Haptic Perception in Virtual Environments Using Proxy Objects: A Usability Study

by

Mina Abdi Oskouie

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Computing Science

University of Alberta

Abstract

Virtual Reality (VR) applications are getting progressively more popular in many disciplines. This is mainly due to (1) technological advancement in rendering images and sounds, and (2) immersive experience of VR which intensifies users' engagement leading to improved joy, sense of presence and success. Although VR creates a realistic experience by simulating visual and auditory sense, it ignores other human senses.

The absence of haptic feedback in Virtual Environments (VE) can impair users' engagement. Our objective is to provide haptic feedback in VE using the most affordable method, proxy haptics. Proxies are physical objects which are aligned with virtual objects in a way that by touching the proxies, users see their virtual hand touching the virtual objects. So, the haptic feedback will be applied to users' hands by touching the proxies.

Before using proxies to employ the human haptic sense in VE, we need to answer some fundamental questions about how this system works in VE. We have designed 3 experiments to find some answers to these questions. Our focus is primarily on the weight attributes.

In Experiment 1, we asked users about their expectation of the weight of 3 everyday objects. Results show an acceptable range for the weight of each of these object. The variation in acceptable weights is not because of inaccuracy in human weight memory since we observed that giving users a chance to evaluate the replicas could not make their expectation more accurate. Weight-size illusion can persuade the users to perceive bigger objects as lighter than

they are, but not under all circumstances.

In the second experiment, we asked users to compare two weights. The majority of our participants were able to detect the differences correctly even when they were as low as 15% of the base weight (the base weights were between 100g and 130g). The results show more accuracy when the comparing weights are lighter.

The last experiment was a usability test comparing using hand and the sense of touch with using VR controllers in a pointing task. Results show that providing haptic feedback can make the experience more believable but it cannot improve the perceived competence or ease of use. That's mainly because the hand tracking system we used was not accurate enough causing problems for users. With more accurate hand tracking the usability might be improved.

Acknowledgements

I would like to thank my supervisor, Professor Pierre Boulanger, for his insightful guidance, kind supports, and encouragement which help me throughout this project and my Master studies.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Humans have always dreamed of experiencing imaginary worlds in which they can define their own rules or experience situations which are hard to achieve in the real-world. Virtual Reality (VR) has made this dream possible. Using a head-mounted display and a pair of earphones, one can change an empty space into any world that one wants. VR can create a unique opportunity for human creativity to go beyond the boundaries of real-life creating in effect an awaken dream.

1.1 Applications of Virtual Reality

VR has been applied in many different ways. Some of the applications include:

- Simulation & Testing
  Virtual testing and simulations are considered an undeniable part of the modern design process. The ability to test new products under different usage or conditions makes VR an ideal platform to test if the design assumptions once translated to the real world can satisfy customers. Virtual testing is considerably cheaper for prototyping as it avoid the construction of expensive mock-ups. Virtual testing is widely used for controlling remote vehicles (mars rover) [73], creating virtual stores [11], and designing/testing new products [3].

- Education
  VR has been widely used for educational purposes. Many companies

use VR applications to train employees faster and more efficiently. For example, [64] proposes a VR training system for children with the psychological disorder. [69] developed a VR balance training program for elderly and [22] introduces a VR-based training system to enhance the functional autonomy of patients with Alzheimer.

Virtual training is especially advantageous for life-threatening situations which are hard to simulate in the real world. For example, traditionally surgical students are trained by practising on cadavers or live pigs and by watching an experienced surgeon operating on a patient gradually allowing them to perform small tasks as their expertise grows. Using VR, one can create similar situations as in an operating room helping medical students to deal with and learn the complexity of surgical procedures. In order to achieve this, VR systems must be able to create a sense of presence and engagement similar to real conditions in an operating room [13], [27], [62].

- Entertainment

  Virtual Reality is becoming more and more popular in the gaming industry. Many companies are introducing VR devices for their game consuls. They are trying to make the devices affordable and convenient for home users. Also, gaming companies are releasing more games for virtual reality environment.

  Coupling the sense of presence in VR with intense engagement of gaming makes virtual reality games motivating for users. They can play for a long time without even noticing the passing time. It is also shown that working memory performance improves in VR games compared to Desktop ones [68].

1.2 Sense of Touch in Virtual Reality

Significant improvements in audio, visual, rendering, and motion tracking technologies has enabled VR systems to create high-quality and low cost immersive environments that can be used to create an intense sense of presence.

Unfortunately, haptics - the sense of touch - remain relatively undeveloped in virtual environments. In current VR systems, the sense of presence disappear when the user's hand goes through an object. Being able to touch objects in VR is critical in order to make the experience significantly more engaging and natural. This is especially important for simulations like virtual surgery or mechanical assembly. Coupling the sense of touch with other senses (audio and visual) not only enables medical students to believe that they are operating on a real-patient but also give them the unique chance to practice important skills such as hand-eye coordination.

1.2.1 Approaches

Many researchers have tried to enrich VR experience with the sense of touch. One can find in the literature two main categories, *Active Haptic* and *Passive Haptic*. In Active Haptic, the haptic feedback is provided by small mechanical robots called haptic interfaces that are controlled by a haptic rendering computer simulating the physical interaction between a virtual tool (registered with a real tool) and the virtual environment. The haptic interface actuation is calculated based on the user's actions in the virtual environment and is returned to the user as a force and a displacement.

Passive Haptic, on the other hand, uses physical objects, called *Proxy haptic objects* or *props*, to simulate the sense of touch [34]. When the user touches a virtual object, its co-registered proxy haptic object will provide a feedback about the texture, weight or temperature of the virtual objects.

1.2.1.1 Active Haptic

Active Haptic systems use mechanical devices to provide the desirable haptic sensation to the user's body based on the simulation on how the device would interact with the real-world using basic laws of physics. These devices can be designed in wearable or non-wearable forms. Most of the available devices are non-wearable [16], [66], [71]. The problem with this approach is that the user is obliged to hold the devices all the time making it hard to create a real-sense of touch.

The other approach to Active Haptic is to use wearable robots to simulate the sense of touch by sending tactile signals to the skin. These robots have a wide range of sizes and shapes. From the small ones just covering finger tips to the huge robots which cover the whole upper part of the body. Prior work by [16], [61], and [6] have proposed wearable haptic devices which create the illusion of touching virtual objects by applying appropriate forces to fingers and fingertips. Other researchers [46] provide the sense of touching walls or lifting heavy objects by sending electrical signals to the arms or hands muscles.

Being forced to wear a robot exoskeleton or receive small electric shocks are not always pleasant for the users. Plus, feeling that there are devices on their skin makes the users aware that he/she is holding a haptic device hence loosing the sense of presence one wants to achieve. Additionally, these haptic robots are very expensive and hard to program.

1.2.1.2 Passive Haptic

Using passive haptic , one can make the tactile VR experience more affordable. In this approach users are actually interacting with real physical objects called ”proxy haptics objects” that are spatially co-registered with virtual objects that may have different appearances and shapes. In other words, proxy haptics objects are providing haptic feedback without the need to build complex robots and making users wear or hold them.

1.3 Problem Definition and Research Methodology

It is not always reasonable to consider that the proxy object should be exactly the same as its VR counterpart. The question then is by how much the proxy object should be similar to its VR counterpart in order to make the haptic perception believable? How much difference between proxy object and the virtual object can go unnoticed by the human haptic system? How accurate do the proxy haptic object’s location, weights, and temperature need to be determined in order for the human haptic system to believe it is real ?

To achieve this goal, we first need to answer fundamental questions about the performance of the human haptic system in VR. Does this system perceive the attributes of the objects in the most accurate way? Finding the limitations of the human haptic system allows us to understand how far one can deviate our haptic feedback from the real one, produced by holding the actual object. This inaccuracy enables us to provide less accurate feedback in VR without harming the delicate sense of presence. As a result, for passive haptic one can use more abstract proxies and in active haptic the robots are not required to be absolutely accurate. The main hypothesis of this thesis is: *Can we use the inaccuracy of the human haptic system to build believable yet affordable VR experiences using proxy haptic?*

To validate this hypothesis, we have designed 3 sets of experiments. In the first experiment, we examine the accuracy of haptic system and haptic memory. By asking users, indirectly, what is the acceptable weight for some of the everyday objects. In the second experience, we investigate how humans compare the weights and how accurate this comparison is. In the last experiment, we question the role of proxy haptics in accuracy and speed of the human aiming and reaching performance in VR. Throughout all experiments we measure the usability of VR experiences enhanced with proxy haptics. A detailed description of each experiment will be presented in Chapter 3.

1.3.1 Experiment 1

The experiments take place in a virtual room shown in Figure 1.1. Participants were asked to wear a HTC Vive head-mounted display. A Leap motion controller attached to the head-mounted display detects participant's hand without requiring them to wear any marker. As a result, they can see a representation of their hand in the virtual reality environment. The physical objects in the experiment room are co-registered with the virtual environment. Therefore, when participants see that they are touching an object in the virtual environment, they are also touching the physical object simultaneously.

The main goal of this experiment is to determine how accurate the weight of

Figure 1.1: Virtual Environment of User Study 1, Phase 1

the proxy haptics objects must meet the user's expectations? This experiment has two phases. In the first phase, we asked users about their evaluation of the weights of the objects and in the second phase we want to know how certain they are about their response.

1.3.2 Experiment 2

In Experiment 1, we examined how accurately the human haptic system can estimate the weight of an object from its appearance. Now the question is how much one can deviate from this estimation without the human haptic system noticing the changes. In other words, replacing actual replica with a proxy haptic object, one must make sure that the weight difference between them is not noticeable.

Weber's law states "the perceived change in stimuli is proportional to the initial stimuli". Fechner completed the law by showing that the Just Noticeable Difference (JND), the smallest difference which is perceivable for human, is proportional to the initial stimuli for any sense. Since what the human senses are precising is different in the virtual environment comparing to the real world,

Figure 1.2: Virtual Environment of User Study 2

we expect the human performances to be different as well. [47] has shown that kinematics of the reaching performance can be altered by availability of virtual and haptic feedback. In experiment 2, our aim is to find the JND for weight in VR and investigate if it remains the same in the virtual and the real world.

In this experiment, we asked the participants to evaluate the cubes' weights relative to each other. They are provided with a basis cube and a comparing cube. We ask them to compare the weights and determine whether the comparing cube is heavier, lighter or has the same weight as the basis cube. Figure 1.2 shows the environment of experiment 2. These cubes will be presented to the user one at a time. They are allowed to lift cubes as many time as they want, just like before. We have repeated the whole experiment with 3 different basis cubes. In order to be able to compare the users' behaviour in VR and real world, we have conducted the same experiment in the real world as well.

1.3.3 Experiment 3

In this experiment, our goal is to investigate the influences of using proxy haptic feedback in a pointing task. We ask users to re-type a word they see on the screen. A keyboard is presented to them just like keyboards on computers. Figure 1.3 shows the environment of experiment 3.

In the first trial, they can see a representation of their hands in the virtual

Figure 1.3: Virtual Environment of User Study 3, First word

environment. In order to type the sentence they need to press keys with their own hands. Using a proxy haptic object, in this case a sponge, we were able to provide a haptic feedback corresponding to pressing a key. In the second trial, users were asked to do the exact same task but now with HTC Vive wand controller.

1.3.4 Questionnaire

After finishing all the experiments, we ask participants to complete a questionnaire which is designed to compare the two trails of experiment 3 in three main categories, perceived competence, ease of use, and believability.

1.4 Thesis Contribution

Based on what we explained in section 1.3, the contributions of this thesis are:

- Designing an experiment to test the human haptic system in a Virtual Environment
- Testing the performance of the human haptic system under different circumstances
- Finding a threshold for Just Noticeable Difference of weights in a Virtual Environment

- Empirically find the advantages and disadvantages of providing haptic feedback using proxies

1.5 Organization of the Dissertation

The rest of this document is separated into five chapters:

- **Chapter 2:*Related word***
  In this chapter, we review previous research related to Virtual Reality, especially those that enriched virtual experiences with simulating the sense of touch.

- **Chapter 3:*Study of Human Weight Expectation Using Proxy Haptic***
  This chapter is devoted to explaining the design and procedure of our first two experiments.

- **Chapter 4:*Study of Human Weight Expectation Using Proxy Haptic: Experimental Results***
  The results of the first two experiments are represented in this chapter.

- **Chapter 5:*Proxy Haptic for Pointing Task in the Virtual World***
  Our usability test, experiment 3, is explained in this chapter along with its results.

- **Chapter 6:*Conclusion***
  We conclude this dissertation with a summary of our findings and the insights they bring for application designers.

Chapter 2

Related Work

Virtual Reality (VR) is a synthetic reality created by a computer simulation. VR provides an immersive experience for human operators in which they have the illusion of interacting with the real world. This interactive illusory environment is called a Virtual Environment (VE) [63].

2.1 VR Taxonomy

The real and virtual worlds can amalgamate in different levels to create environments, in which neither total physicality nor complete synthesis pertains. This merging spectrum establishes the *virtuality continuum*. Milgram et al. [49] proposed a taxonomy of mixed realities. Figure 2.1 illustrates a simplified version of the continuum joining the purely real environments to completely virtual ones. Augmented Reality (AR) refers to environments which represent the real world enhanced by virtual objects. At the other end of the spectrum, Augmented Virtuality (AV) or Virtualized Reality is a virtual environments that take advantage of some real-world objects who are digitized by 3D sensors.

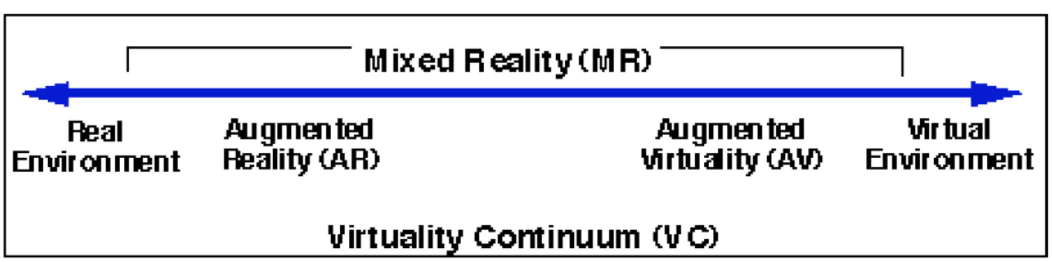

Figure 2.1: Simplified representation of a "virtuality continuum"[49]

2.2 VR Application

The immersive experience of VR makes it a unique platform for simulating situations which are dangerous or hard to achieve in the real world. Flight simulations can be mentioned as one of the earliest applications of VR. These simulations are utilised for training pilots under different circumstances such as airports for taxi, mechanical defects, and landing in various air conditions [55].

The enhanced user engagement, sense of presence, and immersion in VR, are shown to have positive impacts on learning outcomes [9], [19]. Virtual training is widely used in different disciplines: Gordon et al. [28] proposed a way to integrate virtual simulation into the nursing curriculum; Pulijala et al. [56] designed a VR tool for orthognathic surgical training; Filigenzi et al. [21] developed a VR application for mine safety training; Fast et al. [20] created a mixed reality system for training gas metal arc welding.

Many other tools leverage VR for a wide range of applications. Rehabilitation is one of the most well known usages of VR [42], [48], [51], [58], [59]. Some works take advantage of VR for treatment [4], [24], [25], [53]. VR is also used for game and e-shopping industries [5], [40], [44].

2.3 VR and Sense of Touch

VR applications are getting more and more realistic by improving the video quality, advanced rendering, and 3D audio. However, the interaction with virtual worlds remains relatively unnatural. Most of the virtual objects remain untouchable in the way that users' hand go through them while they are touching nothing in the real-world. Interactive objects are always input/output devices which need to be held throughout the whole experience. These devices, usually shown as controllers or gun, allow users to interact with the world. This reduce the delicate sense of presence drastically and deprives users of acting in a natural way in the virtual worlds.

Many works demonstrated the value of adding haptic feedback to the virtual environment. Burdea et al. [10] enhanced VR with sense of touch for tracking and dexterous manipulation task. They were able to observe a significant increase in simulation realism with improvement in task completion time, number of errors, and learning time. Hoffman et al. [35] showed that touching virtual objects makes the virtual environment more realistic for the users. Insko et al. [37] investigated haptic feedback influences on the performance of a navigation task. They trained half of their participants on maze navigation augmented with touchable walls, while the other half were trained in the non-augmented environment. The participants that received augmented training show significantly better time to complete the blindfolded navigation and a reduce number of collisions. Khurshid et al. [41] showed the positive effects of haptic feedback on users' performance in teleoperation.

