. 63-72.

Witmer, B.G., and Singer, M.J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence* 7(3)**,** 225-240.

Yu, R., and Bowman, D.A. (2018a). "Force Push: Exploring Expressive Gesture-to-Force Mappings for Indirect 3D Object Manipulation", in: *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 733-734.

Yu, R., and Bowman, D.A. (2018b). Force Push: Exploring Expressive Gesture-to-Force Mappings for Remote Object Manipulation in Virtual Reality. *Frontiers in ICT* 5**,** 25.

Yu, R., Duer, Z., Ogle, T., Bowman, D.A., Tucker, T., Hicks, D., et al. (2018). "Experiencing an Invisible World War I Battlefield Through Narrative-Driven Redirected Walking in Virtual Reality", in: *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 313-319.

Yu, R., Lages, W.S., Nabiyouni, M., Ray, B., Kondur, N., Chandrashekar, V., et al. (2017). "Bookshelf and Bird: Enabling Real Walking in Large VR Spaces through Cell-Based Redirection", in: *IEEE Symposium on 3D User Interfaces*. 116-119.

Zahorik, P., and Jenison, R.L. (1998). Presence as being-in-the-world. *Presence* 7(1)**,** 78-89.

Zenner, A. (2016). Investigating weight distribution in virtual reality proxy interaction. *Universität des Saarlandes, Masterarbeit*.

Zenner, A., and Krüger, A. (2017). Shifty: A weight-shifting dynamic passive haptic proxy to enhance object perception in virtual reality. *IEEE transactions on visualization and computer graphics* 23(4)**,** 1285-1294.

Zhai, S. (1998). User performance in relation to 3D input device design. *ACM Siggraph Computer Graphics* 32(4)**,** 50-54.

APPENDIX A. USER STUDY MATERIALS

A.1 Study of Bookshelf and Bird Techniques

A.1.1 User Consent Form

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
Informed Consent for Participants
in Research Projects Involving Human Subjects

Title of Project: Locomotion Techniques and Spatial Orientation in Virtual Environments

Investigator(s):

Name: Emmet Hobgood; E-mail/Phone number: ehobgood@vt.edu

Name: Douglas Bowman; E-mail/Phone number: dbowman@vt.edu

I. Purpose of this Research Project

This research project is intended to compare a variety of methods for moving around in virtual reality when the physical space is limited. Virtual reality is a promising technology, with the ability to create virtual environments much larger and more elaborate than the real space. This research will help us understand the best ways to navigate these large environments. In particular, this study will be testing the ability to maintain a sense of direction and location in virtual environments. The results of this study may appear in future publications and presentations building upon this research, though all results will be anonymous. This study involves up to 25 subjects over the age of 18 in good health with good eyesight.

II. Procedures

This study will ask you to perform a series of tasks within a room-scale virtual environment using HTC Vive virtual reality equipment. The first task will put you into a freeform environment you'll be able to explore without pressure or a specific goal. The purpose of this task is to familiarize you with the novel navigation techniques we have developed for testing. This section will like take 10-15 minutes, depending on how quickly you start to feel comfortable with the techniques.

After this, you will be placed into one of four possible virtual test scenarios. These scenarios are structured as a "treasure hunt" through the environment. You will be asked use the assigned navigation technique to move around and collect clues towards the final goal. At various points during your navigation, you will be asked questions regarding the locations of previously visited landmarks. After the final completion of the task, you will be asked to leave the virtual reality and take a short test on the layout of the environment, the paths you've taken, and the locations of landmarks and artifacts. After this first test, you will be asked to complete another treasure hunt and test through one of the other four scenarios. It is expected that each task will take 10-15 minutes to complete, with the test taking 5 minutes, for a total 30-40 minutes for this section.

This study only requires one approximately hour-long session from each participant. Throughout the study, you will be free to leave or end at any time. An investigator will be on-hand to answer any questions or help with anything unclear.

III. Risks

Working in virtual reality presents a set of risks. Those prone to motion sickness - or even those without a history - may experience discomfort arising from the separation between visual information sent to the brain and the experienced motion. Virtual reality sickness presents symptoms such as headaches, nausea, vomiting, fatigue, disorientation, and general discomfort. If at any time while using the virtual reality equipment you experience these symptoms and wish to stop, please inform the investigator and take off the gear. You may take a break or discontinue your participation as you see fit. The research team has calibrated the virtual reality techniques used to minimize discomfort or other issues, but they may still be present as virtual reality sickness is difficult to predict.

The other primary risk involved in virtual reality is the possibility of collision with physical objects you’re unable to see while in the virtual environment. The physical space will be clear of obstacles for the study, and the virtual reality headset provides visual cues for the boundaries of this area. The virtual reality gear is connected via physical wires which present a potential tripping hazard. The investigator will monitor the physical space closely and move all obstacles that may come in your way, and also inform you of potential hazards.

IV. Benefits

This study will provide you no direct benefits, but will assist in developing a shared academic understanding of human-computer interaction techniques and their strengths and weaknesses. Virtual reality is a new, exciting technology and understanding these new ways of interacting with computers will benefit society in being able to best use the technology in the future.

No promise or guarantee of benefits has been made to encourage you to participate.

V. Extent of Anonymity and Confidentiality

All data collected during this study will be done so anonymously. No names, contact information, or any other identifying information will be attached to your responses to an investigator’s questions or to your results from the provided tasks. At no time will the researchers release identifiable results of the study to anyone other than individuals working on the project without your written consent.

The Virginia Tech (VT) Institutional Review Board (IRB) may view the study’s data for auditing purposes. The IRB is responsible for the oversight of the protection of human subjects involved in research.

VI. Compensation

Participating in this study will provide you with no compensation, monetarily or otherwise. All subjects in the study do so of their own free will and a desire to further computing technologies.

VII. Freedom to Withdraw

It is important for you to know that you are free to withdraw from this study at any time without penalty. You are free not to answer any questions that you choose or respond to what is being asked of you without penalty.

Please note that there may be circumstances under which the investigator may determine that a subject should not continue as a subject.

Should you withdraw or otherwise discontinue participation, you will be compensated for the portion of the project completed in accordance with the Compensation section of this document.

VIII. Questions or Concerns

Should you have any questions about this study, you may contact one of the research investigators whose contact information is included at the beginning of this document.

Should you have any questions or concerns about the study's conduct or your rights as a research subject, or need to report a research-related injury or event, you may contact the VT IRB Chair, Dr. David M. Moore at moored@vt.edu or (xxx) xxx-xxxx.

IX. Subject's Consent

I have read the Consent Form and conditions of this project. I have had all my questions answered. I hereby acknowledge the above and give my voluntary consent:

____ Date__________

Subject signature

____

Subject printed name

A.1.2 Questionnaires

A.1.2.1 Background Questionnaire

Subject ID [to be filled by investigator]: ______________

Age: ______

Gender: ____________

Profession: ____________________

If student, what is your major? ______________________

Do you wear glasses or contacts? (circle one) No / Glasses / Contacts

What's your experience with virtual reality, if any?

Do you have a tendency for or history of motion sickness?

A.1.2.2 Post-Experiment Marking Test

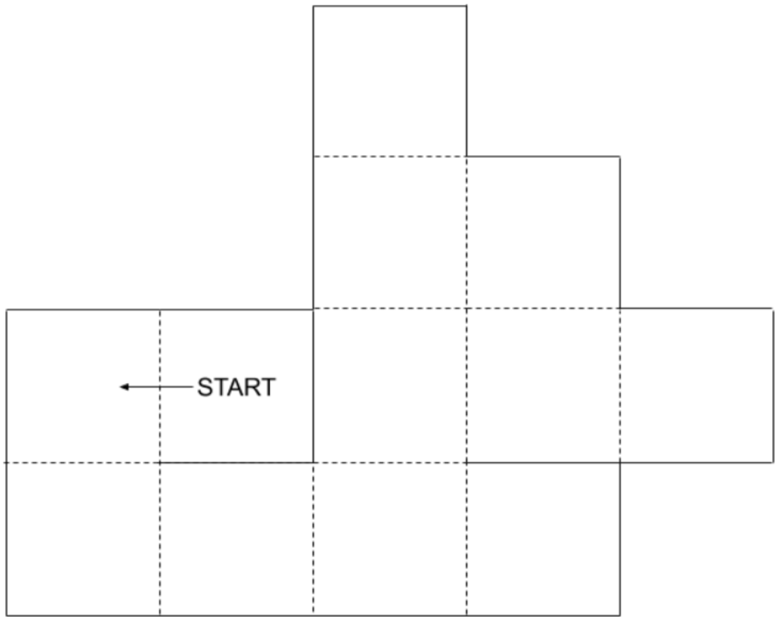
(1) Redirection Post-Session Evaluation

Subject ID: ____________

Technique: Bookshelf / Two-to-One Turn

Treasure: Default / Alternate 1 / Alternate 2

Order: First / Second



a. Mark the location of the Mona Lisa painting with a "P"
b. Mark the location of the stained glass window with a "W"
c. Label the room where the couches were located and facing inwards with a "C"
d. Label the room where you collected the final clue with an "F"
e. Following provided instructions, draw the path you took from the *final clue* you received to the *final goal*.

