Simulating a Virtual Drilling Task Using Audio, Visual, and Mouse Movement Cues During the COVID-19 Pandemic

By

Guoxuan Ning

A thesis submitted to the
School of Graduate and Postdoctoral Studies in partial
fulfillment of the requirements for the degree of

Master of Science in Computer Science

The Faculty of Science
University of Ontario Institute of Technology (Ontario Tech University)
Oshawa, Ontario, Canada
April 2021

THESIS EXAMINATION INFORMATION

Submitted by: **Your Name**

Master of Science in **Computer Science**

Thesis title: Simulating a Virtual Drilling Task Using Audio, Visual, and Mouse Movement Cues During the COVID-19 Pandemic

An oral defense of this thesis took place on June 18, 2021, in front of the following examining committee:

Examining Committee:

Chair of Examining Committee	Dr. Patrick Hung
Research Supervisor	Dr. Bill Kapralos
Research Co-supervisor	Dr. Alvaro Quevedo
Examining Committee Member	Dr. Adam Dubrowski
Thesis Examiner	Dr. Khalil El-Khatib

The above committee determined that the thesis is acceptable in form and content and that a satisfactory knowledge of the field covered by the thesis was demonstrated by the candidate during an oral examination. A signed copy of the Certificate of Approval is available from the School of Graduate and Postdoctoral Studies.

Abstract

This work examined whether a real-world drilling task can be successfully simulated using combinations of appropriate audio, visual cues, and basic haptic stimuli obtained with mouse movements instead of a haptic device. Two experiments were conducted where participants were asked to accomplish a virtual drilling task (drilling a block of virtual wood) under different visual, auditory, and “haptic” conditions using some commonly available computer devices such as a mouse, headphones, and monitor. The results of these experiments indicate that audio, visual, and basic haptic cues can be used to simulate a drilling process without a haptic device. Although greater work remains, this work has shown that skills requiring haptic feedback (e.g., drilling) can be simulated without haptic devices. This is particularly important when considering remote learning where trainees may be able to practice various psychomotor-based skills at home with commonly available computer hardware and devices.

Keywords: Audio; Haptic; Multimodal Interaction; Human-Computer Interaction; Virtual Simulation

AUTHOR'S DECLARATION

I hereby declare that this thesis consists of original work of which I have authored. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize the University of Ontario Institute of Technology (Ontario Tech University) to lend this thesis to other institutions or individuals for the purpose of scholarly research. I further authorize University of Ontario Institute of Technology (Ontario Tech University) to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research. I understand that my thesis will be made electronically available to the public.

The research work in this thesis that was performed in compliance with the regulations of Research Ethics Board/Animal Care Committee under **#16280.**

Guoxuan Ning

Statement of Contributions

Part of the work described in Chapter 3 has been published as:

G. Ning, Q. Daggett, A. Perivolaris, B. Kapralos, A. Quevedo, KC Collins K.Kanev, and A.Dubrowski. Rethinking audio-haptic perceptual immersion from in-person to remote testing during COVID-19. In Proceedings of the International Conference on Interactive Mobile Communication, Technologies and Learning (IMCL 2021), November 4-5, 2021.

G. Ning, B, Kapralos, A. U. Quevedo, K. Collins, K. Kanev, and A. Dubrowski. Examining the perception of drilling depth Using auditory cues. In Proceedings of the IEEE Signature Conference on Computers, Software, and Applications, July 12-16, 2021, virtual conference.

I performed the majority of the testing of experimental and the writing of the manuscript.

Acknowledgements

I would like to express my gratitude to my supervisors Dr. Bill Kapralos and Dr. Alvaro Joffre Uribe Quevedo for their help, guidance, and belief in my ability to achieve my goals. Their feedback, suggestions, and encouragement were critical for me in navigating a tremendous amount of information in the field of computer science, especially during the COVID-19 pandemic. I am also grateful to my family and friends encouraged me to complete this unforgettable journey. I would also like to thank Brianna Grant for helping me with the statistical analysis. Finally, I would like to acknowledge Ontario Tech University for being the right environment to focus and excel in this field.

Contents

List of Figures

Chapter 1

Introduction

Perception is the process of extracting meaning from a multitude of sensory signals (the “senses”) that are unpredictable. The five traditional senses, vision, hearing, touch, smell, and taste [1] [2], provide multimodal feedback to help achieve complex tasks [3]. Perception relies on the connection between the sensory organs and the brain. Most recently, it was discovered that pain, temperature, and balance [4], are additional senses that react and send signals to various parts of the brain, thus analyzed and retained as memory [5]. For example, the five traditional senses are used simultaneously when tasting food. When eating an apple, taste is used to process the flavor of the apple; sight identifies the shape and color of the apple; smell identifies the apple’s odor when biting it; the resulting sounds identify if it is crunchy, while oral touch (the oral haptic stimulation) indicates whether the apple is soft or crispy. The taste of food may change if any of the senses were deprived [6]. Basic perceptions provide humans the ability to sense their environment and to develop necessary survival skills.

Virtual reality is an interactive computer simulation, which senses the user’s state and operation and replaces or augments sensory feedback information to one or more senses in a way that the user obtains a sense of being immersed in the simulation (virtual environment) [7]. Compared with traditional video and audio cues presented in a 2D display, VR can enhance the user’s experience by providing a realistic representation of the real world [8]. VR simulation has primarily been used in entertainment including games [9], although its potential goes far beyond that. Virtual reality simulations are becoming widely embraced in education (and medical education in particular) for teaching or training purposes [10]. Prior work that has shown that virtual reality can improve the efficiency of education and training, and more specifically medical education and training [11].

Due to advances made in computer graphics, software, and hardware technology over the past decade, virtual reality technologies/devices have become more affordable and available at the consumer level. This in turn has allowed them to be incorporated and accepted by the medical industry throughout a variety of areas, including medical education [12]. For example, traditional medical education has employed the apprenticeship model to train surgeons, which follows the concept of “watch, learn and then do” [13]. However, this approach is becoming less acceptable due to patient safety concerns, particularly when considering invasive processes that require high-risk treatment [9]. Therefore, VR-based simulation can provide an effective educational tool for medical training in a safe and cost-effective manner. By recreating clinical experiences, VR-based simulation allows trainees to engage in interactive and immersive activities within virtual reality environments, preventing any associated risks to patients [14]. For instance, VR technology can be used to provide a dentist with a simulated practice scene as a low-cost and safe method; dentists can complete training in cutting and shaping teeth using a VR-based simulation [15]. Another example can be seen in the work of Vapenstad et al., [16] where a VR environment with a surgical simulation and a cylinder-like device with force feedback, the Xitact ITP (Instrument Tracking Port), was used to simulate the laparoscopic surgery procedures. Twenty surgeons with extensive surgical experience participated in an experiment that examined the realism of the virtual simulation. As part of the experiment, the participants were required to accomplish a laparoscopic surgical task using the virtual simulator, and after the task, they were asked to rate the fidelity of the experimental scene. Experimental results showed that the virtual simulation environment was realistic by simulating the visual, auditory, and haptic senses of the surgical procedure [16].

To improve the effect of VR environments (VREs), it becomes necessary to replicate the perceptions that result from the human senses. However, the majority of VREs have focused on replicating the visual, auditory, and to a lesser degree, touch senses, while ignoring taste and smell altogether since they are difficult to simulate due to the technology limitations [9]. In order to simulate the human senses, various VR devices are needed including head mounted displays (HMDs) that provide the user who wears the HMD stereoscopic 3D visuals, headphones worn by the user to output sounds that simulate real-world spatial hearing, and haptic feedback devices.

Haptic devices provide touch-based feedback that is perceived by applied tactile or kinesthetic stimuli to sense and manipulate objects with a user input device [17]. As Melaisi et al. [18] describe, at the consumer-level, haptic feedback is commonly available via actuators that offer basic vibrotactile feedback (e.g., video game “rumble packs” and silent vibrate setting in mobile phones). For instance, the Nintendo Switch is a game console that can provide vibration feedback to users while they play games [19]. Efforts in providing consumer-level force-feedback have led to devices such as the Novint Falcon, a haptic device that can provide force feedback with 3 degrees of freedom [20]. However, even these consumer-level haptic devices are generally restrictive and cannot provide the higher level of fidelity and the range of motion required to realistically simulate many tasks, especially for medical training simulations [21] [18]. For example, in a dental implant simulation [22], the fidelity of the Novint Falcon was not enough to simulate a dental drill in the dental surgery given limitations to associated with the movement (degrees-of-freedom) and force feedback. Therefore, users had to use a more advanced and high-precision haptic feedback device, such as the Omega 6, which is manufactured by Force Dimension, and providing force feedback in six degrees of freedom (DOF); this allows for a highly realistic surgical simulation [22]. Although the Omega 6 haptic device provides higher fidelity haptic feedback than a consumer-level haptic device such as the Novint Falcon, the Omega 6 haptic device is much more expensive (approximately $27,300 USD). Considering whether the fidelity of haptic feedback devices is proportional to their price, a survey of existing haptic devices currently available was conducted (see Table 1.1 for a price comparison of several haptic devices). As shown in Table 1.1, haptic devices that provide high fidelity, such as the Omega series, also have a much higher price, and they are not available at the consumer-level. Consumer-level haptic feedback devices such as the Geomagic 3D Touch generally do not offer high fidelity and can also be limited with respect to degrees of freedom (DOF) and force feedback.

Multimodal Interactions

As described earlier, the human senses relay information from the external world to the brain, then causing the body to respond [4] [5]. For example, people use various senses when walking on a road, such as vision to identify the position of the road, hearing to localize the sound of approaching vehicles, and haptic feedback between the muscles and tendons to help them control their direction and speed. Similarly, a high-quality virtual reality environment requires multimodal interaction of the senses. Kapralos et al. found that multiple sensory cues (e.g., visual and auditory) can influence each other within a virtual environment (VE), and more specifically, sounds may improve the fidelity of the visual scene within a virtual environment [23]. Therefore, the multimodal interaction effect may increase the realism of the VREs.

Table 1.1: Comparison of various haptic device currently available

Haptic Device	DOF	Workspaces	Force (N)	Position Resolution	Price (US dollars)
3D Touch [a]	6	Volumetric 16 × 12 × 7 cm	3.3	451 dpi	$2,800
Touch X [b]	6	Volumetric 16 × 12 × 7 cm	7.9	1100 dpi	$12,400
Omega.3 [c]	3	Translational 160 × 110 mm	12	N/A	$21,200
Omega.6 [d]	6	Translational 160 × 110 mm Rotation 240 × 140 × 320 deg	12	N/A	$27,300
Omega.7 [e]	7	Translational 160 × 110 mm Rotation 240 × 140 × 180 deg	12	N/A	$34,400

---

[a] Link: https://www.3dsystems.com/haptics-devices/touch

[b] Link: https://www.3dsystems.com/haptics-devices/geomagic-touch-x

[c] Link: http://www.forcedimension.com/products/omega-3/specifications

[d] Link: http://www.forcedimension.com/products/omega-6/specifications

[e] Link: http://www.forcedimension.com/products/omega-7/specifications

1.1 Problem Statement

Given the costs associated with properly and accurately simulating the sense of touch and manual dexterity, along with adequately presenting user interfaces that represent the tools used in training, current virtual reality environments, including virtual simulations, focus primarily on cognitive skills (the ability of human brain to process, store and extract information) development, typically ignoring psychomotor skills (the ability to control the movement of body to perform motor tasks) development altogether.

Simulating sense of touch is difficult due to the limitations of current consumer-level technology to properly represent touch and force feedback [17]. Although consumer-grade haptic devices that simulate the sense of touch are readily available, they are limited in terms of the amount of force, and customization required for educational purposes focused on the development of motor skills, in particular those associated with dexterity. For instance, consumer-level haptic devices have been used in training, such as the laparoscopic surgical training describe previously [16], and results indicated that the level of haptic fidelity was not adequate, especially with respect to force direction freedom, friction and resistance.

Therefore, improving the fidelity perception of consumer-level haptic devices may allow for the development of virtual simulations that focus on psychomotor skills development at a cost-effective level [24]. As mentioned above, the prior studies indicated multiple sensory cues can influence each other within a virtual environment (VE) and increase the overall realism of the VE's perception [23]. This raises a question that is it possible using multimodal interaction to increase the overall perception of the virtual environment, more specifically how does sound increase the perception of haptic fidelity in virtual environments?

Furthermore, the need for this research was heightened by the current COVID-19 pandemic and the need to understand the relevance of sound when using a mouse in a non-immersive VR environment for providing haptic feedback. Different from immersive virtual reality using VR devices, such as 3D glasses, headsets and haptic feedback devices, to generate virtual sensory feedback by occluding the user's vision, hearing and touch [9]; non-immersive VR allows for interacting with the environment through a monitor, mouse or headphones; the platform does not fully occlude the user's field of view, and it does not require very specialized equipment to support it [25]. Can non-immersive VR be used to facilitate the simulation of psychomotor-skills in VR-based training simulations? This is an important question and can have important ramifications particularly when considering large-scale remote learning during the COVID-19 pandemic and the lack of access to specialized equipment housed in research centers and university laboratories.

1.2 Goals and Objectives

This thesis builds upon the prior work of Melaisi et al. [18] by examining multimodal interactions within a virtual environment where a drilling task is simulated. More specifically, the thesis will examine the influence of appropriate sound cues on haptic feedback in VREs through an experiment exploring whether a real-world drilling task can be successfully simulated using combinations of sound, visual, and haptic cues obtained with bidimensional mouse movements. Drilling is an important component of many medical procedures, and the simulation of medical drilling requires expensive haptic devices [15]. If the drilling process can be simulated using the appropriate combination of cues, it may allow for the development of novel affordable training strategies for remote online learners. This is particularly important during the current COVID-19 pandemic, where access to educational institutions and training laboratories (and thus haptic devices) is significantly restricted or denied altogether.

The lockdowns (including the shutdown of educational institutions) given the COVID-19 pandemic are likely to continue at least until the middle of 2021 [26], and the difficulties caused associated with conducting experiments involving human participants and specialty devices that are housed within university research laboratories (labs) [27]. After careful consideration, it was decided that although the availability of a haptic device cannot be guaranteed, the assumption that every participant will have access to a computer mouse and thus very basic haptic cues that are associated with grasping the mouse using a hand and moving the mouse back-and-forth movement, along with auditory and visual feedback (in various combinations) can simulate a drilling task.

1.3 Hypothesis Statement

- **Null hypothesis (H0)**: Visual, auditory, and bidimensional mouse movement cues can be used to simulate a drilling task.

- **H1**: Top view (visual condition) will result in better performance (drilling accuracy) than the scenes which only offering the back view (non-visual condition).

- **H2**: Dynamic contextual auditory cues (where the sound changes in response to the drill moving through the material) will influence task performance (accuracy of drilling depth) more than non-dynamic sounds comprised of i) a continuous yet static drilling sound that does not change as the drill moves through the material, and ii) no sound at all.

- **H3**: bidimensional mouse movements that move the virtual drill in and out of the block of wood will result in greater performance (participants will better judge drilling depth) than the keyboard control.

- **H4**: Visual + dynamic auditory + bidimensional mouse movements will lead to the greatest task performance (participants will judge drilling depth most accurately in the presence of these three cues)

1.4 Thesis Organization

The remainder of this thesis is organized as follows. A literature review of human perception and multimodal interactions in order to develop a better understanding of this topic in Chapter 2. In Chapter 3, the first experiment (a preliminary study) that examined the accuracy of which participants can judge drilling depth using sound cues only is presented. The results of this first experiment and a brief discussion of these results are also presented here. Building upon the results of the first experiment, the second experiment, presented in Chapter 4, examined whether a drilling task (drill drilling through a block of wood) could be simulated in a VR environment using only visual (rendering of a drill), sound (dynamic drilling sound comprised of a recording of a drill drilling through the wood in the real world), and very limited haptic cues in the form of bidimensional mouse movement to control the virtual drill. The results, discussion and conclusion of the experiment are presented in Chapter 5, Chapter 6 and Chapter 7.

Chapter 2

Literature Review

Although humans have many senses, such as sight, hearing, touch, taste, smell, as well as pain, or balance, virtual technologies primarily simulate human vision, hearing, and to a lesser degree, touch [28]. The senses interact with each other, and these interactions influence the performance of their own and other perceptions. For example, the sound associated with the crispness of a potato chip can affect the perceived taste (crunchiness or freshness) [29]. Virtual reality technology has been increasingly adopted for medical training and education with industry- and consumer-level hardware and software [30]. The advances in computer-based simulation is providing safe, engaging, cost-effective, and flexible training environments while avoiding exposure to hazards, risks, and limitations associated with live medical training [14]. Moreover, the quality of a virtual environments depends on its fidelity, which is the degree of precision simulating the real-world using models or simulators [31]. This chapter begins with background information on the human senses, fidelity, and multimodal interactions. Then, research that is relevant to multimodal interactions in a virtual reality environment will be discussed. At the end of this chapter, research on the application of VR technology in a medical training simulation, specifically drilling-based medical surgery, is presented.

2.1 Human Senses

Sight, hearing, smell, touch, and taste are the main sensory channels of human communication with the environment [32]. Although there are other senses beyond these five (e.g., pain, temperature, and balance, amongst others), the five senses are still considered the five most significant [1][2]. However, simulating the sense of taste and smell (and the various other senses such as pain, etc.) in VR-based environments is very difficult, and traditionally, emphasis has been placed on simulating visual, auditory, and to a lesser degree, touch [33]. Therefore, for the remainder of this chapter, the focus is placed on the three of the five primary human senses, and more specifically, vision, auditory, and touch.

2.1.1 Visual Sense

The visual system provides humans the ability to see [34], and this ability allows us to navigate effectively through physical spaces and interact with individuals and objects in our surroundings. The majority of the information presented to humans in their everyday lives is transmitted visually, in the form of text, pictures, videos [35]. Therefore, visual perception has become an effective manner of communicating knowledge, and thus, vision is considered an essential sense in virtual VEs [33].

As a significant visual organ, the eye has a very complex structure. Figure 2.1 shows a schematic cross-section of the human eye (reprinted from [36]). The cornea, which is a transparent front part of the human eyes, can refract light to the pupil (a “black hole” located in the center of the eye) as the first barrier of the eye. Depending on the intensity of the light, the muscles connected to the iris (a thin and annular structure in a mammals’ eyes) will control the contraction of the pupil. For example, the pupil will contract to reduce the amount of light coming in when the light is intense and expand when the light is weak. The light passes to a transparent object called the lens, a curved structure, as it goes through the pupil. The lens serves to focus the light, allowing it to fall precisely on the retina, which is the innermost light-sensitive layer of the eye, used for transmitting the image to the brain through electrical neural impulses. Visual information is then sent to the brain through the optic chiasm, which is the optic nerves cross located at the bottom of the brain [36][37].