2.3.1 Approaches

Several studies have been done to augment virtual reality with the sense of touch. They have used different approaches to reach their goal. What follows is a classified summary of some of these works:

2.3.1.1 Active Haptic

In this approach, the force required for haptic feedback is calculated by a computer system and applied to users body through an haptic device. These haptic devices are designed and built in many different forms:

- **Touchable Interfaces**

  Levesque et al. [43] enhanced the screen touch interaction with variable surface friction. They showed that this enhancement is beneficial for low-level targeting activities. Nakagaki et al. [52] proposed a shape-changing interface to represent material properties. Figure 2.2 is showing this interface.

- **Wand-based controllers**

  Wand-based controllers (e.g. HTC Vive controllers) are devices that users need to hold in order to interact with virtual environments. Unlike

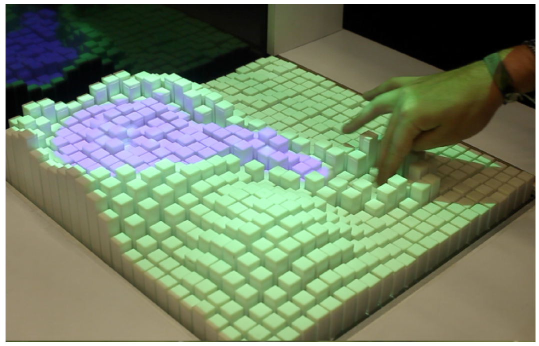

Figure 2.2: Materiable: A shape changing interface for providing haptic feedback [52]

the grounded interfaces, they allow users to move freely. The fact that users' hands are always occupied with these controllers, deprives users of grasping objects naturally or perceiving their physical properties (e.g. shape, temperature or texture). Research conducted with the goal of improving controllers and making them more realistic:

NormalTouch and *TextureTouch* render surface height, orientation and texture [2]. *Shifty* is another device designed to simulate more natural pick-up experience in VR [74]. Weight distribution in *Shifty* change dynamically base on the situation. The *Haptic Revolver* provides a wheel under users' finger to render haptic sensation of texture, shape or interactive elements by spinning [71]. *Haptic Links* attach two VR controllers together to alter forces between users' hands dynamically and mimic the haptic feedback of a variety of two-handed object tasks [66]. Figure 2.3 shows all those devices.

Although these devices are doing a great job at providing realistic haptic feedback, each has a specific application and is not necessarily suitable for all virtual environments. Plus, because a user is obliged to hold a tool throughout the whole VR experience, it may become unpleasant in long session.

- **Wearable Devices**

  Wearable haptic devices come in different sizes and formats. Researchers have designed wearable fingertip-based haptic devices in order to simu-

Figure 2.3: Wand-based Haptic Devices. First Row: left: TextureTouch[2], right: NormalTouch[2]. Second Row: left: Haptic Links[66], middle: Shifty[74], right: Haptic Revolver[71]

late weight[50], contact force[23], texture[72], and kinesthetic grip force for grasping virtual objects[15]. Although performance improvement and enhanced engagement are observed using these devices[6], [54], the fact that the haptic feedback is only applied to a specific body point makes the experience not completely natural.

Many devices have been developed with the goal of providing realistic haptic feedback to the users' entire hand. The *Rutgers Master II* haptic device is applying forces on users palm[7]. The Dexmo *Dexmo* is a light-weight haptic device designed for fingertips and back of the hand[29]. The *Grabity* simulates weight and grasping by applying force on all fingertips[15]. Figure 2.4 shows many of these devices.

Another line of research is to build devices that can apply haptic forces on both arms. This enables them to simulate the weight of heavier objects. Some researchers proposed robotic arms to achieve this goal[12], [26], [30], while others utilise small electrical shocks to arouse neurons[46].

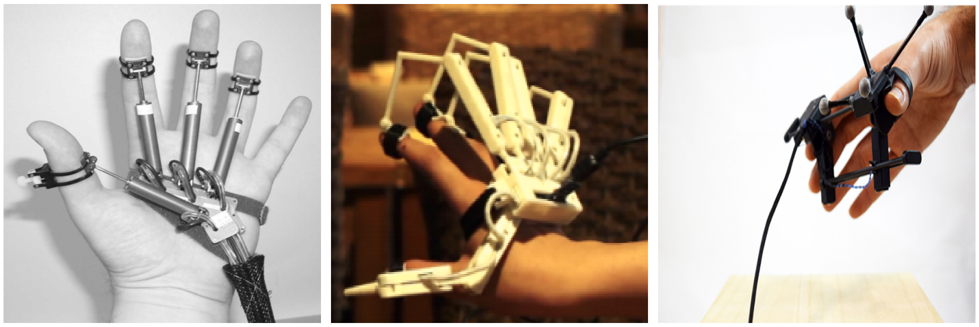

Figure 2.4: Wearable Glove-based Haptic Devices. left: Rutger Mater II [7], middle: Dexmo[29], right: Grabity[15]

The wearable haptic devices are successful in creating realistic haptic feedback. However, they are extremely delicate and hard to build. This makes them very expensive and obstrusive. Additionally, each of these devices is designed for specific task i.e either simulate weight or texture. Plus, users need to wear these tools in order to be able to interact with the virtual world. This obligation can be unpleasant since the devices can be heavy and uncomfortable.

2.3.1.2 Passive Haptic

Passive haptic approach provides the haptic feedback using physical objects or proxies. Unlike the active haptic approach, there is no need to calculate and synthesize the appropriate haptic feedback. In other words, it is leveraging physicality of the real world to create more realistic virtual experiences. The main goal is to trick users' minds to believe they are touching virtual objects while they have a physical object in their hands. The virtual and the physical objects are not necessarily the same. The most important question here is what is the best replacement for virtual objects or how we can safely substitute a virtual object's physical replica without users noticing it.

The work of Hinckley et al. [34] can be considered as one of the first attempts to use passive haptic. They provide neurosurgeons with a more convenient way to plan the surgery. In this way, surgeons can apply the possible cuts on the image they see on screen by manipulating the doll's head they have in their hands. Lindeman et al. [45] made one of the earliest attempts to marry haptic feedback and VR by the means of passive haptic. Henderson et al.[32] created a tangible user interface in AR and Hettiarachchi et al.[33] proposed a way to represent physical objects with non-replica virtual objects in AR.

In order to make sure that the difference between virtual and physical objects will not harm the sense of presence, we need to know how accurately the human haptic system can perceive physical aspects of the objects. Psychologists have shown the dominance of vision when human senses conflict. Using this fact, we can deceive users to believe what they see in VR against what they touch in the real-world. Azmandian et al. [1]proposed a framework, called *Haptic Retargeting*, in which they can redirect users' hands toward the proxy they have provided . As a result, they can have only one proxy for all virtual objects. Cheng et al. [14]used the same technique. They have provided a class of proxy haptics, *Sparse Haptic Proxy*, that consists of surfaces with different orientations. The scene's geometry can then be simulated by guiding users' hand to the appropriate proxy.

Other works have tried to suggest the best substitute for the objects in physical world [18]. Simone et al. [65] introduced a new class of VEs, called *Substitutional Reality*, in which each object of the physical world is aligned with a virtual object . Figure 2.5 shows a physical environment in the middle and two substitutional virtual environments for that on the left and right.

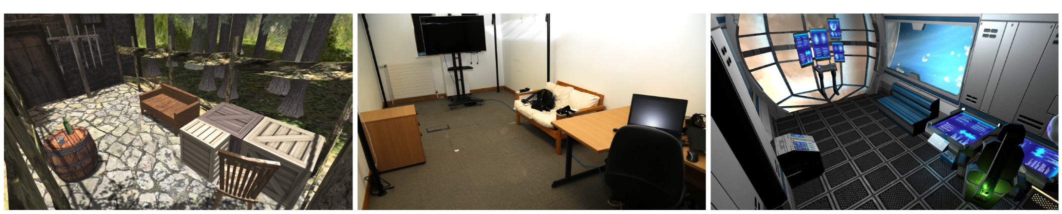

Figure 2.5: Substitutional Reality[65]

They conducted two studies. In the first one, they provided users with a physical object and altered its virtual representation to investigate its influences on participants' suspension of disbelief and ease of use. They considered five levels of mismatch, Replica, Aesthetic, Addition/Subtractions, Function, and Category. In the Replica level, the virtual and the physical object are exactly the same. Aesthetic level represents the situation when the two objects begin to differ in an aesthetic aspect. At Addition/Subtractions level, the physical and the virtual object contain some elements that are absent in the other. The objects differ in Function level to the extent that the possible interactions of the objects are not matched, e.g. a mug is replaced with a lamp which has a different functionality. Category level involves the most mismatch in which there is no connection between the appearance of the two objects.

In the second study, the virtual object stayed the same while different alternatives are presented for it in the real world. for this study, they have used the replica and two other physical objects at the functional level of mismatch and observed the changes in users' level of engagement.

Their results can be summarized as following:

- Substitutive object can be as engaging as a replica if it is easier to work with, i.e. lighter.
- Material of the virtual objects can influence users' expectation about their physical aspects (e.g. plastic box is expected to be lighter than iron one).
- Using methods like *Haptic Retargeting* can improve believability when the level of mismatch is at Addition/Subtraction (i.e. prop is smaller or larger than the virtual object).
- It is important to minimize the differences in the manipulatable parts of the objects while the mismatch of other parts or out of reach objects does not have significant effects on users' engagement.
- The familiar VEs require a higher level of resemblance while they are more believable.

Looking at the literature, many questions have still remained unanswered. How much one can deviate from the exact match between a virtual object and its corresponding physical object without users noticing it? How accurately the human haptic system can perceive the physical aspect of the proxies in VR? In this thesis, we want to focus on the weight aspect. So, we need to first know about the research psychologists have conducted about how humans perceive heaviness.

2.4 Human Weight Perception

Weber's inquiry can be mentioned as one of the earliest attempts to understand the mechanism and properties of perceiving heaviness[70]. Weber's law states the Just Noticeable Difference (JND), defined as the least noticeable difference, for weight depends on the intensity of weight. Weber fraction is found by dividing JND by its corresponding basic weight. It is shown later that Weber fraction varies in the range of 0.02 to 0.13 for lifted weights, when the ratio of weight differences over the base weight is between 0.16 and 1 ($\frac{\Delta W}{W} \in [0.16, 1]$) [36], [57], [67]. Weber fraction highly varies between individuals. One of the factors which influences it is age, which decreases the sense of heaviness [17].

Psychologists are convinced that human heaviness perception depends on many different physical attributes and not necessarily equal to the object's weight. Kawai et al. stated: "*Heaviness perception is not simply a function for sensing object weight. Rather it functions as a form of recognition of the object. That is, all information obtained synchronously when humans grasp and lift an object may be gathered across various modalities, integrated with or subtracted from each other, interpreted by object knowledge, and then formed as perceived heaviness.*" [38]

Kawai devoted his research to investigating human wight perception. In his book, he categorised his experiments into three classes, Real Haptics + Real Vision, Real Haptic + Virtual Vision, and Virtual Haptic + Virtual Vision [38]. Since in this study, we are focusing on providing haptic feedback in VR, we are going to present a more detailed review of the second class of experiments, Real + Virtual Vision.

Kawai's research mainly focus on weight-size illusion. He conducted an experiment in VR and providing passive haptic to the participants. Two cubes

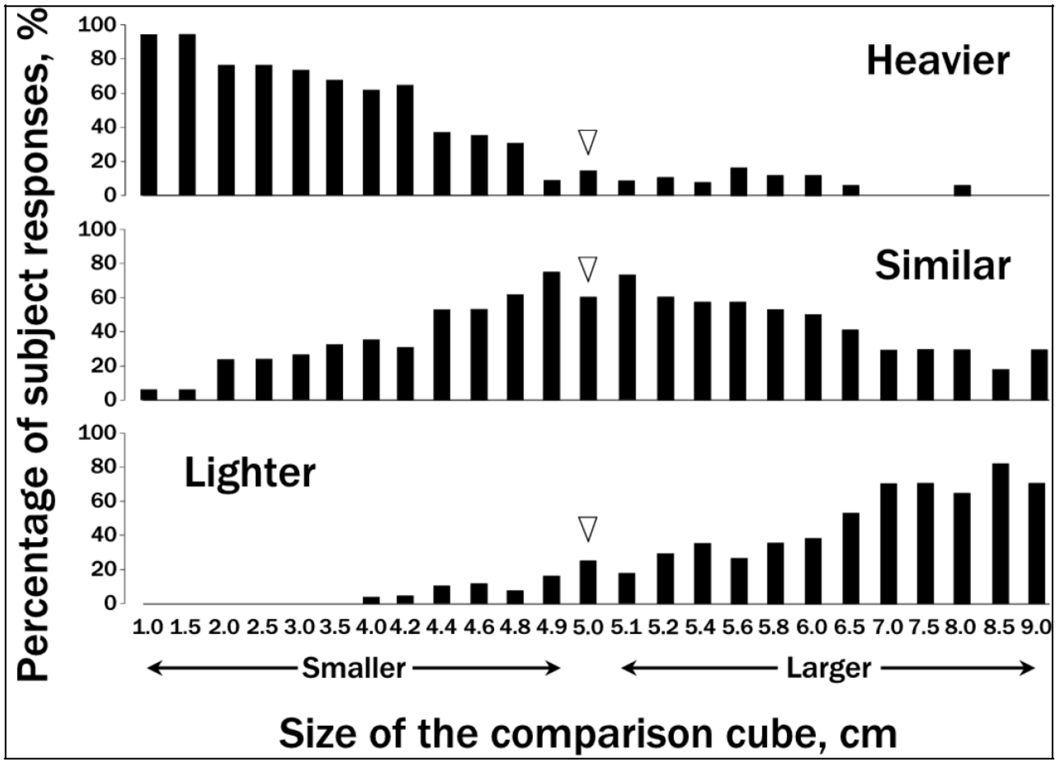

Figure 2.6: The results of weight-size illusion experiment[38]

are represented to the participants, a standard cube and a comparison cube. The physical props are cubes with a size $3_{cm} \times 3_{cm} \times 3_{cm}$ and a weight of 30 grams. In the virtual environment, the standard cube is shown with a size $5_{cm} \times 5_{cm} \times 5_{cm}$ and a size of the comparison cube varies between 1_{cm} to 9_{cm}. The participants are asked to compare the weight of the standard and the comparison cubes and report which one is heavier or they have the same weight. Based on the results in Figure 2.6, visual size significantly influences user perception of weight and users usually perceived smaller cubes to be heavier.

Plenty of questions still require answers. In this thesis, we are investigating user expectation about the weight of everyday objects. We also want to know why this expectation may be inaccurate. Is it because the weight memory in humans is vague and abstract, humans cannot perceive heaviness accurately, or they have seen the object in different weights (i.e. heavier and lighter apples). Additionally, we want to take steps in observing whether the usability of our system can be improved when it is augmented with proxy haptic feedback compared to using VR controllers for a pointing tasks.

Chapter 3

Study of Human Weight Expectation Using Proxy Haptic

It is not always reasonable to consider the proxy to be exactly like its corresponding virtual object since it would require the same environment in the virtual and real worlds. The question which remains unanswered is how much the proxy should resemble its corresponding virtual object? As, the human haptic system can evaluate different attributes of objects (e.g. weight, texture, and temperature), it is necessary to tackle the problem for each of these attributes separately. In this work, we will focus on the perception of weight using proxy haptic.

Imaging we want to represent an everyday object in VR and we want to provide users with a corresponding proxy in the physical world. The questions are: what can be considered as the appropriate weight for the proxy? What is users' expectation about the weight of the proxy when they see an everyday virtual object? What weight is too heavy or too light that can bring them out of their VR illusion?

In this project, we are investigating the human expectation about the weight of some everyday objects.We have designed two sets of experiments to find answers for the aforementioned questions. In the first experiment, we want to find the acceptable weight for a proxy of three everyday objects: an apple, a notepad, and a wireless mouse. Since, all of these objects come in different weights, forms, and appearances, we expect to find an acceptable range rather than a certain weight. So, the follow up question is what is actually creating this range? We have three hypotheses and we will test all of them in our experiments.

3.1 Apparatus

In this section, we will present a description about the physical room the studies took place, the virtual room that participants interacted with, and the equipment used for these experiments.

3.1.1 The Physical Environment

Both of the studies took place in a small area of the AMMI Lab at the University of Alberta, shown in Figure 3.1. Our goal was to use minimum amount of equipment and keep the experiments as simple as possible, therefore we only used HTC Vive head tracking to find the position of participants in the room. The Leap motion controller attached to the head-mounted display detected participants' hand without any markers. All the proxies were presented to the participant either on a table in front of them or on a wall at their left.