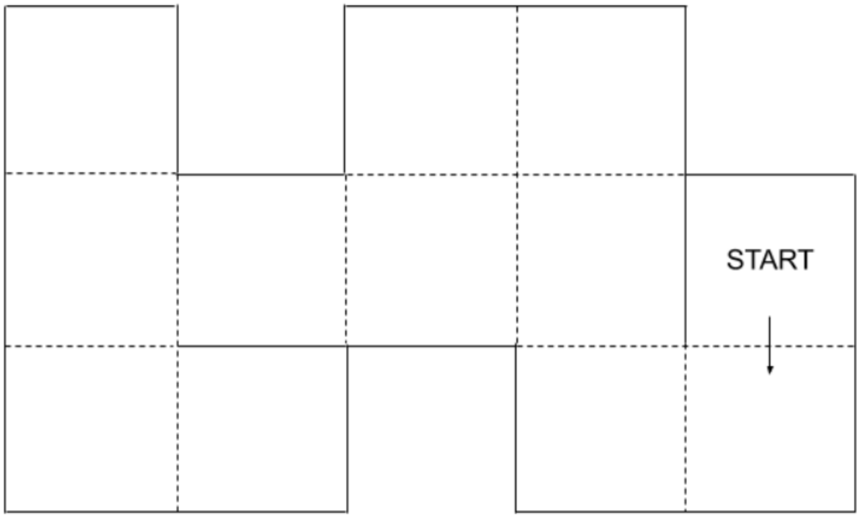
(2) Relocation Post-Session Evaluation

Subject ID: ____________

Technique: Bird / Teleportation

Treasure: Default / Alternate

Order: First / Second



a. Mark the location of the tree with a “T”

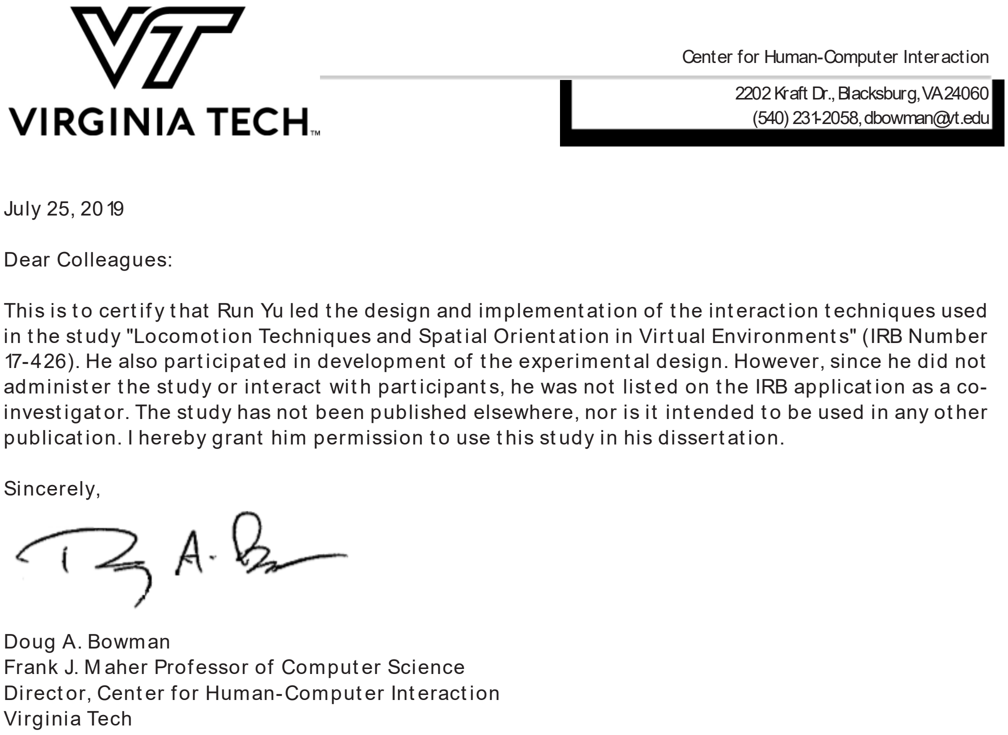
b. Mark the location of the fountain with an “F”

c. Fill in the dotted cell borders which contained the flowing water.

d. Label the cell where you collected the final clue with a “G”

e. Following provided instructions, draw the path you took from the *final clue* you received to the *final goal*.

A.1.3 Permission Letter



VIRGINIA TECH
Center for Human-Computer Interaction
2202 Kraft Dr., Blacksburg, VA 24060
(540) 231-2058, dbowman@vt.edu

July 25, 2019

Dear Colleagues:

This is to certify that Run Yu led the design and implementation of the interaction techniques used in the study "Locomotion Techniques and Spatial Orientation in Virtual Environments" (IRB Number 17-426). He also participated in development of the experimental design. However, since he did not administer the study or interact with participants, he was not listed on the IRB application as a co-investigator. The study has not been published elsewhere, nor is it intended to be used in any other publication. I hereby grant him permission to use this study in his dissertation.

Sincerely,

Doug A. Bowman
Frank J. Maher Professor of Computer Science
Director, Center for Human-Computer Interaction
Virginia Tech

A.2 Study of Usability of Force Push

A.2.1 User Consent Form

Participant #: ___________________________

Date (mm/dd/yyyy): ___________________

Informed Consent for Participant of Investigative Project

Virginia Polytechnic Institute and State University

Title of Project: **Object Manipulation in Virtual Reality**

Investigators: Dr. Doug A. Bowman, Run Yu.

I. THE PURPOSE OF THIS RESEARCH/PROJECT

Object manipulation is one of the fundamental tasks of interaction in Virtual Reality (VR). In this study, we want to compare several interaction designs for moving objects in VR and gain a better understanding of the user experience created by these techniques.

II. PROCEDURES

This study examines different interaction designs for object manipulation in Virtual Reality (VR). You will be asked to move an object to a target pose using several interaction models, all based on hand input.

1. Upon arrival, the participant will be asked to read and sign the consent form. This is expected to take around 5 minutes.

2. The participant will then be asked to complete a background questionaire. This should take no more than 5 minutes.

3. In this step, the participant is introduced to the equipment that s/he will use throughout the experiment, including a head-mounted display (HMD), a hand-held control device and a hand-tracking sensor. This is expected to take 5 minutes.

4. The participant will be asked to put on the HMD and get familiar with the interaction within the virtual environment (VE) through a training session. S/he will be asked to move a virtual object in the VE with four interaction models, all consisting of moving her/his hand and/or interacting with a hand-held device. The training session should last around 15 minutes.

5. After the training session, the participant will be presented with the actual tasks that are similar to what s/he did in the training session, namely moving a virtual object to a target pose through the same four models of hand input. This time the interaction will be timed and the participant is asked to accomplish each task as soon as possible. There will be four sessions totally, each one using one of the four interaction models. After each session, the participant is asked to take off the HMD, take a short rest and answer a survey expressing their perception of the individual interaction model s/he just used. This step is expected to last around 45 minutes.

6. After all interaction tasks are completed, the user is asked to answer a final survey expressing their perception about the interaction. This is expected to take 5 minutes.

Please not that, if you need to take a break at any time during the experiment, please complete the current task, and then inform the experimenter.

III. RISKS

There are no more than minimal risks involved in the study. You might experience mild discomfort such as strain on eyes, dizziness, nausea and arm/hand fatigue.

IV. BENEFITS OF THIS PROJECT

Object manipulation is one of the fundamental tasks in VR. The problem has implications for performance of real, production-grade immersive virtual environments used for educational and training purposes in military, academia and industry. You will, through this study, help us explore effective interaction techniques and provide useful insight into research of interaction design in 3D user interfaces.

V. EXTENT OF ANONYMITY AND CONFIDENTIALITY

The results of this study will be kept confidential. Your written consent is required for the researchers to release any data identified with you as an individual to anyone other than personnel working on the project. The information you provide will have your name removed and only a participant number will identify you during analyses and any written reports of the research.

VI. COMPENSATION

For your participation in any of the procedures of this study, you'll be provide with a compensation of $10.

VII. FREEDOM TO WITHDRAW

You are free to withdraw from this study at any time for any reason.

VIII. APPROVAL OF RESEARCH

This research has been approved, as required, by the Institutional Review Board for projects involving human subjects at Virginia Polytechnic Institute and State University, and by the Department of Computer Science.

Virginia Tech IRB Approval number: 17-XXX

IX. SUBJECT'S RESPONSIBILITIES AND PERMISSION

I voluntarily agree to participate in this study, and I know of no reason I cannot participate. I have read and understand the informed consent and conditions of this project. I have had all my questions answered. I hereby acknowledge the above and give my voluntary consent for participation in this project. If I participate, I may withdraw at any time without penalty. I agree to abide by the rules of this project

Signature Date

Name (please print)

Should I have any questions about this research or its conduct, I may contact:

Investigators: Dr. Doug A. Bowman Phone (xxx) xxx-xxxx
Professor, Computer Science Department (231-6931)
email: bowman@vt.edu

Run Yu
Graduate student, Computer Science Department
Email: runyu@vt.edu

Review Board: Dr. David Moore Phone (xxx) xxx-xxxx
Chair, Virginia Tech Institutional Review Board
For the Protection of Human Subjects
Email: moored@vt.edu

cc: the participant, Dr. Bowman, Run Yu

A.2.2 Questionnaires

A.2.2.1 Background Questionnaire

Please help us to categorize our user population by completing the following items.

Gender (circle one): Male Female

Age: _____________

Occupation (if student, indicate graduate or undergraduate):

____

Major / Area of specialization (if student): _________________________________

Rate how tired you are today: (circle one)

●-----------------------●-----------------------●-----------------------●

very tired somewhat tired a little tired not tired at all

Rate your expertise with computers: (circle one)

●-----------------------●-----------------------●-----------------------●

beginner amateur intermediate advanced

How often do you use computers...

...for work? (circle the best answer)	...for fun? (circle the best answer)
a. not at all	a. not at all
b. once a month	b. once a month

c. once a week
d. several times a week
e. daily

c. once a week
d. several times a week
e. daily

Rate your experience with video games: (circle one)

●----------------------●----------------------●----------------------●

Never Sometimes Often Everyday

Rate your experience with Virtual Reality (VR): (circle one)

●----------------------●----------------------●----------------------●

Never Sometimes Often Everyday

A.2.2.2 Post-Experiment Questionnaire for Evaluating Individual Techniques

Please complete the following questions.

How accurate do you feel about the technique you've just tried? Please rate the accuracy of this technique on a scale of 1 to 5.