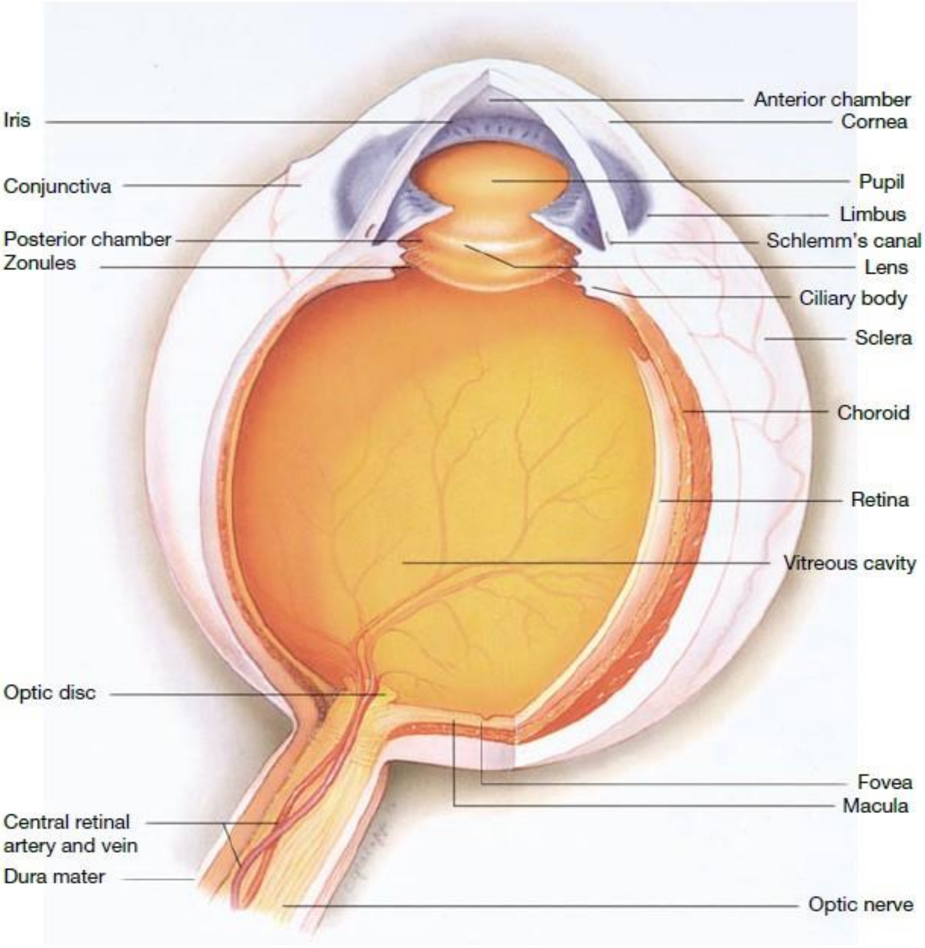


Figure 2.1: A sectional vies of the eye anatomical feature. Reprinted from [36].

When light falls on the retina, photoreceptors in the retina will analyze and give reactions to transmit to the brain after storing the color, brightness, and other information of the light. There are two primary sensory cells in the retina: rods and cones. Cones are more sensitive to daylight, and rods primarily work in dim light [38] [39]. At the same time, color and light details can help to identify the material properties, shadows, transparency, and shapes of the objects/surfaces which reflect the light into the eyes [40]. However, the human eye is more sensitive to changes in brightness than in color, and therefore, it is easier for humans to detect changes in the brightness of the light (the shift between light and dark) than changes in its color [41]. Human eyes can detect light within the visible spectrum, ranging from 400 nm to 76 nm (light within this region is known as visible light). Natural color vision is trichromatic, based on three cone classes with the optimum light sensitivity of approximately 420 nm (blue cones), 530 nm (green cones), and 560 nm (red cones). The absorption of light between the three forms of sensory cone cells makes it possible to detect red, green, and blue individually or in different combinations [39]. Figure 2.2 shows the visible spectrum and corresponding wavelength of color (reprinted from https://www.orcagrowfilm.com/Articles.asp?ID=145).

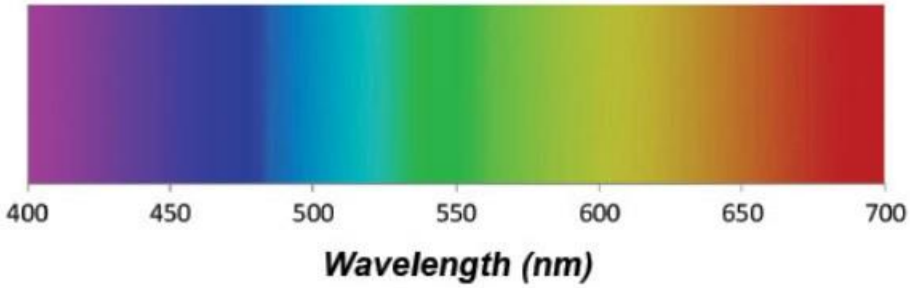


Figure 2.2: Human visible spectrum and corresponding wavelength of color. Reprinted from https://www.orcagrowfilm.com/Articles.asp?ID=145.

The human visual system is more sensitive to the lower frequency of the visible spectrum than the higher frequency portion [42]. For example, if there are spots on a flat surface (low frequency), they can easily be identified, but they are hard to be found if they are on an irregular surface (high frequency). The visual system is also sensitive to moving objects. If anything in the field of vision is moving, even if the observer does not look directly at it, it will be easily noticed [42].

2.1.2 Auditory Sense

The auditory sense of humans is used to receive and provide an understanding of sound-based information. Humans have two ears which are the main organs receiving sounds, one on each side of the head, and given this separation, there will be slight differences with respect to the intensity and time of arrival between the sound arriving at the left and right ear (except when the sound is directly in front, behind, above, or below and when the sound falls within the “cone of confusion,” which is an imaginary cone whose vertex is at the center of the head and extends outward at each ear see) [43]. These subtle differences (binaural effects) are known as the interaural level (or intensity) difference (ILD or IID) and interaural time difference (ITD), respectively, and provide cues to the location of the sound source [44] [45]. Binaural cues (ITD and ILD) are not enough to offer complete spatial localization. For example, binaural cues cannot determine when a sound is directly in front, behind, above, or below a listener or when the sound falls within the cone of confusion).

The ear is a complex structure that plays an essential function in the auditory sense. There are three main components of the human auditory system: i) outer ear, ii) middle ear, and iii) inner ear [46]. The visible portion of the ear is the outer ear, composed of cartilage that folds around the ear canal and is referred to as the pinna. When sound waves reach the pinna, they are reflected and attenuated towards the ear canal of the middle ear. Sounds, particularly those between 2 kHz to 5 kHz are amplified as they travel through the ear canal before reaching the eardrum. The eardrum is a thin and cone-shaped membrane between the outer ear and middle ear. The function of the middle ear is to transfer and transduce the mechanical sound waves to the cochlea, which is a hollow, snail-shaped part. The mechanical waves are then converted to electrical signals. This is accomplished with three tiny bones, the malleus, the incus, and the stapes, which transmit weak vibrations from the eardrum to the oval window and into the cochlea [46]. The cochlea consists of three parts filled with tissue fluid, scala vestibuli, scala media, scala tympani [47]. When the sound signal passes through these areas, it will encounter the organ of corti, which has many hair cells. The movement of the hair cells is ultimately converted into nerve signals and sent to the brain through the auditory nerve [48]. Figure 2.3 provides an illustration of the anatomy of the human ear (reprinted from https://zh.m.wikibooks.org/).

There are three primary auditory sensations in humans: timbre, volume, and pitch [44]. Timbre refers to different sound characteristics in terms of waveform and vibration. For example, musical instruments, like violin and cello, can make different sounds even if they play the same note at the same volume due to their materials and structures [49]. Therefore, timbre can be understood as a feature of the sound [50]. Loudness is an auditory property determined by the intensity of sound, and the relationship between them is positively correlated; moreover, loudness is the perceptual equivalent to intensity. The loudness of a sound is measured in decibels, and the World Health Organization states that prolonged exposure to sound that is greater than 75 dB can cause physical damage [51]. Pitch is the perceptual equivalent to the frequency of a sound. Pitch is also related to the structure of the object that produces the sound. More specifically, humans are more sensitive to sounds with frequencies between 1,000 – 4,000 Hz [52].

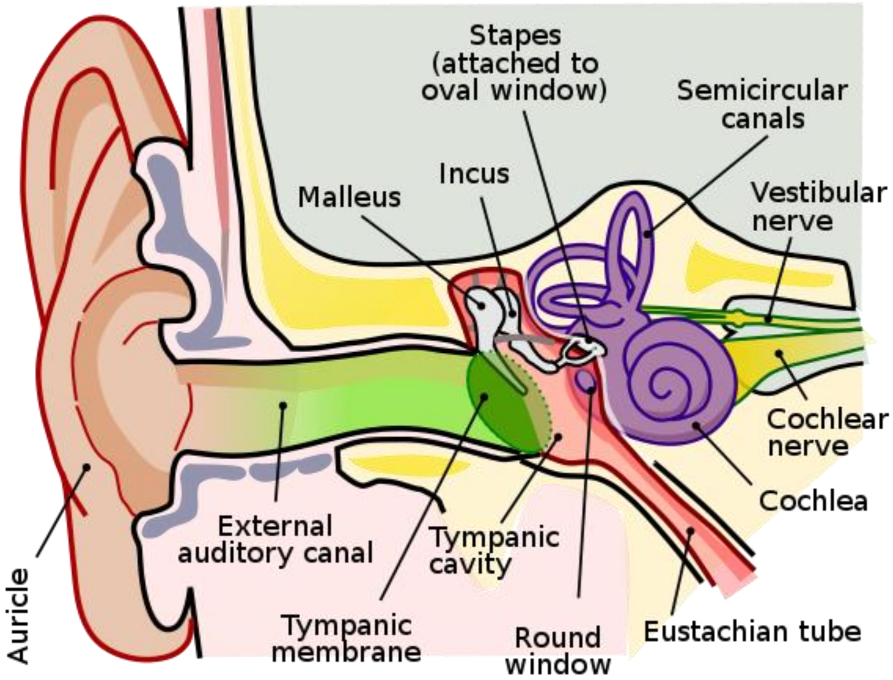


Figure 2.3: Anatomy of human ear (green: outer ear; red: middle ear; purple: inner ear). Reprinted from https://zh.m.wikibooks.org/.

2.1.3 The Sense of Touch

The perception of touch is important for humans to experience their environment. The primary information transmitted by the sense of touch involves the interpretation of a collection of object characteristics, such as shape, size, and surface structure [53]. The sense of touch can be divided into two categories, “tactile sensation” and “kinesthetic sensations” [54]. Tactile sensations are based on receptors in the skin layer to sense vibrations, ambient temperature, object shape and texture, and pain. The kinesthetic sensations are primarily located in the muscles, tendons, and joints of humans and help to feel the body’s strength, motion, and location.

In the deep layers of the skin, tactile corpuscles (essentially a nerve ending) are sensitive to the reaction of slight touch [55]. When a nerve cell senses the pressure of touch, it immediately sends out a tiny electrical signal traveling through the nerve fibers to the brain; then, it leads to the perception of touch. Therefore, the brain is able to determine the touch and the location of the signal [55]. According to research, once people lose their sense of touch, they can no longer feel the surrounding environment [56]. Researchers investigated several patients who lost their sense of touch due to disease and found that these patients are not able to feel their surroundings; in other words, they could only rely on their other senses to make the body respond to the surroundings. For example, in a lighted room, the patients could observe the surrounding environment and controlled their muscles to complete some physical motion (e.g., once they saw a chair, they could move to it and sit on it). However, once the lights in the room were suddenly turned off and the patients were unable to use their vision, they immediately fell since they could not feel the environment and produce the corresponding muscle control. Therefore, the sense of touch is another essential human sense. However, the sense of touch is often overlooked in virtual reality [18] [31] since it is considered less important than visual and auditory feedback.

2.2 Multimodal Interaction and Perception

Multimodal refers to the integration of multiple senses, and multimodal interaction refers to users interacting with a system by using the integration of multiple senses [57]. In everyday life, humans interact with their surroundings, primarily relying on the multimodal interaction of their senses. For example, people typically use both hearing and vision to watch a movie that involves both visual and auditory stimuli [58]. Multimodal interaction is often used in virtual reality environments to simulate real scenes. Due to technical limitations, the current VREs primarily focus on vision, hearing, and touch [59].

The fidelity of the virtual reality environment, which is the degree of precision simulating the real-world using models or simulators, has become an important research direction, given the increasing interest virtual training simulations and the potential issues associated with high fidelity simulations including increased development and computational costs [9]. A high-fidelity virtual reality environment with picture resolution, sound quality and haptic feedback of the scene is necessary to provide a realistic, immersive training environment for the users [14]. The fidelity of a virtual reality environment is determined by the fidelity of the sensory feedback equipment it employs [31].

However, high-fidelity feedback devices are expensive, and the consumer-level equivalents typically provide much less fidelity [18]. Therefore, examining the topic of multimodal interactions and the effect on fidelity perception could determine whether fidelity of virtual environments can be improved by taking advantage of multimodal interactions to offset the low-end feedback provided by consumer-level devices. The interaction of multiple senses is discussed below.

2.2.1 Audio-Visual Feedback

In order to study the interaction between sound and visual factors within a total knee arthroplasty surgical procedure environment, Kapralos et al. [31] studied interaction between sound and visual feedback, on the fidelity of the virtual reality environment. In the experiment, participants were asked to perform a series of actions, such as locating surgical equipment using VR devices (e.g., head-mounted display, headphones, etc.). In each trial, the participants were presented with different resolution visuals and different ambient sounds (such as classical music, heavy metal music, and the sound of an actual operating room, which is including the sound of the machine working and talking between doctors). The result indicated that multiple sensory cues (e.g., visual and auditory) can influence each other within a virtual environment (VE), and more specifically, sounds may improve the perception of fidelity of the visual scene within a virtual environment. In other word, this experiment provided preliminary evidence that multimodal interaction could be a potential way to increase the perception of the user's overall realism of VREs'.

Similarly, exploring the multimodal interaction effects of sound and visual on the overall fidelity of VR environments, and based on the medical training, Cowan et al. also studied the interaction between sound and the visual scene within a medical (operating room) [23]. More specifically, they examined the influence of sound on visual fidelity perception and task performance. This experiment included a surgical procedure scenario, and participants were asked to use the standard arrow keys to control a surgeon to reach a medical device while passing a motionless character and a hospital bed. The participants' time to complete each trial was be recorded. Two variables were cited in the experiment, the visual quality, and the ambient sound. Visual quality was defined with respect to the level of blurring of the visual scene; six levels of blur were considered. Four ambient sound conditions were considered: i) no sound, ii) white noise, iii) typical operating room sounds that included the sounds of the various equipment while they are working and any conversation between doctors and nurses, and iv) hospital operating room ambiance mixed with a drilling sound. There was a total of 24 combinations (trials), and each was randomly presented to the participant three times. Twelve participants participated in the experiment, when the participants completed each trial, they were asked to rate the fidelity of the visual scene on a scale of 1 to 7. Experimental results showed that the quality of the visual scene had significant effects on task performance. However, although the participants completed the task faster in the scene with the operating room ambiance and drilling sounds, there was no evidence to indicate that the sound played a significant role in this experiment. Furthermore, the interaction between sound and visual cues was not significant. Furthermore, sound did affect task performance. For example, based on the participants' feedback, the white noise sound reduced the perceived realism of the visual scene and increased the completion time of the participants. Conversely, the experimental scenes that included relevant sounds were considered more realistic.

Instead of studying the effects of multimodal interaction between sound and visual on the overall perception of VRE realism, Sekuler et al. explored the influence of sound on visual perception [60]. The experiment had two small virtual balls colliding in a VR environment while providing different visual and auditory stimuli to the participants The visual stimuli, t was comprised of three visual conditions during collision: i) continuous motion without stopping, ii) movement paused for one frame, then resumed to move forward without altering its orientation and speed, and iii) movement paused for two frames, then resumed to move forward without changes in orientation and speed. For the auditory stimuli, there were four conditions: no sound at the point of collision; and collision sounds were added during, before, and after the point of contact. Thirteen participants participated in the experiment and in each trial, their task was to determine whether the balls collided. The results indicated that the collisions in the scenes with the visual feedback were more realistic, and that the virtual collision was found to be most realistic when the ball stayed for two frames during the collision. Additionally, in the scenarios where the sound was included, participants believed the collision to be more realistic than in the scenario without sound.

Malpica et al examined whether the existence of colliding sounds in VREs alters the perceived appearance of materials [59]. A virtual scenario for the experiment using the Unity game engine with a head-mounted display (HMD) was developed, where various materials, including metal, fabric, plastic, and phenolic (a kind of yellow transparent organic compound) were introduced, by showing pictures to the participant. twenty-seven participants were shown the properties of these materials used in the experiment, including low-level perceptual characteristics (i.e., soft/hard, glossy/matte, and rough/smooth) and high-level descriptors of appearance (i.e., realistic, metallic-like, plastic-like, fabric-like, and ceramic-like). The study considered two variables: visual and auditory feedback. For the visual effect, a sphere mapped with textures was rendered from different materials. The visual fidelity was represented by the resolution of the rendered material (i.e., higher resolution of the rendered material resulted in a higher fidelity). For auditory feedback, a drumstick allowed participants to strike the surface of a sphere, thus producing sound effects when examining different materials. The participants were asked to identify the four materials in four different combinations: i) material sounds produced by the drumstick striking a material surface, ii) no collision, iii) high-resolution rendering, and iv) low-resolution rendering The participants were asked to rate the perceived appearance (attributes) of the material. The results indicated that there was a significant difference among the participants' visual perception of different materials, but the realistic perception of high-resolution rendering was higher than that of the low-resolution rendering of the materials. Furthermore, for materials rendered at the same resolution, it is better to add the perception of sound and visual effects than to provide visual effects only. In other words, sound can enhance visual perception, which is the recognition of the objects’ material, even just low-resolution rendering offered in this case. Rendering costs could be saved by reducing the resolution while auditory stimuli are added.