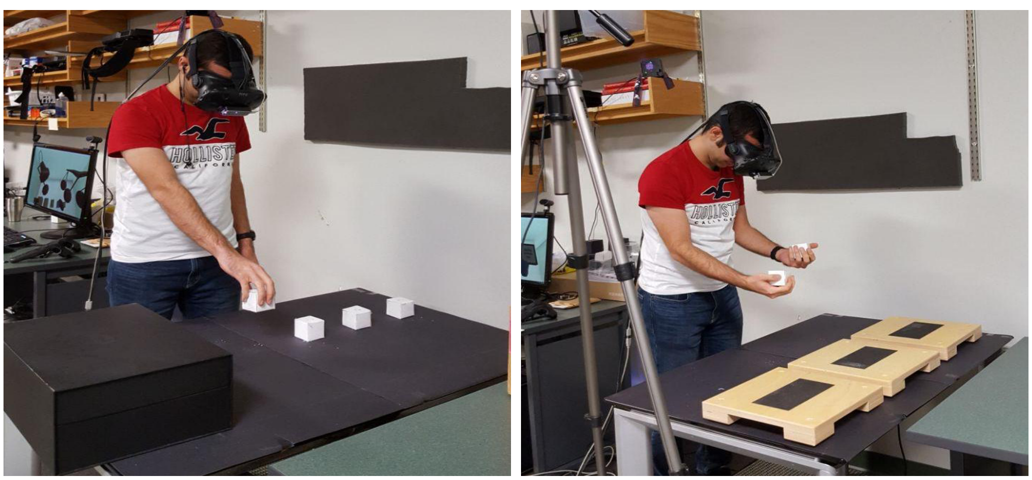

Figure 3.1: Physical Environment for the User Studies, left: User Study 1, right: User Study 2

3.1.2 The Virtual Environment

The experiments took place in a virtual room designed using Unity 3D version 5. The participants wore a tracked HTC Vive head-mounted display (HMD). Using Leap Motion mounted on the HMD one could see a representation of their hand in the virtual environment. The physical table and the left wall in the experiment room were co-registered with their virtual representations. Therefore, when participants touched them, they were actually touching a physical object in the real-world. The co-registration is done based on the position of the virtual and the physical hands, in a way that when we saw our virtual hand touching a virtual object, its corresponding proxy should be touched by our physical hand in the real world. Figure 3.2 shows the virtual environment for both experiments.

3.2 User Study 1: Human Weight Expectation

To make the experiment of virtual touching believable, the attributes of proxy haptics objects should meet the users' expectations. In this experiment, we investigated the accuracy of this expectation in different situations. We have considered three common objects: an apple, a notepad, and a wireless mouse.

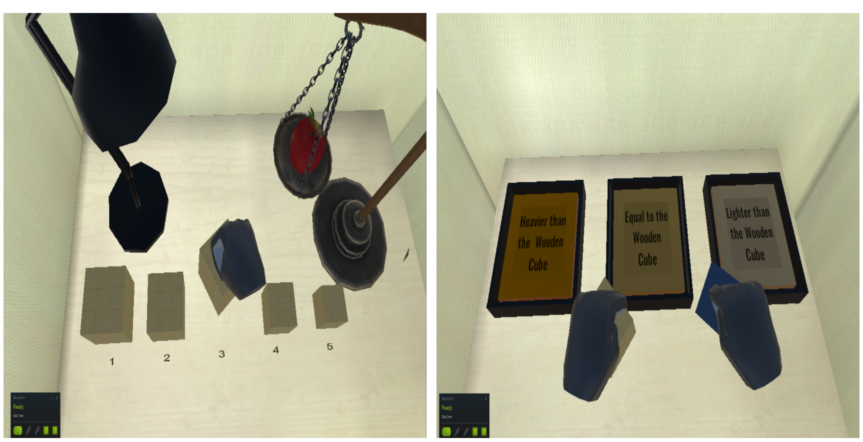

Figure 3.2: Virtual Environment for the User Studies from Participants' Point of View, left: User Study 1, right: User Study 2

The first step was to ask them about their expectation about the weight of these objects.

User study 1 has two phases. In the first phase, we are going to ask users about their evaluation of the objects' weights, and in the second phase, we want to know how certain they are about their response.

Figure 3.3: Replicas of three objects

3.2.1 Phase 1

There were five cubes with different weights on the table. We chose to represent the options as cubes since its flat sides can provide uniformly distributed pressure on users' hand. There was a two sided scale with an ordinary object (e.g. apple) on one of its sides. We asked users to choose the cube which they believed can balance the scale. In this process, the users could lift each cube freely as many time as they want to assess its weight. But they could not touch the object on the scale or use the scale to evaluate the object's weight.

Phase 1 was repeated with 3 different objects on the scale (Apple, notebook, and wireless mouse). Figure 3.3 shows the replicas that we used for these 3 objects. The true weights of the replicas were 131g for the apple, 104g for the notebook, and 117g for the mouse. We wanted to focus our work on a small range of weight. We did not want to examine only one object in that range because this may bias our results. In order to make sure that our results will be generalized enough at least for object in this weight range, we repeated our experiments 3 times with 3 different objects. The five cubes on the table had 70%, 85%, 100%, 115% and 130% of the weight of the object on the scale. The following table shows the weight of the cubes on the table for each object.

	Actual Weight	Conditions				
		A 70%	B 85%	C & F 100%	D 115%	E 130%
Apple (α)	131 g	92 g	111 g	131 g	151 g	170 g
Notepad (β)	104 g	73 g	88 g	104 g	120 g	135 g
Wireless Mouse (γ)	117 g	82 g	99 g	117 g	134 g	152 g

Table 3.1: Proxies' weight in each condition

3.2.2 Phase 2

After users chose one of the cubes and put it on the scale, we asked them to answer a question. The question simply asked how sure they are about their choice. There were four options available and they could choose one by pressing the button. The options were: "Not sure at all", "I could eliminate one option", "I could eliminate two options", and "Absolutely Sure". The buttons are touchable and users should press the right choice using their hands.

3.2.3 Task

Before beginning the study, the concept of proxy haptic was explained to the participants: the virtual table, wall, and boxes are co-registered with a proxy. We explained to the users that the purpose of the project is to evaluate their expectation about the weight so, they can lift the boxes in any way they want but they cannot touch the everyday object on the scale. The actual replica was shown to them before starting the experiment but they could not touch the replicas.

3.2.4 Participants

Before starting the experiments, we needed to estimate the sufficient number of participants for the experiment to be significant. The number should not be too small to draw inaccurate and biased conclusions, or too big to waste a lot of time and energy. Since for all experiments we wanted to compare the results of two settings, power analysis was chosen as an acceptable estimate of the number of participants.

Figure 3.4 shows how increasing the power of the statistical test can change the ideal number of participants. The diagram is achieved by setting alpha = 0.5 and effect size = 0.4. Since we wanted to achieve the power equal to 90%, we needed 34 participants for the experiments.

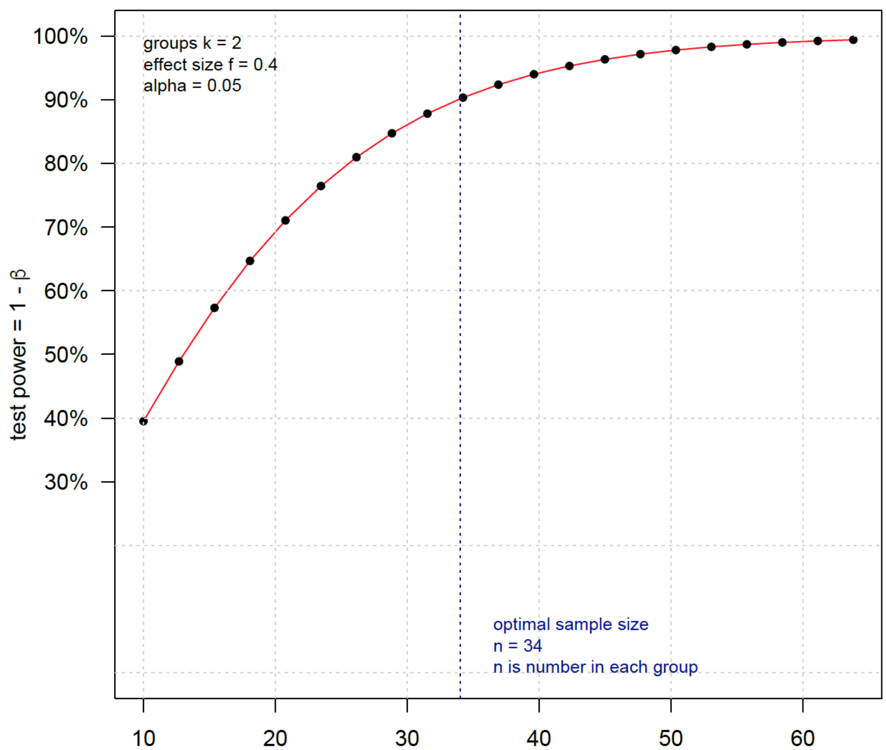

Figure 3.4: Ideal sample size for the studies considering the power of the statistical test.

Thirty three participants, thirty two of them aged between 22 and 34 and one was 53 years old (M= 27.03; SD=5.52), took part in the first study. All of them were graduate students at the University of Alberta, Department of Computing Science. We asked them how frequently they used video games, VR, and VR controllers on a scale of 1 to 4. As the video games generally are more available, we interpret its scales as 1 (not frequently play) to 4 (play everyday). For VR and VR controllers 1 means "Never worked with it" and 4 means "Frequently working with it". Our sample consisted of a group of in-frequent gamer (M=1.94; SD=1.10) with little experience with VR and VR controllers (VR: M=2.33; SD=1.09 , VR controller: M=2.21; SD=1.17).

3.2.5 Situations

Research has shown that human weight perception is highly related to visual clues and the whole environment[31], [39]. Therefore, one needs to test this perception under different circumstances to reach a general conclusion.

We repeated phase 1 of the first study for 4 different settings:

- *Same Size*

  In this setting, the five cubes on the table had the same size while their weight were different. Weight conditions are shown in table 3.1. The size of proxies were $5_{cm} \times 5_{cm} \times 5_{cm}$. We did not want to present the cubes in an ascending or descending weight order because it could give users extra information about the weight without examining each cube independently. Like in case when the right neighbour of a cube is too heavy in users mind and the left neighbour is too light. The user can pick this option without even examining it. We also did not want big weight differences between the neighbours (e.g. putting cubes with weight conditions A and E beside each other). So we presented five cubes on the table in this order: B, A, C, E and D as previously shown in Table 3.1.

- *Different Sizes Correlated with Weight*

  The five cubes presented in this setting are sorted by size. Proxies are square cubes with sides, l, between 4_{cm} and 8_{cm} ($l \in \mathbb{N} : l \in [4_{cm}, 8_{cm}]$). The weight of the proxies are correlated with the sizes in the way that the biggest cube is the heaviest one.

- *Different Sizes Uncorrelated with Weight*

  The sizes of the proxies and the order are exactly like the previous setting, sorted by size, but the weights are not correlated with the sizes. Here, we wanted to present deceiving visual clues and examined whether this information could influence users choices. As it is impossible to try all the uncorrelated size-weight combinations ($5!–2 = 118$ combinations), we considered only 5 samples. We started with the same order as the setting *Same* and used a Latin square of order 5 to find other size-weight combinations. Table 3.2 shows Latin Square for weight-size combinations for the first 5 users. Each row shows the weight conditions assigned to cubes with specific sizes represented for a user. The conditions were the same for all objects but the actual weights were different since the replicas' weight differs. The actual weight can be found in table 3.1.

- *Given Prior Knowledge*

  In this setting, we wanted to observe whether refreshing users' memory by weighting the replicas can make their judgement more accurate. After finishing trials in previous settings, we provided users with the replicas of the three objects and asked them to take off the head-mounted display and examine them naturally. Then, we repeated the experiments under the situation of *Same Size* setting.

3.2.6 Order Effect

The factor that might change the users' responses was order effect. In other words, the order that we presented trials to the users might affect their answer. We assumed that the trials for three different objects are independent, because the object are irrelevant, the five options on the table are totally different and the task is not about comparing these objects. We also did not observe any feedback, thoughts (using think aloud) or questions from users' showing that they have take the trials with other objects into account for choosing between boxes.

	Side Size (cm)				
	4	5	6	7	8
User 1-7	B	A	C	E	D
User 8-14	A	C	E	D	B
User 15-21	C	E	D	B	A
User 22-28	E	D	B	A	C
User 29-33	D	B	A	C	E

Table 3.2: Weight-Size combination for uncorrelated setting

On the other hand, the order of presenting setting might affect users' answers. The setting *Given Prior knowledge* should come last because we want to investigate users' expectation about the weights with no immediate prior knowledge provided. In order to eliminate the order affect for other three settings, we have used a Latin square to randomise the order of trials. Figure 3.3 shows this table.

	Order of Settings		
User 1-12	Same Size	Correlated	Uncorrelated
User 13-24	Uncorrelated	Same Size	Correlated
User 25-33	Correlated	Uncorrelated	Same Size

Table 3.3: Latins Square for Randomizing the Order of Settings in Study 1

3.2.7 Discussion

In Experiment 1, we are examining how accurately the human haptic system can evaluate the weight of the object they see every day. We expected to find an acceptable weight range for each object. Now, the question is why there is an acceptable range and not only one certain weight? The answer would be in these options: (1) People seems to have inaccurate weight expectation because they cannot remember the weight of the objects accurately, (2) People cannot perceive the weight difference less than certain amount, and (3) People have seen each of these three objects with different weights and material (e.g. hardcover and paper cover notebook).

We have already examined option 1 in experiment 1, the *Prior* setting. In experiment 2 we want to investigate Option 2.

3.3 User Study 2: Human Weight Difference Perception

According to Weber's law, "the perceived change in stimuli is proportional to the initial stimuli". Weber defined Just Noticeable Difference(JND)as the least differences which is noticeable for humans and showed that JND of weight is proportional to the basic weight[70]. Weber fraction is equal to the division of JND and its corresponding basic weight. Research showed Weber fraction varies between 0.02 to 0.13 for lifted weight, when the fraction of the weight differences over the base weight is between 0.16 and 1 ($\frac{\Delta W}{W} \in [0.16, 1] \rightarrow Weberfraction \in [0.02, 0.13]$)[36], [67]. JND highly varies for different individuals and depends on many factors. The work of Raj et al. [57] has shown that Weber fraction depend on how the weights are lifted. In their experiments participants measure weights in three different ways:(1)did not lift weights, (2)lifted weights by flexing the metacarpophalangeal joint, or (3) lifted the weight by flexing the elbow joint. The mean of Weber fraction for the base weight of 100g is 25.71%(SD=7.54), 12%(SD=1.76), and 27.43%(SD=7.62) respectively.

In the second study, we wanted investigate whether the differences between options in Experiment 1 were noticeable for majority of the users. In other words, we wanted to find a range for Weber fraction in our specific Virtual Environment (VE) and see if there is a difference between Weber fractions when the comparison weights are heavier than the basic weight comparing to the case that they are lighter. For this matter, we asked the participants to evaluate weights of two cubes relative to each other. They were provided with a standard cube and a comparison cube. We asked them to compare the weights and determine whether the comparison cube is heavier, lighter or has the same weight as the standard cube. Users were allowed to lift cubes as many times as they wanted, just like user study 1. The whole experiment repeated with 3 different standard cubes. The standard cubes' weights were equal to the replicas' in Experiment 1. For each of the standard cubes, we considered 6 different trials with different weights for the comparison cube (18 trials in total). These cubes were presented to the users one at a time.

3.3.1 Task and participants

In VE designed for this study, we asked users to compare the weights of two cubes. Because of some complains about neck pain after the experiment, we gave users a chance to choose either they want to seat down or stand up during the experiment. To make sure that the results for each of the three standard cubes are independent, we announced to the users whenever we changed any cube; So they knew that their responses for the previous standard cubes are not related to the new standard cube. The participants are the same as the user study one.

3.3.2 Order Effect

In order to prevent the order effect to influence our results, we represented the trials to users in different orders. We had to consider two possibly influential orders, three standard cubes' order (α, β, and γ) and six comparison cubes'(A, B, C, D, E, and F). We have considered four Latin Squares, one of order 3 for the standard cubes and three of order 6 for the comparison cubes.