With 1 being not accurate at all and 5 being very accurate

1 2 3 4 5

●----------------------●----------------------●----------------------●----------------------●

Not accurate at all Very accurate

Was this technique easy to use? Please rate the ease of use on a scale of 1 to 5.

With 1 being very hard to use and 5 being very easy to use

1 2 3 4 5

●----------------------●----------------------●----------------------●----------------------●

Very hard to use Very easy to use

How fast do you feel you can complete the task using this technique? Please rate the speed of the technique you've just tried on a scale of 1 to 5.

With 1 being very slow and 5 being very fast

1 2 3 4 5

●----------------------●----------------------●----------------------●----------------------●

Very slow Very fast

Did you feel that you were in control of the virtual object? Please rate the controllability of the technique you've just used on a scale of 1 to 5

With 1 being not in control at all and 5 being very much in control

1 2 3 4 5

●----------------------●----------------------●----------------------●----------------------●

Not in control at all Very much in control

How natural do you feel about the technique you've just tried? Please rate the naturalness of the technique you've just used on a scale of 1 to 5

With 1 being not natural at all and 5 being very natural

1 2 3 4 5

●----------------------●----------------------●----------------------●----------------------●

Not natural at all Very natural

Was the technique fun to use? Please rate it on a scale of 1 to 5

With 1 being not fun at all and 5 being a lot of fun

1 2 3 4 5

●----------------------●----------------------●----------------------●----------------------●

Not fun at all A lot of fun

Was the technique comfortable to use? Please rate the comfort of the technique you've just tried on a scale of 1 to 5

With 1 being very uncomfortable and 5 being very comfortable

1 2 3 4 5

●----------------------●----------------------●----------------------●----------------------●

Very uncomfortable Very comfortable

Any other comments?

A.2.2.3 Post-Experiment Questionnaire for User Preference

These questions ask for preference over the techniques you've just tried based on several criteria. Please pick the technique you prefer based on these criteria accuracy

First technique (Speed switch) ☐

Second technique (Gesture) ☐

ease of use

First technique (Speed switch) ☐

Second technique (Gesture) ☐

speed

First technique (Speed switch) ☐

Second technique (Gesture) ☐

controllability

First technique (Speed switch) ☐

Second technique (Gesture) ☐

naturalness

First technique (Speed switch) ☐

Second technique (Gesture) ☐

fun

First technique (Speed switch) ☐

Second technique (Gesture) ☐

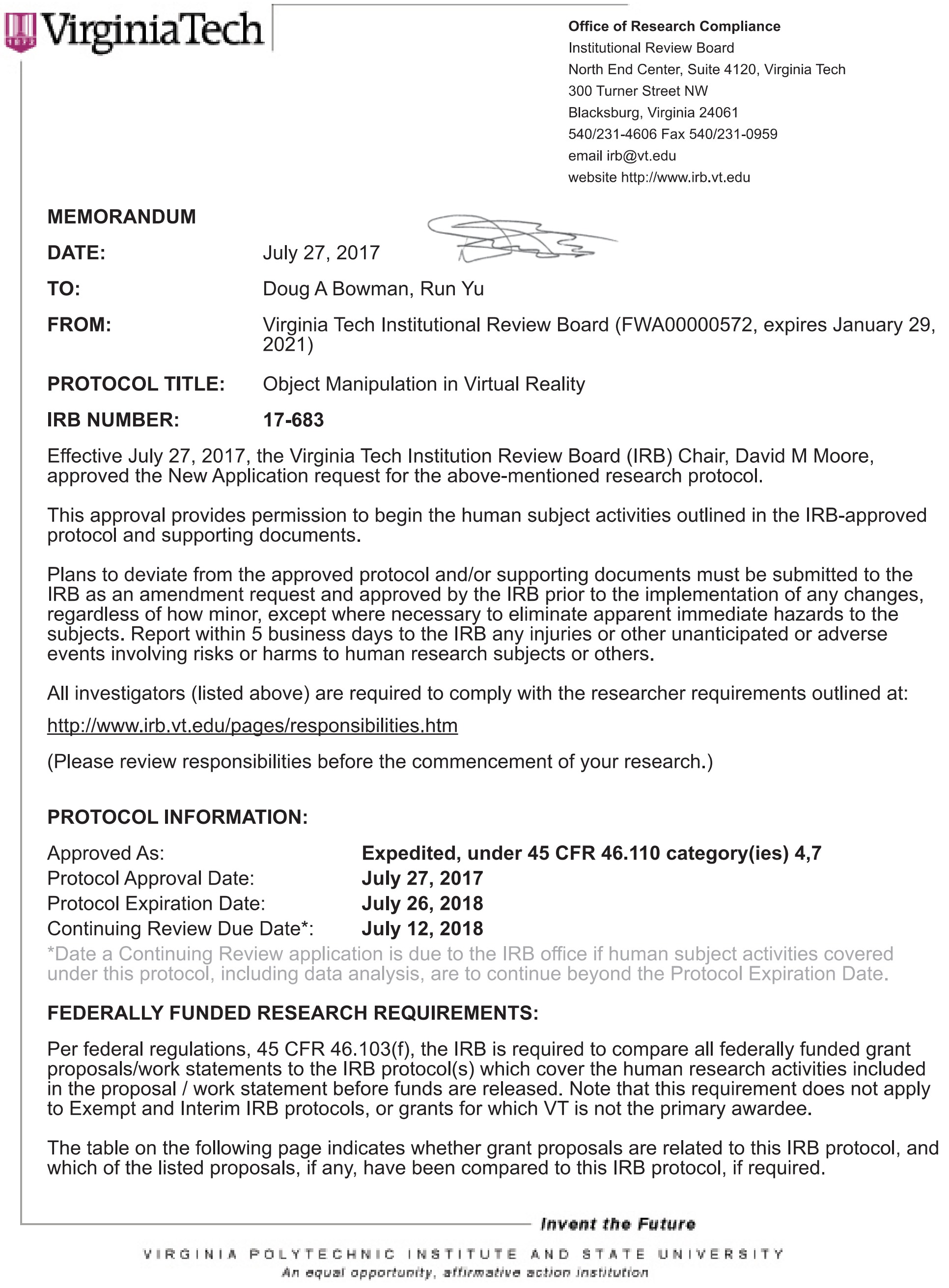
comfort

First technique (Speed switch) ☐

Second technique (Gesture) ☐

Any other comments?

A.2.3 IRB Approval Letter



VirginiaTech

Office of Research Compliance
Institutional Review Board
North End Center, Suite 4120, Virginia Tech
300 Turner Street NW
Blacksburg, Virginia 24061
540/231-4606 Fax 540/231-0959
email irb@vt.edu
website http://www.irb.vt.edu

MEMORANDUM

DATE: July 27, 2017

TO: Doug A Bowman, Run Yu

FROM: Virginia Tech Institutional Review Board (FWA00000572, expires January 29, 2021)

PROTOCOL TITLE: Object Manipulation in Virtual Reality

IRB NUMBER: **17-683**

Effective July 27, 2017, the Virginia Tech Institution Review Board (IRB) Chair, David M Moore, approved the New Application request for the above-mentioned research protocol.

This approval provides permission to begin the human subject activities outlined in the IRB-approved protocol and supporting documents.

Plans to deviate from the approved protocol and/or supporting documents must be submitted to the IRB as an amendment request and approved by the IRB prior to the implementation of any changes, regardless of how minor, except where necessary to eliminate apparent immediate hazards to the subjects. Report within 5 business days to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

All investigators (listed above) are required to comply with the researcher requirements outlined at:

http://www.irb.vt.edu/pages/responsibilities.htm

(Please review responsibilities before the commencement of your research.)

PROTOCOL INFORMATION:

Approved As: **Expedited, under 45 CFR 46.110 category(ies) 4,7**
Protocol Approval Date: **July 27, 2017**
Protocol Expiration Date: **July 26, 2018**
Continuing Review Due Date*: **July 12, 2018**

*Date a Continuing Review application is due to the IRB office if human subject activities covered under this protocol, including data analysis, are to continue beyond the Protocol Expiration Date.

FEDERALLY FUNDED RESEARCH REQUIREMENTS:

Per federal regulations, 45 CFR 46.103(f), the IRB is required to compare all federally funded grant proposals/work statements to the IRB protocol(s) which cover the human research activities included in the proposal / work statement before funds are released. Note that this requirement does not apply to Exempt and Interim IRB protocols, or grants for which VT is not the primary awardee.

The table on the following page indicates whether grant proposals are related to this IRB protocol, and which of the listed proposals, if any, have been compared to this IRB protocol, if required.

Invent the Future

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
An equal opportunity, affirmative action institution

Date*	OSP Number	Sponsor	Grant Comparison Conducted?

* Date this proposal number was compared, assessed as not requiring comparison, or comparison information was revised.

If this IRB protocol is to cover any other grant proposals, please contact the IRB office (irbadmin@vt.edu) immediately.

A.3 Study of Broad User Experience of Physically Coherent Techniques

A.3.1 User Consent Form

Title: Object Manipulation in Virtual Reality

Protocol No.: N/A

Sponsor: N/A

Investigators:

Principle investigator: Doug A. Bowman

Co-investigator: Run Yu

Room 160, Sandbox, Moss Arts Center, 190 Alumni Mall

Blacksburg, VA, 24060

USA

Daytime Phone Contact: Run Yu, xxx-xxx-xxxx

24-hour Phone Contact: N/A

You are being invited to take part in a research study. A person who takes part in a research study is called a research subject, or research participant.

What should I know about this research?

Someone will explain this research to you.

This form sums up that explanation.

Taking part in this research is voluntary. Whether you take part is up to you.

You can choose not to take part. There will be no penalty or loss of benefits to which you are otherwise entitled.

You can agree to take part and later change your mind. There will be no penalty or loss of benefits to which you are otherwise entitled.