2.2.2 Audio-Haptic Feedback

Melaisi et al. [21], explored whether auditory cues affect haptic fidelity perception within a virtual drilling task. The investigation was conducted with eleven participants, requiring them to complete a drilling task where the drill was simulated using the Novint Falcon haptic device under different auditory conditions. The participants were asked to drill through a block of wood to three predetermined depths under four different background sound conditions, more specifically, contextual sound (actual drilling sound) and non-contextual sound (classical music, white noise, and no sound). After completing each trial, participants rated their perception of haptic fidelity. Results indicated that the participants believed the trial with contextual sound condition where the sound corresponded to the task had the highest haptic fidelity perception than other sounds conditions environment. These findings are consistent with the experiment conducted by Moreno et al. [61], where a human brain wave monitor, and an electroencephalogram (EEG) were used to monitor the brain waves of twenty-two participants while they virtually drilled using the Novint Falcon haptic device. The participants were required to complete the drilling task under the same conditions presented in [21] and then, required to rate the haptic feedback fidelity level using a 7-point Likert scale (one being unrealistic and seven being realistic) after each trial. Additionally, the participants completed a usability questionnaire. The results indicated that the scenario involved contextual sound led to the greatest haptic fidelity perception. The EEG showed that the attention of the participants was affected by every sound condition. More specifically, the participants were more concentrated on the drilling task when the sound was played than they were in a soundless environment.

Melaisi et al. [18] studied the effect of sound on haptic feedback by simulating a real drilling task and measuring the drilling depth given by the participants in a virtual drilling simulation. Two consumer-level haptic devices, the Novint Falcon and the Geomagic 3D Touch were used to simulate the resistance/force-feedback provided by a power drill during a drilling task. Fifteen participants performed a virtual drilling task that involved drilling through a predetermined depth into two blocks, one made of wood and the other made of metal. During drilling, one of the following auditory conditions were presented: i) contextual sound, that is the sound of an actual drill going through a block of wood and a block of metal ii) non-contextual sound that was disjoint from the drilling task (e.g., white noise, classic music, and no sound at all). It is worth noting that for the contextual audio, each material triggers different audio recordings that match the corresponding block. the results indicated that there was no significant difference between the contextual and non-contextual sounds, but the accuracy was higher in the presence of sound (contextual or non-contextual) than no sound at all. The results of this experiment do not show any effect of sound on drilling performance, although the researchers conclude that further work is necessary to account for the experimental shortcomings (e.g., lack of dynamic sound stimulus).

Ammi et al. studied the multimodal interactions between haptic and auditory feedback in a shared 2D space with a haptic device (Geomagic 3D Touch) and stereo headphones with a microphone that enabled verbal communication between the participants [62]. In addition to the influence of multimodal interaction between auditory and haptic feedback, the study also focused on understanding the influence efficiency associated with the time to complete the task. Twenty-four participants participated in the experiment, and they were randomly grouped into 12 pairs. A series of 10 different topographical maps were presented to the participants, and the participants were requested to locate and calculate the peaks of various amplitudes, which were some red dots that appear randomly on the map. Two feedback modes, haptic and auditory, were used in the experiment. A haptic feedback arm (Geomagic 3D Touch) provided haptic feedback. When one of the two people in the experimental group used the haptic feedback arm to determine the peak's location (red dots), another person's device would provide force feedback when they reached the same area. The auditory feedback was provided by a stereo headset and worked in the same way as the haptic feedback. When one person determined the peak position, the prompt audio feedback would be provided to the other person when they reached the same area. The experiment was presented to the participants under two different conditions. Condition one included haptic feedback, and condition two offered both haptic and auditory feedback. The pairs were requested to complete the same task but visually separated by a yellow bulkhead. After analyzing the completion time, the number of clicks, and the content of the communication between the pairs of participants, it was observed that compared with a conventional haptic-only and audio-only condition, a significant improvement in output and working efficiency with the audio-haptic condition.

2.2.3 Audio-Haptic-Visual Feedback

Montuwy et al. studied the effects of visual, auditory, and haptic multimodal interactions in a VRE simulating pedestrian navigation [63]. Their experiment involved four navigation guidance methods with the following sensory feedback: i) map navigation (traditional navigation), ii) visual guidance feedback, iii) auditory guidance feedback, and iv) haptic guidance feedback. The visual feedback displayed a green arrow on the screen to indicate direction. The auditory feedback employed a pair of bone-conducting headsets to indicate when to turn left or right. Haptic feedback was provided by a pair of vibrotactile watch-like wristbands and similar to the auditory feedback, vibration of the left wristband indicated a left turn, while vibration of the right wristband indicated a right turn. Fifty-eight participants were invited to participate in the experiment. First, they were asked to complete a navigation guidance task by using a paper map, a visual feedback device, a haptic feedback device, and an auditory feedback device, respectively. Then, they were asked to use a combination of two devices (arrows + sounds; arrow + wristband; voice + wristband) to complete the navigation task after analyzing the time to reach the destination under different conditions. The results showed that a combination of two devices (multimodal interaction) resulted in significantly shorter task completion times than with just one device. The combination of visual and auditory feedback had the best performance. However, this experiment did not explore the simultaneous influence of the visual, sound and haptic feedback on the task.

George et al. studied the influence of multimodal interactions among visual, auditory and haptic in VREs [64]; more specifically, this was achieved by maintaining the awareness of the boundaries with the physical world when immersed in a VRE. Thirty-three participants participated in the experiment, which included a VR game named the "Floor is Lava." The participants were required to avoid the lava and reach the goal zone. The game provided visual feedback, auditory feedback, and haptic feedback using a HMD, a pair of noise-canceling headphones, and a pair of controllers that provided vibrations when touching the lava. Participants were asked to complete the experiment task (escape through a safe zone) in one of three situations: i) earphones + controllers (participants were blindfolded), ii) head-mounted display (HMD) + earphones + controllers, and iii) HMD (added mesh to the original screen). Results indicated that the participants touched the magma the least when all three senses were used simultaneously by measuring the number of times the participants touched the lava. Moreover, visual feedback played the most crucial role in the senses being involved in the experiment because the participants relied more on visual feedback during the experiment.

2.3 Medical Drilling Task in VREs

Drilling is a common task that involves cutting a circular cross-section hole into a solid material with a drill bit. The drill bit presses against the material and rotates at hundreds to thousands of revolutions per minute to cut off the contact surface between the material and the drill bit until it penetrates the material [65]. Wu et al. [66] indicated that accomplishing a drilling task involves the interaction of multiple human senses, such as the vibration (haptic feedback) and drilling sound (auditory feedback) offered by the drilling processing. Audio feedback is one of the most important roles that drilling tasks often rely on. Parsian et al. [67] determined that the sound of drilling (pitch, volume, and frequency) is usually determined by the diameter and the drilling position of the drill bit. In general, a drill bit with a larger diameter will suffer greater friction from the material being drilled in a drilling task, so the frequency and the resulting pitch of the drilling sound will be lower. Moreover, as the drill bit goes deeper into the material, the pitch and loudness of the drilling sound increase since the drill bit bears a stronger pressure from the material in the deeper drilling position [67].

Drilling tasks are used in a variety of therapies in modern medicine, particularly surgery [30], including dental surgery [68], orthopedic surgery [69], and needle insertion [70]. Surgeons require plenty of practice to gain skills and experience before performing an operation that includes surgical drilling on a patient. However, it is not easy to provide a realistic and safe practice environment for doctors who need to practice in the real world [71]. A virtual surgery simulation that offers simulated training for doctors in a virtual reality environment could solve this problem. As Grant et al. describe [72], VR is a multi-sensory experience that offers varying degrees of fidelity in visual, auditory, and haptic sensations. The simulation will decrease training costs by providing a secure and effective learning experience when using virtual reality technology in an educational environment, such as tasks can be easily changed to optimize learning goals for the learner at various levels (e.g., beginner, intermediate, and expert).

The number of surgical VR simulations available has exploded due to advances in virtual reality simulation technology; there are over 400 currently available models [73]. For example, Xia et al. used force feedback equipment and a VRE to simulate dental surgery [15]. The participants (experienced dentists) were asked to perform a dental surgery simulation using a force feedback haptic device (Geomagic 3D Touch) with six degrees of freedom. The participants were required to rate the realism and the availability of the experimental scene for the training use, and the result indicated that the participants felt there was great potential for using VR simulation to provide dental surgery training. Another example was Nguyen et al. who used a similar low-end consumer-level force feedback device to simulate knee bone drilling for training purposes [30]. They believe that virtual simulations provide cost effective and safe medical training opportunities by simulating the real medical drilling, even with the limitations of low-fidelity haptic feedback devices.

K. Khwanngern et al. conducted an experiment to explore the usability of VR simulation in medical education training which was including the drilling task [74]. More specifically, the researchers used a virtual reality environment to simulate the processing of a jaw treatment surgery with respect to craniofacial disorders, a type of birth defects in which the baby's facial and skull bones are malformed. The participants, who had professional surgical experience, were required to use a six-degrees freedom haptic device to accomplish a series of tasks, such as grabbing, cutting and drilling a virtual bone while the vision and sound of cutting and drilling the bone was playing to them, more specifically the participants were hearing the drill sound and some particles generated at the contact point to mimic the bone fragments. Then, they were required to give the feedback on both the positive and negative aspects of the system. The result indicated that the participants believed the interactivity in virtual reality is more realistic than the current teaching materials.

2.4 Concluding Remarks

Virtual simulations are widely accepted across all educational areas, including medical training, such as laparoscopic surgery [16], dental bone drilling [68], and surgical suture [12], as they offer a cost-efficient, safe, and engaging method of training in contrast to traditional medical training that often includes the apprenticeship model [72]. The quality of the virtual simulation environment depends on the fidelity of experimental equipment [23]. However, high-fidelity equipment is often expensive [30]. Previous studies have concluded that multimodal interactions involving multiple human senses can improve a virtual environment’s fidelity [64]. In addition, mixing realistic sensations with high and low fidelity can lead to an enhancement of the overall perception of realism [72]. However, the literature in this field centers around the use of haptic devices and immersive VR approaches for visual and tactile immersion, which requires access to specialized hardware not widespread among users. Furthermore, there is a lack of research exploring the interactions and effects of cross-sensory feedback to determine novel solutions for simulation without the specialized hardware. This is the case during COVID-19 pandemic lockdowns, restricted access to facilities and limited gatherings. As a result, a notorious gap in the literature pertaining to whether a haptic-based virtual simulation task can be accomplished remotely using commonly available computer hardware capable of providing visual and auditory cues, including monitor, headphones coupled with a mouse to provide touch-based feedback.

Chapter 3

Development

The development of the virtual drilling simulation solution proposed in this thesis was completed in two stages. The first stage focused on a preliminary study to inform of potential limitations and opportunities associated with remote online participants and the drilling tasks employing auditory cues only given that sound is an important part of drilling. The second stage aimed to explore the effect of multimodal interactions among visual, auditory, and haptic cues on a virtual drilling task based on the results of the first stage associated with remote online experiment during the COVID-19 pandemic.

3.1 Experiment 1: Preliminary Study

3.1.1 Auditory Stimuli

The auditory stimuli consisted of mono and stereo sound recordings of drilling through a block of wood to a depth of 5 cm and 7 cm leading to four auditory conditions. A Tacklife PCD05B drill was used to drill through a block of wood, and both the mono and stereo recordings were made using a Tascam DR 100 handheld field recorder placed at a distance of 90 cm from the block of wood. The recording setup is shown in Figure 3.1, and the setup was consistent across all four conditions. The participants were exposed to all auditory stimuli.

Figure 3.1: Auditory stimuli recording setup.

3.1.2 Preliminary Study Design

A within-subjects design was chosen due to the anticipated number of participants due to the COVID-19 pandemic. Given the focus on understanding multimodal interactions and the remote online experiment, it is acknowledged that the quality of the headphones and volume settings (set their audio volume to a comfortable level) differ for each user. It is hypothesized that the exposure to mono audio cues does not affect drill depth perception and drilling performance in comparison to stereo drilling sounds in a non-immersive virtual drilling environment. It is also hypothesized that there will be no difference between the mono and stereo recorded conditions.

3.1.3 Scenario Development

The virtual experimental environment was made using the Unity engine. As shown in Figure 3.2, the experimental scene consists of a button marked "Play Sounds" and a text marked "Trials". There were 20 different auditory stimuli, including the different conditions of drilling depth and the type of sound. Each time the button was clicked, the auditory feedback was played in a random order.

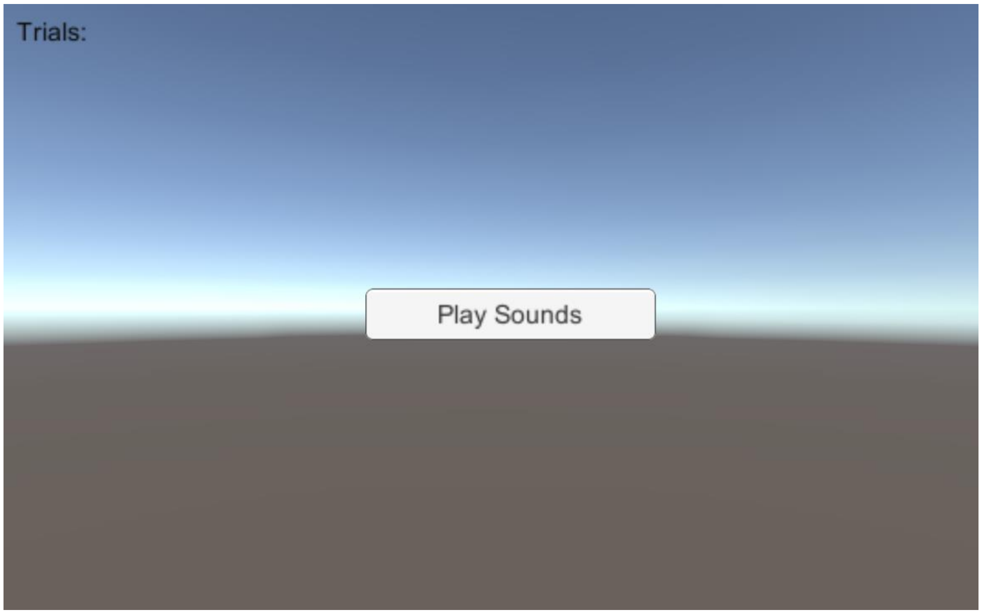


Figure 3.2. Sample screenshot of the experimental application presented

3.1.4 Participants

Participants consisted of thirteen unpaid male volunteers from Ontario Tech University in Oshawa, Canada. Participants were all between 18 and 30 years old. All participants reported having used a drill in the past and reported no known auditory problems.

3.1.5 Procedure

This preliminary study (Experiment 1) was conducted remotely to accommodate the COVID-19 shutdowns which prevented me from conducting experiments with human participants in the university research lab. The participants received an executable file containing the application to facilitate the experiment. All interactions with the participants were facilitated through the Google Meet video communication service.

Since the experiment was conducted remotely, participants were required to prepare a mouse, a headset, and a monitor to complete the experiment. Participants were asked to sit on a chair in front of their computers, to wear their headphones, and to adjust the volume to a comfortable level. During the experiment, participants were always connected to the experimenter via Google Meet.

The experiment consisted of presenting the four auditory conditions previously described to each participant five times for a total of 20 trials. The conditions were presented in random order to minimize carry-over effects. For each trial, participants were prompted to begin by pressing the "Play Sounds" button of the experiment application. After doing so, they were then presented with the auditory stimuli (no visuals were provided). After hearing the stimuli in its entirety, they were prompted to enter (via their keyboard) in a text box their perceived drill depth (their response was saved to a "response file" which to be sent back to the experimenter upon completion of the experiment).

3.2 Experiment 2: Final Study

Experiment 2 builds upon the preliminary study conducted as part of Experiment 1. The study was conducted to explore how accurately a drilling task can be simulated in the virtual domain using visual, auditory, and simple mouse movements. Additionally, the study focused on examining the interaction of these cues with respect to the accuracy of the simulated drilling task. The experiment tests the hypothesis that a drilling task can be simulated in a virtual reality environment by using appropriate auditory cues (real dynamic drilling sound), visual cues (related to the drilling task), and bidimensional mouse movements. Therefore, three main parameters were involved in implementing this experiment, i) visual cues, ii) auditory cues, and iii) mouse movement (simple "pseudo-haptic") cues.

At the time of conducting this study, the participants did not have access to the professional haptic devices housed in laboratory due to existing COVID-19 lockdowns. As a result, the final study was conducted remotely (facilitated using Google Meet) with participants employing their screens, mouse, and keyboard, similarly to the preliminary study.

3.2.1 Experimental Environment

This experimental virtual drilling scenario was created using the Unity3D game engine, as it supports multiple platforms, hardware, and technologies. The experimental scenario was based on a real-world drilling task. As shown in Figure 3.3, two objects were involved in the demo scene, a virtual drill (with a drill bit whose length was 15 cm) and a virtual block of wood (with a thickness of 15 cm). The wooden block of wood was applied with a graphic wood texture to a 3D rectangular object and a virtual model of a common hand drill that was purchased from the Unity asset store. Moreover, the background of the scene was set to the system default "skybox."

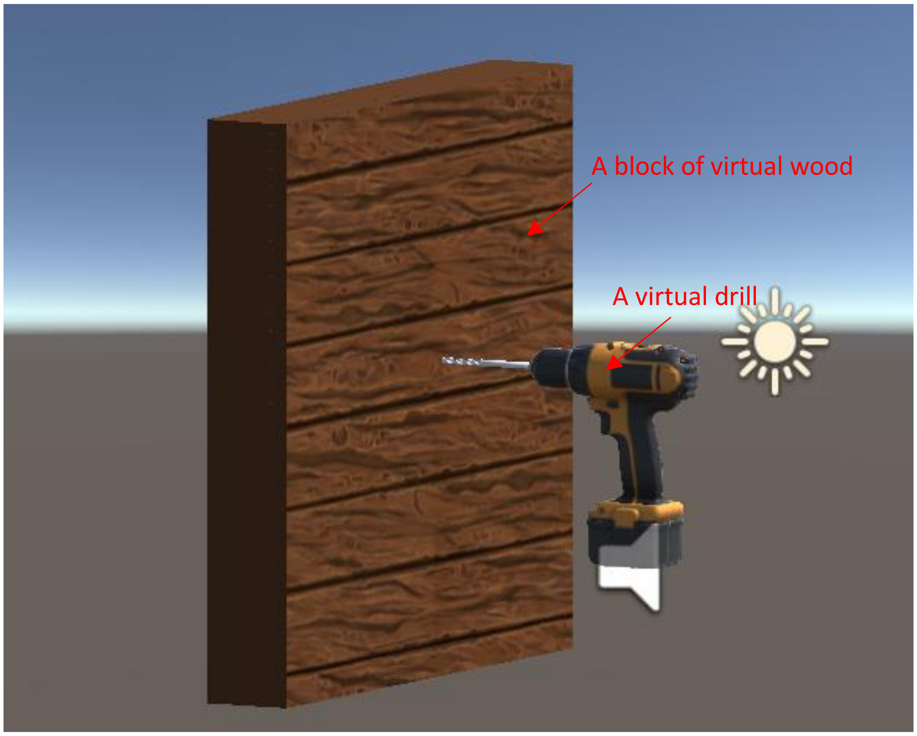


Figure 3.3: The objects in the experimental scenario comprising a block of wood and a drill with drill bit whose length was 15 cm.