If we used only one Latin square for the comparison cubes, one would end up using the same order for representing comparison cubes of all standard cubes. In other words, one would repeat the same order of weight conditions

		Users											
		1-6		7-12		13-18		19-24		25-30		31-33	
		Std	Cmp	Std	Cmp	Std	Cmp	Std	Cmp	Std	Cmp	Std	Cmp
Trial	1	Cα	Bα	Cβ	Aβ	Cγ	Cγ	Cα	Eα	Cβ	Dβ	Cγ	Fγ
	2		Aα		Cβ		Eγ		Dα		Fβ		Bγ
	3		Cα		Eβ		Dγ		Fα		Bβ		Aγ
	4		Eα		Dβ		Fγ		Bα		Aβ		Cγ
	5		Dα		Fβ		Bγ		Aα		Cβ		Eγ
	6		Fα		Bβ		Aγ		Cα		Eβ		Dγ
	7	Cγ	Aγ	Cα	Cα	Cβ	Bβ	Cγ	Dγ	Cα	Fα	Cβ	Eβ
	8		Cγ		Eα		Aβ		Fγ		Bα		Dβ
	9		Eγ		Dα		Cβ		Bγ		Aα		Fβ
	10		Dγ		Fα		Eβ		Aγ		Cα		Bβ
	11		Fγ		Bα		Dβ		Cγ		Eα		Aβ
	12		Bγ		Aα		Fβ		Eγ		Dα		Cβ
	13	Cβ	Cβ	Cγ	Bγ	Cα	Aα	Cβ	Fβ	Cγ	Eγ	Cα	Dα
	14		Eβ		Aγ		Cα		Bβ		Dγ		Fα
	15		Dβ		Cγ		Eα		Aβ		Fγ		Bα
	16		Fβ		Eγ		Dα		Cβ		Bγ		Aα
	17		Bβ		Dγ		Fα		Eβ		Aγ		Cα
	18		Aβ		Fγ		Bα		Dβ		Cγ		Eα

Table 3.4: Latin Square for User Study 2

(Heavier, lighter, equal) three times for each user. This could lead them toward finding the representation pattern, whether it is the correct pattern or not, and they would bias themselves based on that pattern. As a result, we considered a separate Latin square for the order of comparison cubes for each standard cubes.

In the next step, we combined the four Latin squares. For this, we first repeated the order 3 Latin square 2 times in a way that users 1 to 6 and 18 to 24 saw the same order for standard cubes. So, we reached a 3×6 table. We needed to present 6 comparison cubes for each standard cube, or each cell of the table; So, we could consider our table as a 18×6 table for comparison cubes. In order to randomise the order of comparison cubes in this table, we used three different Latin squares of order 6 and put them together.

The final Latin Square is shown in Table 3.4. As it is mentioned, α, β, and γ are representing the three possible weights for the standard cube, shown by Std. These weights are equal to the weight of the replica of an Apple, a notepad, and a wireless mouse respectively. The characters A, B, C, D, E, and F are denoting the six possible weight conditions for the comparison cube, shown as Cmp. Two of the six possible comparison cubes are lighter than the standard one, two of them are equal, and two are heavier. Table 3.1 shows the weight assigned to all conditions of the standard and comparison cubes.

Chapter 4

Study of Human Weight Expectation Using Proxy Haptic: Experimental Results

As we mentioned before, the goal of this work is to investigate the human expectation about the weight of some everyday objects. We wanted to observe how accurate the expectation is and how it changes under different circumstances. In the previous chapter, we explained the design of our two user studies in detail. The results of the experiments will be presented in this chapter.

4.1 Human Weight Expectation

In user study 1, we asked users about their expectation of weight of three everyday objects: an apple, a notepad, and a wireless mouse. We provided them with 5 options, 2 heavier than the replicas, 2 lighter than them and one with the same weight as the replicas. The results for each object are as follows:

4.1.1 Apple

Before looking at the users' responses for the apple under the four situations, we need to make sure that the responses are actually different.

4.1.1.1 Statistical Tests

Let's considered *Same* as our baseline, since no visual clue or prior knowledge is provided in this experiment. We want to compare each of the other 3 experiment with this baseline. In order to check the significance of differences between the results of *Same* and the other experiment, we applied a Wilcoxon signed-rank test. The p-Values for the pairs *Same* vs *Correlated*, *Same* vs *Uncorrelated*, and *Same* vs *Prior* are respectively 0.7400, 0.6069, and 0.0005. We considered 0.01 as our threshold for type I error (α). Since there are 3 different null hypotheses we used Bonferroni's correction method. Therefore, the null hypothesis can only be accepted if the P-Value is less than $\frac{\alpha}{3}$ which is equal to 0.003. So there are strong evidences that the differences between *Same* and *Prior* are significant but there are not enough evidence to show *Correlated* and *Uncorrelated* are different from *Same*.

4.1.1.2 Results

Figure 4.1 shows the percentage frequency of participants' choices for each of the 5 options on the table under 4 different situations. The options has 70%, 85%, 100%, 115%. and 130% of the replicas' weights.

On the top left diagram, *Same*, about one third, 33%, of the participants chose the weight of the replica as the acceptable weight; about 40% accepted the options with weight difference of 15% comparing to replica, and less than 30% chose the heaviest or the lightest options. In this case, the diagram is symmetric where the percentage frequencies are approximately the same for the same weight differences (either heavier or lighter) and the diagram has a shape similar to a normal distribution. So one can conclude that the users' expectation about the apple is not biased. In other words, users expected the apple to have approximately the same weight as our replica.

In the diagram *Prior*, we observe that giving users a chance to examine the replica cannot help them choose more accurate options. On the contrary, the prior knowledge leads 50% of the participants to accept the lightest option.

As it is mentioned before, there are no significant differences between the

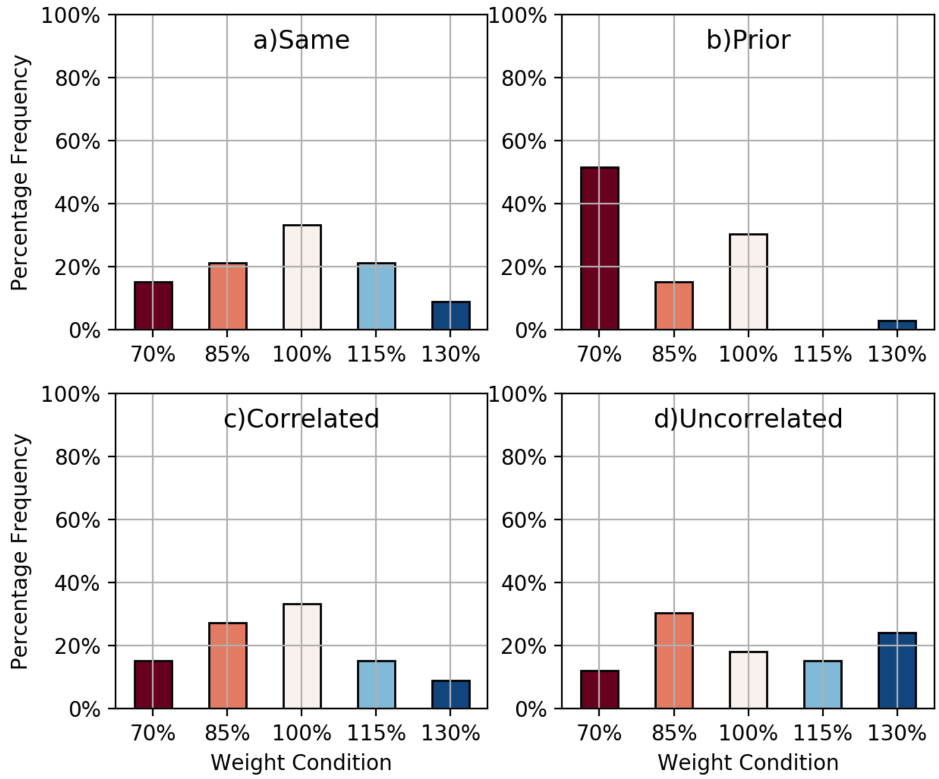

Figure 4.1: Result of User Study1 -Apple

setting *Same* and the two setting *Correlated* and *Uncorrelated*. That means in this case, visual clues did not influence users' choices significantly.

4.1.2 Notepad

Just like the previous section, we considered the setting *Same* as our baseline and compare all other settings with it. we will first present the result of our investigations about the differences using paired t-test.

4.1.2.1 Statistical Tests

Applying Wilcoxon signed-rank test on our three pairs: *Same* vs *Correlated*, *Same* vs *Uncorrelated*, and *Same* vs *Prior* the p-Values are equal to 0.0011, 0.0230, and 0.0480 respectively. Considering Bonferroni's correction method, we consider p-Values less than 0.003 as strong evidence that the two data sets are different. The results are showing that the differences between Same and Correlated are significant, but there is not enough evidence to show that the results for Same is different from Prior or *Uncorrelated*. Looking at the users choices may help us understand these results.

4.1.2.2 Results

Figure 4.2 shows the percentage frequency of the participant's choices in different situations when a notepad is presented to them. The top left diagram, *Same*, shows that participants were prone to choose lighter options. To the extent that about half of them chose the lightest option. This means that the expectation they have about the weight of a notepad in that size is less than our replica's.

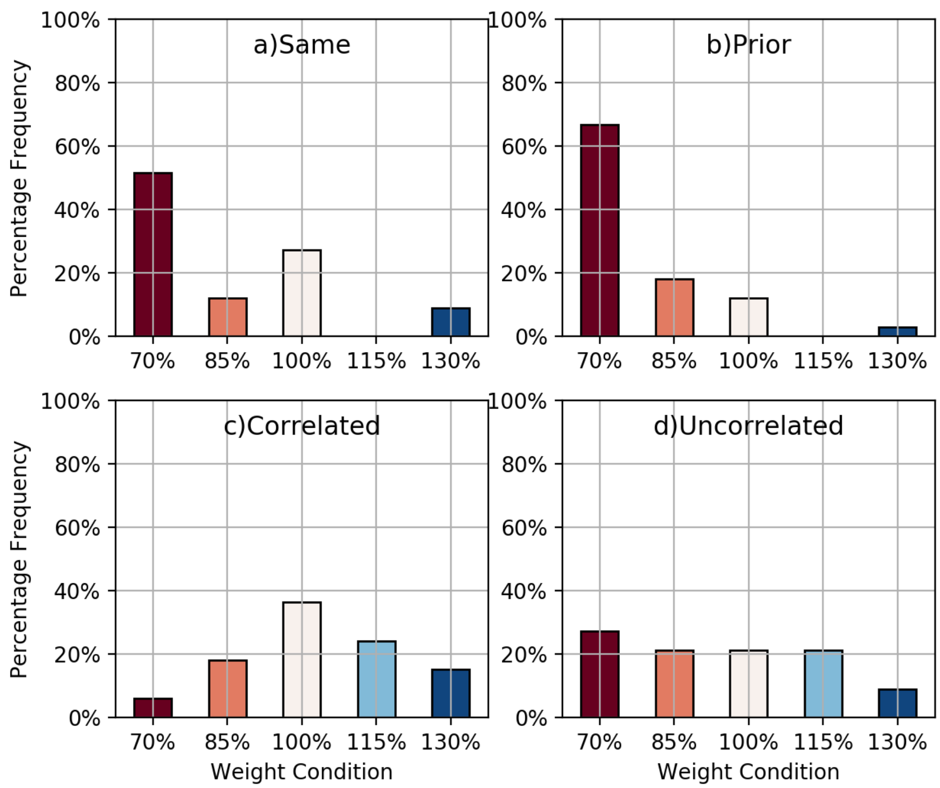

Figure 4.2: Results of User Study1- Notepad

Providing the users with the replica to examine just lead them further toward choosing lighter options. In this setting, shown in the top right diagram, about 70% of users accepted the lightest option as an acceptable weight for a notepad. We believe that the differences between the settings *Same* and *Prior* are not significant because the *Same* setting is already saturated with choices of the lightest option that moving further toward it cannot make considerable differences in data.

In Figure 4.2 diagram c, bottom left diagram, we can see the distribution of choices in the *Correlated* setting. Comparing this diagram with the one for *Same*, we see that visual clue leads people toward choosing heavier options. That is because weight-size illusion persuades people to perceive bigger object lighter. So in the *Correlated* setting, users were still looking for the lighter options, but as the bigger cubes were perceived lighter for them, they chose bigger options which happened to be heavier.

As mentioned before, one cannot say that the results of *Uncorrelated* setting are significantly different. So we can conclude that for *Notepad*, although deceiving visual clue can change users' answers, as one can see the diagrams are different but the changes are not considerable.

4.1.3 Wireless Mouse

Like the previous objects, we will first present the results of the statistical tests and then we will present the distributions of users choices when a wireless mouse is presented for them on the scale.

4.1.3.1 Statistical Tests

The setting *Same* is considered as the baseline and we will compare the results of all other three settings with it. The P-value of Wilcoxon signed-rank test applied on *Same* vs *Correlated*, *Same* vs *Uncorrelated*, and *Same* vs *Prior* the p-Values are equal to $2.5967e^{-5}$, 0.0156, and 0.0020 respectively. Using Bonferroni's correction method, we consider p-Values less than 0.003 showing that there is strong evidence showing that two data sets are significantly different. So *Same* is significantly different from *Correlated* and *prior*, but the differences between *Same* and *Uncorrelated* is not significant.

4.1.3.2 Results Analysis

Figure 4.3 shows how users respond when we showed them a wireless mouse on the scale. The diagram for *Same*, top left diagram, shows us that about 75% of users expected the wireless mouse to be lighter than the replica and about 60% of them chose the lightest option. Talking to users we found out the reason: we were using an old wireless mouse as a replica while they usually work with newer ones which are lighter.

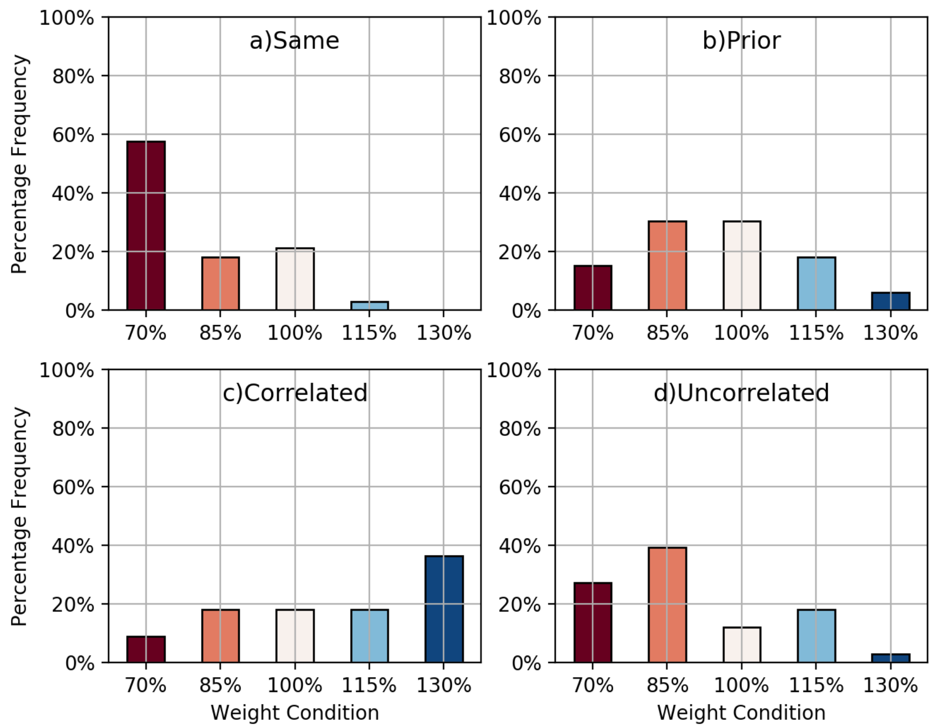

Figure 4.3: Results of User Study1- Mouse

Providing users with the replica and allowing them to examine its weight showed them that the replica is heavier than what they had in mind. As a result, in diagram *Prior*, top right, we can see that users are moved into choosing heavier choices and the distribution is more like the normal distribution.

The *Correlated* diagram shows that the size-weight illusion convinced users to consider the heavier weight acceptable. This illusion makes the bigger objects perceived lighter. The bigger boxes were actually heavier but the illusion influenced users to the extent that they actually consider them lighter than smaller boxes. As diagram *Same* shows that the users expecting the wireless mouse to be light. So they were looking for lighter objects. As they perceive bigger objects as lighter, they chose them.