If you don't understand, ask questions.

Ask all the questions you want before you decide.

Why is this research being done?

The purpose of this research is to understand the user experience of different interaction designs for object manipulation in virtual reality (VR).

About 20 subjects will take part in this research.

How long will I be in this research?

We expect that your taking part in this research will last 80 minutes.

What happens to me if I agree to take part in this research?

You are invited to participant in this experiment in Moss Arts Center at Virginia Tech. In this study, you will move objects in a 3D virtual environment using several interaction techniques all based on hand input. The entire experiment will be audio-recorded and all experimental data will be kept confidential.

Upon arrival, the participant will be asked to read and sign the consent form and complete a background questionnaire. This should take around 5 minutes.

The participant will then enter an introduction session, in which s/he will be introduced to the equipment, different interaction techniques and tasks. All the techniques are meant to move objects using hand input in a 3D virtual environment, while they differ in the design of input-to-output mappings and whether a hand-held controller is used. This is expected to last for 15 minutes.

The participant will then complete a set of interaction tasks with these techniques and evaluate their user experience. The task includes two stages: the user starts with free exploration, in which s/he is free to move objects around with no specific goal, as the purpose of this stage is to let the user feel these techniques through open exploration; the second stage is goal-driven, and the goal is to move objects to their specific target locations. "Think aloud" protocol is applied during the interaction, by which the user is encouraged to freely express their thoughts on the techniques while performing the tasks. Upon completion of the tasks, the user will be asked to answer three questionnaires to evaluate the user experience, including the “user experience questionnaire”, “game experience questionnaire” and “presence in VR questionnaire”. A post-experiment interview will be conducted to gather their final thoughts about the techniques. This is expected to last for 60 minutes.

What are my responsibilities if I take part in this research?

If you take part in this research, you will be responsible to:

Adjust the VR headset to your best comfort under the guidance of the experimenter

Tell the experimenter that you want to rest when you need to take a break.

Tell the experimenter that you want to terminate the experiment if you intend to do so.

Could being in this research hurt me?

There are no more than minimal risks involved in the study. You might experience:

Arm/hand fatigue

Mild discomfort such as strain on eyes, dizziness, nausea

Will it cost me money to take part in this research?

No.

Will being in this research benefit me?

We cannot promise any benefits to you or others from your taking part in this research. However, possible benefits to you include that you'll have an opportunity to experience the latest technology of virtual reality (e.g., 3D display, motion tracking and gesture recognition). Possible benefits to others include that knowledge gained from this study may help academia and industry build for a better experience with VR technology.

What other choices do I have besides taking part in this research?

This research is not designed to diagnose, treat or prevent any disease. Your alternative is to not take part in the research.

What happens to the information collected for this research?

All of the experimental data and results will be kept confidential. Participant data will not be linked to names in any manner. The only place identifying information will be collected is the signature on the consent form. The rest of the data will be tagged with a participant ID, and that we will not keep a key linking participant IDs to names.

The entire experiment will be audio-recorded and all experimental data will be kept confidential. The recording is only meant to gather more information about the user's subjective experience with the interaction techniques. The experimenter will transcribe the audio recording into text within two weeks after the experiment is finished and the audio files will be deleted right after this process is completed.

We may publish the results of this research. Data collected in this research might be used for future research or distributed to another investigator for future research without your consent.

Who can answer my questions about this research?

If you have questions, concerns, or complaints, or think this research has hurt you or made you sick, talk to the research team at the phone number listed above on the first page.

This research is being overseen by an Institutional Review Board ("IRB"). An IRB is a group of people who perform independent review of research studies. You may talk to them at (800) 562-4789, help@wirb.com if:

You have questions, concerns, or complaints that are not being answered by the research team.

You are not getting answers from the research team.

You cannot reach the research team.

You want to talk to someone else about the research.

You have questions about your rights as a research subject.

What if I am injured because of taking part in this research?

There are only minimal risks involved in this study. If you feel sick during the experiment, please inform the experimenter immediately so we can terminate the study session.

Can I be removed from this research without my approval?

The person in charge of this research can remove you from this research without your approval. Possible reasons for removal include:

It is in your best interest

You are under-aged (the experiment requires that you are over 18)

Your physical condition doesn't meet the requirements of this study (this study requires normal vision and normal range of arm/hand motion)

You cannot wear the devices in a proper way that's suitable for conducting this experiment (e.g., your glasses cannot fit in the VR headset while you cannot see the image clearly without glasses)

You have a side effect that requires stopping the research

You or the investigators are unable to keep the scheduled appointments

We will tell you about any new information that may affect your health, welfare, or choice to stay in this research.

What happens if I agree to be in this research, but I change my mind later?

We will terminate the experiment session.

Will I be paid for taking part in this research?

You may be paid up to a total of $10 for taking part in this research, depending on the availability of according funds at the time of this experiment. If you are paid, your compensation will be broken down as follows:

It is a one-time payment

It is paid at the end of your visit

You'll get the same amount of payment if you drop out during the experiment

Statement of Consent:

Your signature documents your consent to take part in this research.

______________________________ ____________

Signature of adult subject capable of consent Date

______________________________ ____________

Signature of person obtaining consent Date

A.3.2 Questionnaires

A.3.2.1 Background Questionnaire

Please help us to categorize our user population by completing the following items.

Gender (circle one): Male Female

Age: _____________

Occupation (if student, indicate graduate or undergraduate):

____

Major / Area of specialization (if student): _________________________________

Rate how tired you are today: (circle one)

●-----------------------●-----------------------●-----------------------●-----------------------●

very tired tired somewhat tired a little tired not tired at all

Rate your expertise with computers: (circle one)

●-----------------------●-----------------------●-----------------------●-----------------------●

beginner amateur intermediate advanced expert

How often do you use computers...

...for work? (circle the best answer)	...for fun? (circle the best answer)
a. not at all	a. not at all
b. once a month	b. once a month
c. once a week	c. once a week
d. several times a week	d. several times a week
e. daily	e. daily

Rate your experience with video games: (circle one)

●-----------------------●-----------------------●-----------------------●-----------------------●

Never rarely occasionally frequently very frequently

Rate your experience with Virtual Reality (VR): (circle one)

●-----------------------●-----------------------●-----------------------●-----------------------●

Never rarely occasionally frequently very frequently

A.3.2.2 Presence Questionnaire

(Slater et al., 1998; Usoh et al., 1999; Usoh et al., 2000)

(1) Please rate your sense of being in the room space with furniture, on the following scale from 1 to 7, where 7 represents your normal experience of being in a place.

●----------------●----------------●----------------●----------------●----------------●----------------●

1 2 3 4 5 6 7

not at all… very much…

(2) To what extent were there times during the experience when the virtual room with furniture became the "reality" for you, and you almost forgot about the "real world" of the laboratory in which the whole experience was really taking place?

●----------------●----------------●----------------●----------------●----------------●----------------●

1 2 3 4 5 6 7

at no time… almost all the time

(3) When you think back about your experience, do you think of the virtual room with furniture more as images that you saw, or more as somewhere that you visited ? Please answer on the following 1 to 7 scale.

●----------------●----------------●----------------●----------------●----------------●----------------●

1 2 3 4 5 6 7

Images that I saw Somewhere I visited

(4) During the time of the experience, which was strongest on the whole, your sense of being in the virtual room, or of being in the real world of the laboratory?

●----------------●----------------●----------------●----------------●----------------●----------------●

1 2 3 4 5 6 7

the real world lab the virtual room

(5) Consider your memory of being in the virtual room. How similar in terms of the structure of the memory is this to the structure of the memory of other places you have been today? By ‘structure of the memory’ consider things like the extent to which you have a visual memory of the room, whether that memory is in color, the extent to which the memory seems vivid or realistic, its size, location in your imagination, the extent to which it is panoramic in your imagination, and other such structural elements.

●----------------●----------------●----------------●----------------●----------------●----------------●

1 2 3 4 5 6 7

not at all… very much so

(6) During the time of the experience, did you often think to yourself that you were actually just standing in a lab wearing a helmet or did the virtual room with furniture overwhelm you?

●----------------●----------------●----------------●----------------●----------------●----------------●

1 2 3 4 5 6 7

A.3.2.3 User Experience Questionnaire

(Laugwitz et al., 2008)

Please make your evaluation now.

For the assessment of the product, please fill out the following questionnaire. The questionnaire consists of pairs of contrasting attributes that may apply to the product. The circles between the attributes represent gradations between the opposites. You can express your agreement with the attributes by ticking the circle that most closely reflects your impression.

attractive	○	⊗	○	○	○	○	○	unattractive

This response would mean that you rate the application as more attractive than unattractive.

Please decide spontaneously. Don't think too long about your decision to make sure that you convey your original impression.

Sometimes you may not be completely sure about your agreement with a particular attribute or you may find that the attribute does not apply completely to the particular product. Nevertheless, please tick a circle in every line.

It is your personal opinion that counts. Please remember: there is no wrong or right answer!

Please assess the product now by ticking one circle per line.