3.2.2 Haptic Feedback

Haptic stimuli were limited to “pseudo-haptics” as a result of lacking access to haptic devices available in the laboratory due to COVID-19 lockdowns. Instead, the mouse movement was chosen to map drilling thrust movements, which in response produce kinesthetic haptic feedback as the user will know the position and movement of the body within the space, in this case, that of the mouse.

Two variants of the pseudo-haptic stimuli (control methods) were included: i) mouse movement whereby the mouse moves the drill to and from the wooden block, and ii) no mouse movement whereby the drill moves automatically without the participants moving the mouse. For the mouse movement condition, the drill is controlled by pressing the left mouse button and dragging the mouse back and forth to move the drill away and towards the wooden block. When the participant releases the left mouse button, the drill stops at its current position. Since the Unity 3D game engine is based on a 3D environment, it uses a Cartesian coordinate system where points in the scene are specified with x, y, and z coordinates.

For the purpose of this experiment, the drill moves forwards and backwards along the z axis (as shown in Figure 3.4). However, the mouse itself allows for movement along two axes (the x and y axes). Therefore, a mapping was implemented to map mouse movements along the y-axis to the z-axis. This was accomplished by locking the drill’s x-axis (i.e., any left and right mouse movements were ignored), while assigning the y-axis to up and down mouse movements to the virtual drill's z-axis. As a result, the mouse can be used to control the drill forward and backward by moving the mouse along its y-axis (up and down). However, since the experiment was conducted remotely, the quality, size and type of mouse used varied across participants, and as a within-subject source of variability (i.e., each participant performed in all conditions using the same mouse), this inconsistency could have an impact on the results.

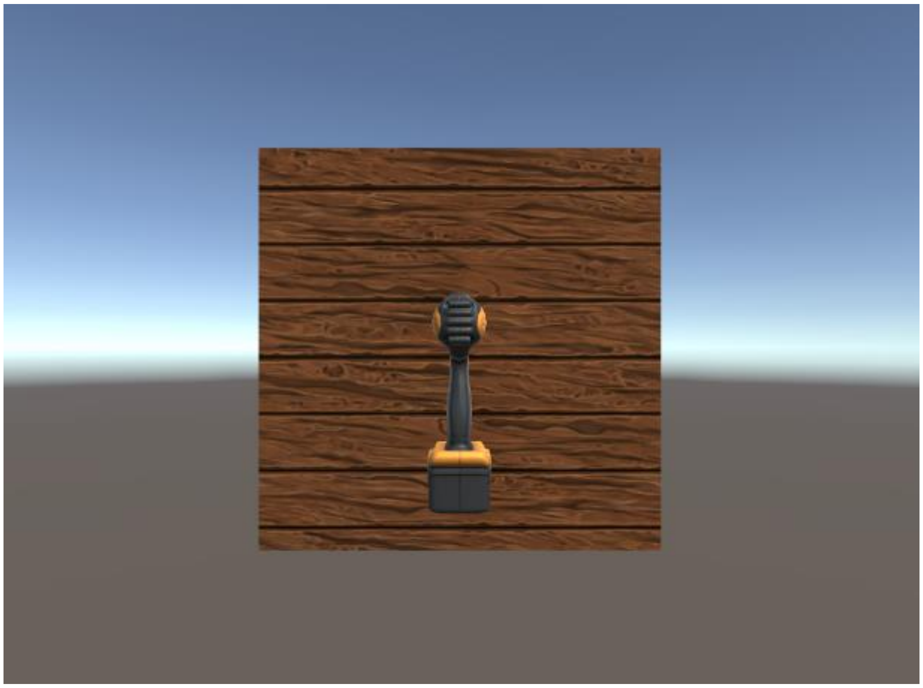


Figure 3.4: The back view of the virtual drilling scene

With respect to the conditions that did not include mouse movements, the mouse was deactivated, and the participants were required to use the keyboard to control the virtual drill. Pressing the “up-arrow” bar on the keyboard starts the drill. Once activated, the drill automatically moved forward, and participants stopped the drill by pressing the “space” bar when they perceived that the drill bit reached the target depth of 12 cm.

3.2.3 Visual Feedback

Visual feedback consisted of two conditions: i) visual, ii) non-visual. In the case of the visual condition, a top view of the experimental scene was provided to the participants, as shown in Figure 3.5. Under this condition, the participants could observe the movement of the drill and observe the drill bit moving into the wooden block during the experiment. As for the non-visual condition, the visual scene offered only a back view of the drill and wooden block, as shown in Figure 3.4. Further, all shadows from any of the objects in the scene were removed to avoid providing the user with any potential cues regarding the depth of the drill as it moved into the wooden block. The participants were required to use a monitor to receive this visual feedback, and the participants were asked to adjust the resolution of the monitor to 1024 × 768 prior to the starting the experiment.

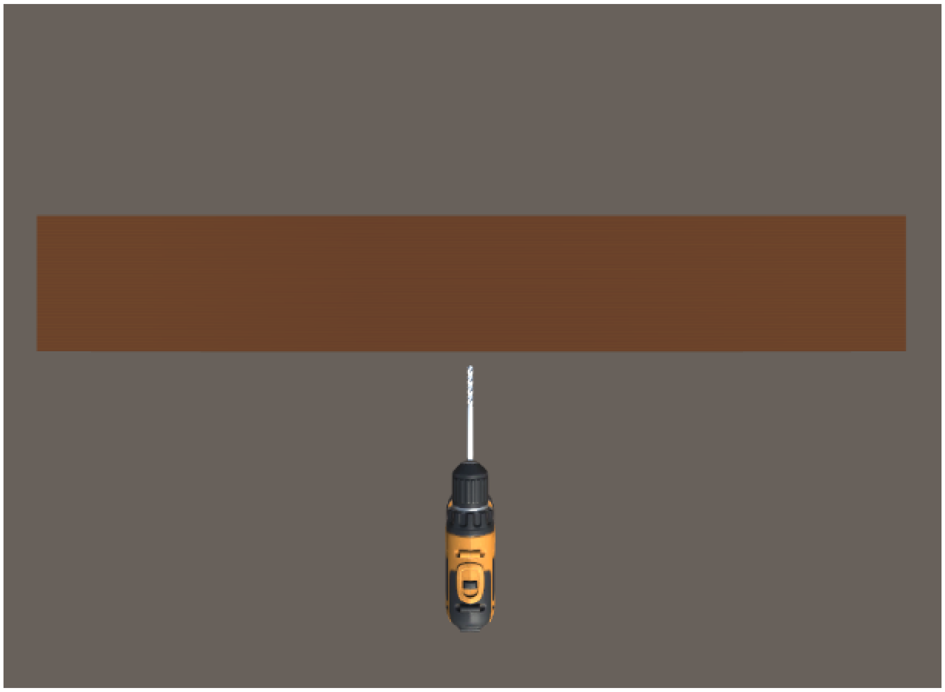


Figure 3.5: The top view of the objects in the scene

Since a drill bit encounters resistance from the material that is being drilled at the moment of contact with the surface of the material, when considering a real-world drilling task, there is a potentially large difference between the speed of the drill moving in the air (as it is moved towards the surface being drilled) and the speed of the drill when it comes into contact with the surface of the material being drilled (i.e., wooden block). More specifically, it slows down upon contact with the surface. Therefore, a trigger point was assigned on the surface of the virtual wooden block, and when the virtual drill bit contacted the surface of the wooden block (the trigger), the speed of the drill was slowed down to simulate the change in speed due to the resistance when the drill bit contacts the surfaces and begins drilling through the material. The decrease in speed was informally determined by observing videos of drilling through various materials and by conducting informal tests with an actual drill drilling through a real block of wood.

3.2.4 Auditory Feedback

The sounds associated with real-world drilling are dynamic, and more specifically, various aspects of the sound, including the pitch and volume change continuously in the process from the moment the drill bit encounters the surface of the material/object to be drilled due to the changing pressure as the drill bit moves through the material (e.g., wooden block of wood) [75]. As a consequence, simulating drilling sounds is a difficult task, and therefore, it was decided to present participants with a recording of the sound associated with real-world drill drilling through a block of wood as opposed to dynamically modeling this change using, for example, procedural audio. This ensures that the sounds presented to the participants include the appropriate dynamic drilling sound stimuli.

The auditory stimuli consisted of three auditory sound conditions: i) no sound, ii) dynamic drilling sound, and iii) continuous (static) drilling sound (drilling in the air). In the preliminary study, participants were presented with the auditory cues associated with drilling through a wooden block to an actual drill depth of either 5 cm or 7 cm and asked to judge how far they perceived the drill to have traveled (e.g., the perceived drill depth). Results indicated that there was no significant difference between drilling to a depth of 5 cm or 7 cm, and it was also determined that both distances were too small and difficult for participants to judge. As a result, it was determined to choose the larger drilling depth (12 cm) for this experiment. Of course, there is a limit to how large the drilling depth could be given that typical general purpose drill bits used around the home are not much bigger. to meet the requirement of this experiment. The auditory stimuli consisted of a stereo recording of a drill drilling in the air and through the wood to a depth of 15 cm leading to the continuous drilling and dynamic drilling auditory conditions. Although participants were asked to drill to a depth of 12 cm, the drilling stimulus recording was made by drilling to a depth of 15 cm providing a small buffer in case the participant's perception of drill depth exceeded 12 cm. Moreover, if the participants drilled more than 15 cm (the maximum thickness and length of the virtual wood board and drill bit), the drill bit would penetrate the wood board in the scene and play the sound of drilling in the air. The recording of the auditory stimulus for the drilling depth of 15 cm was made following the same setup and approach described in Experiment 1. More specifically, a Tacklife PCD05B drill and a drill bit with a total length of 15 cm and a diameter of ¼ inch were used to drill through the wood (see Figure 3.6). The recordings were made using a Tascam DR 100 handheld field recorder which is shown in Figure 3.7. As with the recording of the drilling sounds for Experiment 1, the recorder was placed at a distance of 90 cm from the block of wood. The recording setup is shown in Figure 3.1.

Figure 3.6: The Tracklife drill and a drill bit.

Figure 3.7: Tascam DR 100 handheld field recorder.

Dynamic Sounds Implementation

The dynamic drilling sound implementation in the experimental scene is determined by the different depths that the virtual drill bit reaches into the woodblock. As shown in Figure 3.8, the block of virtual wood was divided evenly into 40 layers, and in each layer, a trigger was added. Similarly, the real-world auditory stimulus recording where the real-world drill was used to drill through a wooden block of wood to a depth of 15 cm depth was also divided into 40 segments and connected to the trigger of the layer with the corresponding depth in the virtual wood block. In other words, as the virtual drill encounters the trigger at each layer of the wooden block, the experiment scenario will provide the drilling sound corresponding to the current depth of the real drilling task to the participant; therefore, the drilling sound changes dynamically according to the depth of the drill.

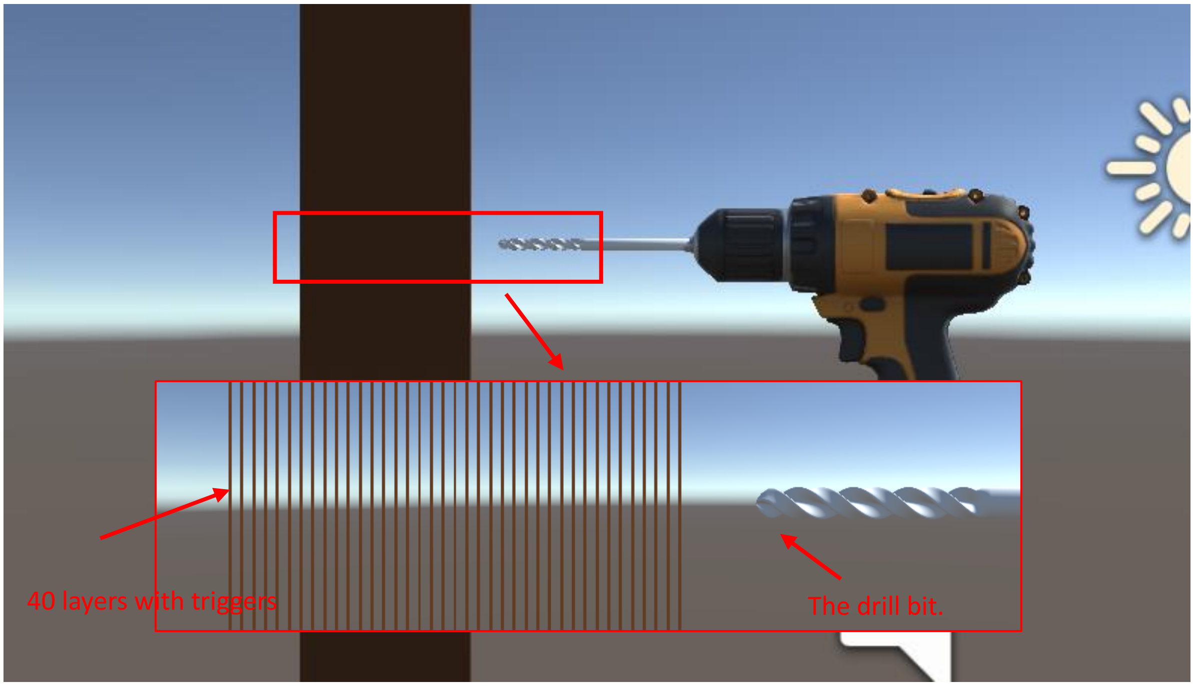


Figure 3.8: The virtual drill bit and the 40 layers in the wood block with corresponding triggers.

Audio Output

All auditory stimuli were presented to the participants through headphones that they were asked to wear. It should be noted that the experiment was conducted remotely, and for this reason, the quality and type of headphones used varied across participants. During the experiment, participants were told they could adjust the volume to a level that they felt comfortable with.

3.2.5 Participants

Thirteen participants were recruited and participated in this experiment. All the participants were students; eight of them were undergraduate students, and five were graduate students in computer science. Seven of the participants were from the Ontario Tech University student community. There were three females and ten males. Three of the participants were between the ages of 21 and 25, eight of them were between the ages 25 to 30, and two of the participants were over 30 years old. Participation in this experiment was entirely voluntary; participants were not compensated for participating in the experiment. Since this experiment aimed to explore the impact of auditory and visual cues on the virtual reality environment, anyone with a hearing problem was not qualified to participate in the experiment. Prior to be beginning the experiment, each participant was asked whether they had any hearing difficulties that they were aware of. None of the participants reported any issues and thus, none were excluded from the experiment. Furthermore, given that drilling was the main task of the investigation, participants without any real drilling experience were also excluded. Once again, prior to beginning the experiment, each participant was asked whether they had prior experience with using a drill. All participants reported using a drill in the past and thus no participants were excluded.

Participants were invited to participate through a recruitment email (see Appendix 1). The experiment was facilitated through the Google Meet online meeting platform. Before the experiment began, the participants were introduced to the experiment, which included an overview of the experiment, the experimental goals, and their required task in this experiment verbally. The verbal recruitment script made it clear that the experiment was unrelated to any of their coursework/classes. After signing the consent form (see Appendix 2), each participant took part in the experiment individually.

Approval for this experiment was granted by the Ontario Tech University Ethics Research Board (Application number: 16280).

3.2.6 Procedure

Participants were tasked with using the virtual drill to penetrate the block of wood until they believed that they have reached a pre-determined depth under various visual, auditory, and haptic feedback conditions. Since the drill depth in the previous experiment did not make a difference, which may be due to the too short pre-determined drilling depth, a drill depth of 12 centimeters (cm) was chosen as the pre-determined drilling depth. This is due to fact that 12 cm is greater than the depths considered in previous experiment (5 cm and 7 cm) since those depths were too small. Greater detail regarding the experimental procedure is provided below.

Training Phase

The participants who accepted the email invitation, had prior experience with using a drill and did not report any hearing issues, took part in the experiment. The experiment was conducted entirely online via Google Meet. The experimenter met with each participant on Google Meet (https://meet.google.com/), and the experiment began by providing the participants with a brief verbal introduction of the experiment that included an overview of the study's purpose and an appreciation of the participation. The experimenter also provided each participant with an executable file (the experiment application developed using the Unity 3D game engine), a consent form (see Appendix 2), and a questionnaire (see Appendix 3) via email. Since the experiment was conducted virtually/remotely, participants were required to have a mouse, headphones/earbuds, and were asked to set the resolution of their monitor/display to 1024 × 768 to complete the experiment. Participants were asked to sit on a chair in front of their display and to adjust the height according to their own habits but to keep their eyes at a similar height to the monitor. During the experiment, participants were required to be connected to the experimenter via Google Meet.

Prior to the formal experiment, participants began by completing a training session to familiarize themselves with the process (i.e., the drilling simulation). This training was divided into two parts for drilling using i) visual cues and drill movement method, and ii) sound cues. Each part was repeated twice to the participants. Greater details regarding the training components are provided below.

1. Visual scene where a top view of the virtual drilling scene was presented (see Figure 3.9). Here, the participants were given a visual top-view of the virtual drilling scenario, and they were asked to move the drill into the block of (virtual) wood by moving the mouse. The drill would stop when the participants reached a depth of 5 cm, at which point they were notified. The participants then repeated this task but for a drill depth of 10 cm. Then, they needed to repeat this task using a keyboard to control the drilling movement. More specifically, the drill moved forward automatically without mouse control, and the participants needed to press the up-arrow bar on the keyboard to activate the virtual drill and to press the space bar to stop it when they were notified that the drilling depth reaches 5 cm. The participants would then repeat this task but for a drill depth of 10 cm. No auditory cues were provided here.