The *Uncorrelated* diagram is not significantly different from the *Same*. We can say that the deceiving visual clues have convinced more people to choose heavier boxes but the influences are not quite considerable.

4.1.4 Final Discussion for Experiment 1

A deeper look at the comparison of the results for the baseline, *Same*, and other settings in all 3 objects can bring us to more general conclusions. In what follows we will talk about our conclusions in detail:

4.1.4.1 Correlated

We have observed that the size-weight illusion can influence users' choices, however, the differences are not always significant. But why the pattern and level of these influences are different for each of the three objects?

In order to find an answer to this question, we took a closer look at the differences between the users' choices in *Same* and *Correlated*. We wanted to know: What percentage of the users actually changed their minds and chose a different option and what is the differences between these choices? Figure 4.4 shows percentage frequency of the differences between *Same* and *Correlated* calculated for each user individually ($choice_{Corr} - choice_{Same}$).

The first diagram, *Apple*, is approximately symmetric around 0. In other words, the number of people whose choice were $x\%$ lighter in their second choice is roughly equal to the number of people whose choice were $x\%$ heavier ($x \in \{15, 30, 45\}$). This makes the differences between *Same* and *Correlated* insignificant. We believe this happens because in *Correlated* users wanted to repeat the choice they made in *Same*.

The second diagram, *Notepad*, shows that about 64% of participants chose a heavier option in *Correlated*. This shows that the difference between *Same* and *Correlated* caused by the majority of the participants not just one group

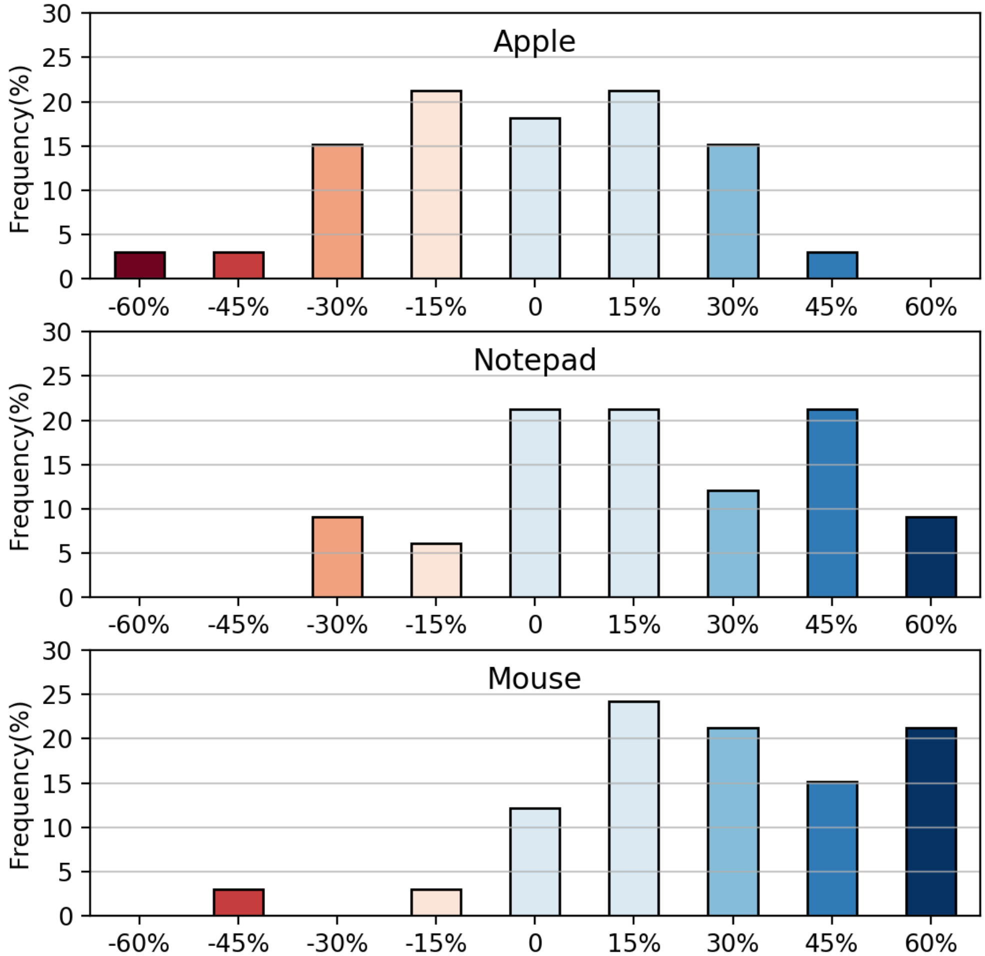

Figure 4.4: Percentage frequency of weight differences between *Same* and *Correlated*

of them. This is because the majority of the users expected the notepad to be lighter than our replica. So in *Correlated*, they want to choose something lighter than what they chose in *Same* and size-weight illusion deceives them to believe bigger options are lighter.

The third diagram, *Mouse*, shows that about 82% of participants chose a heavier option in *Correlated* and only 2 of them, 6%,chose a lighter option. Majority of users expected a wireless mouse to be significantly lighter than our replica so they get affected by size-weight illusion more intensely.

4.1.4.2 Prior Knowledge

Initially, we expected the evaluations to become more accurate by giving users a chance to examine the replicas but it was not the case. Looking at the Prior diagram of the 3 objects, we find different patterns. For *Apple* and *Notepad*, users move toward lighter options after examining the replicas but the differences are significant for *Apple* while it is not for *Notepad*. On the contrary, for *Mouse*, weighting replica leads users to choose heavier options. What is the reason behind this different behaviours?

Our hypothesis is that people consider their choice in *Same* as a baseline. When they are examining the replica, they are actually comparing its weight

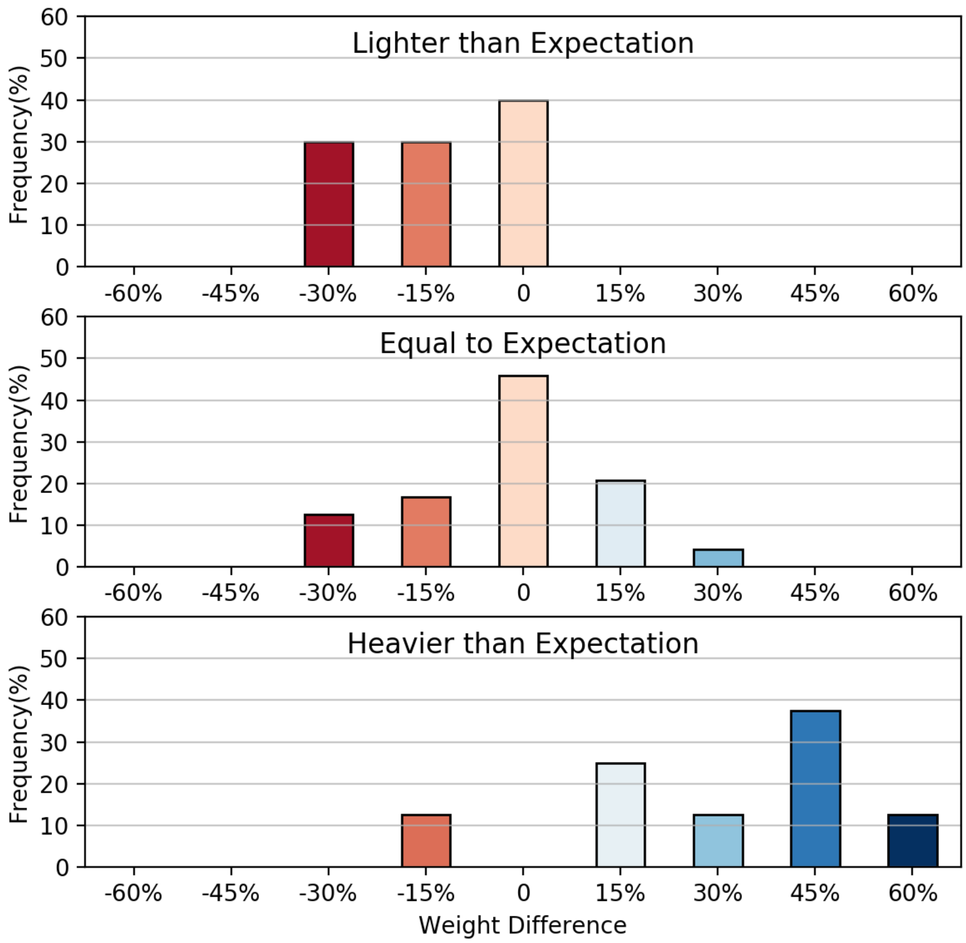

Figure 4.5: Weight Difference between *Same* and Prior Based on Comparison Between Initial Expectation and Replica's Weight Evaluation.

with the expectation they had. Then, they change their *Same* choice base on this comparison.

In order to verify this hypothesis, we asked users to compare the weight of the replicas with their initial expectation when they have the replica in their hands. Their evaluation of replica's weight is either lighter than their expectation, heavier than their expectation or equal to it. Since we came up with this hypothesis halfway through the experiments, we only asked 14 users about this comparison.

Figure 4.5 shows the percentage frequency for each possible weight difference between *Same* and Prior (shown as a percentage of the basic weight). Since the trials for objects are independent, we used are 42 data point (14 users and 3 trials each) for the Figure. This helped us to achieve more generalized results.

Based on Figure 4.5, the majority of participants chose lighter object after they thought that the replica was lighter than their expectation (60%), and 87.5% of them chose heavier options when they found replica heavier. In the case that they reported that replica has approximately the same weight as they expected, the diagram has roughly a normal shape around 0 and about 46% of participants repeated the choice they made for *Same*. This is confirming our hypothesis.

4.1.5 Certainty of Weight Expectations

In Phase 2 of User Study 1, we asked users how certain they are about their choice in Phase 1. In the first step, we applied Wilcoxon signed-rank test to check if the differences between the certainty reported for each object are significant: p-Value of comparing certainty of *Apple* vs. *Notepad*, *Apple* vs. *Mouse* and *Mouse* vs. *Notepad* are correspondingly equal 0.2549, 0.4472, 0.037. Assuming 0.01 as our α and using Bonferroni correction method for our 3 null hypothesis, the difference are only significant if their corresponding p-Value is less than 0.003. Therefore, none of the differences are significant. As a result, we aggregated the data points and Figure 4.6 presents all of them together.

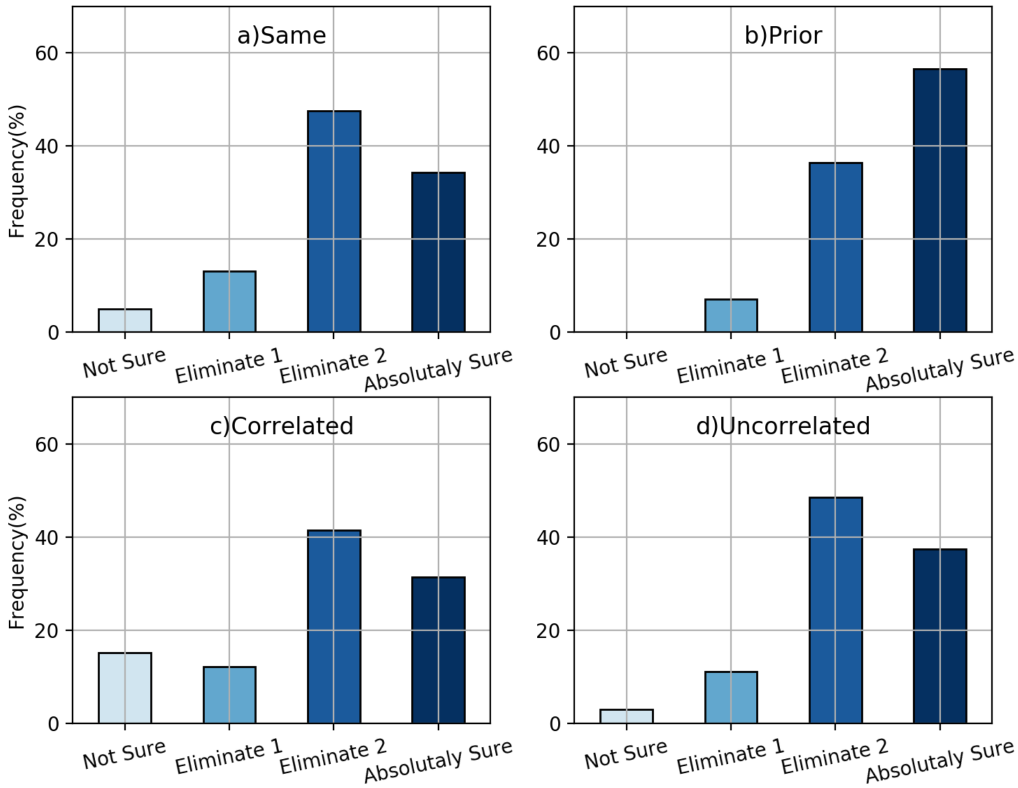

Figure 4.6: Frequency Percentage for level of Certainty

In the second step, we wanted to know if the differences between *Same* and other 3 settings are meaningful. Applying Wilcoxon signed-rank test the p-Values of comparing *Same* with *Prior*, *Correlated*, and *Uncorrelated* are 0.0010, 0.0830, and 0.2859. Just like previous comparison the difference are only significant if their corresponding p-Value is less than 0.003. This shows that either deceiving or correct visual clues cannot make users more certain about their choice. However, giving prior knowledge can do , and at the extent that more than half of the participants became absolutely sure about their choice after examining the replicas. while prior knowledge cannot necessarily make their choice more accurate.

4.2 Human Weight Difference Perception

The question which remained unanswered is whether the differences between the 5 options in User Study 1 is perceivable by the participants. User study 2 is

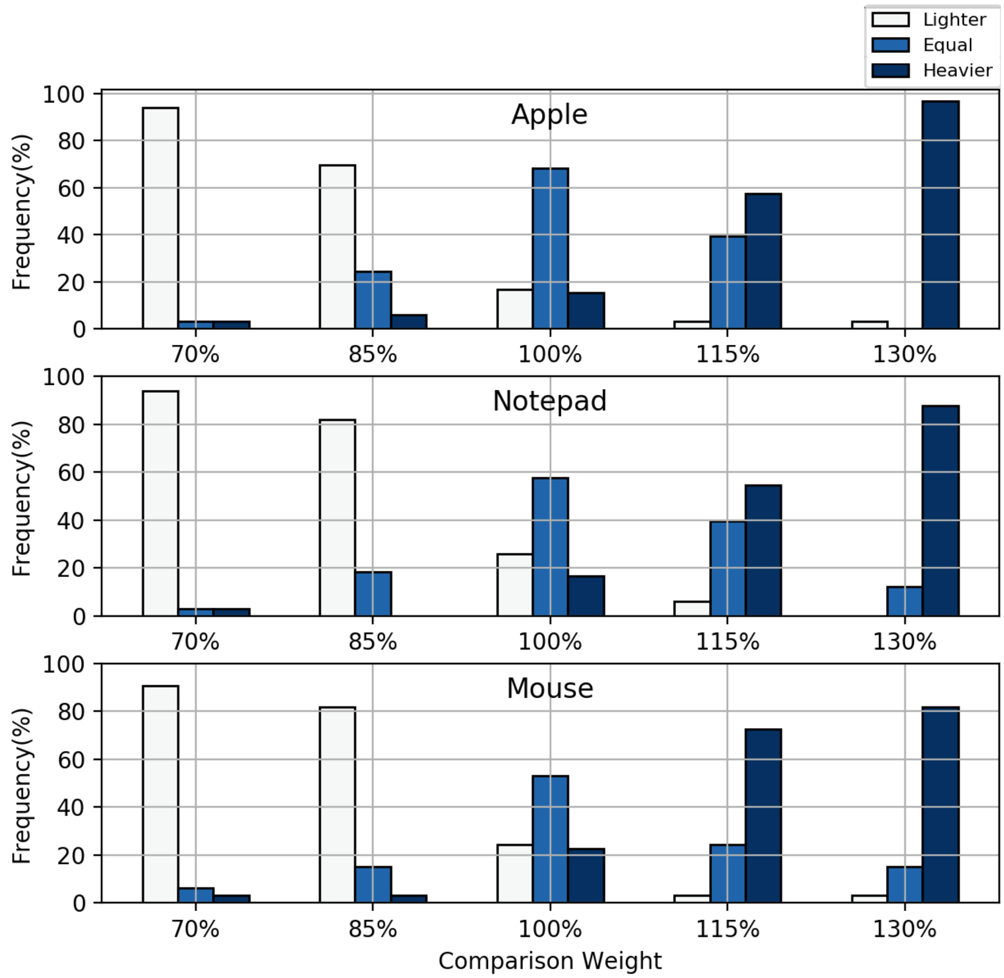

Figure 4.7: Percentage Frequency of the Users' Choices-User Study 2

designed to answer this question by asking participants to compare the weights of 2 cubes, Standard and Comparison, and report whether the Comparison cubes is lighter, heavier, or has the same weight as the Standard cube. Figure 4.7 shows the percentage frequency of the users' choices when the standard cube has the same weight as our replicas for an apple, notepad, or a wireless mouse.