	1	2	3	4	5	6	7	
annoying	○	○	○	○	○	○	○	enjoyable
not understandable	○	○	○	○	○	○	○	understandable
creative	○	○	○	○	○	○	○	dull
easy to learn	○	○	○	○	○	○	○	difficult to learn
valuable	○	○	○	○	○	○	○	inferior
boring	○	○	○	○	○	○	○	exciting
not interesting	○	○	○	○	○	○	○	interesting
unpredictable	○	○	○	○	○	○	○	predictable
fast	○	○	○	○	○	○	○	slow

inventive	○	○	○	○	○	○	○	conventional
obstructive	○	○	○	○	○	○	○	supportive
good	○	○	○	○	○	○	○	bad
complicated	○	○	○	○	○	○	○	easy
unlikable	○	○	○	○	○	○	○	pleasing
usual	○	○	○	○	○	○	○	leading edge
unpleasant	○	○	○	○	○	○	○	pleasant
secure	○	○	○	○	○	○	○	not secure
motivating	○	○	○	○	○	○	○	demotivating
meets expectations	○	○	○	○	○	○	○	does not meet expectations
inefficient	○	○	○	○	○	○	○	efficient
clear	○	○	○	○	○	○	○	confusing
impractical	○	○	○	○	○	○	○	practical
organized	○	○	○	○	○	○	○	cluttered
attractive	○	○	○	○	○	○	○	unattractive
friendly	○	○	○	○	○	○	○	unfriendly
conservative	○	○	○	○	○	○	○	innovative

A.3.2.4 Gaming Experience Questionnaire

(IJsselsteijn et al., 2007; Norman, 2013)

Please indicate how you felt while playing the game for each of the items, on the following scale:

not at all	slightly	moderately	fairly	extremely
0	1	2	3	4

1 I felt content

2 I felt skilful

3 I was interested in the game's story

4 I thought it was fun

5 I was fully occupied with the game

6 I felt happy

7 It gave me a bad mood

8 I thought about other things

9 I found it tiresome

10 I felt competent

11 I thought it was hard

12 It was aesthetically pleasing

13 I forgot everything around me

14 I felt good

15 I was good at it

16 I felt bored

17 I felt successful

18 I felt imaginative

19 I felt that I could explore things

20 I enjoyed it

21 I was fast at reaching the game's targets

22 I felt annoyed

23 I felt pressured

24 I felt irritable

25 I lost track of time

26 I felt challenged

27 I found it impressive

28 I was deeply concentrated in the game

29 I felt frustrated

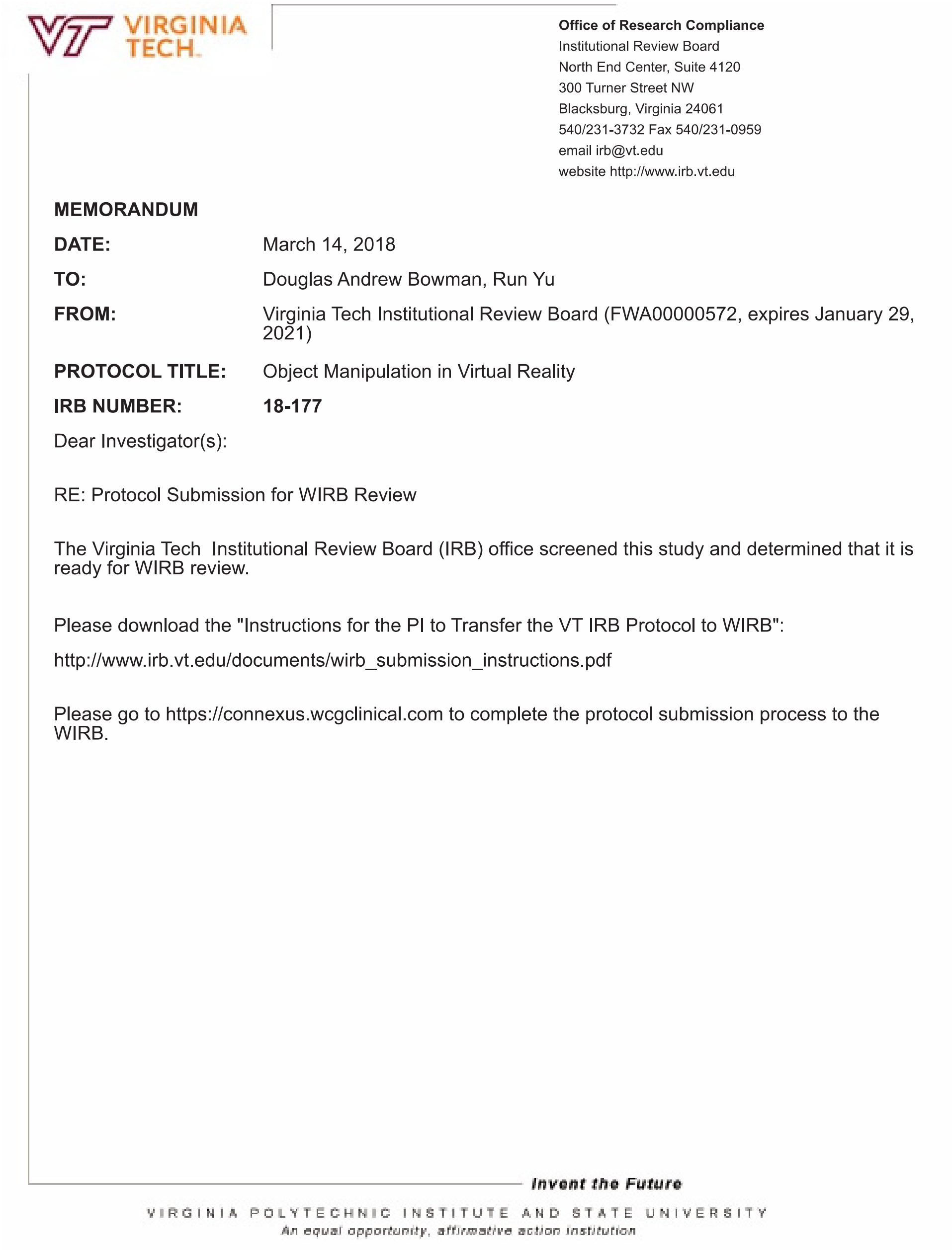
30 It felt like a rich experience

31 I lost connection with the outside world

32 I felt time pressure

33 I had to put a lot of effort into it

A.3.3 IRB Approval Letter



VIRGINIA TECH

Office of Research Compliance
Institutional Review Board
North End Center, Suite 4120
300 Turner Street NW
Blacksburg, Virginia 24061
540/231-3732 Fax 540/231-0959
email irb@vt.edu
website http://www.irb.vt.edu

MEMORANDUM

DATE: March 14, 2018

TO: Douglas Andrew Bowman, Run Yu

FROM: Virginia Tech Institutional Review Board (FWA00000572, expires January 29, 2021)

PROTOCOL TITLE: Object Manipulation in Virtual Reality

IRB NUMBER: **18-177**

Dear Investigator(s):

RE: Protocol Submission for WIRB Review

The Virginia Tech Institutional Review Board (IRB) office screened this study and determined that it is ready for WIRB review.

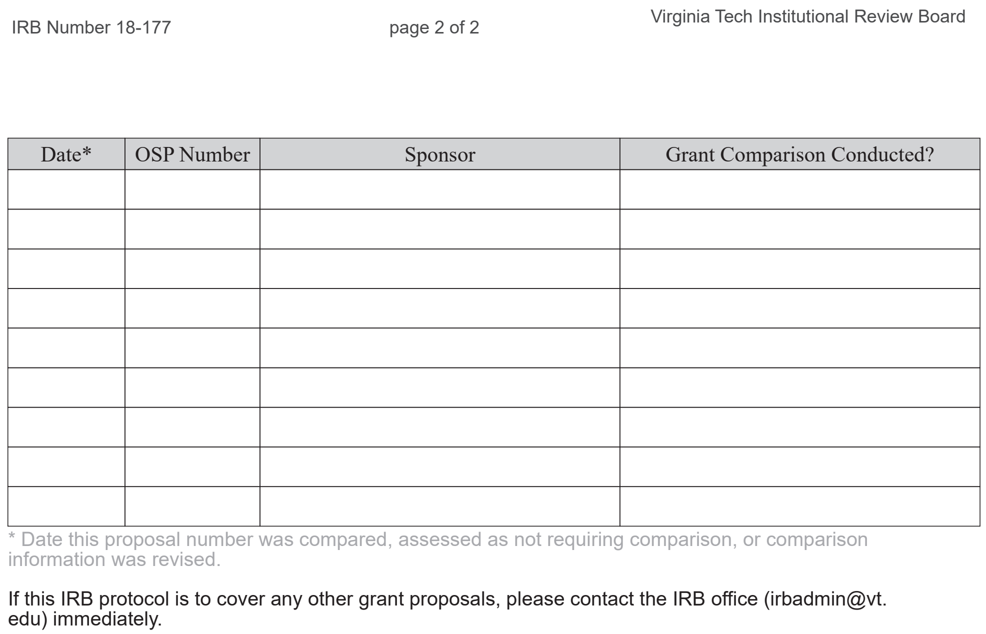
Please download the "Instructions for the PI to Transfer the VT IRB Protocol to WIRB":

http://www.irb.vt.edu/documents/wirb_submission_instructions.pdf

Please go to https://connexus.wcgclinical.com to complete the protocol submission process to the WIRB.