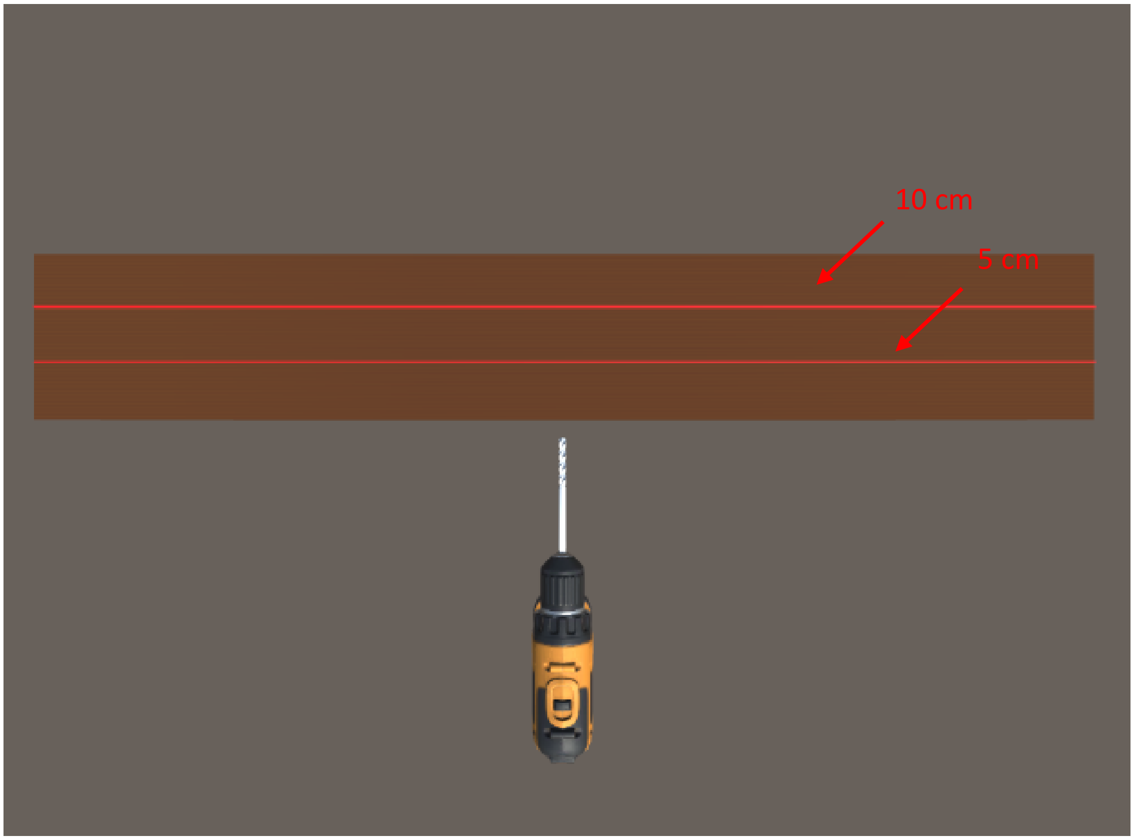


Figure 3.9: Training 1 with a top view of the virtual drilling scene

2. Here, participants were presented with the scene shown in Figure 3.10 (looking at the back end of the drill). In this scenario, the drill would go through the block of wood automatically (the drill would be moved without the participants having to move the mouse), but the participants were presented with auditory cues. They would be notified with a textual message on the screen when the drill reached a depth of 5 cm. This then was repeated, but with a drill depth of 10 cm.

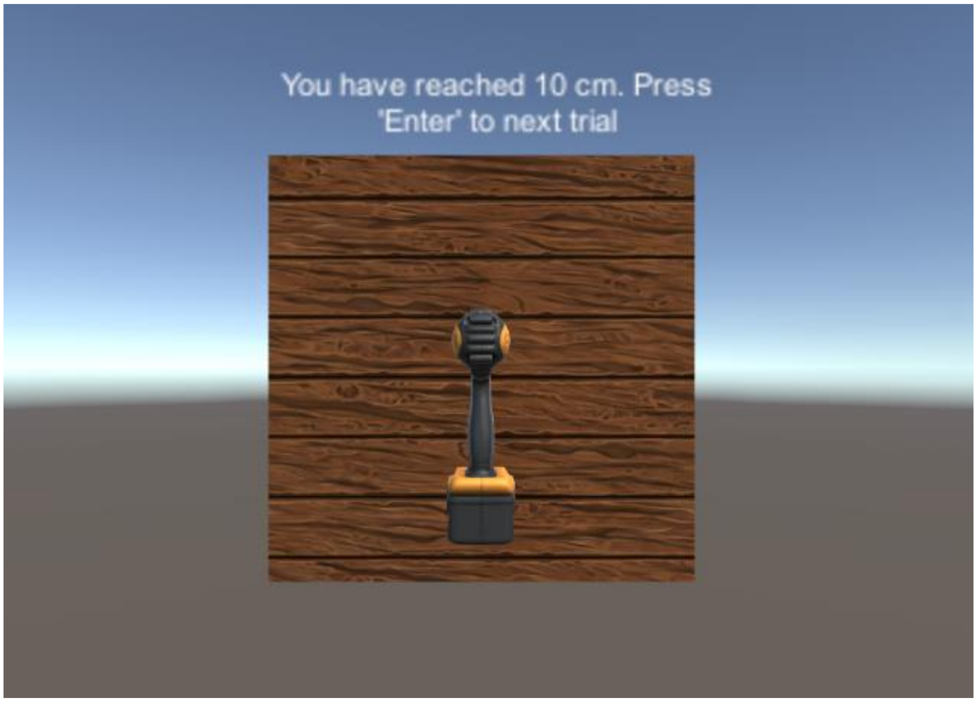

You have reached 10 cm. Press 'Enter' to next trial

Figure 3.10: Training 2 with a back view of the virtual drilling

Experiment Phase

After completing the two training sessions, participants started the formal experiment phase. The actual experiment asked participants to complete a total of 36 trials. For each trial, participants were asked to drill to a depth of 12 cm (depth remained constant throughout the experiment) under various visual, auditory, and mouse-movement conditions. More specifically, the visual conditions are i) visual cues, ii) non-visual cues. The mouse-movement conditions are i) mouse movement, ii) no mouse movement (drill moves automatically without the participant moving the mouse). The three auditory conditions include i) no sound, ii) continuous drilling sound (drilling in the air), iii) dynamic drilling sound. In total, there are 12 conditions (two visual conditions × two movement conditions × three auditory conditions), and each condition was repeated three times for a total of 36 trials. The 36 trials were presented to the participants randomly, and aside from the cues they were provided with, no feedback was given to participants as they completed each trial. However, at the beginning of each scene, the method of drill control was laid out to the participants in the form of text, as shown in Figure 3.11 and Figure 3.12.

For each trial, when the participant completed the drilling task and released the mouse (in the mouse movement condition) or when they hit the space bar (in the no mouse movement condition) to stop the drill, the current drilling depth was recorded in a text-formatted file which was automatically generated on the computer of users at the beginning of the experiment. This file included, for each trial, the experimental conditions (visual, auditory, movement) and the drill depth (the depth that participants perceived they had drilled). Upon completion of the experiment, the experimenter asked participants to email the file to them. Finally, the participants completed a questionnaire that included demographic-type questions and several questions regarding the task completed and submitted it to the experimenter.

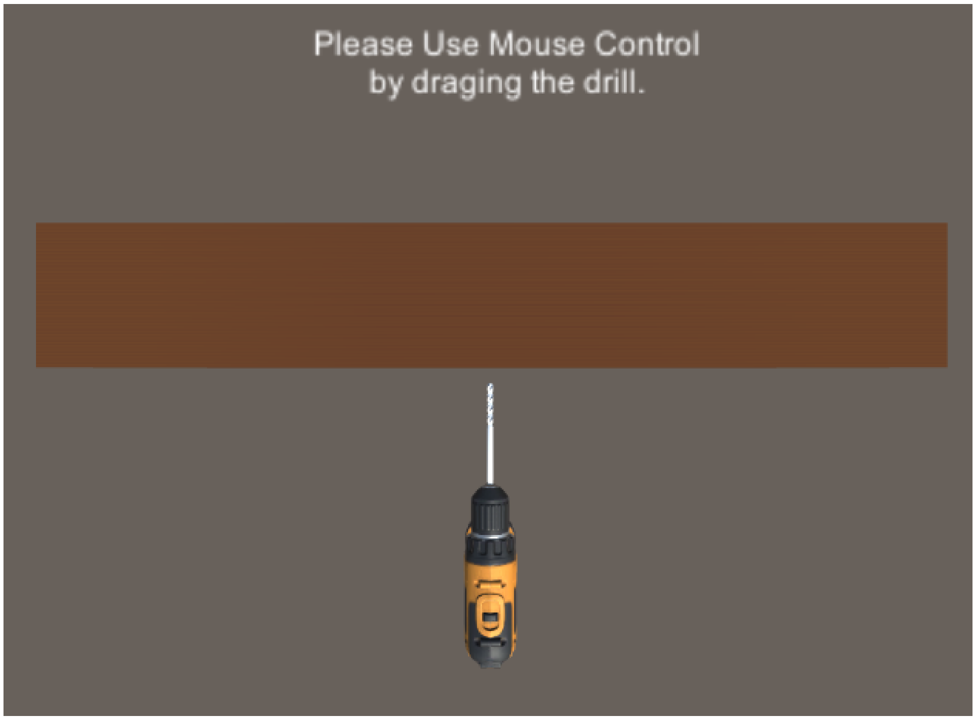


Figure 3.11: A top view of the virtual drilling with mouse control.

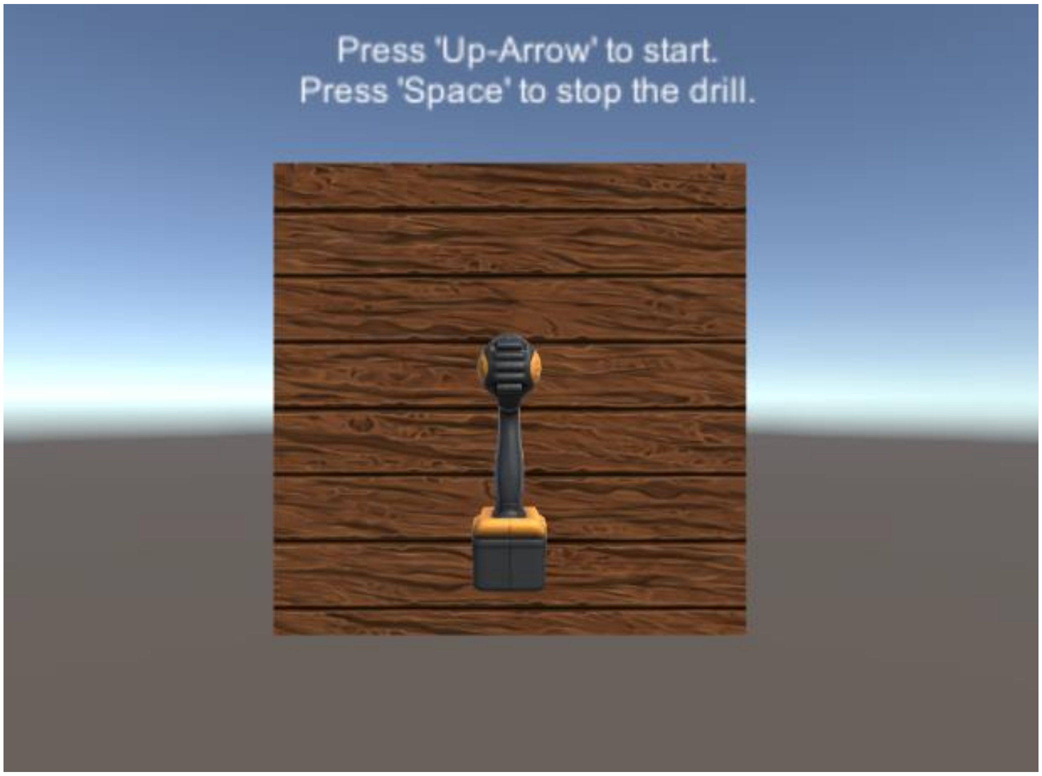

Press 'Up-Arrow' to start.
Press 'Space' to stop the drill.

Figure 3.12: A back view of the virtual drilling with keyboard

Chapter 4

Results

4.1 Preliminary Study Results (Experiment 1)

There were two parameters involved in Experiment 1: sound type (mono, stereo) and drill depth (5 cm, 7 cm) within-subject analysis of variance (ANOVA). Significance was set to less than .01. The 2 × 2 ANOVA test showed no main effect for sound type (mono and stereo) on sound perception F (1, 174) = 3.58, p = 0.06, η^2 = 0.02, implying that the sound type did not impact the perception of drilling depth. Similarly, the results showed no main effect for drilling depth (5 cm and 7 cm) F (1, 174) = 3.47, p = 0.064, η^2 =0.02, implying that the depth of the drilling did not impact the perception of drilling depth. Moreover, there was no interaction effect F (1, 174) = 0.10, p =0.75, η^2 =0.001, implying that the perception of drilling depth in either of the two factors were not modulated by each other (e.g., stereo, and mono sounds differed for 7 cm drilling depth and not for the 5 cm drill depth).

4.2 Final Study Results (Experiment 2)

Thirteen participants (three females and ten males) participated in Experiment 2, facilitated remotely using the Google Meet video-communication service. Five of the participants were graduate students in computer science, and eight were undergraduate students. The participants were asked to complete an experiment involving 12 different conditions, randomly presented three times without repeating in two adjacent trials, thus producing 468 data points in total across 13 participants completing 12 conditions three times. IBM SPSS, which is a software package for statistical analysis, was used to conduct a Repeat Measures Analysis of Variance (Factorial ANOVA) to analyze the differences among means. The Repeat Measures ANOVA was chosen since the experiment required each participant to experience all the within subjects/conditions of the experiment.

The 12 conditions were combined by three parameters, which were included in this analysis as independent variables (IVs), "Visual" (Visual Feedback or Non-Visual Feedback), "Control Method" (Mouse Control or Keyboard Control), and "Auditory Stimulus" (No Sound, Continuous Drilling sound, or Dynamic Drilling Sound), and the graphic representation is shown in the Table 4.1. One dependent variable (DV) was analyzed in this experiment: "drilling depth accuracy." The drilling depth accuracy was calculated by subtracting the target depth (12 cm for each trial) from the actual drilled depth, which was the depth at the point that the participant stopped drilling. In other words,

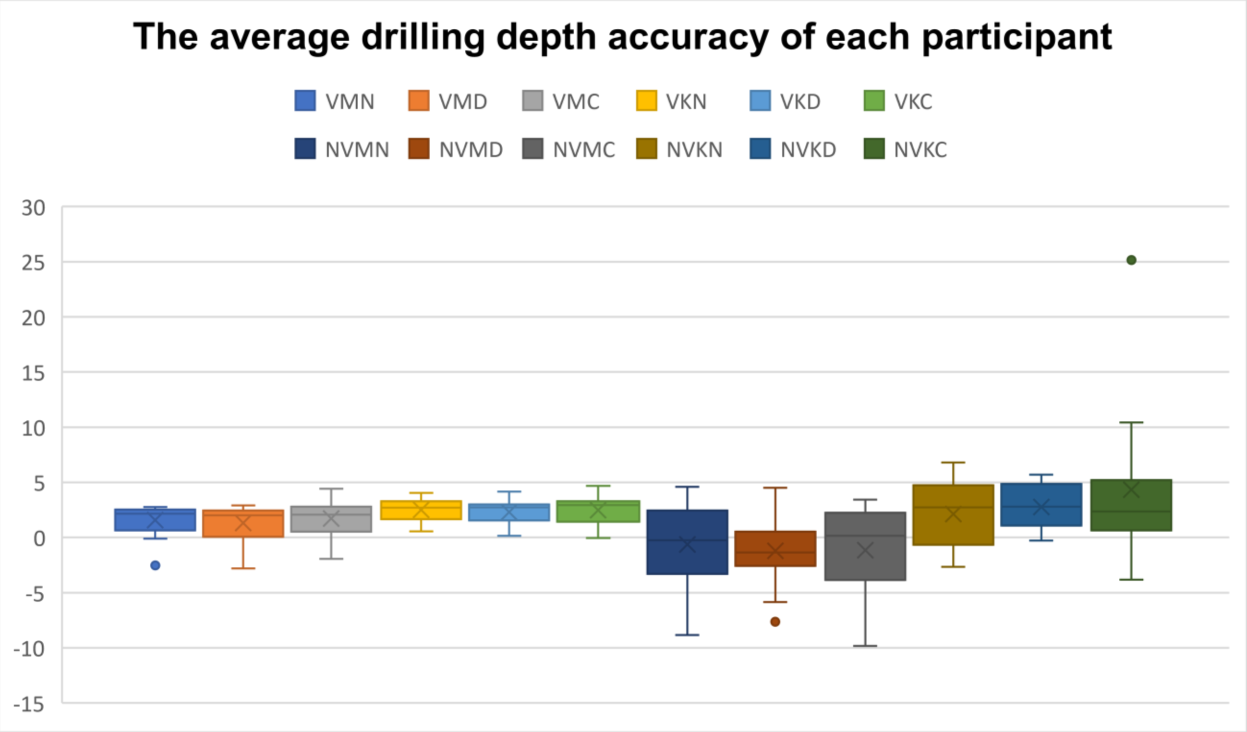
drilling depth accuracy = actual drilled depth – target depth

As shown in Figure 4.1, the average drilling depth accuracy of each participant for each experimental condition was aggregated for the data analysis. For each condition, the drilling depth accuracy was calculated by averaging the results from each trial.

Table 4.1: Graphic representation of the experimental design.

Variables	VISUAL	HAPTIC	AUDITORY	Trials
Levels	Visual Condition (top view)	Mouse Control	No Sound	39
			Dynamic Drilling Sound	39
			Continuous Drilling Sound	39
		Keyboard Control	No Sound	39
			Dynamic Drilling Sound	39
			Continuous Drilling Sound	39
	Non-visual Condition (back view)	Mouse Control	No Sound	39
			Dynamic Drilling Sound	39
			Continuous Drilling Sound	39
		Keyboard Control	No Sound	39
			Dynamic Drilling Sound	39
			Continuous Drilling Sound	39

Figure 4.1: The average drilling depth accuracy of each participant.



Note: V = visual; NV = non-visual; M = mouse control; K = keyboard control; N = no sound; D = dynamic drilling sound; C = continuous drilling sound

Quantitative Analysis

The results of the Repeated Measures ANOVA indicate that there was a statistically significant difference for the “Control Method” (“keyboard control” and “mouse control”) (F (1, 12) = 9.735, p < 0.05, η^2= 0.448) on the drilling depth accuracy. In other words, the “Drill Control Method” was affecting the task performance (drilling accuracy) significantly. More specifically, as shown in Figure 4.2, the participants were more accurate when they were asked to use the mouse to move the drill as opposed to using the keyboard to start and stop the drill, since the “mouse control” (*M* = 0.267 cm, *SD* = 0.599 cm) had a better performance than “keyboard control” (*M* = 2.754 cm, *SD* = 0.388 cm) on the drilling depth accuracy.