All the 3 diagrams show the same patterns: descending frequencies for reporting comparison as lighter, roughly bell-like shape for equal, and ascending for heavier. In all cases, the accuracy is more the 50% meaning more than half of the participants were able to determine which cube is heavier, or they have equal weights.

In the next step, we applied Wilcoxon signed-rank test to investigate whether the results for 3 objects have meaningful differences. The p-Value for comparing *Apple* vs. *Notepad*, *Apple* vs. *Mouse*, and *Mouse* vs.*Notepad* are equal to 0.1197, 0.6856, and 0.2838 correspondingly. Setting α equal to 0.01 and using Bonferroni correction method on our 3 null hypothesis, we found out that the differences are not significant (all p-Values are greater than 0.003). That is because the differences between the weights of the replicas are not big enough for the human haptic perception to discriminate.

Because of insignificant differences, we averaged the 3 diagrams to reach a more generalised conclusion. Figure 4.8 shows the mean and standard deviation of the results. As it shows, for all Comparison Weight's conditions majority of participants were able to correctly spot the differences; the accuracies are equal to 93%, 78%, 59%, 62%, and 89% when the comparison weight is 70%, 85%, 100%, 115%, and 130% of the standard weight correspondingly.

An interesting point is that the accuracy improves when the comparison Weight is lighter than the standard one. So the Weber fraction is smaller when the comparison weight initiates equal to the standard weight and keeps descending comparing to the case that it is ascending.

4.2.1 Effects of Previous Trials

As mentioned in Chapter 3, we considered 6 weight conditions for comparison; 2 of them were lighter than standard, 2 were equal to it and 2 were heavier than it. The cases of the same weight are shown as condition 3 and 6 in Table 3.4. As the Table shows, Condition 3 always comes after a case of lighter Comparison and Condition 6 always comes after a case of heavier except the cases that they come first.

We wanted to investigate the results of condition 3 and 6 to examine whether the previous trial can be influenced by on users' choice or not. For this, we just looked at the trials that condition 3 and 6 are not the first. As a result, we only had 24 users in this investigation.

Figure 4.9 shows the results separated for each object. In all the diagrams, the frequency of choosing 'Lighter' is increased when heavier comparison is

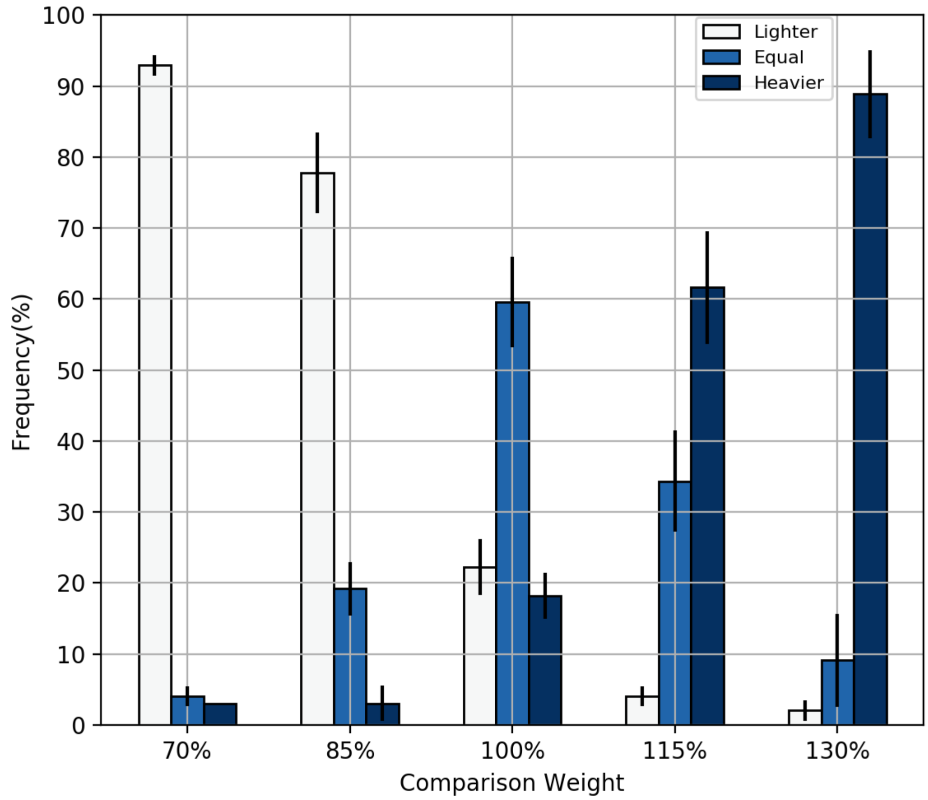

Figure 4.8: Mean and standard deviation of Percentage Frequency of the Users' Choices-User Study 2

provided before the trial. Also, more people chose 'Heavier' when they have examined a lighter comparison beforehand.

The other question is whether the results of 3 objects are meaningfully different from each other. Applying Wilcoxon signed-rank test the p-Values are 0.2163, 0.7166, 0.4678 for *Apple* vs. *Notepad*, *Apple* vs. *Mouse*, and *Mouse* vs. *Notepad* correspondingly. This shows that there are no evidence showing that the results of the 3 objects are different.

In order to reach a more generalised diagram, we aggregated the results of 3 objects. Figure 4.10 shows results after aggregation. The pattern of more choices of 'Heavier' in the case *After Lighter* and more choices of 'Lighter' in the case *After Heavier* is repeated. Generally, previous trials influence users when they are examining the weight of new objects. These influences are based on the differences of objects in the previous and new trial. Lifting

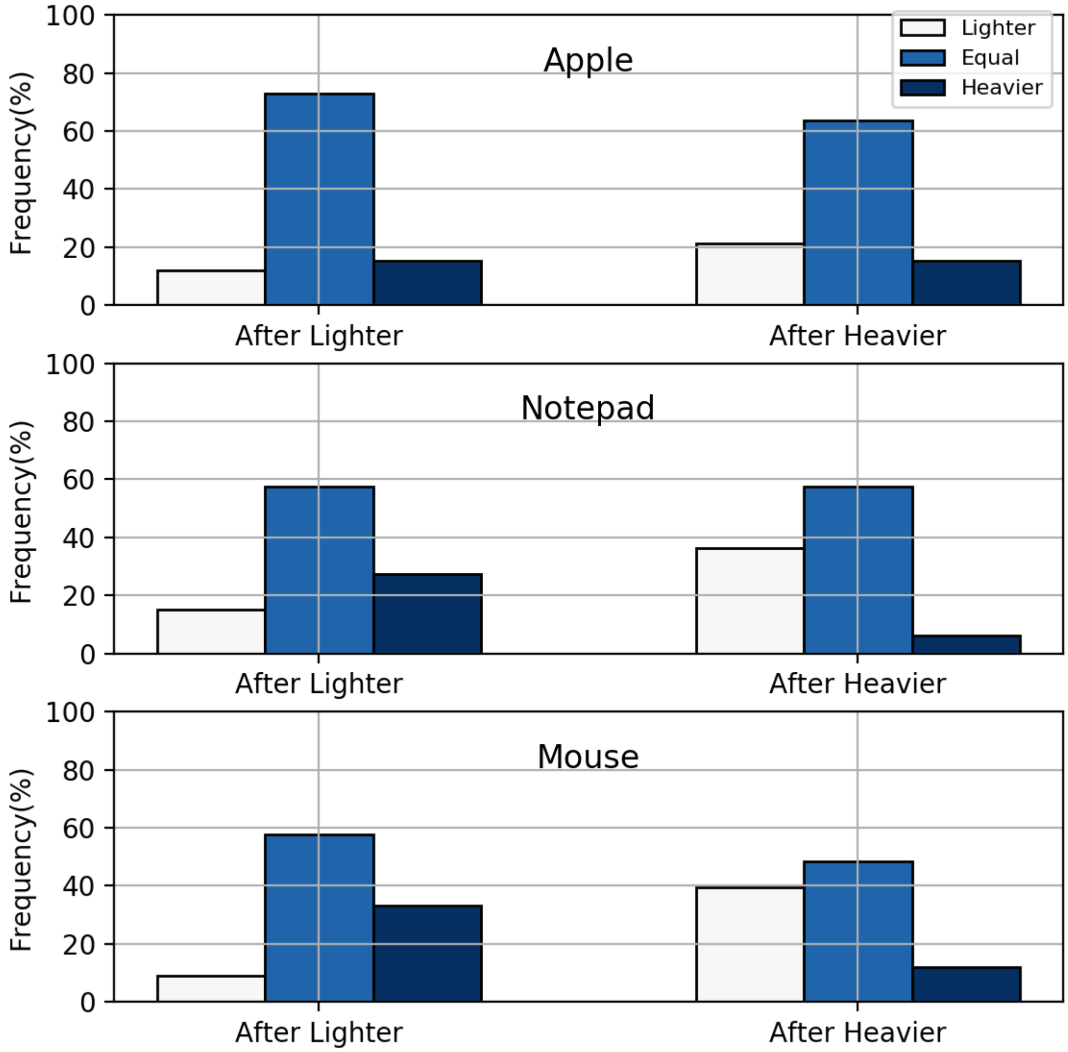

Figure 4.9: Percentage Frequency of users' choices separated for each object when Comparison and Standard have same weight

heavier object in old trials makes the new objects feel lighter for participants and *vice-versa*.

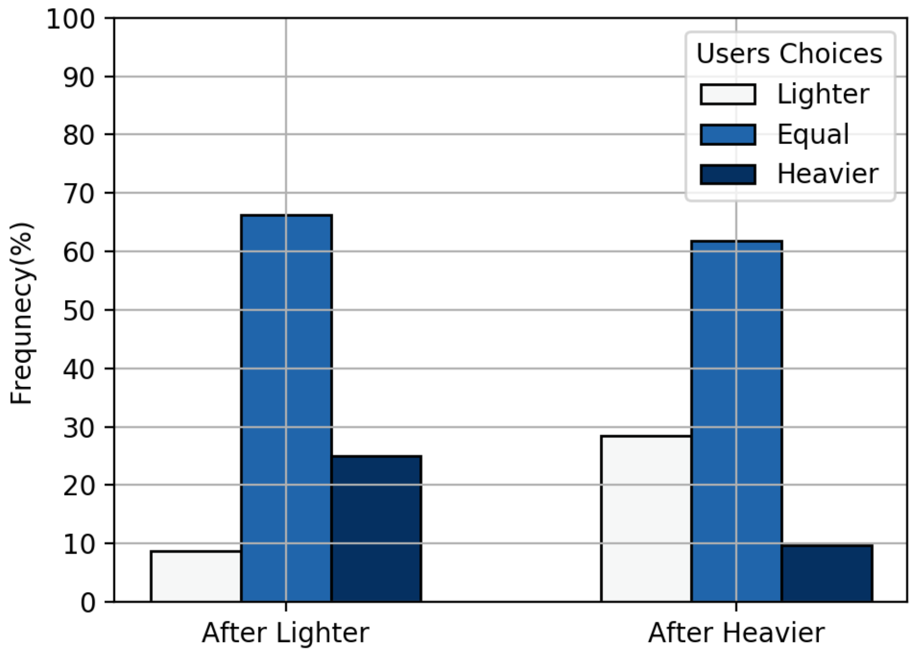

Figure 4.10: Percentage Frequency of users' choices when Comparison and Standard have same weight

Chapter 5

Proxy Haptic for a Pointing Task in the Virtual World

One of the most important advantages of Virtual Reality (VR) is its immersive experience. The intense engagement in VR leads to performance improvement which can be defined as better learning or more user satisfaction. We have designed an experiment to investigate the effects of adding the sense of touch to VR. Since this is an only first step in this path, we simplified the task to a pointing task.

5.1 Design of Experiment

In this experiment, we provided users with a standard computer keyboard and a word that the user needed to type on the keyboard. The experiment was repeated with 3 different words, "BLUE", "CHAIR", and "FAMILY". We wanted to have 3 words with different lengths and also make sure that none of the two consecutive characters in these words are neighbours in the keyboard, so users should aim and reach each character separately. In order to have a baseline to compare the results with, we conducted this experiment twice. First, we asked the participants to type the word using their hands. In this case, the keys were touchable and their hand was detected using Leap Motion controller (No makers were used). Then we repeated the experiment with a the HTC Vive VR controller. We used HTC Vive head-mounted display to present the Virtual Environment (VE) which was created using Unity 5. Figure 5.1

shows the physical environment and the VE.

Figure 5.1: Usability Study Environment: left: Trial with Hand, middle: VE form User's Point of View, Right: Trial with VR Controller

5.2 Measurements

In order to be able to compare the two trials, using hand or controller, we need to define our measurements. The experiment is not designed as a game and also users work with it in a very short time, so we cannot measure users' engagement, excitement or joy as a measure of success.

What we mainly looking to answer: (1) does using bare hands, comparing to controllers, make the task easier and more comfortable for users? (2) does providing haptic feedback make VR more believable? To find answers to these questions, we used some quantitative measures and a questionnaire. In what follows these measures will be explained:

5.2.1 Quantitative Measures

With quantitative measures, we are mainly focusing on measuring how flawless users perform in the task. The less number of errors the easier the task is. The other important factor is the time they consume for typing the words. The measures we collected data for are:

- **Time**
- **Number of Deletes**

- **Accidentally Pressing OK key**

  This one is important because we counted the number of times users pressed keys by mistake but when they accidentally pressed OK, they were automatically transferred to the next level, a new word is presented. So we need to take this factor into consideration.

5.2.2 Questionnaire

It is not possible to observe and measure all the factors during the experiment. Factors like believability of trials are hard to measure by observation. In these cases, we need to ask users about their experience. In our work, we asked the users to answer 27 usability questions, the questionnaire can be found in A. The questions are a mixture of 4 questionnaires each of which will be explained below. We have eliminated 3 questions from the questionnaires, 6 in total for both trials, because we did not want our participants to get bored and answer the questions carelessly.

- **Perceived Competence Scales**

  Self-determination theory (SDT) identifies competence as one of the three innate human needs, if satisfied, leads to optimal function and growth[60]. The theory states that feeling or perception of competence can facilitate goal attainment and make people feel effective in the activity. If users feel competence doing a task in VR that mean the task was doable and they are engaged. Perceived Competence Scales (PCS) is a short 4-item questionnaire designed to assess participant's feeling of competence about doing a task or following through on some commitment. In this work, we used 3 questions of it and we asked the participant to answer them for hand trial and controller trial separately. We considered 7 response options; from "Not at all True" to "Very True".

- **System Usability Scale**

  System Usability Scale (SUS) is an instrument to measure the usability[8]. Originally, it consists of 10 questions. We only used 8 of them and asked the participant to give their answers on a scale of 1 to 7, meaning "Not at all True" to "Very True". The questions were asked once about hand trial and once about the controller.

- **Believability**

  The last factor that we want to ask from participants is about the believability of trials. Unfortunately, we could not find any established instrument to measure it. We have designed 2 questions for this matter. The statements are: "I felt like I am dealing with the real world" and "I felt strange using the system". We asked users to scale their responses from 1 to 7 showing "Not at all True" to "Very True". We asked these question for each of the two trials separately.

- **Preference**

  In the end, we asked users which trial they prefer and why. The explanation can help us understand the score of the 3 aforementioned questionnaires.

5.3 Results

In this section, we will present the results and compare the two trials based on different measures.

5.3.1 Quantitative Measures

- **Time**

  We measured the time consumed to type the words in both trials. Before presenting the data on the plot, we erase the trials in each users' accidentally press OK key because the consumed time in them are shorter but they could not finish typing the words.

  Figure 5.2 shows the average and standard deviation of the time consumed to type the words completely in each of the trials. Users took a longer time to type the word when they were using their hands. That is because the Leap Motion Controller was not always accurate and it sometimes showed the wrong hand gesture or lost the hands completely.