Invent the Future

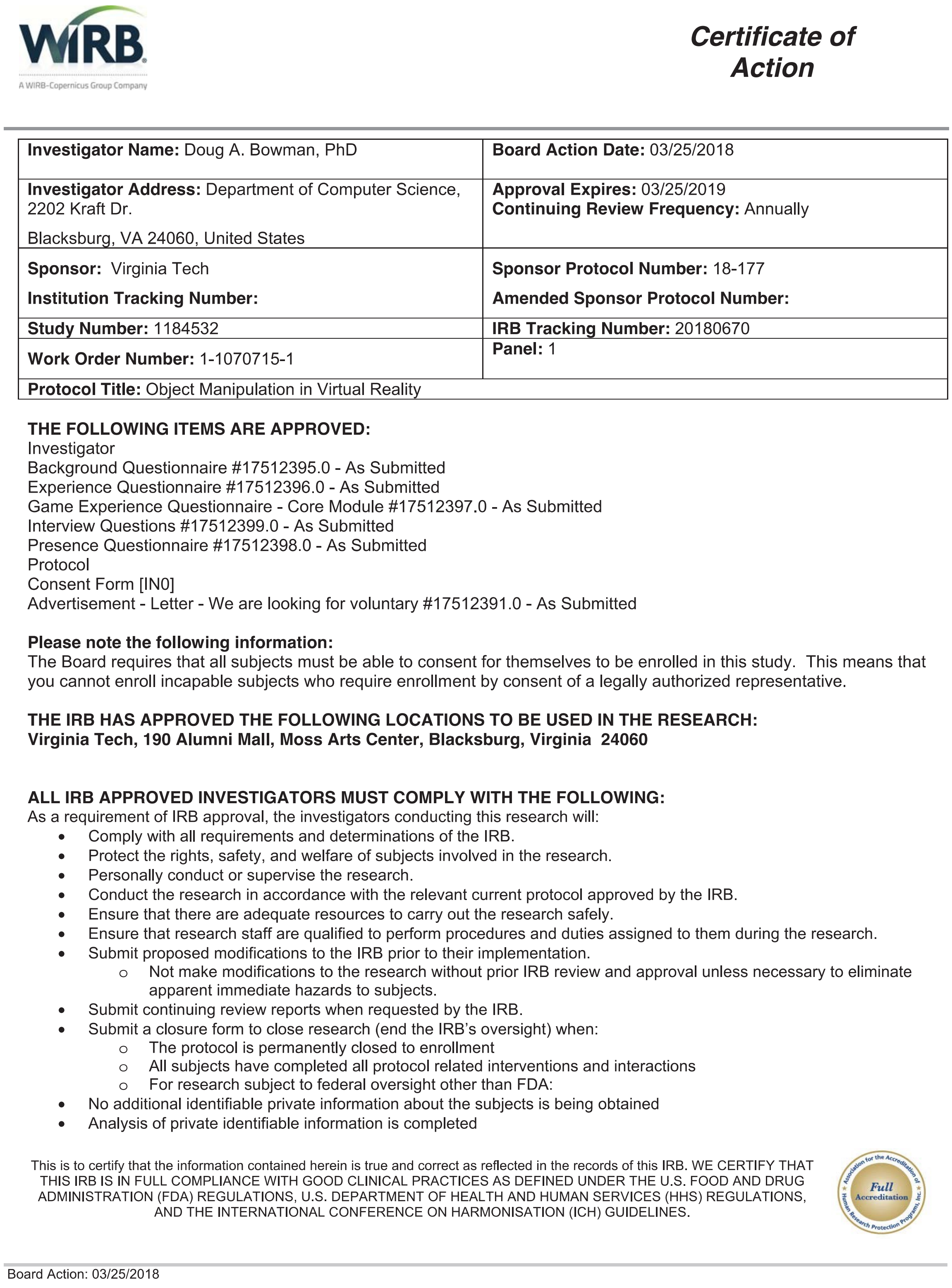
VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
An equal opportunity, affirmative action institution



Date*	OSP Number	Sponsor	Grant Comparison Conducted?

* Date this proposal number was compared, assessed as not requiring comparison, or comparison information was revised.

If this IRB protocol is to cover any other grant proposals, please contact the IRB office (irbadmin@vt.edu) immediately.



WIRB

A WIRB-Copernicus Group Company

Certificate of Action

Investigator Name: Doug A. Bowman, PhD	**Board Action Date:** 03/25/2018
Investigator Address: Department of Computer Science, 2202 Kraft Dr. Blacksburg, VA 24060, United States	**Approval Expires:** 03/25/2019 **Continuing Review Frequency:** Annually
Sponsor: Virginia Tech **Institution Tracking Number:**	**Sponsor Protocol Number:** 18-177 **Amended Sponsor Protocol Number:**
Study Number: 1184532	**IRB Tracking Number:** 20180670
Work Order Number: 1-1070715-1	**Panel:** 1
Protocol Title: Object Manipulation in Virtual Reality	

THE FOLLOWING ITEMS ARE APPROVED:
Investigator
Background Questionnaire #17512395.0 - As Submitted
Experience Questionnaire #17512396.0 - As Submitted
Game Experience Questionnaire - Core Module #17512397.0 - As Submitted
Interview Questions #17512399.0 - As Submitted
Presence Questionnaire #17512398.0 - As Submitted
Protocol
Consent Form [IN0]
Advertisement - Letter - We are looking for voluntary #17512391.0 - As Submitted

Please note the following information:
The Board requires that all subjects must be able to consent for themselves to be enrolled in this study. This means that you cannot enroll incapable subjects who require enrollment by consent of a legally authorized representative.

THE IRB HAS APPROVED THE FOLLOWING LOCATIONS TO BE USED IN THE RESEARCH:
Virginia Tech, 190 Alumni Mall, Moss Arts Center, Blacksburg, Virginia 24060

ALL IRB APPROVED INVESTIGATORS MUST COMPLY WITH THE FOLLOWING:
As a requirement of IRB approval, the investigators conducting this research will:

- Comply with all requirements and determinations of the IRB.
- Protect the rights, safety, and welfare of subjects involved in the research.
- Personally conduct or supervise the research.
- Conduct the research in accordance with the relevant current protocol approved by the IRB.
- Ensure that there are adequate resources to carry out the research safely.
- Ensure that research staff are qualified to perform procedures and duties assigned to them during the research.
- Submit proposed modifications to the IRB prior to their implementation.
    - Not make modifications to the research without prior IRB review and approval unless necessary to eliminate apparent immediate hazards to subjects.
- Submit continuing review reports when requested by the IRB.
- Submit a closure form to close research (end the IRB's oversight) when:
    - The protocol is permanently closed to enrollment
    - All subjects have completed all protocol related interventions and interactions
    - For research subject to federal oversight other than FDA:
- No additional identifiable private information about the subjects is being obtained
- Analysis of private identifiable information is completed

This is to certify that the information contained herein is true and correct as reflected in the records of this IRB. WE CERTIFY THAT THIS IRB IS IN FULL COMPLIANCE WITH GOOD CLINICAL PRACTICES AS DEFINED UNDER THE U.S. FOOD AND DRUG ADMINISTRATION (FDA) REGULATIONS, U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES (HHS) REGULATIONS, AND THE INTERNATIONAL CONFERENCE ON HARMONISATION (ICH) GUIDELINES.

Association for the Accreditation of Human Research Protection Programs, Inc.
Full Accreditation

- If research approval expires, stop all research activities and immediately contact the IRB.
- Promptly report to the IRB the information items listed in the IRB's "Prompt Reporting Requirements" available on the IRB's Web site.
- Not accept or provide payments to professionals in exchange for referrals of potential subjects ("finder's fees.")
- Not accept payments designed to accelerate recruitment that are tied to the rate or timing of enrollment ("bonus payments") without prior IRB approval.
- When required by the IRB ensure that consent, permission, and assent are obtained and documented in accordance with the relevant current protocol as approved by the IRB.
- Promptly notify the IRB of any change to information provided on your initial submission form.

Consistent with AAHRPP's requirements in connection with its accreditation of IRBs, the individual and/or organization shall promptly communicate or provide, the following information relevant to the protection of human subjects to the IRB in a timely manner:

- Upon request of the IRB, a copy of the written plan between sponsor or CRO and site that addresses whether expenses for medical care incurred by human subject research subjects who experience research related injury will be reimbursed, and if so, who is responsible in order to determine consistency with the language in the consent document.
- Any site monitoring report that directly and materially affects subject safety or their willingness to continue participation. Such reports will be provided to the IRB within 5 days.
- Reports from any data monitoring committee, data and safety monitoring board, or data and safety monitoring committee in accordance with the time frame specified in the research protocol.
- Any findings from a closed research when those findings materially affect the safety and medical care of past subjects. Findings will be reported for 2 years after the closure of the research.

If your research site is a HIPAA covered entity, the HIPAA Privacy Rule requires you to obtain written authorization from each research subject for any use or disclosure of protected health information for research. If your IRB-approved consent form does not include such HIPAA authorization language, the HIPAA Privacy Rule requires you to have each research subject sign a separate authorization agreement. "

Federal regulations require that the IRB conduct continuing review of approved research. You will receive Continuing Review Report forms from this IRB when the expiration date is approaching.

Thank you for using this WCG IRB to provide oversight for your research project.

DISTRIBUTION OF COPIES:

Contact, Company

Jennifer Farmer, Virginia Polytechnic Institute and State University (Virginia Tech)
VA Tech WIRB, Virginia Polytechnic Institute and State University (Virginia Tech)
Doug A. Bowman, PhD, Virginia Tech
Run Yu, Virginia Tech

A.4 Study of Pseudo-Haptic Display of Mass and Mass Distribution During Object Rotation

A.4.1 User Consent Form

Title: Object Manipulation and Weight Perception in Virtual Reality

Protocol No.: N/A

Sponsor: N/A

Investigators:

Principle investigator: Doug A. Bowman

Co-investigator: Run Yu

Dept. of Computer Science, Virginia Tech

2202 Kraft Dr.

Blacksburg, VA, 24060

USA

Daytime Phone Number: Run Yu, xxx-xxx-xxxx

24-hour Phone Number: N/A

RESEARCH CONSENT SUMMARY

You are being asked for your consent to take part in a research study. This document provides a concise summary of this research. It describes the key information that we believe most people need to decide whether to take part in this research. Later sections of this document will provide all relevant details.

What should I know about this research?

Someone will explain this research to you.

Taking part in this research is voluntary. Whether you take part is up to you.

If you don’t take part, it won’t be held against you.

You can take part now and later drop out, and it won’t be held against you

If you don’t understand, ask questions.

Ask all the questions you want before you decide.

How long will I be in this research?

We expect that your taking part in this research will last 70 minutes.

Why is this research being done?

The purpose of this research is to understand the perception of weight in Virtual Reality.

What happens to me if I agree to take part in this research?

If you decide to take part in this research study, the general procedures include putting on the VR headset to experience a 3D virtual environment, using hand-held controllers to move some objects in the environment and answer some questions regarding your perception of the objects' weight.

Could being in this research hurt me?