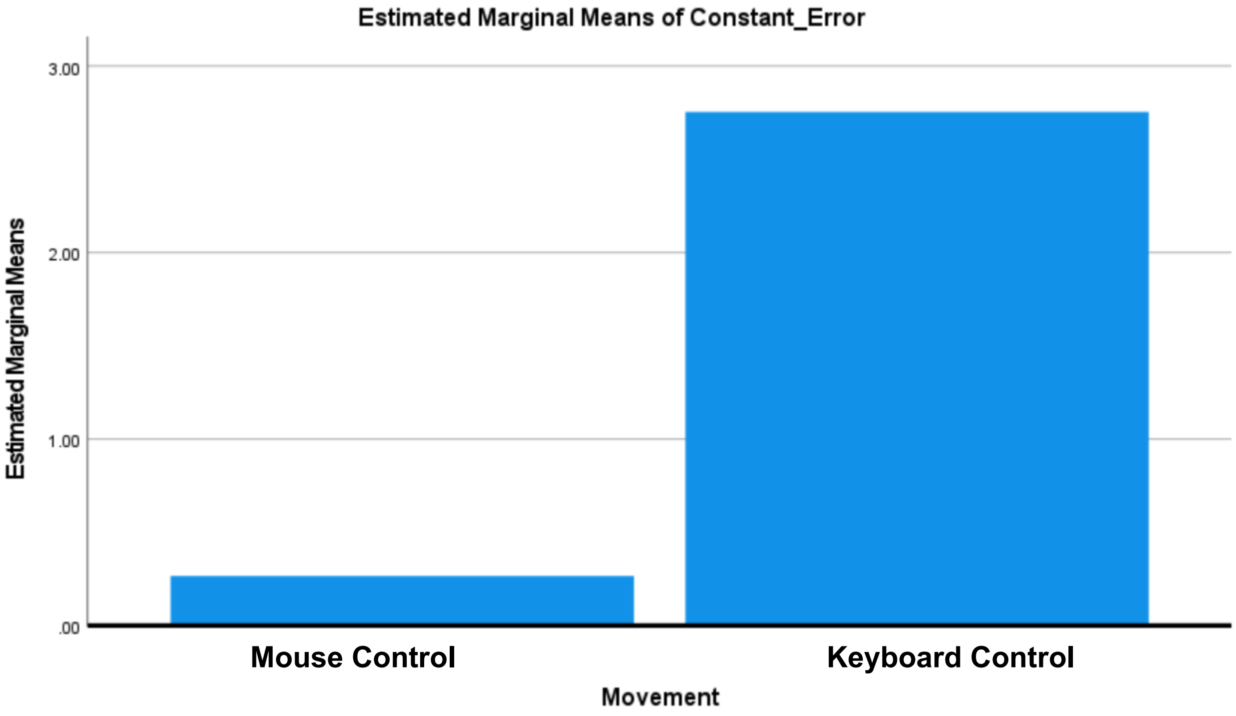


Figure 4.2: The mean value of drilling depth accuracy of control method

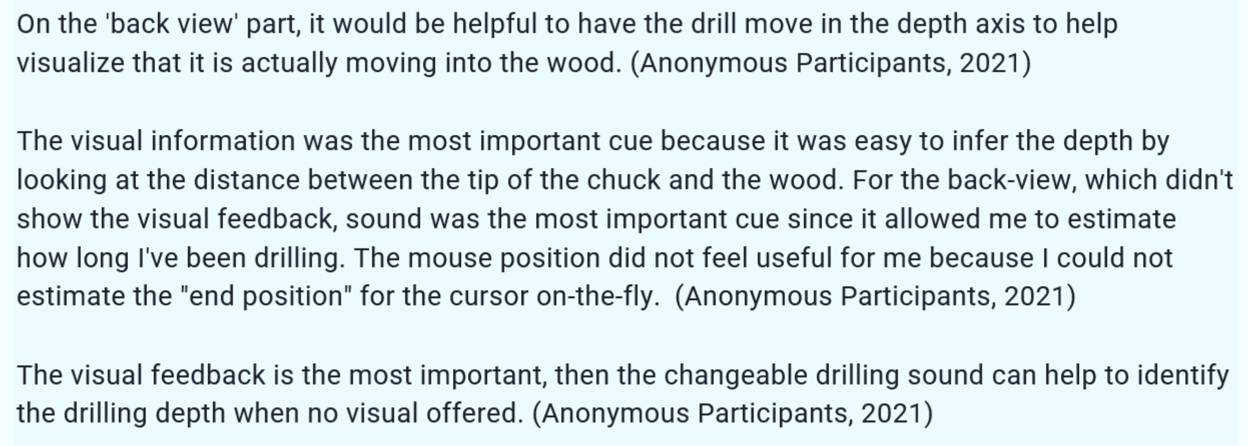
The results of the Repeated Measures ANOVA indicate that there was no statistically significant difference for “Visual” (F (1, 12) = 1.542, p = 0.238, η^2= 0.114), implying that the “visual” and “non-visual” conditions did not influence the participants’ task performance (drilling depth accuracy). Moreover, the “Auditory Stimulus” did not yield any significant differences (F (2, 24) = 0.78, p = 0.47, η^2= 0.061) on the drilling depth accuracy. Therefore, it can be concluded that the auditory feedback (i.e., no sound; dynamic drilling sound; continuous drilling sound) and visual feedback (visual and non-visual conditions) did not influence the task performance. However, there was no statistically significant difference for the interaction between “Auditory” and “Visual” (F (2, 24) = 0.34, p = 0.715, η^2= 0.028) and the interaction between “Auditory” and “Control Method” (F (2, 24) = 1.319, p = 0.286, η^2= 0.099). Moreover, the interaction among “Auditory”, “Visual” and “Control Method” was also not significant different (F (2, 24) = 1.685, p = 0.207, η^2= 0.123) on the drilling depth accuracy. These results indicated that there was no multimodal interaction effect among “Auditory”, “Visual” and “Control Method.”

It can be concluded that visual and auditory feedback did not influence the drilling depth accuracy. However, haptic feedback (Control Method) plays an important role in it. Comparing the mean and standard variance leads to the conclusion that the mouse control (e.g., controlling the drill with the mouse) provided better drilling accuracy in contrast to keyboard control.

Qualitative Analysis

Although the experimental results showed that the auditory feedback did not significantly affect the accuracy of the drilling depth (task performance), the questionnaire survey (see Appendix 3) of participants showed that almost all participants believed that the sound played an important role in the virtual drilling task (some of the participants' comments are shown in Figure 4.3). More specifically, some of the participants mentioned in the questionnaire that the dynamic drilling sound helped to identify the drilling depth (drill’s location), particularly in the condition where the visual feedback was removed, since the drilling sound increased in pitch and volume as the drill moved through the virtual block of wood and thus making the drilling simulation more realistic, and the realistic feeling could help them to visualize that it was moving into the wood.

Figure 4.3: The comments of the open-ended questions from some participants.



On the 'back view' part, it would be helpful to have the drill move in the depth axis to help visualize that it is actually moving into the wood. (Anonymous Participants, 2021)

The visual information was the most important cue because it was easy to infer the depth by looking at the distance between the tip of the chuck and the wood. For the back-view, which didn't show the visual feedback, sound was the most important cue since it allowed me to estimate how long I've been drilling. The mouse position did not feel useful for me because I could not estimate the "end position" for the cursor on-the-fly. (Anonymous Participants, 2021)

The visual feedback is the most important, then the changeable drilling sound can help to identify the drilling depth when no visual offered. (Anonymous Participants, 2021)

Chapter 5

Discussion

5.1 Preliminary Study (Experiment 1) Discussion

The use of non-immersive VR allows for the introduction of interactive virtual worlds to users lacking access to VR (including haptic devices). Here, the results of a preliminary experiment that examined the effects of drilling sound on perceived drilling depth, accuracy, and precision were presented. Results indicate that the effect between the sound type (mono or stereo) and the drilling depth (5 cm or 7 cm) was not significant with respect to the perception of the drilling depth accuracy. It is worth noting that the participants did use their own headphones, and volume levels may have differed (participants were only asked to adjust the volume to a comfortable level). This was not accounted for in the current study, although it will be considered in a more extensive future study. It should also be noted that we did not account for any effects that resulted from the fact that the recorded sounds for the 5 cm and 7 cm drilling depths differed with respect to sound duration. More specifically, the sound associated with the 7 cm drilling depth was longer than the sound associated with the 5 cm drilling depth, and this too may be used as a cue.

Due to Covid-19, face-to-face experiments with human participants in a research lab during the time when this experiment was conducted were not allowed. Thus, the experiment was conducted remotely, over the Internet, which undoubtedly brought many potential challenges, including with respect to the sound quality may be compressed/decreased while being presented remotely because of network data transmission is limited in size and speed. However, in this experiment (audio-only experiment), all the experimental material (sound material) was sent to the participants via the network (e.g., Google Drive and email), and the experiment was conducted on their computer. This approach allows the participants receive sound material that is complete and lossless, which significantly reduces the loss of sound quality through webcasts (such as experiments conducted in real-time via Google Meet).

5.2 Final Study (Experiment 2) Discussion

This thesis research question centers around understanding if the appropriate visual, dynamic drilling sound cues and simple mouse movements would provide a realistic simulation of a drilling task (the drilling through a block of wood). The results indicate that the visual cues had no significant impact on task performance (drilling depth accuracy); therefore, Hypothesis 1 (**H1**): The top view (visual condition) will result in better performance (drilling accuracy) than the scenes which only offering the back view (non-visual condition), was rejected. Compared to the prior studies conducted by Melaisi et al. [18] to explore whether sound cues can be used to enhance the perception of haptics in a virtual reality-based drilling task, a new parameter was included in my work, and more specifically, visual feedback. Visual feedback has long been considered an important part of virtual reality environments, and previous studies found that the interaction of visual and auditory feedback can enhance the fidelity of the overall virtual environment [9]. In this experiment, the visual feedback was divided into two conditions: i) visual feedback (top view) ii) non-visual feedback (back view). With respect to the visual condition, I hypothesized that it provides information that would lead to higher task performance (i.e., improved drilling depth accuracy) since participants could observe the drill's movement from the top view. In contrast, participants were unable to see the movement of the drill bit through the wooden block in the absence of visual cues under non-visual conditions, such as in the absence of any visual feedback at all (the trial with non-visual, no sound, and keyboard control conditions), participants could not determine whether the drill reached the wooden block. However, the results showed there was no statistically significant difference between the two visual feedback conditions. These results could be associated with the online nature of the experiment, as all the participants had similar but diverse equipment, more specifically, the participants had to prepare their monitor to receive the visual feedbacks; therefore, the size or quality of the displays could not be unified. Another situation that could have affected the results of the experiment was that participants were only required to place the monitor where they felt comfortable before the experiment, so the position of the monitor, such as height and distance from the participant's eyes, was not uniform.

The auditory stimulus did not have a significant effect on the drilling depth accuracy among the three sound conditions, no sound, continuous drilling sound, and dynamic drilling sound. Therefore, Hypothesis 2 (**H2**): Dynamic contextual auditory cues will influence task performance better than no sound, and in the presence of a continuous drilling sound, was also rejected. In the previous experiments of Melaisi et al. [18], which was an experiment that explored the effect of sound on haptic fidelity perception within a virtual reality-based drilling task, only continuous/static drilling sound was involved in the experiment; the researchers did not take into account the fact that the pitch and volume of the actual drilling sound increase with the drilling depth getting deeper [67]. In my experiment, a recording of an actual drilling sound was used to compare with the conditions with a static drilling sound or with no sound at all. It was assumed that the experimental condition with dynamic drilling sound would increase/improve task performance (drilling depth accuracy) than the conditions with the other two auditory feedbacks. However, this result showed that there was no significant difference among the three sounds feedback. This may be due to the fact that each participant used their own headphones, mouse, and monitor, to participate in the experiment, but they were only asked to adjust the volume to their comfort level, and therefore, volume was not uniform across participants. This may have resulted in some participants not being able to hear the dynamic drilling sound significantly if they turned the volume down too low. On the other hand, the headphones used by each participant were different, so the quality of these headphones and the resulting sound presented to the participants may also lead to some errors. However, even though the sound did not have a significant effect on drilling depth accuracy in the virtual simulation drilling experiment, almost all participants believed that sound played an important role in the virtual simulation drilling experiment, according to a questionnaire survey of the participants. More specifically, some of the participants stated in the survey that the dynamic drilling sound assisted them in determining the drilling depth (drill location), especially in the absence of visual feedback since the drilling sound became more realistic as the drill was drilled with the increase in pitch and volume to make the drilling simulation more realistic, and the realistic feeling could help them envision the drill moving into the wood.

The control method (i.e., use of keyboard vs. mouse to start/stop and move the drill) affected the drilling depth accuracy significantly. More specifically, using the mouse to control the drill resulted in improved performance (drilling depth accuracy) than using the keyboard to control. Therefore, Hypothesis 3 (**H3**), which states that simple mouse movements will result in better performance than keyboard control, was confirmed. Due to the COVID-19 lockdowns, haptic devices could not be used in the experiment, limiting participants' ability to receive more realistic haptic feedback during the drilling simulation experiment. Instead, the mouse, which is not an ideal haptic device but common in the home, was chosen to offer the "pseudo-haptic" feedback to the participants (e.g., the user was able to grasp the mouse and move the virtual drill by moving the mouse). As a haptic feedback device, the mouse is not ideal because it does not provide any force feedback or vibration feedback (that would be available during the operation of a drill), to the participants; it simply allowed the participants to physically hold the mouse and control (move) the virtual drill. Moreover, with respect to the mouse movement control, the participants required practice in order to control the virtual drill with the mouse (e.g., the need to click and drag a specific area (the position of the drill handle) and become familiar with the drill movement speed under the mouse control), and thus may have required greater cognitive effort/load. Although the mouse control only provides limited (“pseudo-haptic”) feedback which is not ideal, compared with the keyboard control, in which participants had little involvement aside from pressing a key to start and stop the drill, the mouse control brings a better drilling task accuracy. Haptic feedback, even if limited and of low fidelity, is important in a VR environment where psychomotor skills are involved.

Results show that the interaction effect between visual, sound, and control methods was not significant on task performance. Therefore, Hypothesis 4 (**H4**): Visual + dynamic auditory + mouse movements will lead to the most excellent task performance, was rejected. The Null Hypothesis (**H0**): Visual, auditory, and mouse movement cues that can be used to simulate a drilling task, was also rejected. Furthermore, the interaction effect between auditory and haptic had no significant difference on the drilling task accuracy; therefore, it cannot be concluded that the multimodal interaction between the audio and haptic feedback can enhance the fidelity of virtual reality environments which may be due to the limitation given by the online experiment during the COVID-19 lockdowns.

Chapter 6

Conclusion

Within this thesis, I explored the influence of dynamic sound on virtual drilling perception and whether a haptic-based virtual environment could be simulated using only the appropriate auditory and visual cues and standard computer devices in the home (headphones, mouse, monitors, etc.) in the absence of haptic devices. The first experiment presented the results of a preliminary experiment that was conducted to examine drilling depth accuracy within a non-immersive VR environment employing audio cues only. The second experiment explored whether appropriate visual and auditory cues with simple mouse movements would influence the drilling task performance, more specifically, whether they could enhance the drilling depth accuracy.

Collectively, the results from these two experiments showed that auditory feedback alone is not enough to provide a realistic virtual reality environment that involves a multimodal task that includes haptic, and auditory stimuli/feedback. However, it is still possible to simulate a psychomotor-based task that is heavily dependent on haptic cues (drilling in this work) in the virtual domain, using sound in conjunction with very primitive haptic feedback provided with a computer mouse, since many participants mentioned in the questionnaire survey that the dynamic drilling sound could help them locate the drill bit to complete the experiment even the experimental results showed the auditory did not influence the drilling depth accuracy, which may be affected by the limitation of the online experiment. Although greater work is required before more concrete statements can be made, the ability to simulate psychomotor-based tasks using low fidelity haptic cues can have large consequences in the virtual training domain. More specifically, it may allow trainees to practice such tasks outside of research/training labs as currently done given the use of complex and expensive equipment to simulate haptic cues.

6.1 Challenges and Limitations

It is believed that most of the experimental errors were due to the COVID-19 pandemic and resulting lockdowns, access to teaching facilities and laboratories has been greatly restricted, and education has moved to a remote (e-learning) model [76]. More specifically, in the process of conducting the experiment and the preparation before the experiment, several problems and challenges were encountered:

1. Creation of the Experiment Demo

The two experiments that were conducted as part of this thesis required the use of dynamic drilling sounds. However, it was difficult to simulate the real changing drilling sound by some sound modification software; therefore, it was decided to record the sound of a real drilling task as a reference to reference the experiment. However, given the COVID-19 shutdowns, it was not possible to access the GAMER Lab in order to use the professional recording equipment and audiometric recording room in the laboratory. Therefore, the drill recordings were made in a room of my home using a Tascam portable recorder. It should be noted that the resulting recordings included any (even if minimal) reverberations resulting from the room (their reverberations would have been greatly reduced if the recordings were made in the audiometric room contained in the GAMER Lab.

2. Lack of Access to Specialized Equipment

Another problem that was encountered was the lack of laboratory equipment and the inability to unify experimental equipment. The original experiment involved haptic feedback using a haptic device housed in the GAMER Lab, but access to the GAMER Lab was not permitted due to the COVID-19 shutdowns, and the experiment was suitably modified and conducted remotely via the Internet; therefore, I had to substitute the haptic device with a common household/office device instead of a professional haptic feedback device. The mouse was chosen to provide very basic haptic feedback (e.g., pseudo-haptic feedback) to the user, although it should be noted that the typical mouse is not capable of providing haptic feedback in response to the user's interactions. Rather, the user held the mouse and moved it to simulate holding the drill and moving the drill.

Aside from the lack of a haptic device and using a mouse instead, participants were required to use their mouse, keyboard, monitor, and headphones to complete the experimental task. However, there was no consistency of these devices across all of the participants, and this may have impacted the results.

3. **Online Remote Data Collection**

   As the experiment was carried out remotely, facilitated using the Google Meet videoconferencing platform, the state of the participants during the experiment and the process of the experiment could not be observed and controlled. For example, in the absence of face-to-face communication between participants and researchers, some participants may not be very serious about the experiment. Therefore, there was a large number of outliers in the experimental data which were removed before the data analyses.