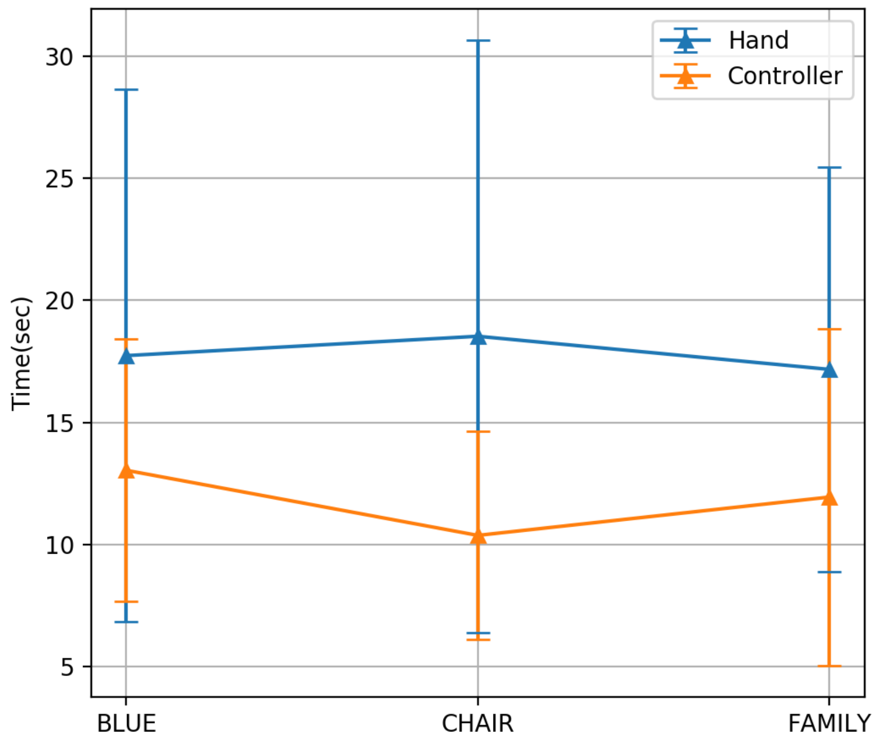

Figure 5.2: Time Consumed for Typing Words

This made it impossible to press keys. Users had to wave their hand in front of the Leap Motion Controllers and gave it some time to restart. While the VR controller was always trustworthy.

- **Number of Deletes**

  The number of times participants pressing delete key is an indicator of how many time they pressed a key by mistake or accuracy of their performance. Figure 5.3 shows the average and standard deviate number of deletes for the hand and the controller trials.

  Based on the diagram, people can type more accurately when they are using their hands comparing to controllers. That is mainly because pressing the trigger button on the controller make the controller shake and that leads to pressing neighbour keys by mistake. While in the hand trial users have more control over the movement of their hand. As mentioned before, the inaccuracy of Leap Motion Controller in detecting hand sometimes cause a problem as well. Using a more accurate sys-

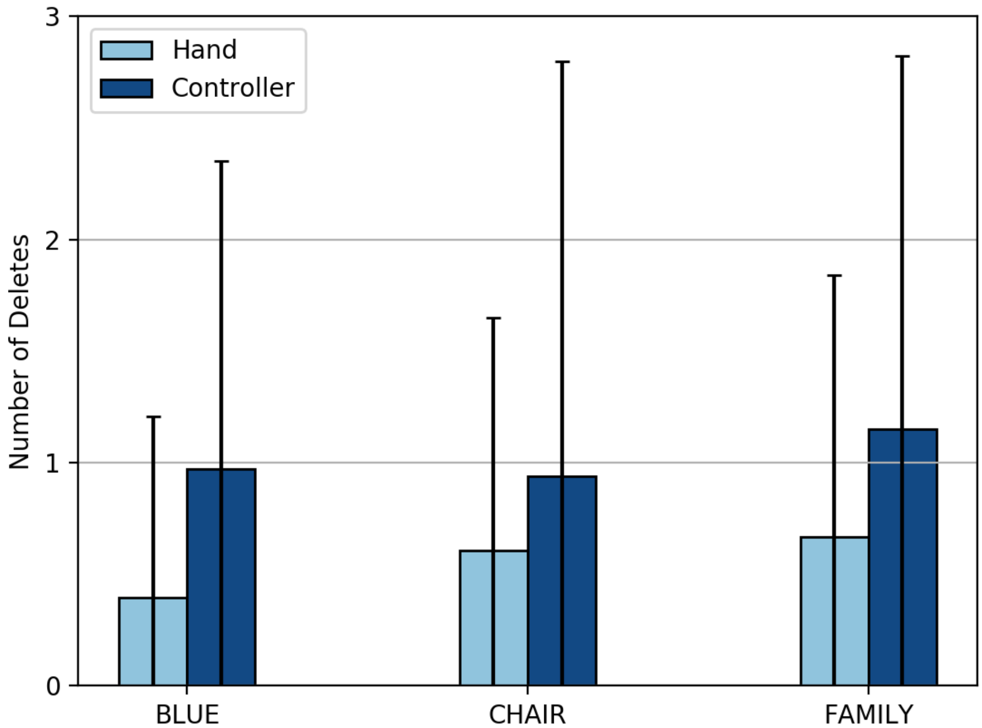

Figure 5.3: Average Number of Deletes in the Hand and the Controller Trials

tem to detect hand gesture and position can help us achieve even more flawless performance in hand trial.

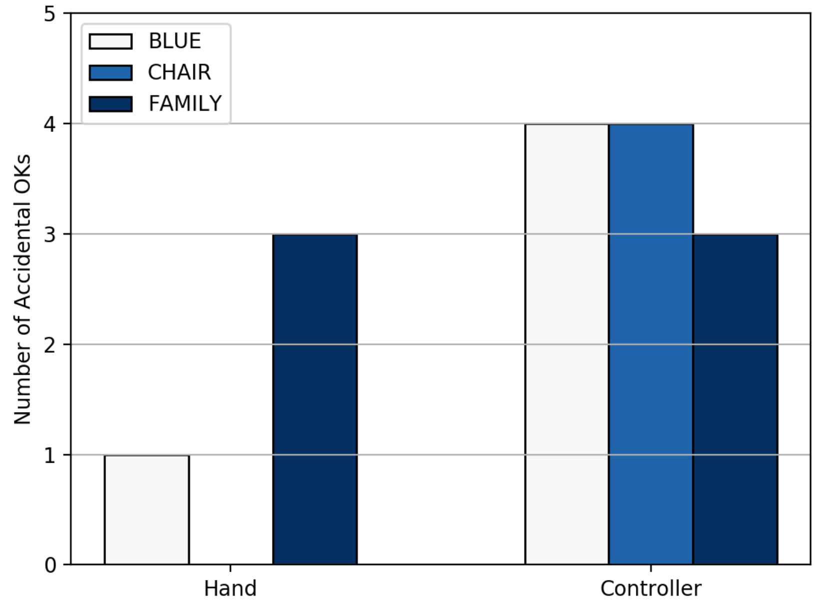

Figure 5.4: Number of Pressing OK by Mistake in each Trial

- **Accidentally Pressing OK**

  Accidentally Pressing OK is another indicator of the accuracy of users' performances. Figure 5.4 shows the number of people who accidentally pressed OK during typing each word.

  Fewer users accidentally pressing OK key in the hand trial. Just like the "Number of Deletes", it is because people have generally better control over their fingers comparing to the VR controllers.

5.3.2 Questionnaire

As mentioned before, we used a mixture 4 different questionnaires, 27 questions in total. 32 participants answered these questions and the result of each will be presented below:

- **Perceived Competence Scales**

  The average PCS score for the hand and the controller trial are 5.41 and 5.80 out of 7 correspondingly. That means users felt more competent and successful when they typed the word with the controllers.

- **System Usability Scale**

As mentioned before, 8 questions were selected from SUS questionnaire, 3 of them are negative questions meaning giving a lower score to them is showing better usability. In order to find the total SUS score, we reversed the scale for the negative question. That means if they chose 1 for a negative question their score is 7. While for positive questions choosing 1 is equal to 1 point. The average SUS scores calculated for the hand and the controller trial are equal to 5.03 and 5.64 out of 7 correspondingly showing better usability for controller trial.

- **Believability**

  We had 2 question for believability, one negative and one positive. We reversed the results of the negative question and get the average of the scores. The Believability score for hand and controller trials are equal to 5 and 4.93 out of 7 correspondingly. This shows that using hands for typing can make the experience more natural for users.

- **Preference**

  In the last question, we asked users which system they preferred to work with and why. 34.38% of the participants chose the hand trial and the other 65.63% preferred VR controller. Their reasons for this choice shed light on the advantages and disadvantages of these two systems. In what follows we will summarise users' feedback:

  - The hand trial is more immersive and natural. Some users felt more control over it because they could do the task by their hand instead of a tool.
  - Inaccurate hand tracking and hand gesture detection made this trial time consuming, cumbersome and hard.
  - The size of the keys on the keyboard were the same for both trials, while the fingers were much bigger than the laser came out of the controller. This made it harder for the users to press keys by their fingers.

- Participants found the controller trial to be faster and easier to use. Some of them expressed that this is because they are used to work VR controller or tools like it.
- The controller trial was more unrealistic and erratic. It was hard for users to avoid pressing the neighbours of a key accidentally.

5.4 Discussion

Although the hand trial was more realistic it failed to improve the system usability and ease of use. That was mainly because of the poor performance of the Leap Motion interface we used in this work.

Providing the passive haptic feedback needs the virtual objects to be exactly aligned with their proxies. In the typing task, presenting the keyboard deep into the physical wall made users push the sponge, keyboard's proxy, too hard which was not pleasant for them. On the other hand, showing the virtual keyboard outside the wall cannot provide any haptic feedback because users cannot touch the sponge anymore. The co-registration was too fragile and we needed to redo it for each user. It is because the Leap Motion controller was not accurate in showing the position of users' hand and also since the sizes of fingers are different for each individual, the system needed to be calibrated for each user.

Better equipment for hand and object tracking can lead to better usability of the hand trial. We tried using Optitrack Trio Camera to tracking hand and object. We used the software Motive to see the results of Optitrack Trio on a computer. Unfortunately, Motive can only detect either the whole skeleton or a rigid body. A hand cannot be considered as a rigid body because the distance between the marks attached to it can change. We also did not want to detect the whole skeleton for just find the position of hands and fingers. That is because detecting a skeleton needs more than just one Optitrack Trio and we wanted to the experiment to be as affordable as possible.

We also tried to track objects in the scene using Optitrack Trio and detect the hands using Leap Motion controller. But having HTC Vive, Optitrack

Trio and Leap Motion Controller together lead to the signals collide and none of the equipment worked properly.

Chapter 6

Conclusion

There are different methods attempting to enrich Virtual Reality (VR) experiences with the sense of touch. Among all, providing passive haptic by means of proxy haptic is the most affordable one. In order to be able to substitute physical replicas with proxies, we need to know how much these two need to resemble.

In this work, we conducted two user studies to investigate how the human haptic system perceives weight in Virtual Environments (VE). We also compare the usability of typing a word in VE using hands vs. using a VR controller.

Our results bring some insights for VR application designers who want to use proxy haptic:

- For objects which appeared in different weights, we need to estimate an acceptable weight for the replicas based on our users and their background.
- Small deviation from the weight of the replica is considered acceptable for the majority of users.
- We can use weight-size illusion to convince users to believe VE.
- The human haptic system works more accurately when it is evaluating lighter weights.
- For objects with weight in the range of 100g to 150g, the majority of

people are able to notice the differences more than or equal to 15% of the base weight.

- As the previous weight evaluations affect users' weight judgment, we need to be careful about the sequence of events representing for users.

Overall, we conducted these experiences under the circumstances in which the users were primarily focused on evaluating weight and they were obliged to choose the best option. This condition is much more restricted comparing to the case that users touch proxies while they are focusing on the purpose of the application (e.g. training). Additionally, choosing one option as the best does not necessarily mean that other options are not acceptable at all. All in all, what we found is how people perceive weights in VE, yet another interesting question is how far we can deviate from users' expectation without harming the sense of presence.

Our usability studies show that using hands for typing words and providing haptic feedback can reduce the number of errors and make the experience more natural. While it can neither make the system easier to use nor improve the perceived competence. Since we have observed users struggling a lot with the hand tracking system (Leap Motion Controller) to detect the hand gesture correctly, we believe this is the reason people did not find the hand trial easy to use. further investigations are needed to see if using better equipment for hand detection is able to improve the usability of the hand trial. Additionally maintaining the delicate alignment between virtual objects and their proxies are cumbersome and time consuming. Utilizing an object tracking system can fix this problem for us.

References

[1] M. Azmandian, M. Hancock, H. Benko, E. Ofek, and A. D. Wilson, "Haptic retargeting: Dynamic repurposing of passive haptics for enhanced virtual reality experiences," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ACM, 2016, pp. 1968–1979. 16

[2] H. Benko, C. Holz, M. Sinclair, and E. Ofek, "Normaltouch and texturetouch: High-fidelity 3d haptic shape rendering on handheld virtual reality controllers," in *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, ACM, 2016, pp. 717–728. 13, 14

[3] V. Bharath and P. Rajashekar, "Virtual manufacturing: A review," in *National Conference Emerging Research Areas Mechanical Enggineering Confrence Proceedings*, vol. 3, 2015, pp. 355–364. 1

[4] J. Blaha and M. Gupta, "Diplopia: A virtual reality game designed to help amblyopics," in *Virtual Reality (VR), 2014 iEEE*, IEEE, 2014, pp. 163–164. 11

[5] F. Bonetti, G. Warnaby, and L. Quinn, "Augmented reality and virtual reality in physical and online retailing: A review, synthesis and research agenda," in *Augmented Reality and Virtual Reality*, Springer, 2018, pp. 119–132. 11

[6] I. Bortone, D. Leonardis, N. Mastronicola, A. Crecchi, L. Bonfiglio, C. Procopio, M. Solazzi, and A. Frisoli, "Wearable haptics and immersive virtual reality rehabilitation training in children with neuromotor impairments," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 7, pp. 1469–1478, 2018. 4, 14

[7] M. Bouzit, G. Burdea, G. Popescu, and R. Boian, "The rutgers master ii-new design force-feedback glove," *IEEE/ASME Transactions on mechatronics*, vol. 7, no. 2, pp. 256–263, 2002. 14, 15

[8] J. Brooke *et al.*, "Sus-a quick and dirty usability scale," *Usability evaluation in industry*, vol. 189, no. 194, pp. 4–7, 1996. 51

[9] B. Brooks, F. Rose, E. Attree, and A. Elliot-Square, "An evaluation of the efficacy of training people with learning disabilities in a virtual environment," *Disability and rehabilitation*, vol. 24, no. 11-12, pp. 622–626, 2002. 11

[10] G. Burdea, P. Richard, and P. Coiffet, "Multimodal virtual reality: Input-output devices, system integration, and human factors," *International Journal of Human-Computer Interaction*, vol. 8, no. 1, pp. 5–24, 1996. 12

[11] R. R. Burke, "Virtual reality for marketing research," in *Innovative Research Methodologies in Management*, Springer, 2018, pp. 63–82. 1

[12] C. Carignan, J. Tang, and S. Roderick, "Development of an exoskeleton haptic interface for virtual task training," in *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*, IEEE, 2009, pp. 3697–3702. 14

[13] M. J. Casuso-Holgado, R. Martın-Valero, A. F. Carazo, E. M. Medrano-Sánchez, M. D. Cortés-Vega, and F. J. Montero-Bancalero, "Effectiveness of virtual reality training for balance and gait rehabilitation in people with multiple sclerosis: A systematic review and meta-analysis," *Clinical rehabilitation*, p. 0 269 215 518 768 084, 2018. 2

[14] L.-P. Cheng, E. Ofek, C. Holz, H. Benko, and A. D. Wilson, "Sparse haptic proxy: Touch feedback in virtual environments using a general passive prop," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ACM, 2017, pp. 3718–3728. 16

[15] I. Choi, H. Culbertson, M. R. Miller, A. Olwal, and S. Follmer, "Grabity: A wearable haptic interface for simulating weight and grasping in virtual reality," in *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, ACM, 2017, pp. 119–130. 14, 15

[16] I. Choi, E. Ofek, H. Benko, M. Sinclair, and C. Holz, "Claw: A multifunctional handheld haptic controller for grasping, touching, and triggering in virtual reality," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ACM, 2018, p. 654. 3, 4

[17] A. J. Dijker, "Why barbie feels heavier than ken: The influence of size-based expectancies and social cues on the illusory perception of weight," *Cognition*, vol. 106, no. 3, pp. 1109–1125, 2008. 18