The most important risks or discomforts that you may expect from taking part in this research include arm/hand fatigue and potentially mild discomfort that's common for VR experiences, such as strain on eyes, dizziness, nausea.

Will being in this research benefit me?

The most important benefits that you may expect from taking part in this research is that you'll have an opportunity to experience the latest technology of virtual reality. It is not expected that you will personally benefit from this research.

Possible benefits to others include that knowledge gained from this study may help academia and industry build for a better experience with VR technology.

DETAILED RESEARCH CONSENT

You are being invited to take part in a research study. A person who takes part in a research study is called a research subject, or research participant.

What should I know about this research?

Someone will explain this research to you.

This form sums up that explanation.

Taking part in this research is voluntary. Whether you take part is up to you.

You can choose not to take part. There will be no penalty or loss of benefits to which you are otherwise entitled.

You can agree to take part and later change your mind. There will be no penalty or loss of benefits to which you are otherwise entitled.

If you don't understand, ask questions.

Ask all the questions you want before you decide.

Why is this research being done?

The purpose of this research is to understand the perception of weight in Virtual Reality.

About 20 subjects will take part in this research.

How long will I be in this research?

We expect that your taking part in this research will last 70 minutes.

What happens to me if I agree to take part in this research?

You are invited to participant in this experiment in Moss Arts Center at Virginia Tech. In this study, you will move objects in a 3D virtual environment and answer questions regarding your perception of the objects' weight and weight distribution. All experimental data will be kept confidential.

1. Upon arrival, you will be asked to read and sign the consent form and complete a background questionnaire. This should take around 5 minutes.

2. The participant will then enter an introduction session, in which you will be introduced to the equipment (VR headset and hand-held controllers), the 3D virtual environment and the task of identifying weight difference and weight distribution. You will learn how to control the virtual objects' movement and complete the task using the hand-held controllers. This is expected to last for 15 minutes.

3. Then you will will enter the actual experiment, which will repetitively ask her/him to move some virtual objects in 3D and identify weight difference and weight distribution you felt during the manipulation. A set of questionnaires will be presented to you during and after these trials and an interview will be conducted to gather more information about your perception of the interaction. This is expected to last for 50 minutes.

What are my responsibilities if I take part in this research?

If you take part in this research, you will be responsible to:

1. Adjust the VR headset to your best comfort under the guidance of the experimenter
2. Tell the experimenter that you want to rest when you need to take a break.
3. Tell the experimenter that you want to terminate the experiment if you intend to do so.

Could being in this research hurt me?

There are no more than minimal risks involved in the study. You might experience:

1. Arm/hand fatigue
2. Mild discomfort such as strain on eyes, dizziness, nausea

Will it cost me money to take part in this research?

No.

Will being in this research benefit me?

We cannot promise any benefits to you or others from your taking part in this research. However, possible benefits to you include that you'll have an opportunity to experience the latest technology of virtual reality (e.g., 3D display and motion tracking). Possible benefits to others include that knowledge gained from this study may help academia and industry build for a better experience with VR technology.

What other choices do I have besides taking part in this research?

This research is not designed to diagnose, treat or prevent any disease. Your alternative is to not take part in the research.

What happens to the information collected for this research?

All of the experimental data and results will be kept confidential. Participant data will not be linked to names in any manner. The only place identifying information will be collected is the signature on the consent form. The rest of the data will be tagged with a participant ID, and that we will not keep a key linking participant IDs to names.

We may publish the results of this research. Data collected in this research might be used for future research or distributed to another investigator for future research without your consent.

Who can answer my questions about this research?

If you have questions, concerns, or complaints, or think this research has hurt you or made you sick, talk to the research team at the phone number listed above on the first page.

This research is being overseen by an Institutional Review Board (“IRB”). An IRB is a group of people who perform independent review of research studies. You may talk to them at (800) 562-4789, help@wirb.com if:

You have questions, concerns, or complaints that are not being answered by the research team.

You are not getting answers from the research team.

You cannot reach the research team.

You want to talk to someone else about the research.

You have questions about your rights as a research subject.

What if I am injured because of taking part in this research?

There are only minimal risks involved in this study. If you feel sick during the experiment, please inform the experimenter immediately so we can terminate the study session.

Can I be removed from this research without my approval?

The person in charge of this research can remove you from this research without your approval. Possible reasons for removal include:

It is in your best interest

You are under-aged (the experiment requires that you are over 18)

Your physical condition doesn’t meet the requirements of this study (this study requires normal vision and normal range of arm/hand motion)

You cannot wear the devices in a proper way that’s suitable for conducting this experiment (e.g., your glasses cannot fit in the VR headset while you cannot see the image clearly without glasses)

You have a side effect that requires stopping the research

You or the investigators are unable to keep the scheduled appointments

We will tell you about any new information that may affect your health, welfare, or choice to stay in this research.

What happens if I agree to be in this research, but I change my mind later?

We will terminate the experiment session.

Will I be paid for taking part in this research?

You may be paid up to a total of $10 for taking part in this research, depending on the availability of according funds at the time of this experiment. If you are paid, your compensation will be broken down as follows:

It is a one-time payment

It is paid at the end of your visit

You'll get the same amount of payment if you drop out during the experiment

Statement of Consent:

Your signature documents your consent to take part in this research.

Signature of adult subject capable of consent	Date
Signature of person obtaining consent	Date

A.4.2 Questionnaires

A.4.2.1 Background Questionnaire

Please help us to categorize our user population by completing the following items.

Gender (circle one): Male Female

Age: _____________

Dominant hand (circle one): Left Right

Occupation (if student, indicate graduate or undergraduate):

____

Major / Area of specialization (if student): _________________________________

Rate how tired you are today: (circle one)

●----------------------●----------------------●----------------------●----------------------●

very tired	tired	somewhat tired	a little tired	not tired at all

Rate your expertise with computers: (circle one)

●-----------------------●-----------------------●-----------------------●-----------------------●

beginner amateur intermediate advanced expert

How often do you use computers...

...for work? (circle the best answer)	...for fun? (circle the best answer)
a. not at all	a. not at all
b. once a month	b. once a month
c. once a week	c. once a week
d. several times a week	d. several times a week
e. daily	e. daily

Rate your experience with video games: (circle one)

●-----------------------●-----------------------●-----------------------●-----------------------●

Never rarely occasionally frequently very frequently

Rate your experience with Virtual Reality (VR): (circle one)

●-----------------------●-----------------------●-----------------------●-----------------------●

Never rarely occasionally frequently very frequently

A.4.2.2 Questionnaire to Evaluate Pseudo-Haptic Illusion

(Rietzler et al., 2018a; Rietzler et al., 2018b)

(For Questions in Group A and B, we use Likert scale)

On a scale of 1 to 7, rate how much you agree with each of these statements:

●--------------●--------------●--------------●--------------●--------------●--------------●

1: Strongly disagree 2: disagree 3: more or less disagree 4: undecided 5: more or less agree 6: agree 7: strongly agree

Questions regarding overall weight

1. I could feel that the two objects had different weight
2. The representation of the weight difference felt realistic
3. I had the feeling of manipulating real objects
4. The representation of weight difference was sufficient for VR applications
5. I liked this representation of weight difference
6. I felt I could control the object with my actions with ease

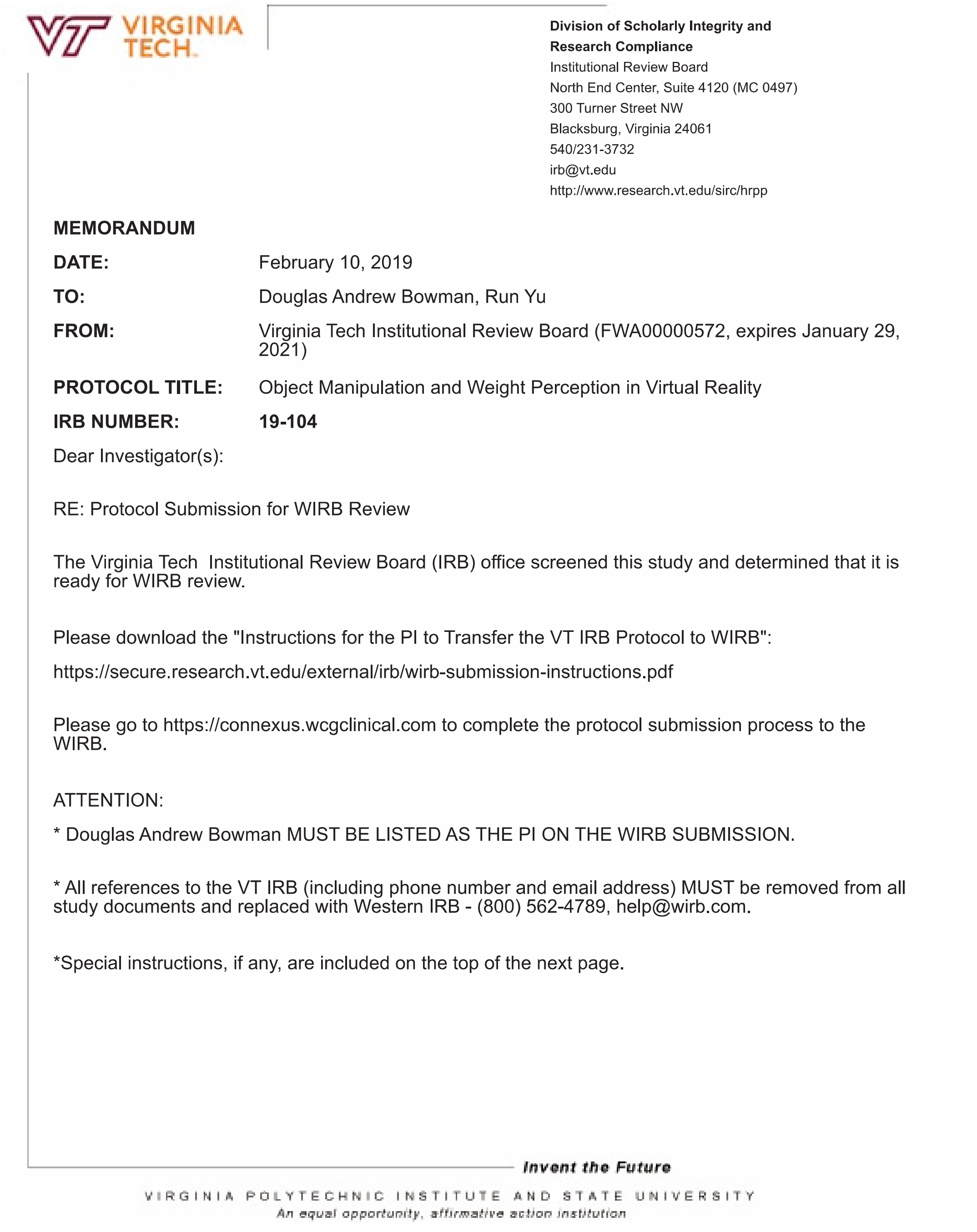
Questions regarding weight distribution

1. I could feel the weight distribution of the object’s body (e.g., the weight was not evenly distributed on the body and one side was heavier than the other).
2. The representation of weight distribution felt realistic
3. I had the feeling of manipulating real objects
4. The representation of weight distribution was sufficient for VR applications
5. I liked this representation of weight distribution
6. I felt I could control the object with my actions with ease