6.2 Future Work

This researching topic is worthy to explore deeper because the virtual reality-based simulation that employs virtual reality environments (VREs) training is being widely embraced in medical and educational use. This technology has been maturing to offer a secure and efficient training environment using consumer-level hardware and devices. However, due to the COVID-19 lockdowns, lacking appropriate equipment and more importantly, the lack of equipment consistency across participants may have affected the experimental results. Therefore, once the COVID-19 pandemic ends and access to the university and research labs is once again permitted, future work will focus on conducting experiments that examine the effect of sound on haptic fidelity perception using consumer-level haptic feedback devices to simulate the drill to offer the participants more realistic haptic feedback. In other words, the experiments that were planned prior to the COVID-19 shutdowns will be conducted. That being said, regardless of when the current COVID-19 rules are relaxed, and research labs reopen, future work will also build upon the work described in this thesis by focusing on providing better haptic feedback with common devices. For example, the experiments conducted here can be repeated with the use of a smartphone that will provide vibration (a basic feature built into the phone) as the haptic feedback (vibration from the phone) when the participant controls the drill using the phone's touch screen. Given the ubiquity of smartphones and their ability to provide simple haptic feedback in the form of vibrations, it is worth exploring their potential use in virtual simulations.

Chapter 7

Bibliography

[1] S. Coren, “Sensation and Perception,” in *Handbook of Psychology*, American Cancer Society, 2003, pp. 85–108.

[2] “Sensation and Perception,” *Noba*. https://nobaproject.com/modules/sensation-and-perception (accessed Mar. 14, 2021).

[3] C. R. Sims, “Rate–distortion theory and human perception,” *Cognition*, vol. 152, pp. 181–198, Jul. 2016, doi: 10.1016/j.cognition.2016.03.020.

[4] B. Copits*,* M. V. Valcheva, S. Davidson, T. Sheahan, “(344) Human sensory neurons: a novel strategy for translational approaches to pain and itch,” *The Journal of Pain*, vol. 17, no. 4, pp. S61–S62, Apr. 2016, doi: 10.1016/j.jpain.2016.01.251.

[5] J. H. Kaas, “The evolution of the complex sensory and motor systems of the human brain,” *Brain Research Bulletin*, vol. 75, no. 2, pp. 384–390, Mar. 2008, doi: 10.1016/j.brainresbull.2007.10.009.

[6] J. H. F. Bult, R. A. de Wijk, and T. Hummel, “Investigations on multimodal sensory integration: Texture, taste, and ortho- and retronasal olfactory stimuli in concert,” *Neuroscience Letters*, vol. 411, no. 1, pp. 6–10, Jan. 2007, doi: 10.1016/j.neulet.2006.09.036.

[7] Mihelj, M., Novak, D., Beguš, S.: Virtual Reality Technology and Applications. Springer Science & Business Media, Dordrecht, Netherlands (2014)

[8] P. Bourdot, T. Convard, F. Picon, M. Ammi, D. Touraine, and J.-M. Vézien, “VR–CAD integration: Multimodal immersive interaction and advanced haptic paradigms for implicit edition of CAD models,” *Computer-Aided Design*, vol. 42, no. 5, pp. 445–461, May 2010,

doi: 10.1016/j.cad.2008.10.014.

[9] B. Kapralos, A. Uribe-Quevedo, and A. Dubrowski, “Immersive Technologies for Medical Education,” in *Encyclopedia of Computer Graphics and Games*, N. Lee, Ed. Cham: Springer International Publishing, 2018, pp. 1–8.

[10] B. I. Edwards, K. S. Bielawski, R. Prada, and A. D. Cheok, “Haptic virtual reality and immersive learning for enhanced organic chemistry instruction,” *Virtual Reality*, vol. 23, no. 4, pp. 363–373, Dec. 2019, doi: 10.1007/s10055-018-0345-4.

[11] N. E. Seymour, et al., “Virtual reality training improves operating room performance,” Ann. Surg., vol. 236, no. 4, pp. 458–464, 2002

[12] P. Yu, J. Pan, H. Qin, A. Hao, and H. Wang, “Real-time suturing simulation for virtual reality medical training,” *Computer Animation and Virtual Worlds*, vol. 31, no. 4–5, p. e1940, 2020, doi: https://doi.org/10.1002/cav.1940.

[13] K. Rassie, “The apprenticeship model of clinical medical education: time for structural change,” vol. 130, no. 1461, p. 7, 2017.

[14] W. I. M. Willaert, R. Aggarwal, I. Van Herzeele, N. J. Cheshire, and F. E. Vermassen, “Recent Advancements in Medical Simulation: Patient-Specific Virtual Reality Simulation,” *World J Surg*, vol. 36, no. 7, pp. 1703–1712, Jul. 2012, doi: 10.1007/s00268-012-1489-0.

[15] P. Xia, A. M. Lopes, and M. T. Restivo, “Virtual reality and haptics for dental surgery: a personal review,” *Vis Comput*, vol. 29, no. 5, pp. 433–447, May 2013, doi: 10.1007/s00371-012-0748-2.

[16] C. Våpenstad, E. F. Hofstad, T. Langø, R. Mårvik, and M. K. Chmarra, “Perceiving haptic feedback in virtual reality simulators,” *Surg Endosc*, vol. 27, no. 7, pp. 2391–2397, Jul. 2013, doi: 10.1007/s00464-012-2745-y.

[17] K. Collins and B. Kapralos, “Pseudo-haptics: leveraging cross-modal perception in virtual environments,” *The Senses and Society*, vol. 14, no. 3, pp. 313–329, Sep. 2019, doi:

10.1080/17458927.2019.1619318

[18] M. Melaisi, D. Rojas, B. Kapralos, A. Uribe-Quevedo, and K. Collins, “Multimodal Interaction of Contextual and Non-Contextual Sound and Haptics in Virtual Simulations,” *Informatics*, vol. 5, no. 4, p. 43, Nov. 2018, doi: 10.3390/informatics5040043.

[19] C.-M. Wu, C.-W. Hsu, T.-K. Lee, and S. Smith, “A virtual reality keyboard with realistic haptic feedback in a fully immersive virtual environment,” *Virtual Reality*, vol. 21, no. 1, pp. 19–29, Mar. 2017, doi: 10.1007/s10055-016-0296-6.

[20] N. Karbasizadeh, M. Zarei, A. Aflakian, M. Tale Masouleh, and A. Kalhor, “Experimental dynamic identification and model feed-forward control of Novint Falcon haptic device,” *Mechatronics*, vol. 51, pp. 19–30, May 2018, doi: 10.1016/j.mechatronics.2018.02.013.

[21] M. Melaisi, M. Nguyen, A. Uribe, and B. Kapralos, “The effect of sound on haptic fidelity perception,” in Global Engineering Education Conference (EDUCON), 2017 IEEE, pp. 714–717, IEEE, 2017.

[22] X. Chen, P. Sun, and D. Liao, “A patient-specific haptic drilling simulator based on virtual reality for dental implant surgery,” *Int J CARS*, vol. 13, no. 11, pp. 1861–1870, Nov. 2018, doi: 10.1007/s11548-018-1845-0.

[23] B. Cowan, D. Rojas, B. Kapralos, F. Moussa, and A. Dubrowski, “Effects of sound on visual realism perception and task performance,” *Vis Comput*, vol. 31, no. 9, pp. 1207–1216, Sep. 2015, doi: 10.1007/s00371-014-1006-6.

[24] C. M. Serrano, P. R. Wesselink, and J. M. Vervoorn, “First experiences with patient-centered training in virtual reality,” *Journal of Dental Education*, vol. 84, no. 5, pp. 607–614, 2020, doi: https://doi.org/10.1002/jdd.12037.

[25] R. Bevilacqua *et al.*, “Non-Immersive Virtual Reality for Rehabilitation of the Older People: A Systematic Review into Efficacy and Effectiveness,” *J Clin Med*, vol. 8, no. 11, p. 1882, Nov. 2019, doi: 10.3390/jcm8111882.

[26] M. Scudellari, “How the pandemic might play out in 2021 and beyond,” *Nature*, vol. 584, no. 7819, Art. no. 7819, Aug. 2020, doi: 10.1038/d41586-020-02278-5.

[27] X. St-Denis, “Sociodemographic Determinants of Occupational Risks of Exposure to COVID-19 in Canada.,” *Canadian Review of Sociology*, vol. 57, no. 3, pp. 399–453, Aug. 2020.

[28] M. Mihelj, D. Novak, and S. Begus, “Introduction to Virtual Reality,” in *Virtual Reality Technology and Applications*, M. Mihelj, D. Novak, and S. Beguš, Eds. Dordrecht: Springer Netherlands, 2014, pp. 1–16.

[29] J. Simner, C. Cuskley, and S. Kirby, “What Sound Does That Taste? Cross-Modal Mappings across Gustation and Audition,” *Perception*, vol. 39, no. 4, pp. 553–569, Apr. 2010, doi: 10.1068/p6591.

[30] M. Nguyen, M. Melaisi, B. Cowan, A. J. Uribe Quevedo, and B. Kapralos, “Low-end haptic devices for knee bone drilling in a serious game,” *W J of Sci, Tech and Sus Dev*, vol. 14, no. 2/3, pp. 241–253, Apr. 2017, doi: 10.1108/WJSTSD-07-2016-0047.

[31] B. Kapralos, F. Moussa, K. Collins, and A. Dubrowski, “Fidelity and Multimodal Interactions,” in *Instructional Techniques to Facilitate Learning and Motivation of Serious Games*, P. Wouters and H. van Oostendorp, Eds. Cham: Springer International Publishing, 2017, pp. 79–101.

[32] M. Turk, “Multimodal interaction: A review,” *Pattern Recognition Letters*, vol. 36, pp. 189–195, Jan. 2014, doi: 10.1016/j.patrec.2013.07.003.

[33] B. Kapralos, “The senses and virtual environments,” p. 8.

[34] S. Reichelt, R. Häussler, G. Fütterer, and N. Leister, “Depth cues in human visual perception and their realization in 3D displays,” Orlando, Florida, Apr. 2010, p. 76900B, doi: 10.1117/12.850094.

[35] H. Knoblauch, A. Baer, E. Laurier, S. Petschke, and B. Schnettler, “Visual Analysis. New

Developments in the Interpretative Analysis of Video and Photography," *Forum : Qualitative Social Research*, vol. 9, no. 3, 2008, Accessed: Mar. 17, 2021. [Online]. Available: http://www.proquest.com/docview/869615944/abstract/20F89D55896140AEPQ/1.

[36] H. J. Kaplan, "Anatomy and Function of the Eye," in *Chemical Immunology and Allergy*, J. Y. Niederkorn and H. J. Kaplan, Eds. Basel: KARGER, 2007, pp. 4–10

[37] W. Wichmann and W. Müller-Forell, "Anatomy of the visual system," *European Journal of Radiology*, vol. 49, no. 1, pp. 8–30, Jan. 2004, doi: 10.1016/j.ejrad.2003.11.001.

[38] D. Loe, "Light, vision and illumination: The interaction revisited," *Lighting Research & Technology*, vol. 48, no. 2, pp. 176–189, Apr. 2016, doi: 10.1177/1477153515572240.

[39] S. S. Deeb, "The molecular basis of variation in human color vision," *Clinical Genetics*, vol. 67, no. 5, pp. 369–377, 2005, doi: 10.1111/j.1399-0004.2004.00343. x.

[40] C. Breuil, B. J. Jennings, S. Barthelmé, N. Guyader, and F. A. A. Kingdom, "Color improves edge classification in human vision," *PLoS Comput Biol*, vol. 15, no. 10, p. e1007398, Oct. 2019, doi: 10.1371/journal.pcbi.1007398.

[41] S. Mathôt and S. Van der Stigchel, "New Light on the Mind's Eye: The Pupillary Light Response as Active Vision," *Curr Dir Psychol Sci*, vol. 24, no. 5, pp. 374–378, Oct. 2015, doi: 10.1177/0963721415593725.

[42] S. Thorpe, D. Fize, and C. Marlot, "Speed of processing in the human visual system," *Nature*, vol. 381, no. 6582, pp. 520–522, Jun. 1996, doi: 10.1038/381520a0.

[43] B. Kapralos, M. R. Jenkin, and E. Milios, "Virtual Audio Systems," *Presence: Teleoperators and Virtual Environments*, vol. 17, no. 6, pp. 527–549, Dec. 2008, doi: 10.1162/pres.17.6.527.

[44] C. Alain, S. R. Arnott, S. Hevenor, S. Graham, and C. L. Grady, "'What' and 'where' in the human auditory system," *Proceedings of the National Academy of Sciences*, vol. 98, no. 21, pp. 12301–12306, Oct. 2001, doi: 10.1073/pnas.211209098.

[45] L. Riecke, I.-A. Marianu, and F. D. Martino, “Effect of Auditory Predictability on the Human Peripheral Auditory System.,” *Frontiers in Neuroscience*, p. NA-NA, Apr. 2020, doi: 10.3389/fnins.2020.00362.

[46] A. de Castro Rodrigues, A. L. Shinohara, J. C. Andreo, R. L. Buchaim, and F. J. Ahmed, “Surgical anatomy of the human middle ear: an insight into cochlear implant surgery,” *Surg Radiol Anat*, vol. 34, no. 6, pp. 535–538, Aug. 2012, doi: 10.1007/s00276-012-0947-6.

[47] N. Gerber *et al.*, “A multiscale imaging and modelling dataset of the human inner ear,” *Sci Data*, vol. 4, no. 1, p. 170132, Dec. 2017, doi: 10.1038/sdata.2017.132.

[48] K. Braun, F. Böhnke, and T. Stark, “Three-dimensional representation of the human cochlea using micro-computed tomography data: Presenting an anatomical model for further numerical calculations,” *Acta Oto-Laryngologica*, vol. 132, no. 6, pp. 603–613, Jun. 2012, doi: 10.3109/00016489.2011.653670.

[49] J.- Aucouturier, F. Pachet, and M. Sandler, “‘The way it Sounds’: timbre models for analysis and retrieval of music signals,” *IEEE Transactions on Multimedia*, vol. 7, no. 6, pp. 1028–1035, Dec. 2005, doi: 10.1109/TMM.2005.858380.

[50] A. M. C. Isaac, “Prospects for timbre physicalism,” *Philos Stud*, vol. 175, no. 2, pp. 503–529, Feb. 2018, doi: 10.1007/s11098-017-0880-y.

[51] F.-G. Zeng, Q.-J. Fu, and R. Morse, “Human hearing enhanced by noiseq,” *Brain Research*, p. 5, 2000.

[52] D. A. Hall and C. J. Plack, “Pitch Processing Sites in the Human Auditory Brain,” *Cereb Cortex*, vol. 19, no. 3, pp. 576–585, Mar. 2009, doi: 10.1093/cercor/bhn108.

[53] C. Batty, “The First Sense: A Philosophical Study of the Sense of Touch,” *Philosophical Psychology*, vol. 29, no. 1, pp. 138–146, Jan. 2016, doi: 10.1080/09515089.2015.1010197.

[54] M. Fulkerson, “The unity of haptic touch,” *Philosophical Psychology*, vol. 24, no. 4, pp.

493–516, Aug. 2011, doi: 10.1080/09515089.2011.556610.

[55] A. W. Goodwin and H. E. Wheat, “Physiological mechanisms of the receptor system,” in *Human Haptic Perception: Basics and Applications*, M. Grunwald, Ed. Basel: Birkhäuser, 2008, pp. 93–102.

[56] G. Robles-De-La-Torre, “The Importance of the Sense of Touch in Virtual and Real Environments,” *IEEE Multimedia*, vol. 13, no. 3, pp. 24–30, Jul. 2006, doi: 10.1109/MMUL.2006.69.

[57] A. Jaimes and N. Sebe, “Multimodal human–computer interaction: A survey,” *Computer Vision and Image Understanding*, vol. 108, no. 1–2, pp. 116–134, Oct. 2007, doi: 10.1016/j.cviu.2006.10.019.

[58] A. B. M. Gerdes, M. J. Wieser, and G. W. Alpers, “Emotional pictures and sounds: a review of multimodal interactions of emotion cues in multiple domains,” *Front. Psychol.*, vol. 5, Dec. 2014, doi: 10.3389/fpsyg.2014.01351.

[59] S. Malpica, A. Serrano, M. Allue, M. G. Bedia, and B. Masia, “Crossmodal perception in virtual reality,” *Multimed Tools Appl*, vol. 79, no. 5–6, pp. 3311–3331, Feb. 2020, doi: 10.1007/s11042-019-7331-z.

[60] Sekuler R, Sekuler AB, Lau R (1997) Sound alters visual motion perception. Nature 385(6614):308

[61] D. Moreno, M. Melaisi, S. MacDonald, A. Uribe-Quevedo, B. Kapralos, and M. V. Martin, “A brain-computer interface to examine the effects of sound on a haptic-based virtual drilling task,” in *2017 8th International Conference on Information, Intelligence, Systems & Applications (IISA)*, Larnaca, Aug. 2017, pp. 1–4, doi: 10.1109/IISA.2017.8316403.

[62] M. Ammi, B. F. Katz, and G, “Intermodal audio-haptic intermodal display: improvement of communication and interpersonal awareness for collaborative search tasks,” *Virtual Reality*, vol. 19, no. 3–4, pp. 235–252, Nov. 2015, doi: http://dx.doi.org.uproxy.library.dc-uoit.ca/10.1007/s10055-015-0273-5.

[63] A. Montuwy, A. Dommes, and B. Cahour, “Helping older pedestrians navigate in the city: comparisons of visual, auditory and haptic guidance instructions in a virtual environment,” *Behaviour & Information Technology*, vol. 38, no. 2, pp. 150–171, Feb. 2019, doi: 10.1080/0144929X.2018.1519035.

[64] C. George, P. Tamunjoh, and H. Hussmann, “Invisible Boundaries for VR: Auditory and Haptic Signals as Indicators for Real World Boundaries,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 12, pp. 3414–3422, Dec. 2020, doi: 10.1109/TVCG.2020.3023607.

[65] J. Lee, B. A. Gozen, and O. B. Ozdoganlar, “Modeling and experimentation of bone drilling forces,” *Journal of Biomechanics*, vol. 45, no. 6, pp. 1076–1083, Apr. 2012, doi: 10.1016/j.jbiomech.2011.12.012.

[66] Wu, G. Yu, D. Wang, Y. Zhang, and C. C. L. Wang, “Voxel-based interactive haptic simulation of dental drilling,” in 29th Computers and Information in Engineering Conf., Parts A and B, 2009, pp. 39–48.

[67] A. Parsian, M. Magnevall, T. Beno, and M. Eynian, “Sound Analysis in Drilling, Frequency and Time Domains,” *Procedia CIRP*, vol. 58, pp. 411–415, 2017, doi: 10.1016/j.procir.2017.03.242.

[68] H. Kinoshita *et al.*, “Development of a Drilling Simulator for Dental Implant Surgery,” *Journal of Dental Education*, vol. 80, no. 1, pp. 83–90, 2016, doi: https://doi.org/10.1002/j.0022-0337.2016.80.1.tb06061.x.