[18] J. G. Estrada and A. L. Simeone, "Recommender system for physical object substitution in vr," in *Virtual Reality (VR), 2017 IEEE*, IEEE, 2017, pp. 359–360. 16

[19] E. Fassbender, D. Richards, A. Bilgin, W. Thompson, and W. Heiden, "Virschool: The effect of background music and immersive display systems on memory for facts learned in an educational virtual environment," English, *Computers and Education*, vol. 58, no. 1, pp. 490–500, Jan. 2012, ISSN: 0360-1315. DOI: `10.1016/j.compedu.2011.09.002`. 11

[20] K. Fast, T. Gifford, and R. Yancey, "Virtual training for welding," in *Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality*, IEEE Computer Society, 2004, pp. 298–299. 11

[21] M. T. Filigenzi, T. J. Orr, and T. M. Ruff, "Virtual reality for mine safety training," *Applied occupational and environmental hygiene*, vol. 15, no. 6, pp. 465–469, 2000. 11

[22] D. A. Foloppe, P. Richard, T. Yamaguchi, F. Etcharry-Bouyx, and P. Allain, "The potential of virtual reality-based training to enhance the functional autonomy of alzheimer's disease patients in cooking activities: A single case study," *Neuropsychological rehabilitation*, vol. 28, no. 5, pp. 709–733, 2018. 2

[23] V. Frati and D. Prattichizzo, "Using kinect for hand tracking and rendering in wearable haptics," in *World Haptics Conference (WHC), 2011 IEEE*, IEEE, 2011, pp. 317–321. 14

[24] S. A. Freedman, E. Dayan, Y. B. Kimelman, H. Weissman, and R. Eitan, "Early intervention for preventing posttraumatic stress disorder: An internet-based virtual reality treatment," *European journal of psychotraumatology*, vol. 6, no. 1, p. 25 608, 2015. 11

[25] D. Freeman, S. Reeve, A. Robinson, A. Ehlers, D. Clark, B. Spanlang, and M. Slater, "Virtual reality in the assessment, understanding, and treatment of mental health disorders," *Psychological medicine*, vol. 47, no. 14, pp. 2393–2400, 2017. 11

[26] A. Frisoli, F. Rocchi, S. Marcheschi, A. Dettori, F. Salsedo, and M. Bergamasco, "A new force-feedback arm exoskeleton for haptic interaction in virtual environments," in *Eurohaptics Conference, 2005 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2005. World Haptics 2005. First Joint*, IEEE, 2005, pp. 195–201. 14

[27] A. Gallagher, N. McClure, J. McGuigan, I. Crothers, and J. Browning, "Virtual reality training in laparoscopic surgery: A preliminary assessment of minimally invasive surgical trainer virtual reality (mist vr)," *Endoscopy*, vol. 31, no. 04, pp. 310–313, 1999. 2

[28] R. M. Gordon, D. McGonigle, *et al.*, *Virtual Simulation in Nursing Education*. Springer Publishing Company, 2018. 11

[29] X. Gu, Y. Zhang, W. Sun, Y. Bian, D. Zhou, and P. O. Kristensson, "Dexmo: An inexpensive and lightweight mechanical exoskeleton for motion capture and force feedback in vr," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ACM, 2016, pp. 1991–1995. 14, 15

[30] A. Gupta and M. K. O'Malley, "Design of a haptic arm exoskeleton for training and rehabilitation," *IEEE/ASME Transactions on mechatronics*, vol. 11, no. 3, pp. 280–289, 2006. 14

[31] E. Heineken and F. P. Schulte, "Seeing size and feeling weight: The size-weight illusion in natural and virtual reality," *Human factors*, vol. 49, no. 1, pp. 136–144, 2007. 27

[32] S. J. Henderson and S. Feiner, "Opportunistic controls: Leveraging natural affordances as tangible user interfaces for augmented reality," in *Proceedings of the 2008 ACM symposium on Virtual reality software and technology*, ACM, 2008, pp. 211–218. 16

[33] A. Hettiarachchi and D. Wigdor, "Annexing reality: Enabling opportunistic use of everyday objects as tangible proxies in augmented reality," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ACM, 2016, pp. 1957–1967. 16

[34] K. Hinckley, R. Pausch, J. C. Goble, and N. F. Kassell, "Passive real-world interface props for neurosurgical visualization," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, 1994, pp. 452–458. 3, 15

[35] H. G. Hoffman, "Physically touching virtual objects using tactile augmentation enhances the realism of virtual environments," in *Virtual Reality Annual International Symposium, 1998. Proceedings., IEEE 1998*, IEEE, 1998, pp. 59–63. 12

[36] A. Holway and C. Pratt, "The weber-ratio for insensitive discrimination," *Psychol Rev*, vol. 43, pp. 322–340, 1963. 18, 30

[37] B. E. Insko, M. Meehan, M. Whitton, and F. Brooks, "Passive haptics significantly enhances virtual environments," PhD thesis, University of North Carolina at Chapel Hill, 2001. 12

[38] S. Kawai, P. H. Faust, C. L. MacKenzie, *et al.*, "Computer graphic and phantom haptic displays: Powerful tools to understand how humans perceive heaviness," in *Haptics Rendering and Applications*, InTech, 2012. 18, 19

[39] S. Kawai, F. Henigman, C. L. MacKenzie, A. B. Kuang, and P. H. Faust, "A reexamination of the size–weight illusion induced by visual size cues," *Experimental Brain Research*, vol. 179, no. 3, pp. 443–456, 2007. 27

[40] V. K. Ketoma, P. Schäfer, and G. Meixner, "Development and evaluation of a virtual reality grocery shopping application using a multi-kinect walking-in-place approach," in *International Conference on Intelligent Human Systems Integration*, Springer, 2018, pp. 368–374. 11

[41] R. P. Khurshid, N. T. Fitter, E. A. Fedalei, and K. J. Kuchenbecker, "Effects of grip-force, contact, and acceleration feedback on a teleoperated pick-and-place task," *IEEE transactions on haptics*, vol. 10, no. 1, pp. 40–53, 2017. 12

[42] K. E. Laver, S. George, S. Thomas, J. E. Deutsch, and M. Crotty, "Virtual reality for stroke rehabilitation," *Cochrane database of systematic reviews*, no. 2, 2015. 11

[43] V. Levesque, L. Oram, K. MacLean, A. Cockburn, N. D. Marchuk, D. Johnson, J. E. Colgate, and M. A. Peshkin, “Enhancing physicality in touch interaction with programmable friction,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2011, pp. 2481–2490. 12

[44] J.-H. T. Lin, “Fear in virtual reality (vr): Fear elements, coping reactions, immediate and next-day fright responses toward a survival horror zombie virtual reality game,” *Computers in Human Behavior*, vol. 72, pp. 350–361, 2017. 11

[45] R. W. Lindeman, J. L. Sibert, and J. K. Hahn, “Hand-held windows: Towards effective 2d interaction in immersive virtual environments,” in *Virtual Reality, 1999. Proceedings., IEEE*, IEEE, 1999, pp. 205–212. 16

[46] P. Lopes, S. You, L.-P. Cheng, S. Marwecki, and P. Baudisch, “Providing haptics to walls & heavy objects in virtual reality by means of electrical muscle stimulation,” in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ACM, 2017, pp. 1471–1482. 4, 14

[47] A. H. Mason, M. A. Walji, E. J. Lee, and C. L. MacKenzie, “Reaching movements to augmented and graphic objects in virtual environments,” in *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, 2001, pp. 426–433. 7

[48] D. Meldrum, A. Glennon, S. Herdman, D. Murray, and R. McConn-Walsh, “Virtual reality rehabilitation of balance: Assessment of the usability of the nintendo wii® fit plus,” *Disability and rehabilitation: assistive technology*, vol. 7, no. 3, pp. 205–210, 2012. 11

[49] P. Milgram and F. Kishino, “A taxonomy of mixed reality visual displays,” *IEICE TRANSACTIONS on Information and Systems*, vol. 77, no. 12, pp. 1321–1329, 1994. 10

[50] K. Minamizawa, S. Fukamachi, H. Kajimoto, N. Kawakami, and S. Tachi, “Gravity grabber: Wearable haptic display to present virtual mass sensation,” in *ACM SIGGRAPH 2007 emerging technologies*, ACM, 2007, p. 8. 14

[51] C. D. Murray, E. Patchick, S. Pettifer, F. Caillette, and T. Howard, “Immersive virtual reality as a rehabilitative technology for phantom limb experience: A protocol,” *CyberPsychology & Behavior*, vol. 9, no. 2, pp. 167–170, 2006. 11

[52] K. Nakagaki, L. Vink, J. Counts, D. Windham, D. Leithinger, S. Follmer, and H. Ishii, “Materiable: Rendering dynamic material properties in response to direct physical touch with shape changing interfaces,” in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ACM, 2016, pp. 2764–2772. 12, 13

[53] M. M. North, S. M. North, and J. R. Coble, "Virtual reality therapy: An effective treatment for the fear of public speaking," *International Journal of Virtual Reality (IJVR)*, vol. 3, no. 3, pp. 1–6, 2015. 11

[54] C. Pacchierotti, G. Salvietti, I. Hussain, L. Meli, and D. Prattichizzo, "The hring: A wearable haptic device to avoid occlusions in hand tracking," in *Haptics Symposium (HAPTICS), 2016 IEEE*, IEEE, 2016, pp. 134–139. 14

[55] F. Persiani, L. Piancastelli, and A. Liverani, "Virtual flight simulator," *Diem Alma Mater*, pp. 07–17, 1997. 11

[56] Y. Pulijala, M. Ma, M. Pears, D. Peebles, and A. Ayoub, "An innovative virtual reality training tool for orthognathic surgery," *International journal of oral and maxillofacial surgery*, vol. 47, no. 9, pp. 1199–1205, 2018. 11

[57] D. V. Raj, K. Ingty, and M. Devanandan, "Weight appreciation in the hand in normal subjects and in patients with leprous neuropathy," *Brain*, vol. 108, no. 1, pp. 95–102, 1985. 18, 30

[58] D. Rand, R. Kizony, and P. L. Weiss, "Virtual reality rehabilitation for all: Vivid gx versus sony playstation ii eyetoy," in *5th Intl. Conf. On Disability, Virtual Environments and Assoc. Technologies*, 2004, pp. 87–94. 11

[59] A. " Rizzo and G. J. Kim, "A swot analysis of the field of virtual reality rehabilitation and therapy," *Presence: Teleoperators & Virtual Environments*, vol. 14, no. 2, pp. 119–146, 2005. 11

[60] R. M. Ryan and E. L. Deci, "Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being.," *American psychologist*, vol. 55, no. 1, p. 68, 2000. 51

[61] S. Scheggi, L. Meli, C. Pacchierotti, and D. Prattichizzo, "Touch the virtual reality: Using the leap motion controller for hand tracking and wearable tactile devices for immersive haptic rendering," in *ACM SIGGRAPH 2015 Posters*, ACM, 2015, p. 31. 4

[62] N. E. Seymour, A. G. Gallagher, S. A. Roman, M. K. O'brien, V. K. Bansal, D. K. Andersen, and R. M. Satava, "Virtual reality training improves operating room performance: Results of a randomized, double-blinded study," *Annals of surgery*, vol. 236, no. 4, p. 458, 2002. 2

[63] L. Shapiro and G. C. Stockman, "Computer vision. 2001," *ed: Prentice Hall*, 2001. 10

[64] S. Shema-Shiratzky, M. Brozgol, P. Cornejo-Thumm, K. Geva-Dayan, M. Rotstein, Y. Leitner, J. M. Hausdorff, and A. Mirelman, "Virtual reality training to enhance behavior and cognitive function among children with attention-deficit/hyperactivity disorder: Brief report," *Developmental neurorehabilitation*, pp. 1–6, 2018. 2

[65] A. L. Simeone, E. Velloso, and H. Gellersen, "Substitutional reality: Using the physical environment to design virtual reality experiences," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ACM, 2015, pp. 3307–3316. 16

[66] E. Strasnick, C. Holz, E. Ofek, M. Sinclair, and H. Benko, "Haptic links: Bimanual haptics for virtual reality using variable stiffness actuation," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ACM, 2018, p. 644. 3, 13, 14

[67] R. Teghtsoonian, "On the exponents in stevens' law and the constant in ekman's law.," 1971. 18, 30

[68] L. TOKARCHUK, D. Gabana, E. HANNON, H. GUNES, *et al.*, "Effects of valence and arousal on working memory performance in virtual reality gaming," 2018. 2

[69] A. W. de Vries, G. Faber, I. Jonkers, J. H. Van Dieen, and S. M. Verschueren, "Virtual reality balance training for elderly: Similar skiing games elicit different challenges in balance training," *Gait & posture*, vol. 59, pp. 111–116, 2018. 2

[70] E. Weber, "The sense of touch (he ross, ed. & trans.)," *London: Academic.(Original work published 1834)*, 1978. 18, 29

[71] E. Whitmire, H. Benko, C. Holz, E. Ofek, and M. Sinclair, "Haptic revolver: Touch, shear, texture, and shape rendering on a reconfigurable virtual reality controller," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ACM, 2018, p. 86. 3, 13, 14

[72] V. Yem, R. Okazaki, and H. Kajimoto, "Fingar: Combination of electrical and mechanical stimulation for high-fidelity tactile presentation," in *ACM SIGGRAPH 2016 Emerging Technologies*, ACM, 2016, p. 7. 14

[73] C.-q. Yu, H.-h. Ju, and Y. Gao, "3d virtual reality simulator for planetary rover operation and testing," in *Virtual Environments, Human-Computer Interfaces and Measurements Systems, 2009. VECIMS'09. IEEE International Conference on*, IEEE, 2009, pp. 101–106. 1

[74] A. Zenner and A. Krüger, "Shifty: A weight-shifting dynamic passive haptic proxy to enhance object perception in virtual reality," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 4, pp. 1285–1294, 2017. 13, 14

Appendix A

Usability Test: Questionnaire

Hand & touch

Please answer the question 1 to 13 about your experience with typing using your own hand:

Perceived Competence Scales

	Not at all True			Some-What True			Very True
1. I feel confident in my ability to type words correctly.	1	2	3	4	5	6	7
2.I am capable of typing the words in VR using my own hands.	1	2	3	4	5	6	7
3.I feel able to meet the challenge of typing with no error in VR using my own hands.	1	2	3	4	5	6	7

System Usability Scale

	Not at all True			Some-What True			Very True
4. I think that I would like to use this system frequently	1	2	3	4	5	6	7
5. I found the system unnecessarily complex	1	2	3	4	5	6	7
6. I thought the system was easy to use	1	2	3	4	5	6	7
7. I found the various functions in this system were well integrated	1	2	3	4	5	6	7
8. I would imagine that most people would learn to use this system very quickly	1	2	3	4	5	6	7
9. I found the system very cumbersome to use	1	2	3	4	5	6	7
10. I felt very confident using the system	1	2	3	4	5	6	7
11. I needed to learn a lot of things before I could get going with this system	1	2	3	4	5	6	7

Believability

	Not at all True			Some-What True			Very True
12. I felt like I am dealing with the real world	1	2	3	4	5	6	7
13. I felt strange using the system.	1	2	3	4	5	6	7

VR Controller:

Please answer the question 14 to 26 about your experience with typing using VR controller:

Perceived Competence Scales

	Not at all True			Some-What True			Very True
14. I feel confident in my ability to type words correctly.	1	2	3	4	5	6	7
15.I am capable of typing the words in VR using VR controllers.	1	2	3	4	5	6	7
16.I feel able to meet the challenge of typing with no error in VR using VR controllers.	1	2	3	4	5	6	7

System Usability Scale

	Not at all True			Some-What True			Very True
17. I think that I would like to use this system frequently	1	2	3	4	5	6	7
18. I found the system unnecessarily complex	1	2	3	4	5	6	7
19. I thought the system was easy to use	1	2	3	4	5	6	7
20. I found the various functions in this system were well integrated	1	2	3	4	5	6	7
21. I would imagine that most people would learn to use this system very quickly	1	2	3	4	5	6	7
22. I found the system very cumbersome	1	2	3	4	5	6	7

to use							
23. I felt very confident using the system	1	2	3	4	5	6	7
24. I needed to learn a lot of things before I could get going with this system	1	2	3	4	5	6	7

Believability

	Not at all True			Some-What True			Very True
25. I felt like I am dealing with the real world	1	2	3	4	5	6	7
26. I felt strange using the system.	1	2	3	4	5	6	7

Preference

27. Which setting do you prefer (Using hands vs using controllers)? why?