Questions comparing interaction techniques

1. Among the approaches you’ve just tried, which one gave you a stronger sense of weight distribution (or weight difference)?
2. Among the approaches you’ve just tried, which representation of weight distribution (or weight difference) felt more realistic?
3. Among the different approaches you’ve just tried, which gave you a sensation that’s closer to the feeling of manipulating real objects?
4. Among the different approaches you’ve just tried, which representation of weight distribution (or weight difference) felt more sufficient for VR applications?
5. Among the different approaches you’ve just tried, which do you prefer?
6. Among the different approaches you’ve just tried, which felt easier to control?

A.4.3 IRB Approval Letter



VIRGINIA TECH.

Division of Scholarly Integrity and
Research Compliance
Institutional Review Board
North End Center, Suite 4120 (MC 0497)
300 Turner Street NW
Blacksburg, Virginia 24061
540/231-3732
irb@vt.edu
http://www.research.vt.edu/sirc/hrpp

MEMORANDUM

DATE: February 10, 2019

TO: Douglas Andrew Bowman, Run Yu

FROM: Virginia Tech Institutional Review Board (FWA00000572, expires January 29, 2021)

PROTOCOL TITLE: Object Manipulation and Weight Perception in Virtual Reality

IRB NUMBER: **19-104**

Dear Investigator(s):

RE: Protocol Submission for WIRB Review

The Virginia Tech Institutional Review Board (IRB) office screened this study and determined that it is ready for WIRB review.

Please download the "Instructions for the PI to Transfer the VT IRB Protocol to WIRB":

https://secure.research.vt.edu/external/irb/wirb-submission-instructions.pdf

Please go to https://connexus.wcgclinical.com to complete the protocol submission process to the WIRB.

ATTENTION:

* Douglas Andrew Bowman MUST BE LISTED AS THE PI ON THE WIRB SUBMISSION.

* All references to the VT IRB (including phone number and email address) MUST be removed from all study documents and replaced with Western IRB - (800) 562-4789, help@wirb.com.

*Special instructions, if any, are included on the top of the next page.

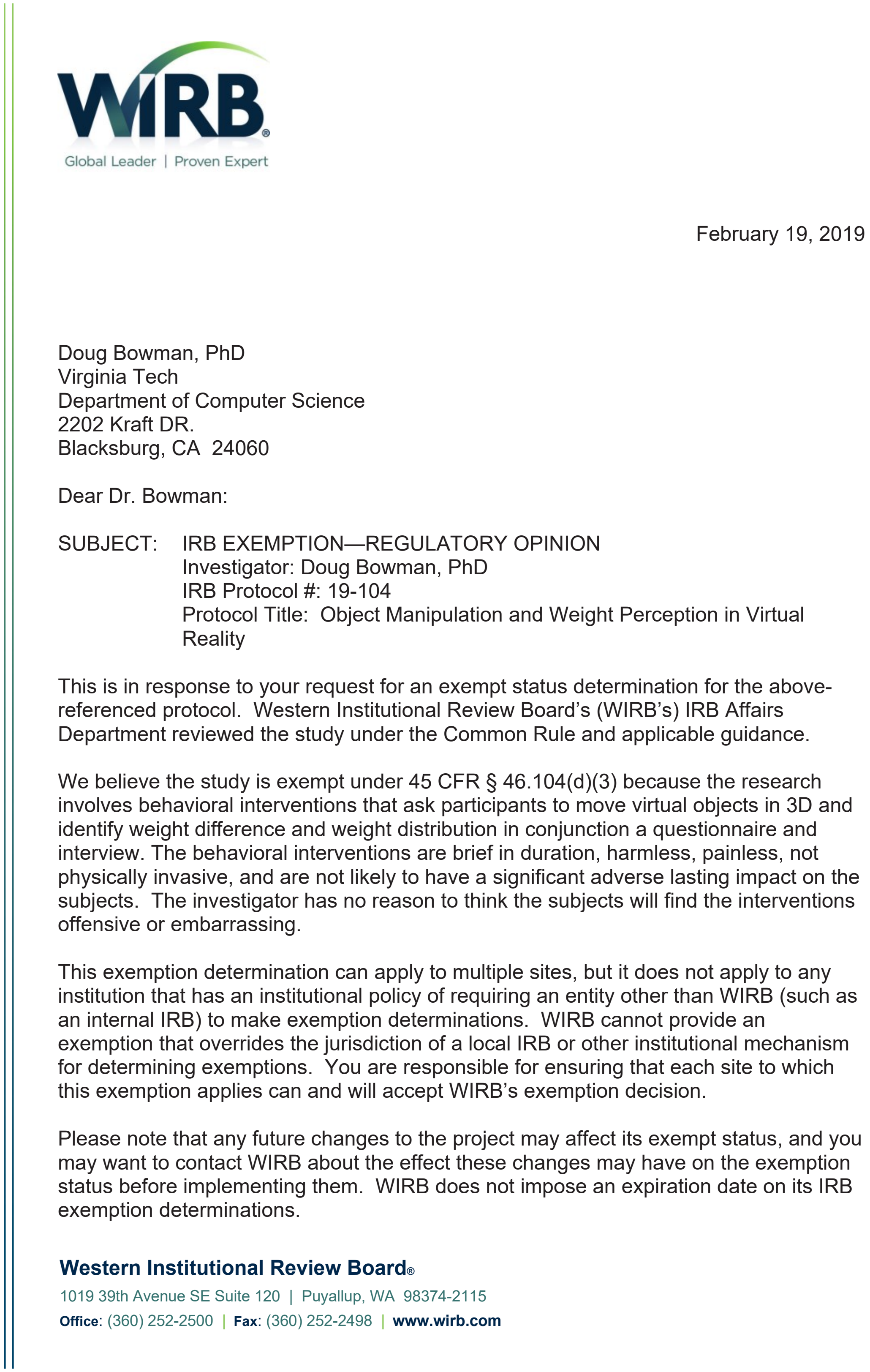
Invent the Future

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
An equal opportunity, affirmative action institution

Date*	OSP Number	Sponsor	Grant Comparison Conducted?

* Date this proposal number was compared, assessed as not requiring comparison, or comparison information was revised.

If this IRB protocol is to cover any other grant proposals, please contact the HRPP office (irb@vt.edu) immediately.



WIRB
Global Leader | Proven Expert

February 19, 2019

Doug Bowman, PhD
Virginia Tech
Department of Computer Science
2202 Kraft DR.
Blacksburg, CA 24060

Dear Dr. Bowman:

SUBJECT: IRB EXEMPTION—REGULATORY OPINION
Investigator: Doug Bowman, PhD
IRB Protocol #: 19-104
Protocol Title: Object Manipulation and Weight Perception in Virtual Reality

This is in response to your request for an exempt status determination for the above-referenced protocol. Western Institutional Review Board's (WIRB's) IRB Affairs Department reviewed the study under the Common Rule and applicable guidance.

We believe the study is exempt under 45 CFR § 46.104(d)(3) because the research involves behavioral interventions that ask participants to move virtual objects in 3D and identify weight difference and weight distribution in conjunction a questionnaire and interview. The behavioral interventions are brief in duration, harmless, painless, not physically invasive, and are not likely to have a significant adverse lasting impact on the subjects. The investigator has no reason to think the subjects will find the interventions offensive or embarrassing.

This exemption determination can apply to multiple sites, but it does not apply to any institution that has an institutional policy of requiring an entity other than WIRB (such as an internal IRB) to make exemption determinations. WIRB cannot provide an exemption that overrides the jurisdiction of a local IRB or other institutional mechanism for determining exemptions. You are responsible for ensuring that each site to which this exemption applies can and will accept WIRB's exemption decision.

Please note that any future changes to the project may affect its exempt status, and you may want to contact WIRB about the effect these changes may have on the exemption status before implementing them. WIRB does not impose an expiration date on its IRB exemption determinations.

Western Institutional Review Board®
1019 39th Avenue SE Suite 120 | Puyallup, WA 98374-2115
Office: (360) 252-2500 | **Fax**: (360) 252-2498 | **www.wirb.com**

If you have questions, please contact WIRB Regulatory Affairs at 360-252-2500, or e-mail RegulatoryAffairs@wirb.com.

AI:jca
D3-Exemption-Bowman (02-19-2019)
cc: VA Tech, WIRB, Virginia Polytechnic Institute and State University
Run Yu, Virginia Tech
WIRB Accounting
WIRB Work Order #1-1158022-1