[69] Q. Wang *et al.*, “Haptic rendering of drilling process in orthopedic surgical simulation based on the volumetric object,” in *2015 IEEE International Conference on Digital Signal Processing (DSP)*, Jul. 2015, pp. 1098–1101, doi: 10.1109/ICDSP.2015.7252049.

[70] C. G. Corrêa, D. M. Tokunaga, E. Ranzini, F. L. S. Nunes, and R. Tori, “Haptic interaction objective evaluation in needle insertion task simulation,” in 31st Annual ACM Symposium on Applied Computing 2016, pp. 149–154.

[71] K. W. van Dongen *et al.*, “Virtual reality training for endoscopic surgery: voluntary or obligatory?,” *Surgical Endoscopy*, vol. 22, no. 3, pp. 664–7, Mar. 2008, doi: http://dx.doi.org.uproxy.library.dc-uoit.ca/10.1007/s00464-007-9456-9.

[72] B. L. Grant, P. C. Yielder, T. A. Patrick, B. Kapralos, M. Williams-Bell, and B. A. Murphy, “Audiohaptic Feedback Enhances Motor Performance in a Low-Fidelity Simulated Drilling Task,” *Brain Sciences*, vol. 10, no. 1, p. 21, Dec. 2019, doi: 10.3390/brainsci10010021.

[73] J. D. Bartlett, J. E. Lawrence, M. E. Stewart, N. Nakano, and V. Khanduja, “Does virtual reality simulation have a role in training trauma and orthopaedic surgeons?,” vol. 100, no. 5, p. 7, 2018.

[74] K. Khwanngern *et al.*, “Jaw Surgery Simulation in Virtual Reality for Medical Training,” in *Advances in Networked-based Information Systems*, vol. 1036, L. Barolli, H. Nishino, T. Enokido, and M. Takizawa, Eds. Cham: Springer International Publishing, 2020, pp. 475–483. doi: 10.1007/978-3-030-29029-0_45.

[75] K. S. Kulkarni, S. C. Yaragal, K. S. Babu Narayan, and H. Vardhan, “Assessment of thermally deteriorated concrete by drilling resistance test and sound level,” *Russ J Nondestruct Test*, vol. 53, no. 11, pp. 805–815, Nov. 2017, doi: 10.1134/S1061830917110031.

[76] M. Al-Balas, H. L. Al-Balas, H. M. Jaber, K. Obeidat, H. Al-Balas, E. A. Aborajooh, R. Al-Taher, and B. Al-Balas, “Distance learning in clinical medical education amid COVID-19 pandemic in Jordan: Current situation, challenges, and perspectives,” BMC Medical Education, vol. 20, no. 1, pp. 1-7, 2020.

Appendices

Appendix 1: Recruitment Email

Date to be sent:
Sender: Dr Bill Kapralos and Guoxuan Ning
Target audience: Current students at Ontario Tech University only
Subject line: Voluntary study: Appropriate Audio and Visual Cues to Simulate Drilling

Email copy:

This message is on behalf of Dr. Bill Kapralos (Principal Investigator) and Guoxuan Ning (Student Lead). Participation is entirely **voluntary** and there is no obligation or need to participate. Please direct inquiries to Guoxuan Ning, guoxuan.ning@ontraiotechu.net.

You are invited to participate in a voluntary research study, **Appropriate Audio and Visual Cues to Simulate Drilling**, which examines the simulation of a drilling task (drilling through a block wood using a drill) in the virtual domain using visual and auditory cues in addition to simple mouse movements. The experiment will be carried out remotely via the Google Meet platform, and using standard home computer equipment (e.g., mouse, keyboard, headphones, and monitor).

Participation in this research study will take approximately **20 minutes to complete**. Your task as a participant in this experiment will require you to perform a drilling task that involves drilling through a piece of wood to a pre-defined depth under various conditions including the use mouse movements, the presence of auditory and visual cues in a virtual environment. The experiment consists of 36 trials followed by an online questionnaire where you will be asked to answer 10 question about the experimental experience. **Since the experiment includes auditory cues, you must not have any issues with your hearing in order to participate in the experiment. Furthermore, If you are currently taking any course with Dr. Bill Kapralos or have never used a drill to drill through some material in the past, you are not eligible to participate in the experiment.**
Participation in this study is completely voluntary (you will not be paid for your participation), and this experiment does not have any bearing or influence on your physical health, privacy/reputation, and/or academic evaluation/standing. At any time during the study, you may decline to answer a question and may withdraw from the experiment altogether for whatever reason without any explanation or fear of repercussion. You may also withdraw from the experiment at any time within seven days of completing the experiment. If you do choose to withdraw from the experiment for any reason, your data will be deleted and not considered further. Your data will not be analyzed or viewed until seven days after the last participant has completed the experiment. Your data will be anonymized. Only Dr. Kapralos will have access to the data upon completion of this study. Every effort will be made on behalf of the facilitators to avoid any invasion of your privacy.
If you have any questions regarding your rights as a participant or have any concerns about this study, please contact the Research Ethics Office at researchethics@ontariotechu.ca. This study has been reviewed by the Ontario Tech University Research Ethics Board, assigned REB #16280 on 19th February 2021.

Download the study's consent form (the attaching file), which includes the names and information for the researchers, the purpose of the research, confidentiality, conflicts of interest, etc. If you are interested in participating or have any further questions, contact Guoxuan Ning, at guoxuan.ning@ontraiotechu.net.

Sincerely,
Dr. Bill Kapralos and Guoxuan Ning

Appendix 2: Consent Form

Consent Form to Participate in a Research Study

Title of Research Study: Appropriate Audio and Visual Cues to Simulate Drilling
Name of Principal Investigator (PI): Bill Kapralos
PI's contact number(s)/email(s): 905-721-8668 x2882 (or email: bill.kapralos@ontariotechu.ca)
Name(s) of Co-Investigator(s), Faculty Supervisor, Student Lead(s), etc., and contact number(s)/email(s):

Guoxuan Ning (Student Lead), guoxuan.ning@ontariotechu.net

Departmental and institutional affiliation(s): Faculty of Business and IT, Ontario Tech University

Introduction:

You are invited to participate in a research study (experiment) entitled *Appropriate Audio and Visual Cues to Simulate Drilling*. Please read the information about the study presented in this form below. This information includes details on the study's procedures, risks and benefits that you should know before you decide if you would like to take part in it. You should take as much time as you need to carefully read the provided information and make your decision. You should ask the Principal Investigator (PI) or study team to explain anything that you do not understand and make sure that all of your questions have been answered before agreeing to participate and signing this consent form. Before you make your decision, feel free to talk about this study with anyone you wish including your friends and family. Participation in this study is voluntary.

This study has been reviewed by the University of Ontario Institute of Technology (Ontario Tech University) Research Ethics Board, assigned REB #16280 on 19th February 2021.

Purpose and Procedure:

Purpose:

Current virtual simulations and serious games focus primarily on cognitive skills development, typically ignoring psychomotor/technical skills development altogether, given the complexities and costs associated with simulating the sense of touch. Although consumer-grade haptic devices (used to simulate the sense of touch) are readily available, they are restrictive with respect to their available degrees-of-freedom (DOF), the electro-mechanical characteristics of sensors and actuators they include. These limitations do not allow such devices to be used in many haptic-based applications, such as applications where high-fidelity haptic interaction is essential, including medical education. Further complicating matters is the current COVID-19 pandemic and the resulting lockdowns and shutdowns of colleges and universities, making access to laboratories and simulation facilities where such haptic devices are housed difficult, if not impossible. Therefore, we aim to explore whether a haptic based virtual environment could be simulated using only the appropriate auditory and visual cues, as well as common computer devices in the home (headphones, mouse, monitors, etc.) in the absence of haptic devices.

In this study, we are focusing on a drilling task, and more specifically, the simulation of drilling through a block of wood using a virtual drill. The purpose of this study/experiment is to examine the simulation of the drilling task using auditory and visual cues, and simple mouse movements which can be delivered with common computer devices available to most people in their home (e.g., headphones, mouse, etc.). In order to participate in this study/experiment, we require that you have prior experience using a drill and that you do not have any problems/issues with your hearing (at least no issues that you are aware of).

Procedures:

Since the experiment will be conducted virtually/remotely, you must have a mouse, headphones/earbuds, and a monitor/display with the resolution set 1024 x 768 to complete the experiment. The experimenter will meet you remotely via Google Meet (https://meet.google.com/) and provide a brief verbal explanation of the experiment and required tasks to you via this online meeting. At the same time, you will receive an executable file (the experiment application developed using the Unity 3D game engine) via email from the experimenter. Prior to the formal experiment, you will start by completing a training session to familiarize yourself with the process (i.e., the drilling simulation). This training is divided into two parts, each focusing on simulated drilling using i) visual cues and drill movement method, ii) sound cues. Each part will be repeated twice. The training session will involve the following: i) you will be presented with a visual top-view of the virtual drilling scenario and you will be asked to move the drill into the block of (virtual) wood by moving the mouse. The drill will stop when you reach a depth of 5 cm, at which point you will be notified. You will then repeat this task but for a drill depth of 10 cm. Then, you need to repeat this task using a keyboard to control the drilling movement. More specifically, the drill will move forward automatically without mouse control and you will need to press the up-arrow bar on the keyboard to activate the virtual drill and press the space bar to stop it when you are notified that the drilling depth reaches 5 cm. You will then repeat this task but for a drill depth of 10 cm. No auditory cues are provided here. ii) You will be presented with a rear view of the drilling scenario. In this scenario, the drill will go through the block of wood automatically (the drill will be moved without you having to move the mouse), but you will be presented with auditory cues. You will be notified with a textual message on the screen when the drill reaches a depth of 5 cm. This will then be repeated, but with a drill depth of 10 cm.

After completing the two training sessions, you will start the formal experiment phase. The actual experiment will ask you to complete a total of 36 trials. For each trial, you will be asked to drill to a depth of 12 cm (depth will remain constant throughout the experiment) under various visual, auditory, and mouse-movement conditions. More specifically, the visual conditions are i) visual cues, ii) non-visual cues. The mouse-movement conditions are i) mouse movement, ii) no mouse movement (drill moves automatically without moving the mouse). The three auditory conditions include i) no sound, ii) continuous drilling sound (the sound associated with operating the drill in the air under no load), iii) dynamic drilling sound where the sound corresponds to a drill drilling through wood. In total, there are 12 conditions (2 visual conditions x 2 movement conditions x 3 auditory conditions), and each condition will be repeated three times for a total of 36 trials. The 36 trials will be presented to you randomly. Feedback regarding your performance will not be offered to you. For each trial, when you complete the drilling task and releases the mouse (in the mouse movement condition) or when you hit the space bar (in the no mouse movement condition) to stop the drill, the current drilling depth (the drill bit into the wood) is recorded. A text-formatted will be generated on your computer. This file will include the experimental conditions (visual, auditory, movement) used for each trial and the drill depth you drilled into the wood board under these settings. At the end of each trial, information regarding the trial will be recorded in the file. After completion of the experiment, the experimenter will instruct you to email the file back to them. After completing the testing stage, you will be asked to complete an online questionnaire. The total estimated time to complete the experiment to be approximately 20 minutes.

Potential Benefits:

By participating in this project you will gain an understanding and appreciation for the work being completed by graduates in the Computer Science program at Ontario Tech University. Furthermore, you may gain new knowledge on immersive technologies in general. You may also find the study rewarding, given that the work may be a tool used to provide better, more effective educational tools that will ultimately lead to better-trained professionals.

Potential Risks or Discomforts:

You will be asked to use a pair of headphones, mouse, and monitor to complete this experiment. During this experiment, drilling sounds will be presenting, and sounds will be presented to them over headphones. The sound level will be set to 65 dB initially (the level of typical conversation which falls between the range of

60-70 dB) and you will be asked to adjust the volume of their computer to a comfortable level. We do not anticipate any issues with this, but you will be able to remove yourself from the experiment at any point without any repercussions.

Use and Storage of Data:

The data you generate in the experiment will be collected in order to analyze the impact of auditory and visual feedback in a virtual reality environment. Your data will be anonymized. Anonymized data implies that identifying information is completely removed and your responses/results will only be identified with a code not linked to you personally. All anonymized raw data will be collected by Student Lead Guoxuan Ning on his computer. After analyzing the data and drawing any conclusions, Guoxuan will send the original (anonymized) data to Dr. Bill Kapralos via email. Then, this data will be kept in an encrypted zip folder on Bill Kapralos' computer, which further requires a password to log in only known by Professor Kapralos. It will also be "backed-up" on a storage disk drive (owned/maintained by Dr. Bill Kapralos). This back-up disk is currently kept in Bill Kapralos' home office in a secure filing cabinet and is not accessed by anyone else. At the end of the process, Guoxuan will delete all the raw data about the experiment from his computer to ensure that the data only exists on Dr. Bill Kapralos' computer and the storage disk drive. Please note that although confidentiality cannot be guaranteed while data is transferred over email, your data will be anonymized when transferred and therefore, you cannot be identified.

All information collected during this study, including your personal information (e.g., your signature on the consent form) will be kept confidential and will not be shared with anyone outside the study unless required by law. You will not be named in any reports, publications, or presentations that may come from this study.

Confidentiality:

Your privacy shall be respected. No information about your identity will be shared or published without your permission, unless required by law. Confidentiality will be provided to the fullest extent possible by law, professional practice, and ethical codes of conduct. The experiment will be conducted by a graduate student working under the supervision of Dr. Bill Kapralos, and any information collected will be used to develop a thorough understanding of the interaction of sound and haptic fidelity perception. All data will be anonymized and will be kept by Dr. Bill Kapralos on his computer and backed-up on a hard disk that is accessed only by Dr. Bill Kapralos and stored in a secure filing cabinet. Every effort will be made on behalf of the experimenters to avoid any invasion of your privacy.

Voluntary Participation:

Your participation in this study is voluntary and you may partake in only those aspects of the study in which you feel comfortable. You may also decide not to be in this study, or to be in the study now, and then change your mind later. You may leave the study at any time without affecting your grades in a course. You will be given information that is relevant to your decision to continue or withdraw from participation. Such information will need to be subsequently provided. You may refuse to answer any question you do not want to answer, or not answer an interview question by saying, “pass”.

Right to Withdraw:

At any time during the experiment, you may decline to answer a question and may withdraw from the study altogether at any point for whatever reason without any fear of repercussion. You may also choose to withdraw after completing the study. If you choose to withdraw, you may do so by letting the experimenter know that you wish to withdraw. If you withdraw from the research project at any time, any data or human biological materials that you have contributed (via the participant number) will be removed from the study and you do not need to offer any reason for making this request."

Conflict of Interest:

Researchers have an interest in completing this study. Their interests should not influence your decision to participate in this study." If you are currently taking any courses with Dr. Bill Kapralos, or taking any courses in which, the Student Lead/experimenter (Guoxuan Ning) is the Teaching Assistant for during the experimental period (or you know you will be taking a course with Dr. Bill Kapralos in the future), you will not be invited to participate in the experiment.

Compensation, Reimbursement, Incentives:

There will be no compensation for participating in the experiment as participating in the experiment is completely on a voluntary basis (i.e., you will not receive any monetary reward at the end of the experiment). If you opt out, you will not suffer any financial consequences.

Debriefing and Dissemination of Results:

The results of the experiment will eventually be reported in the defense of Student Lead Guoxuan Ning's thesis in the form of a presentation and presented in Guoxuan's Master's thesis. If you find the information obtained from this experiment interesting, you can request a copy of the final report by emailing Dr. Kapralos (bill. kapralos@ontariotechu.ca) at any time.

Participant Rights and Concerns:

Please read this consent form carefully and feel free to ask the researcher any questions that you might have about the study. If you have any questions about your rights as a participant in this study, complaints, or adverse events, please contact the Research Ethics Office at (905) 721-8668 ext. 3693 or at researchethics@ontariotechu.ca. If you have any questions concerning the research study or experience any discomfort related to the study, please contact Dr. Bill Kapralos at 905-721-8668 x2882 (or email: bill.kapralos@ontariotechu.ca).

By signing this form, you do not give up any of your legal rights against the investigators, sponsor or involved institutions for compensation, nor does this form relieve the investigators, sponsor or involved institutions of their legal and professional responsibilities."

Consent to Participate:

I (please print your participant number here), participant number ________________ have read the consent form and understand the study being described; I have had an opportunity to ask questions and those questions have been answered. I am free to ask questions about the study in the future; I freely consent to participate in the research study, understanding that I may discontinue participation at any time without penalty. A copy of this consent form has been made available to me.

I understand that I have been selected to participate in this study. Participation involves me participating in the study described above. I understand that my honesty and openness is very important to further development of the virtual drilling simulation described earlier. I also understand that as a participant in this experiment, I am not waiving any of my legal rights. I agree to participate in this study and will keep a copy of this consent form for my personal records.

Participant Signature: ______________________________, Date: ______________________________.

Witness Signature: ________________________________, Name: _________________________________,

Appendix 3: Questionnaire

Simulate a Drilling Task with Visual, Audio, and Mouse Movement Cues

Section 1: General Demographic Information

1. Gender (please enter here): __________.

2. Age (mark only one circle):
   - ○ Under 18
   - ○ 18-20
   - ○ 21-25
   - ○ 25-30
   - ○ Over 30

Section 2: Drilling Scenario Information

3. The drilling scenario was understandable (mark only one circle).

	1	2	3	4	5	6	7	
Strongly Disagree	○	○	○	○	○	○	○	Strongly Agree

4. The drilling scenario was enjoyable (mark only one circle).

	1	2	3	4	5	6	7	
Strongly Disagree	○	○	○	○	○	○	○	Strongly Agree

5. The drilling task was easy to learn (mark only one circle).

	1	2	3	4	5	6	7	
Strongly Disagree	○	○	○	○	○	○	○	Strongly Agree

6. The virtual drill was easy to operate (mark only one circle).

	1	2	3	4	5	6	7	
Strongly Disagree	○	○	○	○	○	○	○	Strongly Agree

7. The visual cues were an important part of the virtual drilling task (mark only one circle).

	1	2	3	4	5	6	7	
Strongly Disagree	○	○	○	○	○	○	○	Strongly Agree

8. The auditory cues were an important part of the virtual drilling task (mark only one circle).

	1	2	3	4	5	6	7	
Strongly Disagree	○	○	○	○	○	○	○	Strongly Agree

9. The mouse movement cues were an important part of the virtual drilling task (mark only one circle).

	1	2	3	4	5	6	7	
Strongly Disagree	○	○	○	○	○	○	○	Strongly Agree

Section 3: Open

10. Please provide any comments regarding the virtual drilling tasks you may have: