TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technologies

Daud Suleman IASM194332

Measuring Audio Visual Elements in Immersive Virtual Reality

Master's thesis

Supervisor: Saleh Ragheb Saleh Alsaleh

Msc

Early Stage Researcher

Co-supervisor: Aleksei Tepljakov

PhD

Senior Researcher

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond

Daud Suleman IASM194332

Audiovisuaalsete efektide mõju hindamine immersiivses virtuaalreaalsuse keskkonnas

Master's thesis

Juhendajad: Saleh Ragheb Saleh Alsaleh

Msc

Early Stage Researcher

Aleksei Tepljakov

PhD

Senior Researcher

Author's declaration of originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature, and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author: Daud Suleman

Acknowledgments

The author of this thesis would like to sincerely express his kind appreciation to his supervisor Saleh Ragheb Saleh Alsaleh, MSc, and co-supervisor Dr. Aleksei Tepljakov for their encouragement, responsiveness, guidance, and supervision throughout the whole thesis work. This study was performed in Alpha Control Lab, which is within the Centre for Intelligent Systems, Department of Computer Systems, Tallinn University of Technology. My big appreciation goes to the people working in Alpha Control Lab for making things easier to learn.

Abstract

Though virtual reality (VR) is a promising approach for psychological research, there is still a lot to learn about how different aspects of VR environments impact people [1]. The primary goal of this research is to analyze the two primary characteristics of virtual reality audio and visual aspects. While there is some research demonstrated about immersion, it is still uncertain how broadly the effect of scale and audio can affect the level of immersion.

To solve this challenge, a digital twin of a laboratory-scale replica of a gantry crane (referred to as “3D crane”) was developed. The developed digital twin was complemented with audio features (mechanical noise) and the ability to scale its physical dimensions. The goal of constructing a bigger scale digital twin of this crane was to do experiments with subjects to see how scale affects immersion. The laboratory-scale replica of a gantry crane was utilized to record actual sound as it moved, which was then applied to the digital twin built in Unreal Engine.

Different types of experiments were performed; participants interacted with a 3D crane in a virtual environment with audio, without audio, and with the varied scale of the 3D crane.

The participants of the experiment were also asked to fill a specialized survey. The corresponding results are discussed in the present thesis.

Keywords Digital Twin, Presence, Virtual reality, Immersion, Audio-Visual, Human Behaviour,

This thesis is written in English and consists of 84 pages, including 7 chapters, 27 figures and 16 tables.

List of abbreviations and terms

DT	Digital Twin
VR	Virtual Reality
HMD	Head-mounted Display
BP	Blueprint
UE	Unreal Engine
PQ	Presence Questionnaire
VE	Virtual Environment
IVR	Immersive Virtual Reality
IQR	Interquartile Range
XR	Extended Reality
SD	Standard Deviation

Table of contents

List of figures

List of Tables

1 Introduction

Virtual reality (VR) is an advanced technology that has the capability to transform our lives in ways which no other technologies have. Our bodies are tricked into adopting a different version of reality by artificially boosting our senses. VR is like a waking dream that might send us to another area of the Earth or space, or it could transfer us to a wonderful cartoon-like environment. It's the next stage in a series of steps that covers anything from paintings to movies to video games. We can even interact with others in a new environment, which may or may not be real or fictional [2].

The level of engagement a person can have when in a VR environment is referred to as immersion [1][14]. According to Brown et al. [14], immersion may be viewed on three levels. They are absolute immersion, engrossment, and involvement. The term “engagement” refers to a basic level of immersion that is defined as being stimulated by VR environment involvement. Engrossment is a medium degree of immersion in which the person connects emotionally with the characters. The moment at which the person constructs an alternate reality in the VR environment is called total immersion, or a high level of immersion [15]. As a result, it's critical to figure out a good approach to evaluate immersion. This study tries to move in that direction [1].

According to the literature, immersion can be measured objectively or subjectively. To calculate eye movements and bodily reactions, the objective technique uses an eye-tracking equipment or a galvanic skin response device [16]. This information from the devices will be used to determine immersion. Subjective methods include using post-experiment questionnaires and interviews to extract participants' perceptions and then measuring immersion [14],[15]. It has been suggested that combining the two techniques might result in a more accurate assessment of immersion [16]. This study adds to the enhancement of subjective immersion measurement [1].

Research performed by Stanislav Jeršov and Aleksei Tepljakov as a part of their activities in XR laboratory, Department of Computer Systems, Taltech, created an extended reality digital twin for a control object, complete with the essential interaction mechanics, and connected it to the real plant for real-time hardware-in-the-loop (HIL) tests. The theoretical foundations and technical development are described in-depth in the paper, as well as the results of an illustrated experiment, which are presented and discussed. And describes DTs as software implementations of their physical counterparts that interact with, monitor, and perhaps control physical components via Application Program Interfaces (API) [8].

The research team look into the prospect of employing a data-driven strategy that allows researchers to get insight into user behavioural data. A virtual reality data replayed and annotation approach is presented, which enables the analysis and classification of VR acquired data using a graphical user interface. This method is used to analyze data from VR lab research that included people with varying levels of VR experience. Finally, the system is employed to detect undesired user contact, and machine learning methods are studied as a potential option for the classification of user interactions. The authors demonstrated a replay and annotation system that allows data-driven approaches to be employed instead of traditional rule-based algorithms in virtual reality. Two examples of the technology being effectively applied to data from a virtual reality study were demonstrated. The technology was first employed in the VE to recognize, categorize, and acquire information from users' object-grabbing actions. Based on this new knowledge, an alternative algorithm was developed that increased the rate of success of grabbing actions by around 23%. In the second case, the system was utilized for training an ML classifier that can accurately anticipate a user's throwing/moving actions to the tune of 95%. Authors believe that strategies similar to those provided in this research will aid in the development of the next generation of data-driven VEs capable of successfully reusing data obtained from experiments [43].

1.1 Overview of VR

Lavalle, a professor in the Department of Computer Science at the University of Illinois in Urbana-Champaign, came up with a useful definition of VR. “Using artificial sensory stimulation to induce targeted behavior in an organism while the organism has minute or no awareness of the interference.”

In the definition, there are four main elements.

Targeted behaviour: The organism is going through an “experience” that the designer has planned for it. Flying, strolling, exploring, watching a movie, and interacting with other creatures are just a few examples.

Organism: You, someone else, or even a fruit fly, cockroach, fish, rodent, or monkey (scientists have utilized VR technology on all of these!) might be the organism.

Artificial sensory stimulation: One or all of the organism's senses are co-opted, at least in part, by engineering, and their normal inputs are modified or boosted by sensory pleasures.

Awareness: During the experience, the organism appears to be ignorant of the disturbance, leading to a feeling of being involved in a virtual world. This state of unawareness leads to a perception of being present in a different or alternate reality. It is thought to be natural.

Terminology for various "realities" Immanuel Kant, a German philosopher, invented the term virtual reality [3], despite the fact that it was not used with technology. The word was invented by Kant to distinguish the "reality" that exists in one's mind from the exterior physical world, which is also a reality. Jaron Lanier introduced the phrase "virtual reality" in the 1980s. However, the word virtual reality appears to be self-contradictory, which is a philosophical challenge solved in [4] by suggesting the term *virtuality* as an alternate term. Several competing terminologies connected to virtual reality are already in use, despite the fact that the word VR is already extremely broad. Most university academics prefer the phrase *virtual environments*, which predates the extensive use of VR [2][5].

1.2 Overview of the Digital Twin concept

Digital Twin technology is a relatively new technology that has captured the interest of business and, more recently, academics. Advances in industry 4.0 principles have supported its expansion, notably in the manufacturing sector. The establishment of an interconnected physical and virtual twin (Digital Twin) helps address the challenges of seamless connectivity between IoT and data analytics. Instant assessment and real-time choices may be made in a Digital Twin environment by reliable analytics [6].

"A comprehensive multi-physics, multi-scale, deterministic simulation of a vehicle or system that utilizes the best possible prototype, sensor feeds, fleet records, etc., to replicate the life of its fly twin," NASA defined the DT first. It is ultra-realistic and may involve one or more essential and interrelated vehicle systems: this definition made its first appearance in the draft then in NASA Modelling, Simulation, Information Technology, and Process Roadmap drafting in 2010[7]. In research [9], it is defined as the simulations of a real object in order to forecast future system states [9]. In another paper, it is defined as a unifying framework that can integrate models in diverse supplier tools and configuration-controlled repositories to manage architecture, mechanical, electrical, software, verification, and another discipline-specific modelling all over the life cycle of the system [10].

DTs can be handled and configured in real-time from a computer, and they can make operational adjustments in the physical thing if needed [11][12]. The physical twin can employ knowledge via digital twins to enhance its qualities as a result of two-way interaction between the digital and actual objects [12].

1.3 Problem Statement

There is no doubt that virtual reality (VR) is a potential tool for psychological study. There's still a lot to learn on how various aspects of VR sceneries affect people. Although much study has been done on various aspects of virtual reality, more research is still needed on two aspects of virtual reality, namely audio and scale (visual). According to research [13], audio does not play a significant role in the VR world since the user is already immersed in the sensory environment [13]. It is preferable to combine these two aspects and conduct experiments on actual people in order to obtain more precise results.

The goal of this research is to determine the effects of audio-visual elements on the level of presence in the IVR environment.

1.4 Structure of Thesis

The current thesis is organized in such a manner that each following section describes a new element of progress while taking into account and referencing prior sections, allowing to follow a logically consistent development path.

Chapter 1

The first chapter introduces the current thesis and outlines the goals that must be met as well as the challenges to be overcome in order to accomplish the stated aim.

Chapter 2

A complete literature review is presented in the second chapter, laying the groundwork for future study and development.

Chapter 3

The experimental system is described in detail in this chapter that has been constructed, including descriptions of the key components, their interconnections, and purposes, as well as a complete workflow.

Chapter 4

The fourth chapter's main goal is to describe the process of creating a virtual reality environment, development of questionnaires.

Chapter 5

This chapter discusses the methods used to conduct VR experiments and collection of data.

Chapter 6

The sixth chapter illustrates the results and comparison of those results.

Chapter 7

The last chapter draws a conclusion, summarizing all that has been accomplished during the development process.

2 Literature Review

The definition of a literature review is the synthesis and analysis of prior research in a certain topic area. It is used to describe, summarize, analyze, and integrate the material of primary studies throughout the early stages of a research project. It is carried out to learn how research in the chosen field has progressed, and this information aids in the establishment of new research that advances prior research [39].

Several research works have already been done in the field of virtual reality immersion. Many ideas and models have been developed in virtual reality by the use of different technologies to analyze the audio-visual effects of the proposed model. This section is a review of recent interesting developments and research on the audio and visual elements of virtual reality immersion.

2.1 VR Technologies

VR has been around since the 1950s. Morton Heilig [17], also known as the "Father of Virtual Reality," was a U.S Born cinematographer. Sensorama, a single-user entertainment console that he developed in the 1960s. It gathered data from a variety of sensors. The output includes stereo speakers, a stereoscopic screen, fans, a vibrating chair, and an aroma emitter. Users might obtain an immersive environment similar to riding a bike along a Brooklyn street using the mix of sensor outputs. It was a simple, immersive technology for the users with many restrictions, but it was a tremendous endeavor, to begin with [18]. VR has been around since the 1950s. Morton Heilig [17], also known as the "Father of Virtual Reality," was a U.S Born cinematographer. Sensorama, a single-user entertainment console that he developed in the 1960s. It gathered data from a variety of sensors. The output includes stereo speakers, a stereoscopic screen, fans, a vibrating chair, and an aroma emitter. Users might obtain an immersive environment similar to riding a bike along a Brooklyn street using the mix of sensor outputs. It was a simple, immersive technology for the users with many restrictions, but it was a tremendous endeavor, to begin with [18].

Following that, plenty of inventive firms and people set out to create their own VR gadgets. The head-mounted display (HMD) appears to be a recent breakthrough for virtual reality missions, although this is not the case. Philco Corporation invented the first iteration of this gadget in 1961. In 1968, Iva Sutherland, a well-known computer scientist, invented the finest display called "Sword of Damocles." In terms of UI (user interface) and realism, it was a basic kind. The technology used the output from the computer software to create a stereoscopic display. Tracking the user's head might modify the user's perception as soon as they rotate their head. The apparatus was so weighty that it was held in place by a mechanical arm suspended from the ceiling. Jason Lanier, a computer philosophy writer, contributed positively by making the word virtual reality available to our common language when it was first used in 1989.

Nevertheless, the principle remained the same: computer science developed a virtual reality system that people could control and experience. Virtual reality is now applied in a variety of sectors. Entertainment, health, athletics, and construction are the most commonly used fields. Security officials are also utilizing it to teach their armed services and aviation personnel [18].

Nonetheless, the application areas are not restricted; in fact, they are expanding. Virtual reality apps are being used for a variety of fascinating and productive purposes. VR, for example, has been shown to reduce work-related sadness and stress [19]. Virtual reality, on the other hand, is typically a representation of a real-world environment that can be classified as artificial simulations. It improves a visual condition or reality. For instance, (mimicking a character in a Minecraft gaming adventure) [20]. "Artificial simulations, generally recreations of real-life environments, that improve an imagined reality or scenario" is one of the most common definitions of VR [21]. Burdea and Coiffet provide a more precise definition: "a simulation in which computer graphics are employed to generate a realistic-looking world."

2.2 Aspects of VR

Some of the most important aspects of a virtual reality experience are as follow.

Virtual world A virtual environment that exists in its own way, separate from the real world. The medium employed to construct this area is, of course, a computer-generated simulation made up of visual components. The creator's rules establish the relationships and interactions between these aspects [22].

Immersion On a perceptual level, the users are immersed in a virtual area that is cut off from the actual world. VR headsets do this by dominating the user's entire field of view, whilst headphones provide the same effect with sound, completely immersing the user in another environment.

Sensory Feedback Virtual reality headsets monitor the user's location inside a specific environment, allowing the computer to depict an object's position. Users will get the sensation that they are walking in the virtual environment by moving their heads or body. The input is as realistic as possible; people do not push a button to go around.

Interactivity A simulation world should have virtual components that we can engage with in order to seem real: trying to pick up an item, striking an axe to kill a monster, smashing a mug, pressing the ignition button of an aeroplane in a virtual simulation, and so on [22].

2.3 Immersion and presence in VR

Immersion and presence are two of the most important aspects of virtual reality. The following variables are in charge of establishing the sensation of being within the simulation. Steuer even claims that the experience of being there is a key aspect of virtual reality [53]. Despite the fact that immersion and presence are commonly used interchangeably in literary works, they are treated as distinct concepts in academics. In a VE, presence is seen only as a subjective psychological state. Immersion, on the other hand, is thought to be a description of hardware and software capabilities. One of the most important factors that leads to the cognitive state of presence is immersive stimuli.

Various studies have presented methods for determining the level of presence based on a variety of hypotheses [49]. After that, there was a summary presented by [50] a range of possible factors. The following is the list:

Slater and Usoh [51]:

- Information of excellent quality and resolution.
- Uniformity in all displays.
- Interaction with the natural world.
- Virtual body, which is the VE's depiction of the user's body.
- It is necessary to anticipate the outcome of an activity.

Witmer and Singer [45]:

- Sensory aspects - include the depth and uniformity of the exhibited information.
- Distraction factors - The amount of time the user is diverted from the VE.
- Factors affecting realism in the VE include visual and social realism.
- Users' control factors.

Sheridan [52]:

- The amount of sensory data available.
- Controlling the relationship between sensors and their environment.
- The ability to change one's physical surroundings.

Steuer [53]:

- Technology's capacity to create a sensory-rich mediated environment is referred to as vividness.
- The ability of users of a medium to change the form or content of the mediated environment is referred to as interactivity.
- Individual variances among users are referred to as user characteristics [49].

The list above is simply meant to serve as an example of a number of viable explanations for presence. The majority of today's research is done through subjective questionnaires. As a result, a variety of data insufficiencies prevent the development of a basic method, as the research described above also found [50].

2.4 Background and Work-Related

Following are recent interesting developments and research on the audio and visual elements of immersive virtual reality.

Effects of innovative technologies on the sensation of immersion

This research paper by authors of [41] presents a comparison of how these innovative displays affect gaming Quality of Experience (QoE) in comparison to traditional 2D PC screens. They show that Oculus enhances the sensation of immersion in the 3D world as well as perceived usability in an experiment where they offered 22 people to control a virtual forklift driving in both contexts. This platform (Oculus), which increases levels of wonder, astonishment, and exhilaration, has a significant influence on affective variables. Interesting findings have been obtained after evaluating the data collected through questionnaires in terms of three main QoE factors: perceived presence, perceived usability, and emotions. The Oculus Rift enhances the sensation of immersion in the virtual world. Users sense increased realism and naturalness. More research should determine how much of the good feelings produced by Oculus are a result of the novelty and uniqueness of a gadget that is not yet part of our daily lives. Longitudinal QoE(Quality of Experience) tests can be used to get answers to this question. In studies, physiological sensors (EEG (electroencephalogram) and ECG (electrocardiogram) signals) can be utilized to provide automated measurements of brain cognitive load and emotions [41].

Examining audio-visual effects of the proposed project on subjects in Immersive VR.

A highway project was proposed to examine audio–visual effects of a proposed highway project on individuals was examined using immersive VR technology in this study. Participants were shown 3D reconstructions of a realistic landscape but without the proposed highway (ante operam condition) and the same landscape with the proposed motorway (post operam condition). Individuals' responses to noise were also examined using objective cognitive tests (short-term verbal memory and executive functioning) as well as subjective evaluations (noise and visual annoyance). However, there is no research on the characteristics of road traffic. Here it is needed to identify what are the special characteristics of road traffic noise that cause it to have such a quick negative impact on short-term verbal memory. Also, not enough research on the difference between unimodal (audio or video only) and multisensory method presentations [23].

Identifying the influence of audio in Immersive VR experience

The research carried by Rogers Katja et al. l tried to address a gap in the research on the influence of audio on immersive VR experiences. They carried out two studies: (I) a 12-person within-subjects experiment involving a commercial horror-adventure game to compare the impact of virtual reality and monitor-display versions of the same game on player experience (PX). (II) 40-person between-subjects research on the impact of audio dimensionality on PX in VR. With comparison, they found that audio does not play an effective role in VR as participants are already occupied by the experience of sensors. Overall, their research shows that visual and creative factors are important in HMD-VR. On the other hand, in monitor display conditions, participants focus on game progression. This research would be more trustworthy if the participants were more expert VR users, and to make it more interesting, audio can be customized for participants according to their preference, as some people prefer background music to ambient noises [13].

Investigation of visual and auditory elements of a wind farm on participants in Immersive VR

The research by authors investigated some visual and auditory elements of the influence of a wind farm on a sample of participants and was examined and evaluated using the Immersive Virtual Reality method. The participants were engaged in a virtual environment that simulated a typical rural outdoor scene that they encountered at various distances from the wind turbines. The impact of wind turbine number and color on global, visual, and aural judgement was examined. The main findings revealed that the visual component has a minimal effect on individual reactions when it comes to the number of wind turbines, but color affects both visual and aural individual replies. Researchers focus on three factors: distance, color, and no of turbines, but there is no research on type weather, windy or rainy, etc. will have different effects on individuals' reactions [26].

Overview of spatial audio recording and reproduction techniques

The authors of this research paper did research to study and give a broad overview of several spatial audio recording and reproduction techniques that may be used in soundscape studies and applications. Researchers studying soundscapes should be aware of the advantages and disadvantages of various spatial audio recording and reproduction techniques and use the most appropriate approach for their research. In particular, new VR/AR technologies, in combination with enhanced spatial audio recording and reproduction techniques, offer a more engaging and immersive auditory and visual experience and would be an excellent fit for soundscape designing with high ecological validity. However, no study has been conducted on the spatial aspects of soundscape design. To build more realistic soundscape models, further study on spatial elements is needed [27].

Audio rendering method to achieve audio quality, spatial accuracy, and performance in VR

In this paper, the authors show a spatialized audio rendering method that may be used in virtual reality. Using off-the-shelf audio hardware, the system is designed for generating a sufficient number of dynamically shifting sound sources in multi-speaker setups. Using simple physics-based models, researchers were able to achieve a reasonable balance of audio quality, spatial accuracy, and performance. Integrating common hardware reverberation devices used in the professional audio and broadcast allow access for convincing acoustic room modelling. There is no research on different sound emission characteristics based on the angle of an item towards the listener using directional sound sources [28].

Analyzing noise-related features of various barriers with different visual qualities in immersive VR

This research provided a case study in which a group of people living near a railway line used an IVR laboratory test to analyze noise-related features of various barriers with varying visual qualities. Three main factors were investigated: the barrier type, which affected the noise source's visibility through the screens; the barrier's visual aspect, which affected some aesthetic concerns; and the noise level at the receiver, which affected the barrier's acoustic performance and the magnitude of the sound source. The study's main findings indicated that transparent barriers were evaluated to have lower Perceived Loudness and Noise Annoyance than opaque barriers. No research on the noise level at source and different types of tests can be added to test the normal hearing and vision of participants [29].

Investigation about spatial audio recording and reproduction techniques

In this research, paper authors have discussed spatial audio recording and reproduction strategies in this article. Binaural recording and rendering, sound field recording and reproduction, and multi-zone reproduction are among the approaches that have been investigated. Binaural audio, which uses a pair of headphones or a few loudspeakers to control sounds within a region, is more easily integrated into personal audio products, whereas sound field reproduction techniques, which use a large number of loudspeakers to control sounds within a region, are primarily used in commercial and professional audio applications. Both approaches aim for the most immersive experience possible; The spatial audio impact in sound field reproduction is not limited to a single user or a few spatial locations. To make this research more interesting, Real-time room acoustic simulations for a natural AR/VR listening experience and Sound source can be separated for augmented audio reality [30].

Subjective assessment based on affecting factors of VR audio experience quality

A research paper of 2019 represents the subjective assessment criteria are presented in this study, based on the affecting factors of VR audio experience quality, and include four indicators: VR-Audio Quality (VAQ), VR-Audio Orientation (VAO), VR-Audio sound field reality (VAR), and VR-Audio reverberation effect (VAE). The test situations for each indicator are revised, and the corresponding evaluation concepts for each section are determined using the present evaluation techniques; subjective experiments were designed, and test software projects were built. The results of the tests indicate that this approach may be used to quantitatively analyze various test indicators and examine the effectiveness of various VR audio renderers or rendering techniques. VR audio experience can be more interesting by adding human perception study. This model can be a complete VR audio subjective evaluation model, which may help with VR audio system algorithm optimization, advancements, and hardware and software combination [31].

Examining the approaches to determine the ecological validity of immersive VR

This article aims to examine the methods for determining the ecological validity of IVR for the perception of urban sound environments, as well as the technologies required for assuring ecological validity during audio-visual replication. The qualitative findings suggest that immersive virtual reality techniques have a lot of potential as an ecologically viable tool for soundscape or noise evaluations. Virtual reality's ecological validity in assessing urban sound environments is multimodal, dynamic, and contextual. More study is needed, in particular, on the impact of Ambisonics orders of complexity at the recording and reproduction phases, as well as concerns like encoding and decoding Ambisonics formats, on soundscape perception in order to offer a dynamic virtual experience [32].

Evaluation of the audio-visual interaction effects in immersive VR

This research is based on a subjective evaluation of the audio-visual interaction effects inside diverse urban spaces with varying environmental circumstances; this research studied the connection between overall satisfaction with the urban environment and its soundscape and landscape. To replicate actual site circumstances using immersive virtual reality technology, Participants were given visual information via a head-mounted display (HMD) and auditory information via head-tracking technology utilizing first-order ambisonics (FOA) of headphone-based three-dimensional auralization. Subjective reactions to auditory and visual aspects such as cognitive response, semantic expression in audio-visual as well as audio- and visual-only environments, and reported satisfaction were examined. As all of the participants were in their twenties, the findings may not be relevant to all cities or age groups. Evaluation across a wide range of age groups in urban areas with varied functions would be necessary to generalize the study results to a more diversified urban space and demography [33].

Comparison of user experience in immersive modes and non-immersive modes

Authors of this research paper tried to investigate the differences in player experience between games played in immersive modes and games played in non-immersive modes (e.g., on a desktop monitor). A game, Smash Hit, was tested in both immersive (virtual reality) and non-immersive (desktop) modes by 24 young adults. Usability, emotional reaction, and perceived feeling of presence were assessed using self-report questionnaires Visual Analogue Scale of Anxiety (VAS-A), Visual Analogue Scale of Happiness, (VAS-HP), Visual Analogue Scale of Surprise (VAS-SP), SUS, SUS-II), and psycho-physiological measurements (heart rate and skin conductance). (a) Using virtual reality to play a video game was no more challenging than using a desktop monitor; (b) players had a more extremely intense emotional response after playing in virtual reality compared to playing through a desktop display, as evaluated by a self-report questionnaire and psycho-physiological indexes (heart rate and skin conductance); (c) the perceived sense of presence was higher after playing in virtual reality versus after playing through a desktop display. Their research and results are related to a specific VR system (Samsung Gear VR Oculus); it would be more interesting to do research by using other VR systems like off-the-shelf VR systems such as HTC Vive. Only one type of game is studied, which is FPS. It would be more interesting to do experiments on other genres of games, e.g., racing games, etc [34].

Influence of illusion in immersive VR

The authors of this research paper investigated if the illusion of a significant increase in belly fat may be induced in men using immersive virtual reality through the following Prospective position of first-person viewpoint Self-induced synchronous visual-tactile stimulation in the stomach area, as well as synchronous visual-motor correlation between actual and virtual arm motions. The findings demonstrate that synchronous multisensory stimulation combined with a first-person view of a virtual body that replaces one's own body in virtual reality might cause temporary alterations in body representation toward a bigger belly size. This was proven by questionnaire findings, the difference in self-estimated belly size after and before the experimental manipulation, evaluated from a first-person perspective, and strong positive relationships between these two factors. This research is related to only the illusion of thin people to fat, but there is no research or experiment on the illusion of fat people to thin. Also, to make it more useful in the treatment of body size distortions, there is a need to research body ownership illusion [35].

Creating virtual simulator for training of crane operator

In this research, the authors tried to create a virtual simulator for crane operator training. It features a motion chair to add to the vibrating impact and an operational chamber to represent the driving chamber of a truck crane. A 3D projector displays the virtual sceneries developed by VR software on a screen in front of the operating chamber. This research is done by using geometric objects to develop the background environment to make it more realistic. A 3D model of a crane and a realistic scene can be designed through Unreal Engine [36].

Assessing the learning and efficiency as measured by self-efficacy in VR training system of different types of cranes

In this research work, researchers added three distinct types of cranes to a series of VR training systems employing innovative technology, namely HMD, and performed an experiment to assess their efficiency as measured by self-efficacy. They also looked at why and how the system may improve learning results. The experiment involved a total of 108 participants. According to the findings of a paired sample t-test, the participants' self-efficacy increased after the Virtual experience, with statistically significant differences between the levels of self-efficacy before and after the experiment. More research can be done on the difference between usage of this system by trained and untrained people and further research needed to assess the effectiveness of this simulation system, e.g., how much the participant has learned [37].

Examining the relationship b/w immersion and user experience

The primary objective of this study is to use the five criteria of immersion to examine the connection between immersion and user experience. In addition, the research will look into several techniques for measuring immersion. Following the experiment, regression analysis was used to characterize the relationship between immersion and user experience using the data acquired from the outcomes. Following the analysis of the data, it was discovered that immersion levels drop as the user's experience increases, based on the observed inverse linear regression. More research can be done to explore the relationship between the level of experience of the user and immersion. It would be more interesting to find out the relation of immersion related to the satisfaction level of user experience [39].

Analyzing the influence of different aspects of immersion on presence and memory

The research of this paper [40] analyzes how three different aspects of immersion influenced presence and memory: field of view, unimodal (visual only) vs. bimodal (visual and aural) environments, and lighting realism (e.g., the presence or absence of shadows).

Individual differences (using a within-subjects design) and differences across VR devices (by using identical equipment for all 85 tests) were both addressed in this study. Furthermore, the degrees of interactive immersion in this study were consistent across all circumstances. The findings show that the effects of immersion on presence and memory are varied and appear to be reliant (in part) on the specific feature of immersion being investigated. It would be fascinating to see if altering all three factors of immersion at the same time (essentially resulting in Very High- and Very Low-Immersion situations) produces outcomes that are similar to or different from the individual components of immersion. More research should also look at how different aspects of experimental design might affect the outcomes of a study's empirical interest [40].

The research presented in this thesis aimed to add to the findings of prior investigations. Previous work has been done on presence, memory, user experience, and interaction (as discussed in the 2.4 section), and some of it has been done on measuring audio effects in IVR. Other research has looked at the impact of audio-visual components on memory, therefore further work could be done in this area to accurately measure presence in IVR with audio and visual elements. The goal of this research was to evaluate audio visual elements in IVR.

3 Overview of Developed Experimental System

The industrial crane's three-dimensional model is a highly nonlinear MIMO system with a specific sensor system and a unique 2D angle measurement unit. The system works in real-time and is completely integrated with MATLAB/Simulink. RTG crane model is shown below in Figure 1.

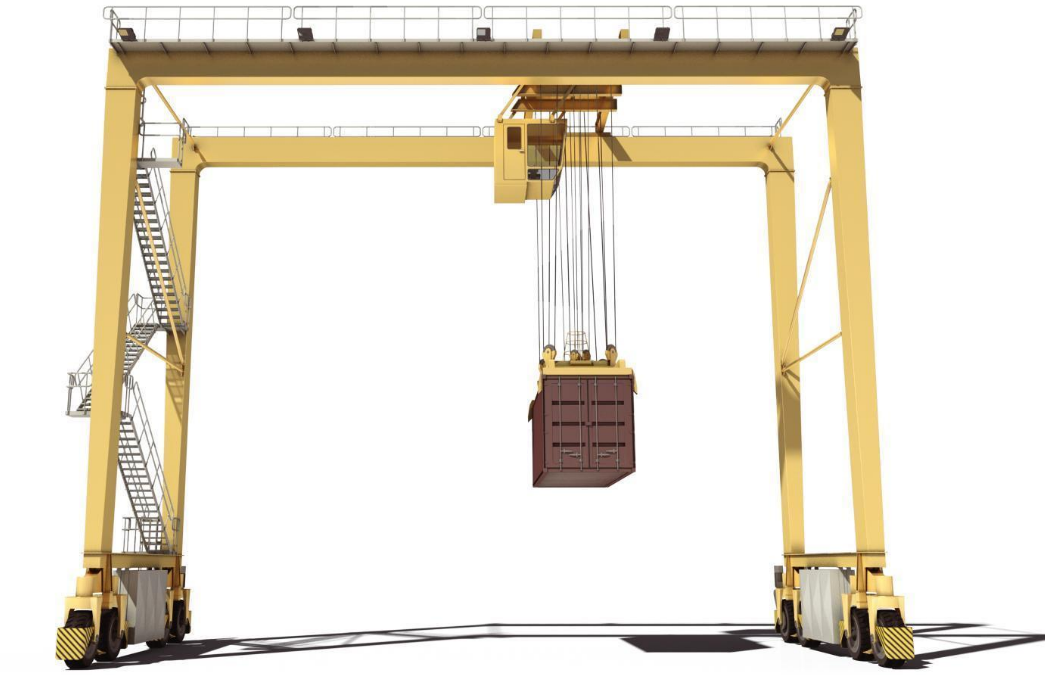


Figure 1. RTG Crane Model Image source: [47].

Figure 2. Laboratory Scale RTG Crane Model. Image source: [46].

The program enables real-time control algorithms to be prototyped quickly. It is not necessary to write C-code. The library of basic controllers that comes with 3D Crane is rather extensive. Three control DC motors and five-position measurement wheel encoders are included in the model. The P controller for the 3D crane is shown in Figure 3. The RTW routines are executed by selecting Real-Time Workshop and Build Model from the pull-down options. As a result of linking it to Windows Target, one receives an executable file that functions as a real-time controller [42].

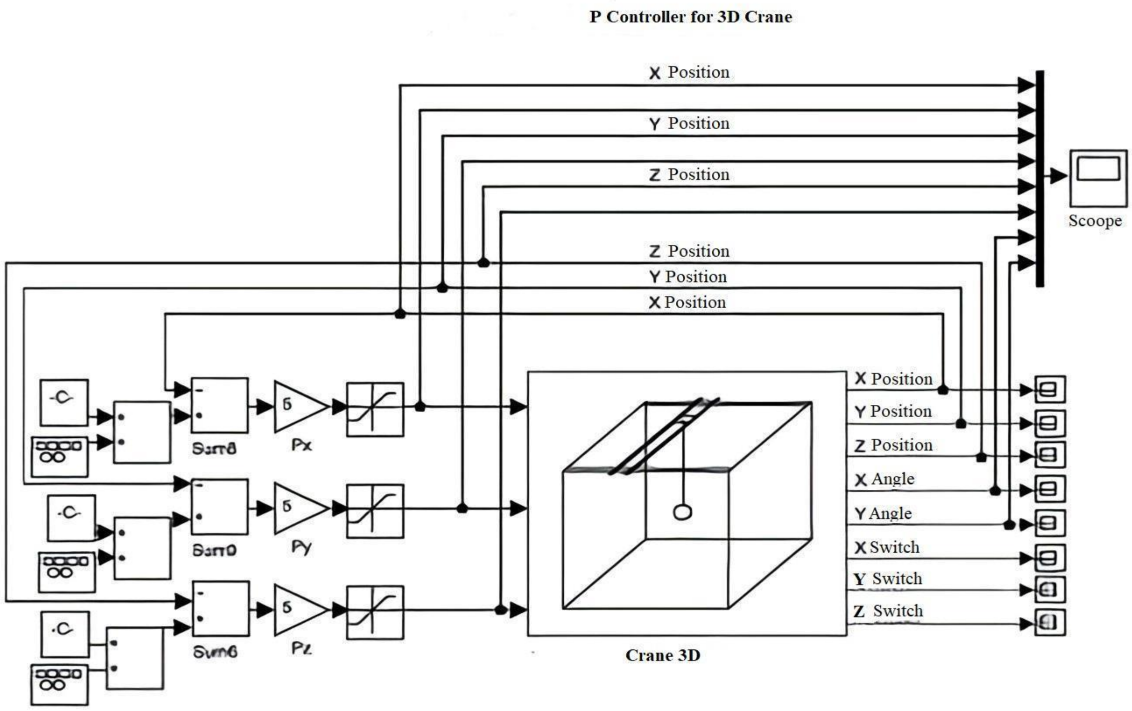


Figure 3. A Simulink model implementing a P controller to control the position of the payload in three-dimensional space for the 3D Crane control object investigated in this thesis.

3.1 Hardware

- Three motors 24V DC PWM controlled
- power interface
- Five incremental encoders
- RT-DAC I/O internal PCI or external USB board (PWM control and encoder logics are stored in a XILINX chip) [42].

3.2 Digital Twin

To show several automated control ideas, a digital twin of a 3D crane is employed. The digital twin of the 3D-Crane is already developed using the Unreal Engine.

The project was started from a template that generated some project content automatically, e.g., virtual reality blueprints, materials, and maps, etc. later motion controller pawn was edited according to project requirements. After that, a feature toggle to enable/disable teleportation was added to the motion controller pawn.

Then go over the functionality in BP MotionController, which includes the spawning and attaching motion controllers, receiving input from motion controllers, managing teleportation, constructing the teleportation arc, and rumbling the controller when colliding with static meshes, among other things. Because both blueprints include vital functionality, communication between them is critical (for example, executing a function from the BP MotionController blueprint from MotionControllerPawn, altering the value of a variable in BP MotionController from MotionControllerPawn, and so on).

The functionality created for the 3DCrane BP was then analyzed, as well as how it interacts with MATLAB through the UDP connection. Figure 4 shows the summary of this process.

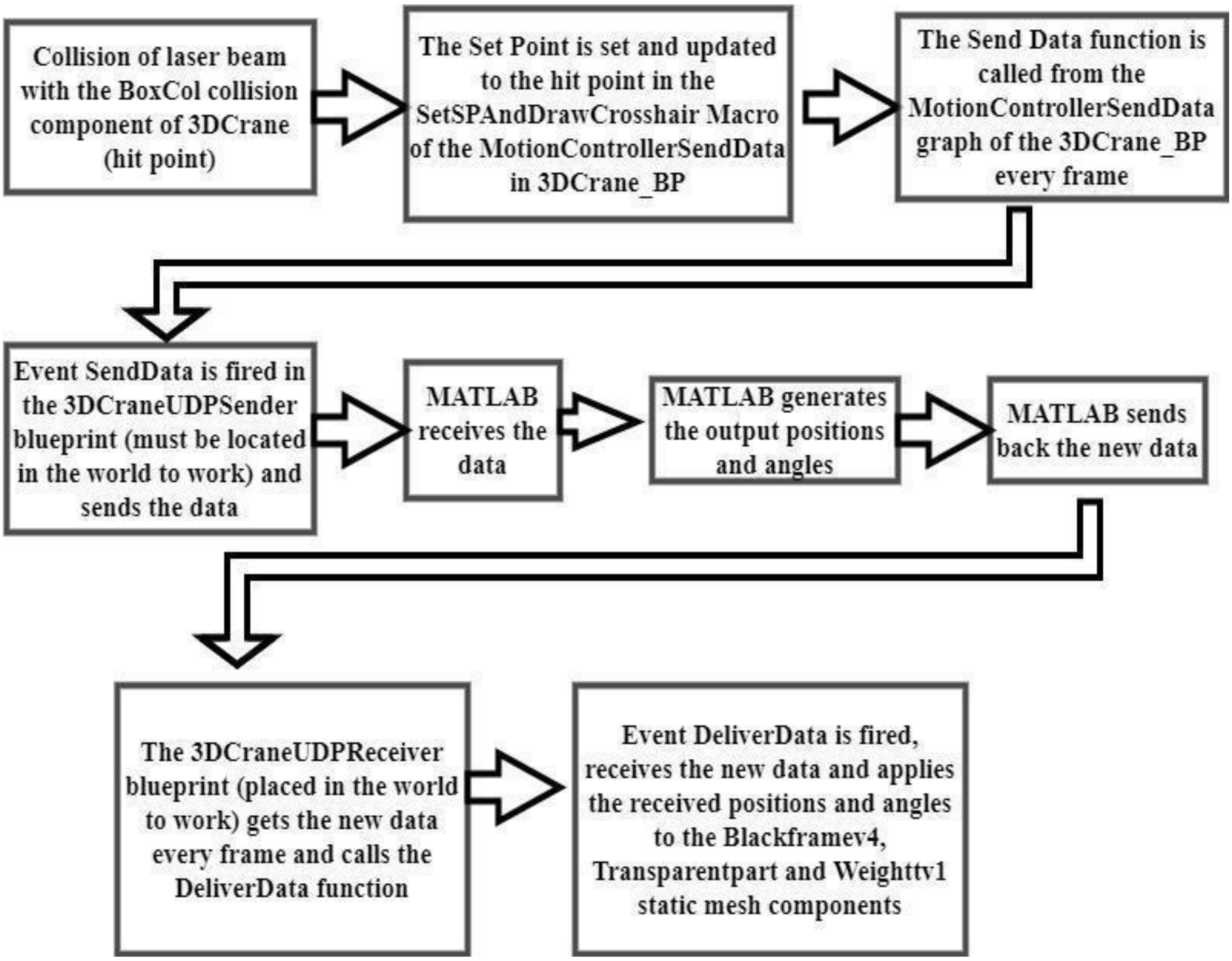


Figure 4. Implementation in UE4 for the 3DCrane_BP.

After that, the MATLAB Simulink model was reviewed that generates the output positions and angles. Later on, so many modifications were made for further improvements; for example, modifications were made to Add features like Automatic activation and deactivation of the visibility of the load projection and Crosshair with Animation, updating 3D Crane position and rotation with Vive Tracker, Plotting data with different colors in 3DCrane_AnglesGraphWidget, and take recorder, etc.

4 Methodology and Design

The explanatory technique was applied in this research, while quantitative methods were used to obtain the results data. The goal of this study was to solve a real-world research challenge. This approach was chosen because it provides us with a wide range of data since each individual perceives things differently.

The idea of environment design was vital to the entire process. Epic Games Technology's Unreal Engine was selected for this project. Outside of games, the Unreal Engine is frequently utilized in architecture, product innovation, and cinematography. The level of immersion and the realism aspect, which was preserved as near to the real world as possible, were taken into consideration.

The following is a list of the methodologies employed in this research.

1. Literature review

2. Subjective measurement

3. Data collection

4. Data presentation

5. Comparison and analysis

Comparison and analysis: In the research process, comparing and analysis are the two critical strategies. It will be useful to compare various methodologies used in human interaction with VE and evaluate the data acquired at the conclusion of the experiment. Additionally, the effect of the data-retrieved information may be assessed [44]. Since the topic is primarily statistical, data acquired from the documentation may be displayed in a various form, including line graphs and bar charts in item 4.

4.1 System Design

Tools used to design VR experiment

Unreal Engine

UE was used to construct a digital twin of 3D crane and implement sound to its constructed digital twin also used to make a larger scale of DT of 3D crane.

MATLAB

MATLAB was used to design the control environment for DT; it was used to make a logic to filter data according to different variables and to represent data.

Microphone

Figure 5. Zoom H1 Handy Recorder. Image source: [48].

We used a Zoom H1 handy recorder to record sound, as shown in figure 5. We recorded sound in 16-bit Wav files because UE4 supports them. The ultra-compact H1 Handy Recorder offers professional-quality stereo recording. The X/Y technique is ideal for all types of live stereo recording because it covers a large area while still capturing sound sources in the centre with clarity and definition.

The built-in X/Y microphone on the H1 delivers two matched unidirectional microphones at a 90-degree angle to one another, which is ideal for most stereo recording applications.

Audacity

Audacity is a multi-track audio editor and recorder for Windows that is simple to use. It was used to edit and improve the quality of recorded audio. In figure 6 user interface of this software is shown.

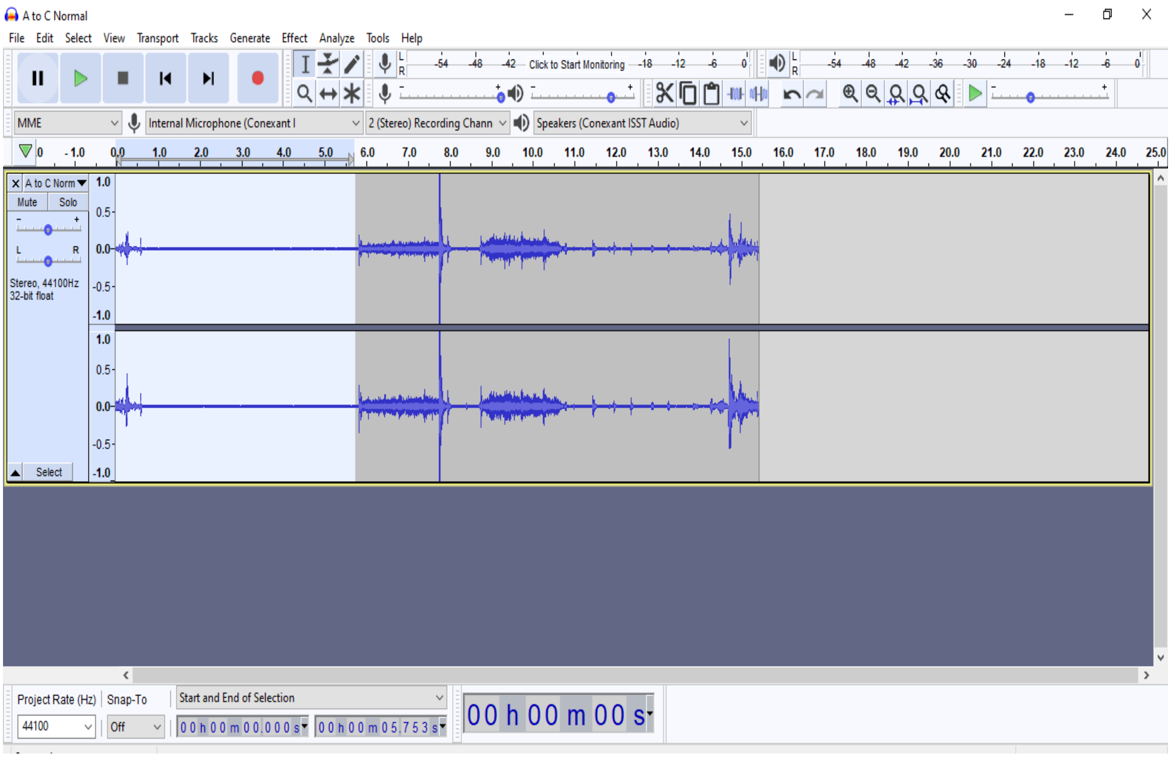


Figure 6. The user interface of Audacity Software: the process of recording a stereo audio clip.

Two VR systems were designed in such a way that they can help to get results on audio and visual aspects of VR. For this purpose, a digital twin of a laboratory-scale replica of a gantry crane (referred to as “3D crane”) was developed. The developed digital twin was complemented with audio features (mechanical noise) and the ability to scale its physical dimensions. The goal of constructing a bigger scale digital twin of this crane was to do experiments with subjects to see how scale affects immersion.

Unreal Engine is the system's development platform. The data collecting plugin is an Unreal Engine blueprint class. That plugin is written in the C++ programming language. The plugin is then integrated with the VR experiment when it has been developed. The Unreal Engine's event BeginPlay function fired when the experiment began. The plugin begins to save user movement data from HDM, motion controller, and custom data associated with the user action performed to operate the crane.

4.2 Questionnaire Development

To get the most accurate responses from participants, we decided to use three different types of a standard questionnaires. The first section of the questionnaire was background, which contains questions related to the background information of participants. The second section was about the system usability scale developed by James R. Lewis. The third section was about presence. For this section, we used a questionnaire developed by Witmer and Singer based on the outcomes of prior studies into relevant areas of presence. The PQ assesses how strongly people feel present in a VE, as well as the impact of relevant contributing factors on the intensity of that experience [45]. A comparison of 15 published presence questionnaires is performed in the paper [55] is depicted in the following table 1.

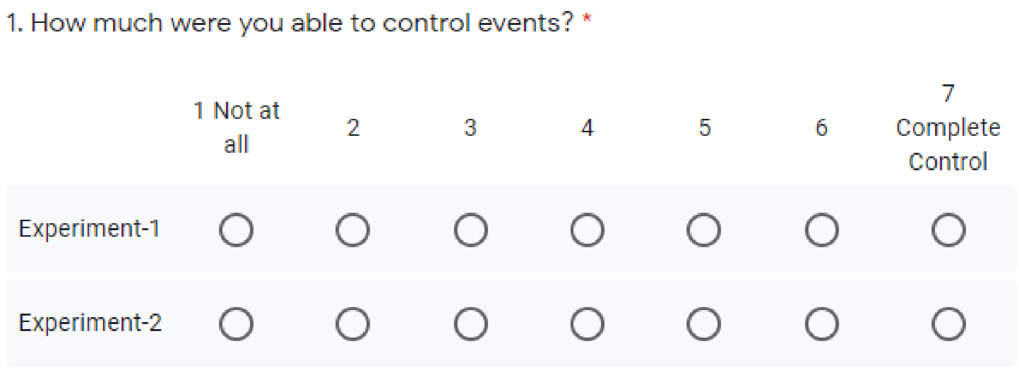
Table 1. Overview and Comparison of Published Presence Questionnaires.

Authors	Year	Citations	Items	Usage
Banos et al.	1998	146	77	VE
Barfield & Hendrix	1995	186	5+1	VE
Cho et al.	2003	34	4	VE
Dinh et al.	1999	365	13+1	VE
Gerhad et al.	2001	57	19+4	SVE
Kim & Biocca	1997	664	8	VE
Krauss et al.	2001	8	42	VE
Lombard & Ditton	2000	205	103	NA
Lombard & Weinstein	2009	120	4-8	CM
Lessiter et al.	2001	861	44	CM
Nichols et al.	2000	158	9	VE
Nowak & Frank	2003	569	9	SVE
Schubert et al.	2001	758	14	VE
Usoh/Slater et al.	1994	853	3	VE
Witmer & Singer	1998	3569	32	VE

VE = Virtual Environment, CM = Cross-Media, SVE = Shared Virtual Environment, NA = items not listed Determined by [55] using Google Scholar, Sept 2018

As the questionnaire from Witmer and Singer possesses the qualities that are applicable to the research presented in this thesis and because it has been widely adopted by research scholars (which is also evident from the citation count), we decided to use this questionnaire in the present work.

The PQ is focused on the semantic differential concept and uses a seven-point scale format (Dyer, Matthews, Stulac, Wright, Yudowitch, 1976). Each item is connected at the ends by opposing descriptors, just as the semantic differential. The scale, unlike the semantic differential, has a midpoint anchor. The anchors are dependent on the content of the question stem and are similar to the anchors used in typical rating scales in that regard. The PQ instructions instructed respondents to select the appropriate scale box based on the content of the question and the explanatory labels. Figure 7 is an example item from the PQ [45].


1. How much were you able to control events? *

	1 Not at all	2	3	4	5	6	7 Complete Control
Experiment-1	○	○	○	○	○	○	○
Experiment-2	○	○	○	○	○	○	○

Figure 7. Example of Presence Questionnaire.

5 Experiments and Data Collection

This dissertation examines the data collected from participants regarding their level of immersion and control during their experiment. The focus of the experiment was on the subject's interaction with a 3D crane and level of immersion. Since it has three degrees of freedom, the crane is known as a “3D crane.” It can move the load along three separate axes. While the weight is being moved by a 3D crane, it will naturally swing. The goal of the study is to teach participants about swing compensation and assess their level of immersion depending on the audio and scale of a 3D crane.

5.1. Preparation

During the planning stage, we decided on the hardware that would be used, which included an HTC Vive Pro headset, HTC motion controllers, and HTC trackers. The headset serves as an interface, providing participants with a visual image. Controllers are utilized to interact with any VR actor. Trackers are in charge of collecting data from controllers and a headset in order to ensure a satisfying experience. It is feasible to gather time-stamped information on the position of the hands and head during the experiment, as well as information about eye movement, using these tools, and then analyze the data to extract useful information.

In November 2021, the preparation stage started. First and foremost, a registration form for the purpose of contacting subjects was designed and shared among friends, classmates, and colleagues to choose a suitable time slot when they can participate in the experiment. Data collection takes place inside the lab due to the restricted resources and portability of VR equipment. Individually, the participants are invited to the Virtual Reality Lab. One experiment took an estimated 40 minutes to complete.

The following would have to be done or assured to be done before the subject arrived at the laboratory in which the experiments were conducted:

- Get the area ready.
- Clean and sanitize surfaces that the last participant came into touch with to ensure participant safety.

- Assign an ID value to the subject.

- Prepare MATLAB/STEAM VR.

- Restart the software if it was previously opened.
- Set all of the settings to their default values.

- Get the VR Equipment ready

- Charge the controllers.
- Set the headset and controllers to their default settings.
- Make sure the trackers are in the right spot and that there are no obstacles between the trackers and the experimental area.

- Reopen the survey form.

5.1.1 Preparation for the experiment

The subjects had minimal prior knowledge of the topic. As soon as the subject entered the laboratory, they were given a brief training via slides that informed them of the conditions and provided them with an understanding of the environment, controls, and how to use the controls and VR. The headset is put on and adjusted such that the display is in the centre of the field of view.

The controllers have been located and picked up.

The eye-tracker has been calibrated as follows:

a) The headset is vertically adjusted to center the display on the eyes;

b) The distance between the lenses is adjusted depending on the participants' eyes;

c) The participants were asked to follow a series of dots with their eyes without moving their heads.

Following that, each participant was allocated a unique ID, and two experiments were assigned to each participant based on the ID. For instance, the first experiment could be a small-scale crane with sound, while the second experiment could be a large-scale crane without sound. The experiment was started once everything was in place and they were comfortable and ready to go.

5.1.2 Performing the control experiment with the 3D crane

Once everything is in place, participants are transported to the simulated virtual world, which in this case is a laboratory. They see the “Walk here” marking on the ground and walk there. They would be able to see a virtual 3D crane model from that location. The purpose is to use a controller to point on the ground underneath the 3D crane, causing the crane to transport the payload from arbitrary location A to arbitrary position B. The participant eventually activates swing compensation and examines the outcomes. Subjects can see graphs with flowing data about angles during the experiment and try to grab the graph and place it anywhere they like in the surrounding area.

The major goal of this stage was to collect data regarding the participants' level of immersion and control during the experiment. As long as this goal is met, the subject is free to explore the 3D crane model or graphs for the rest of the experiment, as long as he or she stays within the boundaries of the area assigned. After each trial, the data acquired during the experiment was ZIP-archived and password-protected, and uploaded to a private cloud repository.

5.1.3 Feedback

Subjects were requested to fill out a survey form to provide feedback once the second stage was completed. Feedback aids in determining which method is more user-friendly and engageable. Sound has or hasn't an influence on the immersion level of the subject, whether using a 3D crane with normal size or a larger scale.

5.2 Participants

A total of 58 people contributed data to our research. The participants range in age from 20 to 55 years old. The participants worked in a variety of fields, with the majority of them being engineers or students. Around 53% of participants were experienced with VE equipment, while others were using it for the first time. The following figures depict the distribution of participants.

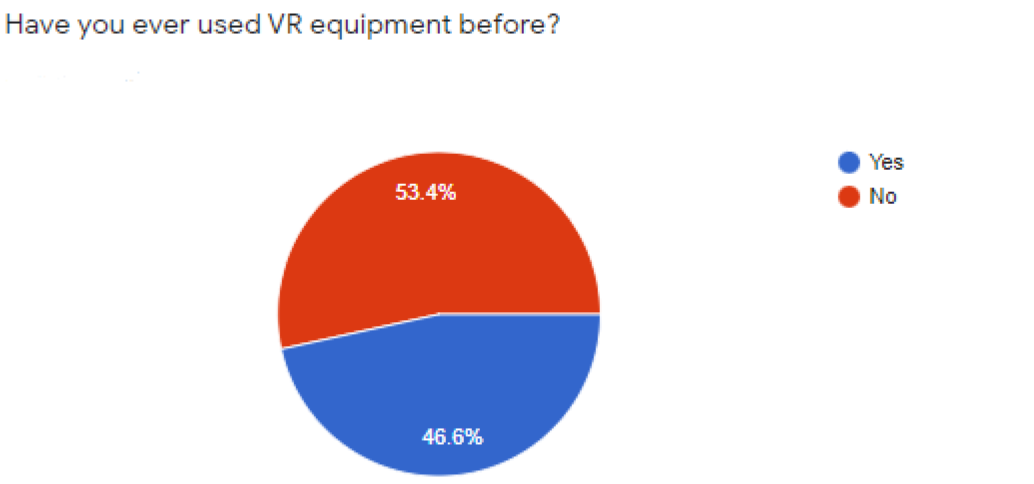


Figure 8. Participants Distribution Based on Experience.

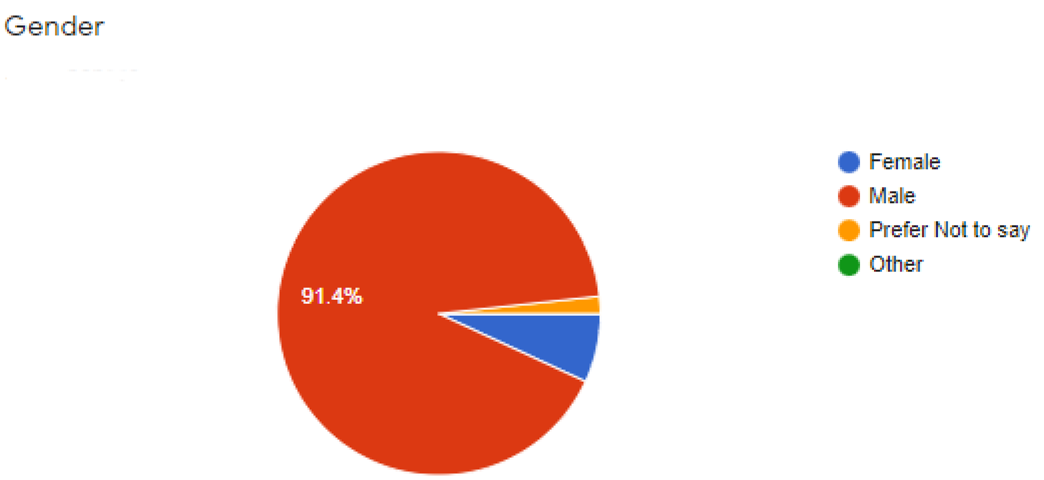


Figure 9. Participants Distribution Based on Gender.

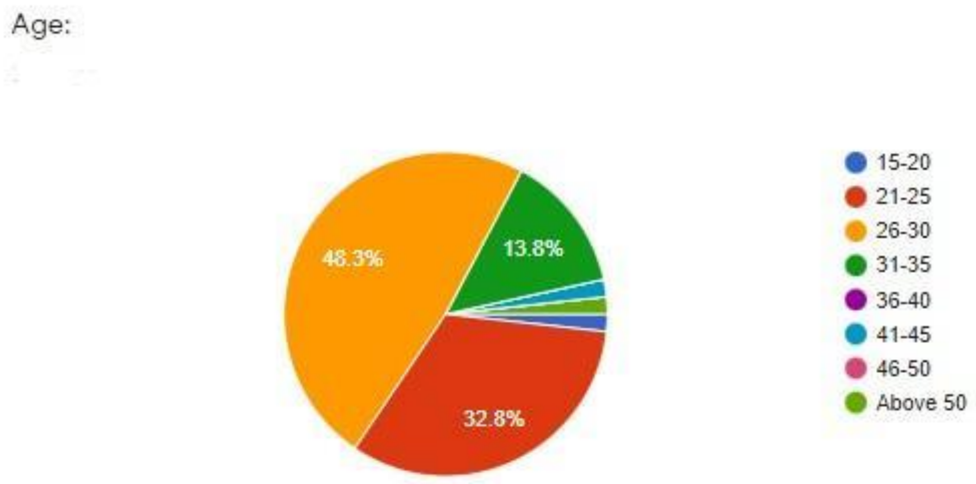


Figure 10. Participants Distribution Based on Age.

5.3 Data Collection

The procedure of data collecting and storage during the experiments is described in this section. It also specifies the type of data collected. The participants completed a consent form outlining the protection of their data, as well as the risks and advantages of participating.

In the experiment in question, there were four key actions.

- Walking from the starting position to the target.
- Pointing on the target point accurately
- Activates the swing compensation.
- Grabbing the graphs depicting time series information about the dynamic states of the 3D crane and moving them around.

Sensors or lighthouses (in our case, HTC Base Station) placed at a reasonable distance track the motions of the controllers and headset. Both the controller and the headset provide data (such as controller inputs) to the computer, which is processed according to the program in use. While the lighthouses are self-contained, they still require the controller and headset to provide meaningful data, and while the controller and headset do transfer data to the computer, the data would be incomplete if it were not calculated with the lighthouse data [44].

The coordinates of their actions were tracked as X, Y, and Z coordinates after they started controlling the crane in a VR environment. The method utilized here is that the Unreal Engine 4 has a function in the Blueprint feature that allows us to print the string while the project is operating. This piece of information is produced in real-time and can be used to extract whatever information is needed from the project. The coordinates are kept in CSV files, and they will be extracted as a graphical representation of the data that we have collected. Each participant got their own CSV file, which included a graphical representation of how they behaved in the situation they were given. Figure 11 depicts how the project's methodology works in order to provide a clear picture of the workflow.

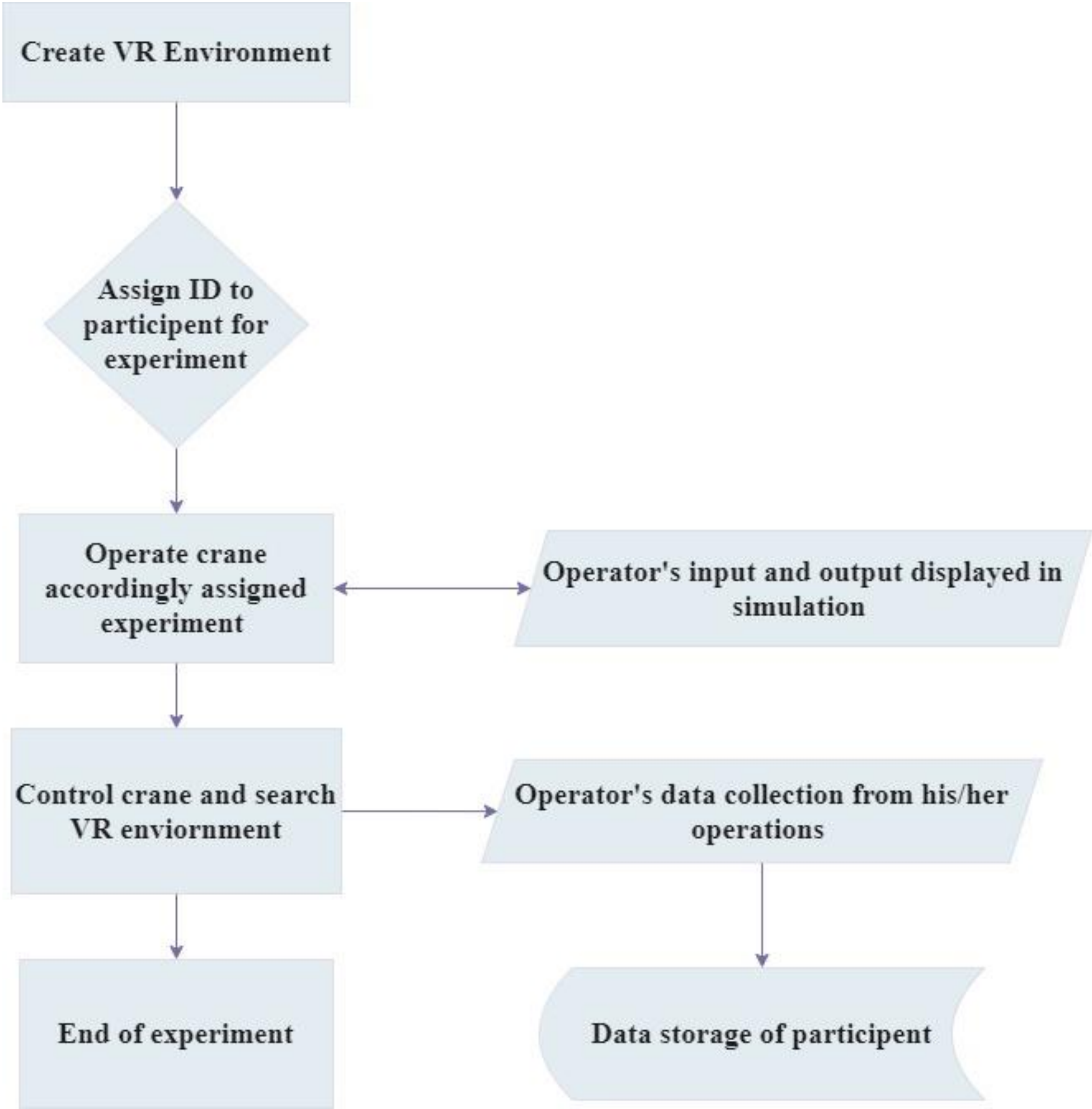


Figure 11. Flow Chart of Applied Methodology.

6 Results

The feedback file of the survey was processed in such a way that data could be filtered according to requirements. To filter data, we imported a CSV file of feedback into MATLAB and programmed it in a way that it can search or filter data according to different variables. The output of the data filtration program is shown in figure [12]. And results of different variables are visualized through box plots.

```
Command Window
Do you want to filter data? (Type 0 for no and 1 for yes) = 1
Hi,You can filter the results based on following parameters
Press 1 to filter using Experiment No
Press 2 to filter using AgeGroup
Press 3 to filter using Experiance
Press 4 to filter using Scale
Press 5 to filter using Sound
If you want to use multiple filters type more then one digit like 153 means it has 3 filters No(1,5,3)
Your Choice = 4
In order to filter results based on Scale
Type 1 to filter with Big scale and 0 for Normal Scale
fx Your Choice = |
```

Figure 12. Output of Data Filtration Program.

6.1 Boxplot

A box plot or box plot (also known as box and whisker plot) is a form of chart commonly employed in explanatory data analysis in descriptive statistics. By presenting the data quartiles (or percentiles) and averages, box plots visually depict the distribution of numerical data and skewness.
Box plots display the minimum score, first (lower) quartile, median, third (upper) quartile, and the maximum score of a collection of data as a five-number summary [56].

Box plots split the data into parts, each of which contains about a quarter of the total data in the set as shown in figure 13.

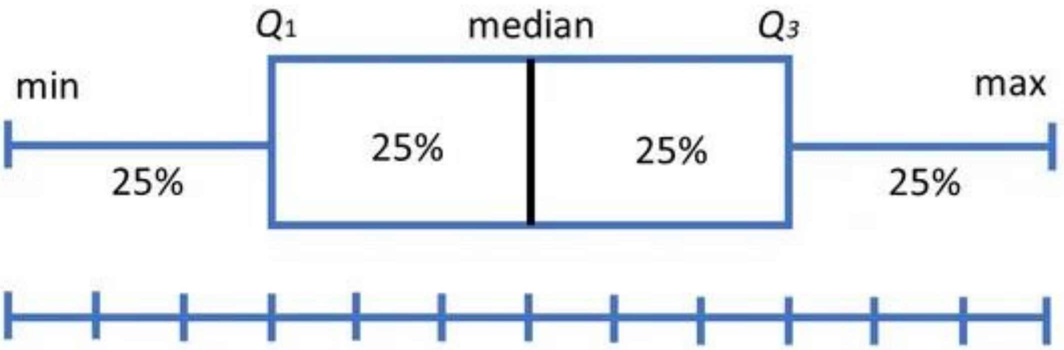


Figure 13. Parts of Boxplot. Image source: [56].

Minimum Score
Excluding outliers, the lowest score (shown at the end of the left whisker).

Lower Quartile (Q1)
Twenty-five percent of the scores are below the lower quartile value.

Median

The median indicates the data's midpoint and is represented by the line that divides the box into two parts (also known as the second quartile). Half of the scores are higher than or equal to this value, while the other half are lower.

Upper Quartile (Q3)

Seventy-five percent of the results are below the upper quartile (also known as the third quartile). As a result, 25% of the data is above this limit.

Maximum Score

Excluding outliers, the highest score (shown at the end of the right whisker).

Whiskers

Outside the middle 50%, the upper and lower whiskers denote scores (i.e., the lower 25 percent of scores and the upper 25 percent of scores).

The Interquartile Range (IQR)

The range between the 25th and the 75th percentiles) [56].

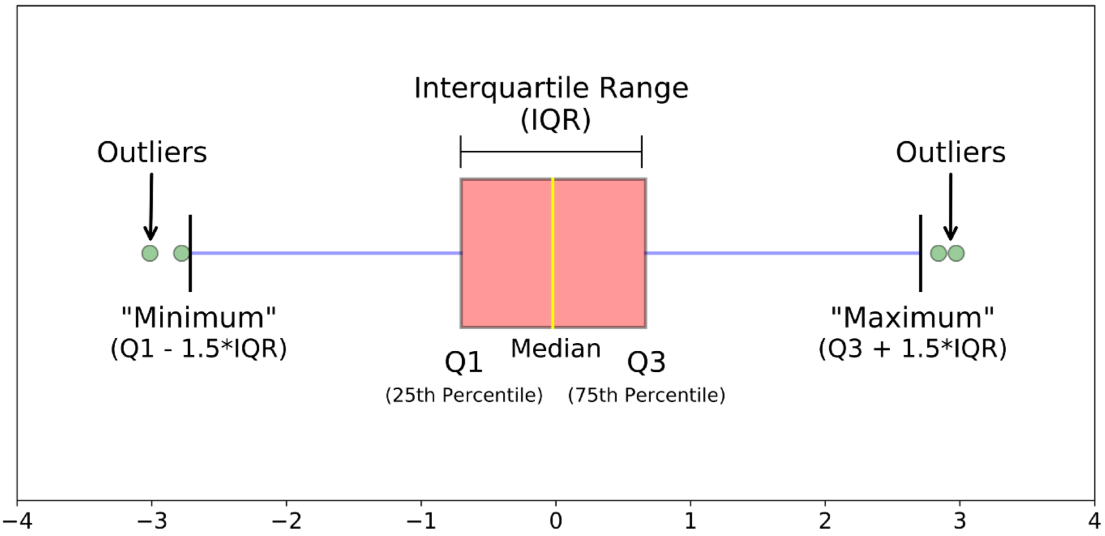


Figure 14. Parts of Boxplot with Outliers. Image source: [54].

Outliers

An outlier is a data point that lies outside the box plot's whiskers while examining it. It is usually referred to as unusual data. Following are the equations to calculate outliers

$$Low\ Outlier = Q1 - 1.5\ X\ IQR$$

$$Upper\ Outlier = Q3 + 1.5\ X\ IQR$$

6.2 PQ Item Analysis

A 32-item presence questionnaire (PQ) was developed by Wittmer and Singer [45]. Different subscales were determined by the authors. The reason for using this questionnaire for our research is discussed in the 4.2 questionnaire development section. To measure the presence, we used subscales of Items(questions) which are shown in the table. On the basis of the subscale, there were fourteen questions in INV/C, and three questions were related to NATRL, Item 6,15,16 belongs to AUD, item 17,21 related to HAPTC, items 28,29,30 related to IFQUAL, and item 19,20 belongs to RESOL. Item 22, which assesses the extent to which users get disoriented while switching from VE to the actual world, indirectly addresses involvement that is not used in our results. Items 8,9,11 are related indirectly to INV/C; that's why we didn't use them in our subscale category. Item 31 is related to learning outcome, which was not directly related to presence.

Table 2. Subscale of Presence Questionnaire Based on Different Factors.

Nr.	Item Stems	Factors	Subscale
1	How much were you able to control events?	CF	INV/C
2	How responsive was the environment to actions that you initiated (or performed)?	CF	INV/C
3	How naturally did your interactions with the environment seem?	CF	NATRL
4	How completely were all of your senses engaged?	SF	INV/C
5	How much did the visual aspects of the environment involve you?	SF	INV/C
6	How much did the auditory aspects of the environment involve you?	SF	AUD
7	How natural was the mechanism which controlled movement through the environment?	CF	NATRL
8	How aware were you of events occurring in the real world around you?	DF	
9	How aware were you of your display and control devices?	DF	
10	How compelling was your sense of objects moving through space?	SF	INV/C
11	How inconsistent or disconnected was the information coming from your various senses?	RF	
12	How much did your experiences in the virtual environment seem consistent with your real-world experiences?	RF,CF	NATRL
13	Were you able to anticipate what would happen next in response to the actions that you performed?	CF	INV/C
14	How completely were you able to actively survey or search the environment using vision?	RF,CF, SF	INV/C
15	How well could you identify sounds?	RF,SF	AUD

16	How well could you localise sounds?	RF,SF	AUD
17	How well could you actively survey or search the virtual environment using touch?	RF,SF	HAPTC
18	How compelling was your sense of moving around inside the virtual environment?	SF	INV/C
19	How closely were you able to examine objects?	SF	RESOL
20	How well could you examine objects from multiple viewpoints?	SF	RESOL
21	How well could you move or manipulate objects in the virtual environment?	CF	HAPTC
22	To what degree did you feel confused or disoriented at the beginning of breaks or at the end of the experimental session?	RF	
23	How involved were you in the virtual environment experience?		INV/C
24	How distracting was the control mechanism?	DF	INV/C
25	How much delay did you experience between your actions and expected outcomes?	CF	INV/C
26	How quickly did you adjust to the virtual environment experience?	CF	INV/C
27	How proficient in moving and interacting with the virtual environment did you feel at the end of the experience?	CF	INV/C
28	How much did the visual display quality interfere or distract you from performing assigned tasks or required activities?	DF	IFQUAL
29	How much did the control devices interfere with the performance of assigned tasks or with other activities?	DF,CF	IFQUAL
30	How well could you concentrate on the assigned tasks or required activities rather than on the mechanisms used to perform those tasks or activities?	DF	IFQUAL

31	Did you learn new techniques that enabled you to improve your performance?	CF	
32	Were you involved in the experimental task to the extent that you lost track of time?		INV/C

Major Factor Category: CF = Control Factors, SF = Sensory Factors, DF = Distraction Factors, RF = Realism Factors.
Subscales: INV/C = Involvement/Control, NAT = Natural, AUD = Auditory, HAPTC **=** Haptic**,** RES **=** Resolution, IFQUAL **=** Interface Quality [45].

6.3 Results of different variables

The following plot shows results when sound was off in both experiments (Big scale, Normal scale), and it depicts the impact of sound off on different subscales of measurement.

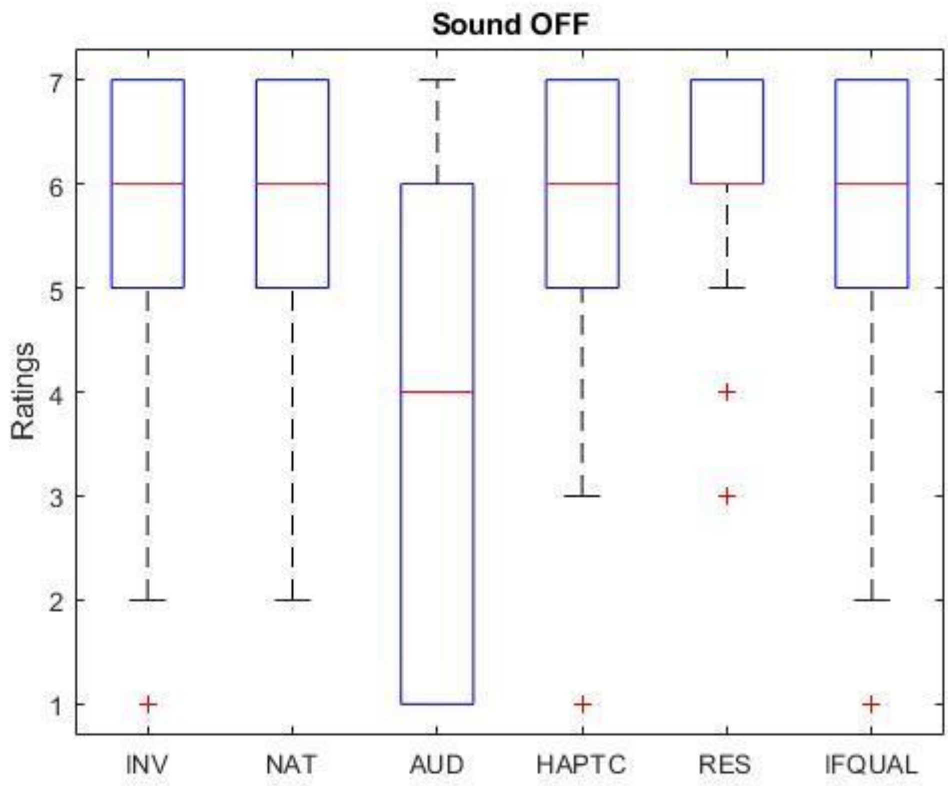


Figure 15. Results of Sound OFF.

The comparison of subscales when sound was off is shown in the following table.

Table 3. Comparison of Subscales when Sound was OFF.

Subscale	**Min**	**Q1**	**Median**	**Q3**	**Max**
INV/C	2	5	6	7	7
NAT	2	5	6	7	7
AUD	1	1	4	6	7
HAPTC	3	5	6	7	7
RES	5	6	6	7	7
IFQUAL	2	5	6	7	7

This summary shows that the majority of the total participants were involved or had better control; only one-fourth of the total gave feedback under 5, which represents that 25% of the total remained neutral or they were not involved and also didn't have better control. The completely natural environment was experienced or felt by 25% of the total, and 50% felt good natural or to some level natural environment, and one-fourth of the total remained neutral or didn't feel the natural environment at all when sound was off. About auditory aspects, 50% of the total didn't experience sound at all, and the remaining 50% experienced some kind of sound or they were engaged to visual elements in such a way that they didn't pay attention to audio elements and forgot that sound was off. Most of them experienced the high quality of interface and sense of interaction or touch in IVR environment; only one-fourth of the total remained neutral or didn't feel a sense of touch, or they experienced the low quality of the interface and experienced the high quality of resolution as well only two participants of the total responded under five which represents that overall majority of the participants experienced good quality or felt more involved in each category except auditory as it was off.

The figure below displays the influence of sound on different subscales of measurement when sound was turned on in both tests.

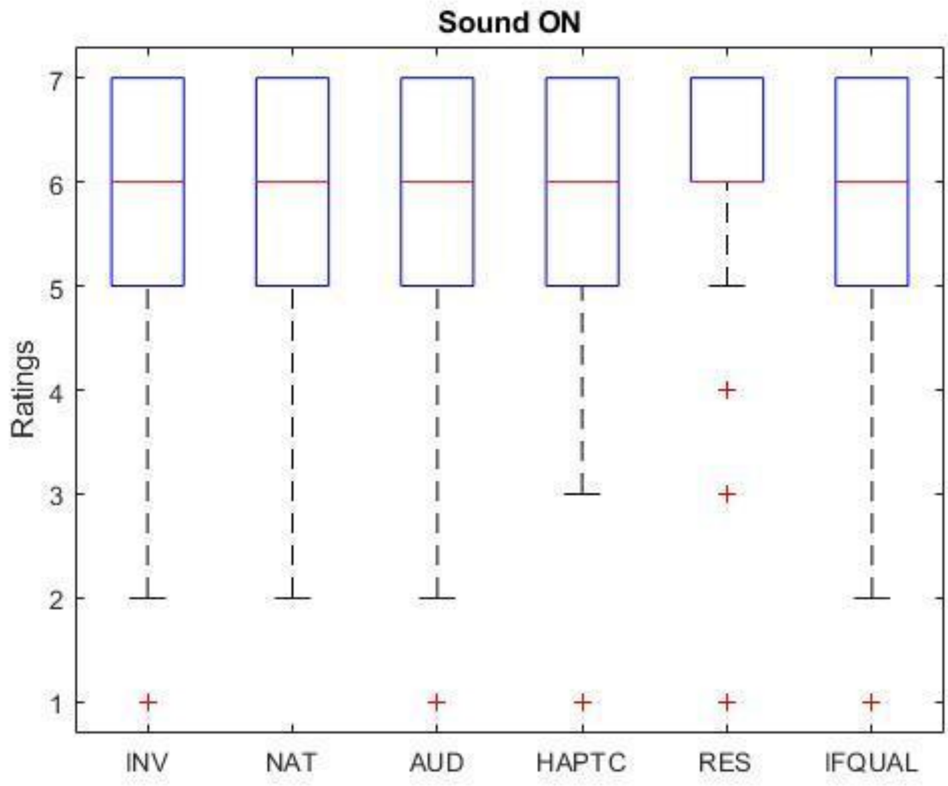


Figure 16. Results of Sound ON.

The following table shows the subscale comparison when the sound was turned on.

Table 4. Comparison of Subscale when Sound was ON.

Subscale	**Min**	**Q1**	**Median**	**Q3**	**Max**
INV/C	2	5	6	7	7
NAT	2	5	6	7	7
AUD	2	5	6	7	7
HAPTC	3	5	6	7	7
RES	5	6	6	7	7
IFQUAL	2	6	6	7	7

There is a striking resemblance that can be seen in each category when sound is on; the only difference is on the experience of touch, resolution, and interface quality; this difference is at min value, median which shows more positive feedback in these categories. In each category median is six, which shows that majority of the participants felt immersed, had better control, or experienced the high quality of interface, resolution, and interaction in IVR environment when sound was on.

The following plot shows the results of the normal scale in both experiments and represents the impact of the normal scale on presence in the IVR environment.

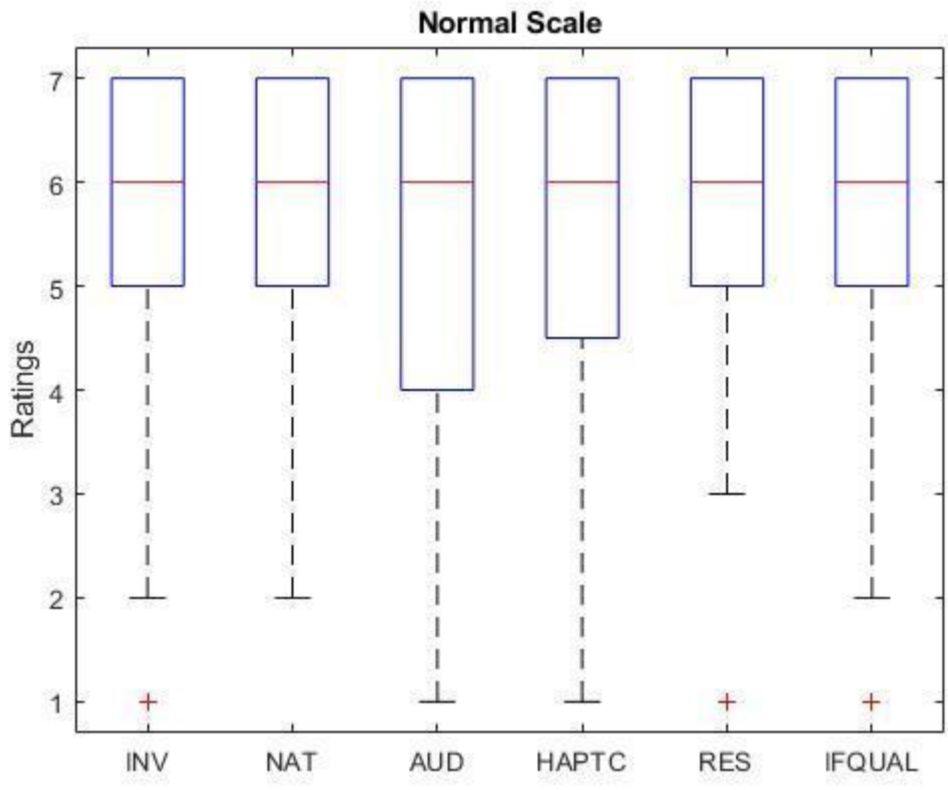


Figure 17. Results of Normal Scale.

When there was a normal scale in the IVR environment, the subscale comparison is shown in the table below.

Table 5. Comparison of Subscale when Scale was Normal.

Subscale	**Min**	**Q1**	**Median**	**Q3**	**Max**
INV/C	2	5	6	7	7
NAT	2	5	6	7	7
AUD	1	4	6	7	7
HAPTC	1	5	6	7	7
RES	3	5	6	7	7
IFQUAL	2	5	6	7	7

When they experienced the normal scale of a 3D crane in IVR, the majority of total participants had similar views in each category. On a normal scale, nearly 75% of total participants experienced high quality, better control, or felt more immersed in the IVR environment, while the remaining 25% remained neutral or had a bad experience. The only difference is in auditory aspects, where 50% of total participants experienced the high quality of sound and 25% experienced good quality of audio to some level or remained neutral. At the same time, the remaining 25% did not hear sound at all or experienced poor audio quality.

The following plot depicts the big-scale results in both experiments and the influence of big-scale on presence in the IVR environment.

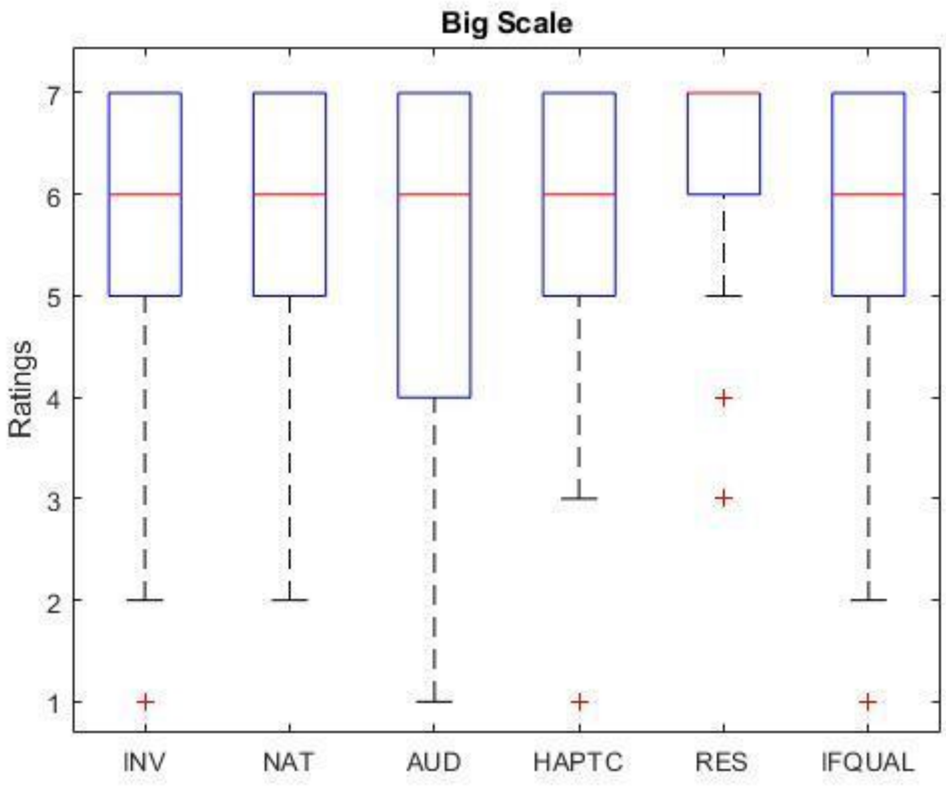


Figure 18. Results of Big Scale.

The subscale comparison is presented in the table below when there was a big scale in the IVR environment.

Table 6. Comparison of Subscale when Scale was Big.

Subscale	**Min**	**Q1**	**Median**	**Q3**	**Max**
INV/C	2	5	6	7	7
NAT	2	5	6	7	7
AUD	1	4	6	7	7
HAPTC	3	5	6	7	7
RES	5	6	7	7	7
IFQUAL	2	5	6	7	7

This summary represents that majority of the participants were involved and experienced great control, quality, sense of interaction, and we can say that about 97% of the total participants experienced high quality or good quality of resolution in IVR environment on a big scale as only couple of response below 5 one is at neutral value and one below neutral. Overall, the big scale improves the experience of the resolution.

6.4 Comparison of Results

The plot below shows a comparison of operating/interacting with a 3D crane in an IVR environment with and without sound. It also demonstrates how sound affects each of the presence measurement categories.

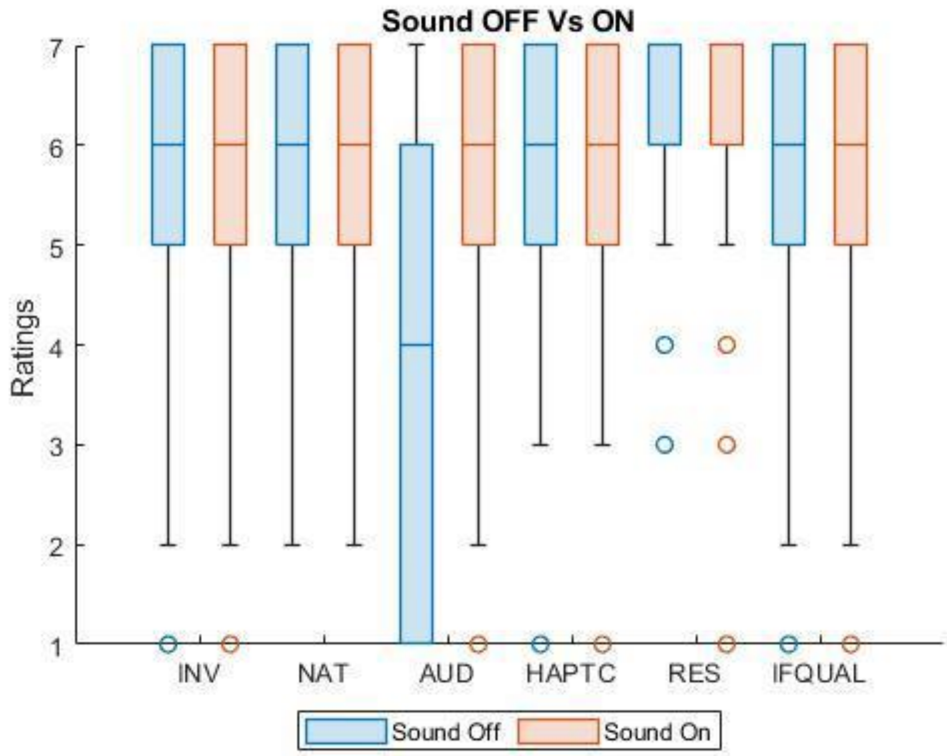


Figure 19. Sound OFF Vs Sound ON.

The following plot demonstrates how sound affects each of the presence measurement categories exempt AUD.

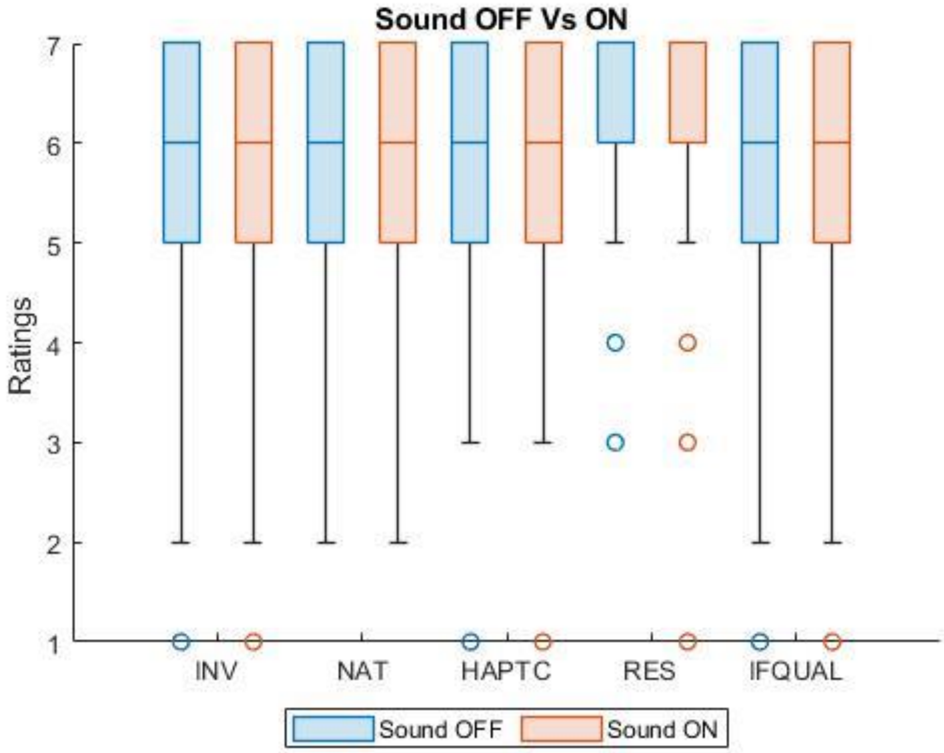


Figure 20. Sound OFF Vs ON without AUD.

It can be seen from above plot that if we ignore AUD subscale to measure presence, the results are same for each other subscales which indicates that sound doesn't affect participant's immersion or they or they responded questions other than AUD subscale without paying full attention to questionnaire as it was bit lengthy.

The subscale comparison is shown in the table below when the sound was switched on and off. In the upper section of each cell, it represents the value of sound off, and in the lower section, it represents the value when sound was on.

Table 7. Comparison of Subscale Sound ON Vs Sound OFF.

Subscale	Min Off/On	Q1 Off/On	Median Off/On	Q3 Off/On	Max Off/On
INV/C	2 2	5 5	6 6	7 7	7 7
NAT	2 2	5 5	6 6	7 7	7 7
AUD	1 2	1 5	6 6	6 7	7 7
HAPTC	3 3	5 5	6 6	7 7	7 7
RES	5 5	6 6	6 6	7 7	7 7
IFQUAL	2 2	5 5	6 6	7 7	7 7

This summary shows that sound has no effect on any other category of presence measurement; participants in each category had the same experience whether the sound was on or off. The only category that is influenced by sound is AUD, which should be affected by sound because it is associated with sound. In section AUD, 50% of the answers are below 4, indicating that the majority of people did not hear any sound when it was turned off, while on the other hand, only 25% of the responses are below 5, indicating that 75% of people experienced good quality of sound in the IVR environment when it was switched on.

The plot below compares operating/interacting with a 3D crane in an IVR environment on a big scale vs. a small scale. It also shows how each of the presence measurement categories is affected by scale.

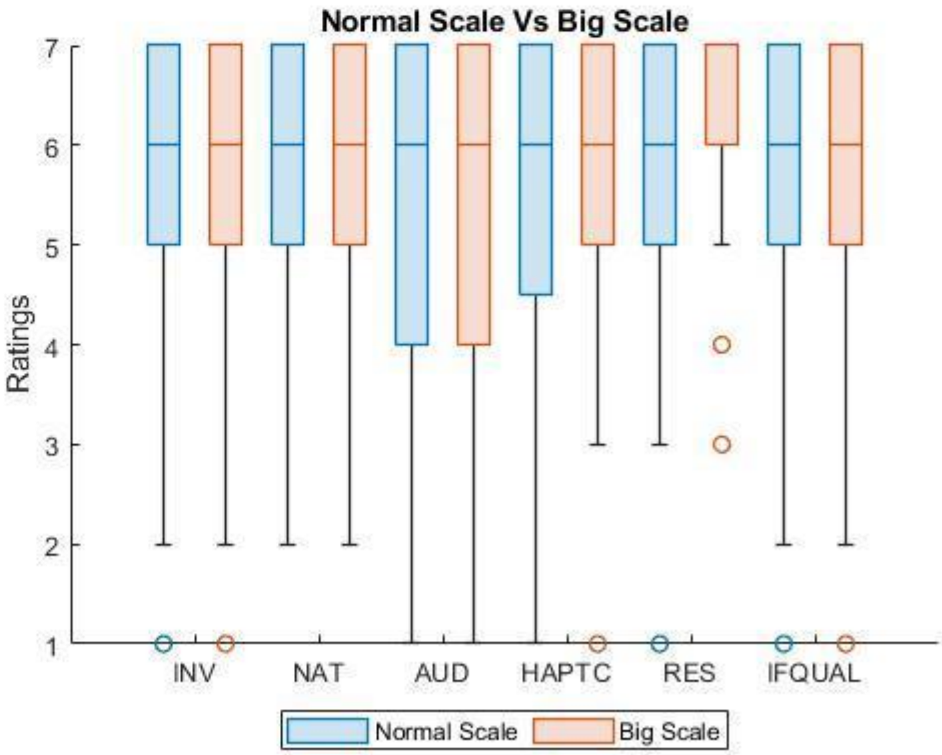


Figure 21. Normal Scale Vs Big Scale.

When the scale was big and normal, the subscale comparison is displayed in the table below. The value of the big scale is represented in the lower portion of each cell, while the value of the normal scale is represented in the upper section.

Table 8. Comparison of Subscale Normal Scale Vs Big Scale.

Subscale	**Min** Normal/ Big	**Q1** Normal/ Big	**Median** Normal/ Big	**Q3** Normal/ Big	**Max** Normal/ Big
INV/C	2 / 2	5 / 5	6 / 6	7 / 7	7 / 7
NAT	2 / 2	5 / 5	6 / 6	7 / 7	7 / 7
AUD	1 / 1	4 / 4	6 / 6	7 / 7	7 / 7
HAPTC	1 / 3	5 / 5	6 / 6	7 / 7	7 / 7
RES	3 / 5	5 / 6	6 / 7	7 / 7	7 / 7
IFQUAL	2 / 2	5 / 5	6 / 6	7 / 7	7 / 7

This summary reveals that big scale has affected the sense of interaction or touch and resolution in a positive way. On a big scale sense of interaction, whiskers' ends at three while on the normal scale, its end at one, which represents low ratings at normal scale for the HAPTC category, and we can say that resolution feedback on a big scale is improved by 25% compared to normal scale in IVR environment.

When sound was on and scale was big vs. small, the plot below compares operating/interacting with a 3D crane in an IVR environment. It also illustrates how the scale affects each of the presence measurement categories when the sound is turned on.

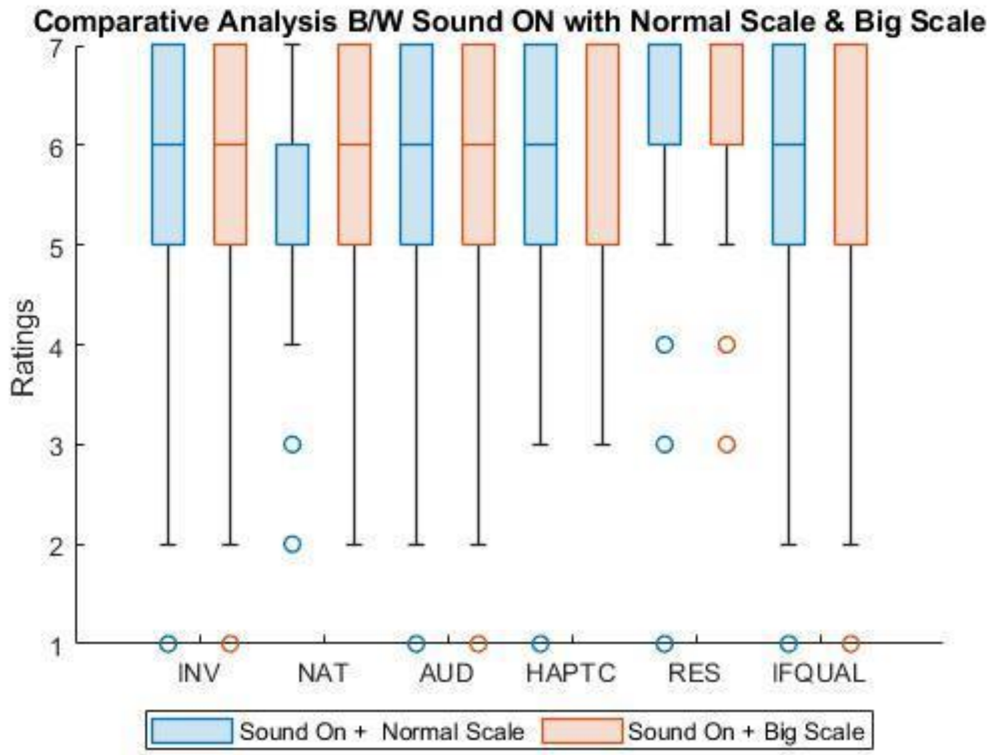


Figure 22. Comparison of Sound ON + Normal Scale Vs Sound ON + Big Scale.

The subscale comparison is shown in the table below when the scale was big, normal, and the sound was on. The value of the small scale with sound on is indicated in the upper half of each cell, while the value of the big scale with sound on is represented in the lower half.

Table 9. Subscale Comparison when Sound was ON at Normal Scale + Big Scale.

Subscale	Min	Q1	Median	Q3	Max
	SoundOn+ small scale/ big scale	**SoundOn+ small scale/ big scale**	**SoundOn+ small scale/ big scale**	**SoundOn+ small scale/ big scale**	**SoundOn+ small scale/ big scale**
INV/C	2	5	6	7	7
	2	5	6	7	7
NAT	4	5	6	6	7
	2	5	6	7	7
AUD	2	5	6	7	7
	2	5	6	7	7
HAPTC	3	5	6	7	7
	3	5	6	7	7
RES	5	6	7	7	7
	5	6	7	7	7
IFQUAL	2	5	6	7	7
	2	5	6	7	7

This summary represents that participants had the same level of involvement or control on a small scale and big scale when sound was on. They also have the same views about auditory, sense of interaction, resolution, and interface quality. The difference is in only one aspect, which is natural when sound was on, and scale was small 75% of the total responded under six while on big scale 75% percentile is at seven which represents that overall majority of the participants felt more natural environment on a big scale when sound was on.

The graph below shows how operating/interacting with a 3D crane in an IVR environment differs when the sound is turned off and the scale is big vs. small. It also shows how the scale affects each of the presence measurement categories when the sound is off.

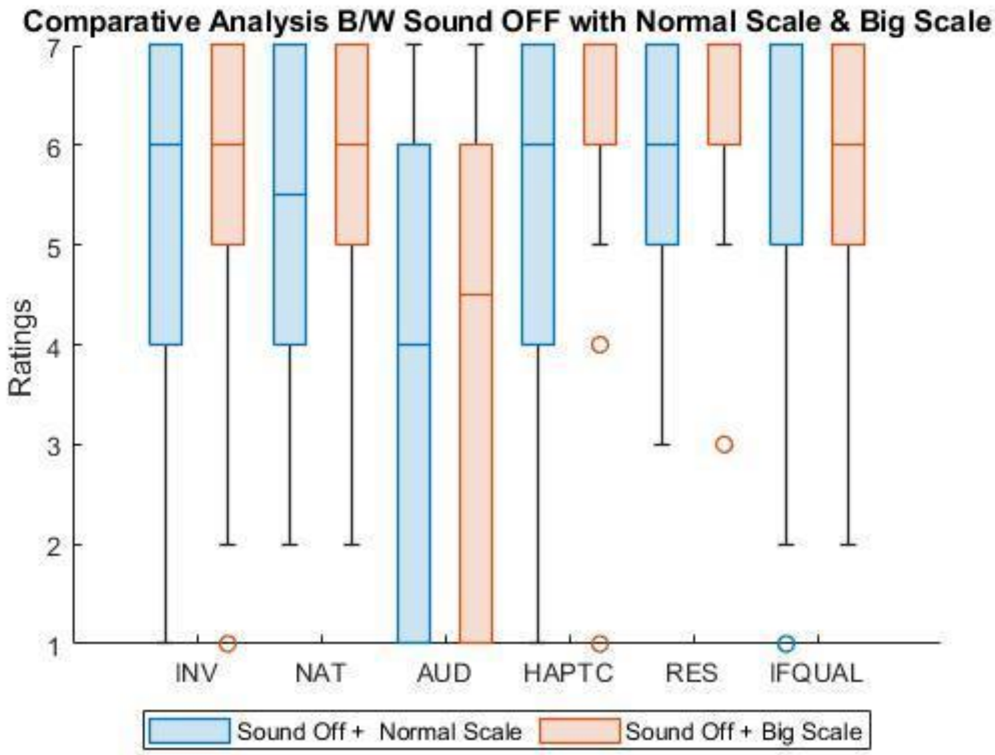


Figure 23. Comparison of Sound OFF + Normal Scale Vs Sound OFF + Big Scale.

The subscale comparison is given in the table below when the scale is big, normal, and sound is on. The value of the small scale with sound off is indicated in the top section of each cell, while the value of the big scale with sound off is represented in the below section.

Table 10. Subscale Comparison when Sound was OFF at Normal Scale + Big Scale.

Subscale	Min SoundOff+ small scale/ big scale	Q1 SoundOff+ small scale/ big scale	Median SoundOff+ small scale/ big scale	Q3 SoundOff+ small scale/ big scale	Max SoundOff+ small scale/ big scale
INV/C	1 2	4 5	6 6	7 7	7 7
NAT	2 2	4 5	6 6	7 7	7 7
AUD	1 1	1 1	4 5	6 6	7 7
HAPTC	1 5	4 6	6 7	7 7	7 7
RES	3 5	5 6	6 7	7 7	7 7
IFQUAL	2 2	5 5	6 6	7 7	7 7

This table represents that in the INV/C category minimum value on the big scale is 2, and 25% of participants responded under five while on a small scale, 25% of responses are below four which indicates that participants experienced good control and felt more involved on the big scale when sound was off. Also, the majority of the participants felt a more natural environment on a big scale. The sense of interaction or touch and high-quality resolution experience was increased by 25% on a big scale when sound was off.

The plot below demonstrates how gender affects how well a 3D crane is operated and interacted with in an IVR environment.

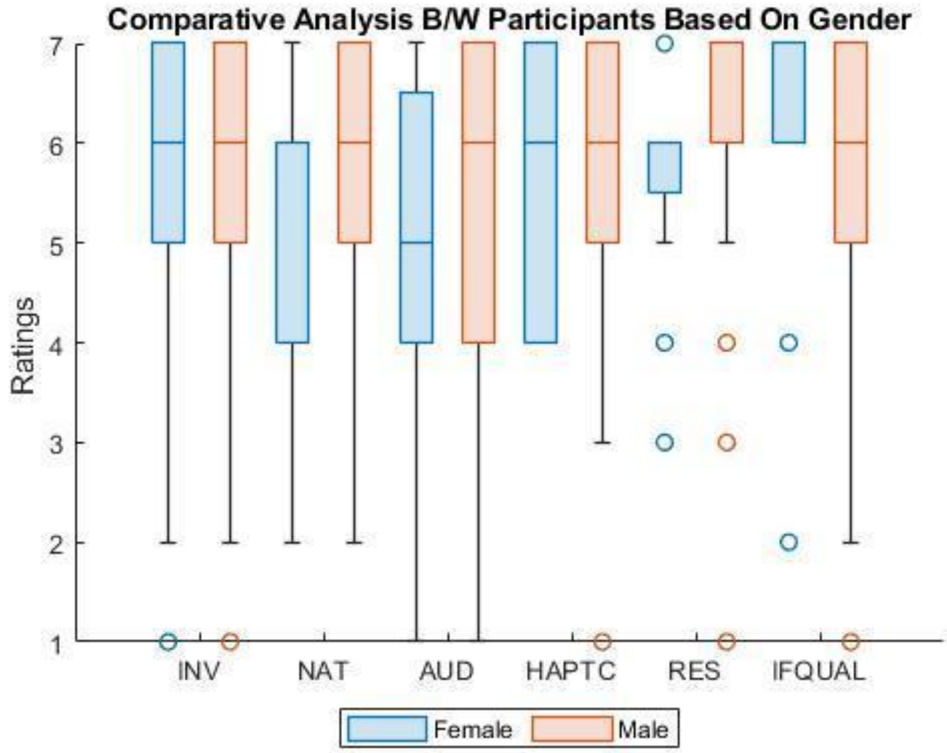


Figure 24. Comparative Analysis B/W Participants Based on Gender.

The table below shows a gender-based comparison. The female participants' value is shown in the top area of each cell, while the male participants' value is shown in the bottom section.

Table 11. Subscale Comparison Based on Participant's Gender.

Subscale	**Min** Female/ Male	**Q1** Female/ Male	**Median** Female/ Male	**Q3** Female/ Male	**Max** Female/ Male
INV/C	2 / 2	5 / 5	6 / 6	7 / 7	7 / 7
NAT	2 / 2	4 / 5	5 / 6	6 / 7	7 / 7
AUD	1 / 1	4 / 4	5 / 6	7 / 7	7 / 7
HAPTC	4 / 3	4 / 5	6 / 6	7 / 7	7 / 7
RES	5 / 5	5 / 6	6 / 6	6 / 7	6 / 7
IFQUAL	6 / 2	6 / 5	7 / 6	7 / 7	7 / 7

This summary depicts that there is no difference among INV/C both male and female participants experienced same level of involvement or control. In comparison to female participants, male participants felt more natural environment and experienced better auditory aspects around 25% more than female participants. Females had a greater feeling of interaction than males, with a minimum value of four for females and three for men. The male participants had a richer experience with resolution, whereas the female participants had a better experience with interface quality, with a minimum value of six for female participants and two for male participants.

The boxplot below shows how the age of a person influences how successfully a 3D crane is controlled and interacts with in an IVR environment.

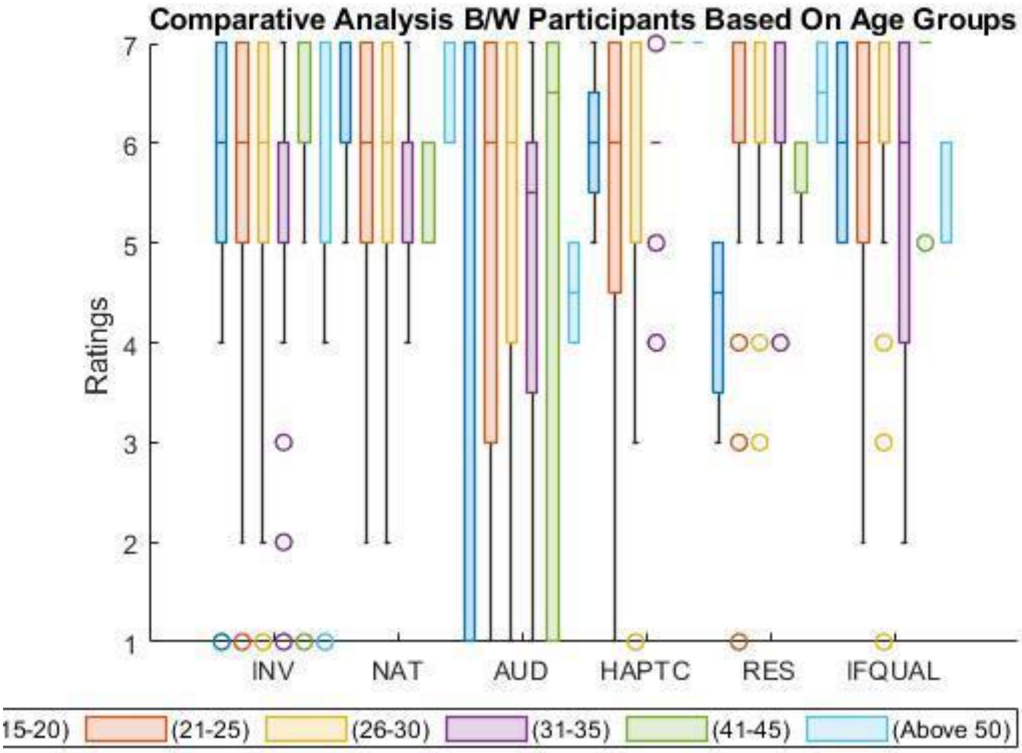


Figure 25. Comparative Analysis B/W Participants Based on Age Group.

The table below demonstrates how a person's age affects how well they can operate and interact with a 3D crane in an IVR environment.

Table 12. Subscale Comparison Based on Participant's Age Group.

Involvement/Control	This representation represents that from age 15-30 almost had same level control or involvement while operating the 3D crane in IVR environment, overall 50% of them experienced better or complete control and involvement. The level of presence was decreased by 25% for age group 31-35 and 25% increased for age group 41-45 comparing to 15-30 age group participants. Over the age of 50, 75% of the participants gave responses that were above neutral, and the remaining 25% of them remained neutral.
Natural	There is more diversity in responses of natural section for age 15-20 all of them responded above neutral value. For 21-30 age group they had same experience 50% of them experienced good or complete natural environment. The age group of 31-35 one fourth of them remained neutral remaining responded above neutral value. All of the

	responded between five to six for age group 41-45 which means that they experienced good natural environment and above 50 felt complete natural environment.
Auditory	This section shows that at least 25% of the total participants responded with minimum value 1 and upper quartile is at maximum value seven which indicates that 75% of them gave feedback under max value for age group 15-20. For age group 21-30 the feedback is almost same the difference is only at lower quartile value by one, overall 50% of them high quality of auditory elements. The median value for age group 31-35 is around 5.5 which indicates that most of them experienced good audio quality and median for 41-45 is around 6.5 which indicates that 50% of them experienced high quality audio and above 50 age the responses range is between four to five.
Haptic	The high quality of sense of interaction was experienced by 50% of the age group of 15-20. The age group 21-30 almost had same experience only 25% of them responded under five. The median for 31-35 age group is six which represents that most of them experienced high quality of interaction. And above 40 all of the responded with maximum value.
Resolution	For resolution 75% of the age group 15-20 gave feedback under five which means that they remained neutral or didn't experienced high quality resolution. For age group 21-35 they had same experience regarding resolution only 25% of them responded under six. For 41-45 the most of the responses are between five to six and participants above 50 responded between six to seven which indicates that they experienced high quality of resolution.
Interface Quality	The age group 15-25 and 31-35 had almost same experience of interface quality as the median value is six for these groups although there is difference in lower quartile value of age group 31-35 compared to age group 15-25. For age group 26-30 only 25% of them responded under six which represents that majority experienced high quality of interface. For age group most of the responses are at maximum value seven only one response is at five. The participants above 50 responded between five to six which indicates that they experienced average quality of interface.

The plot below illustrates how a person's previous VR experience affects how well a 3D crane is controlled and interacted with in an IVR environment.

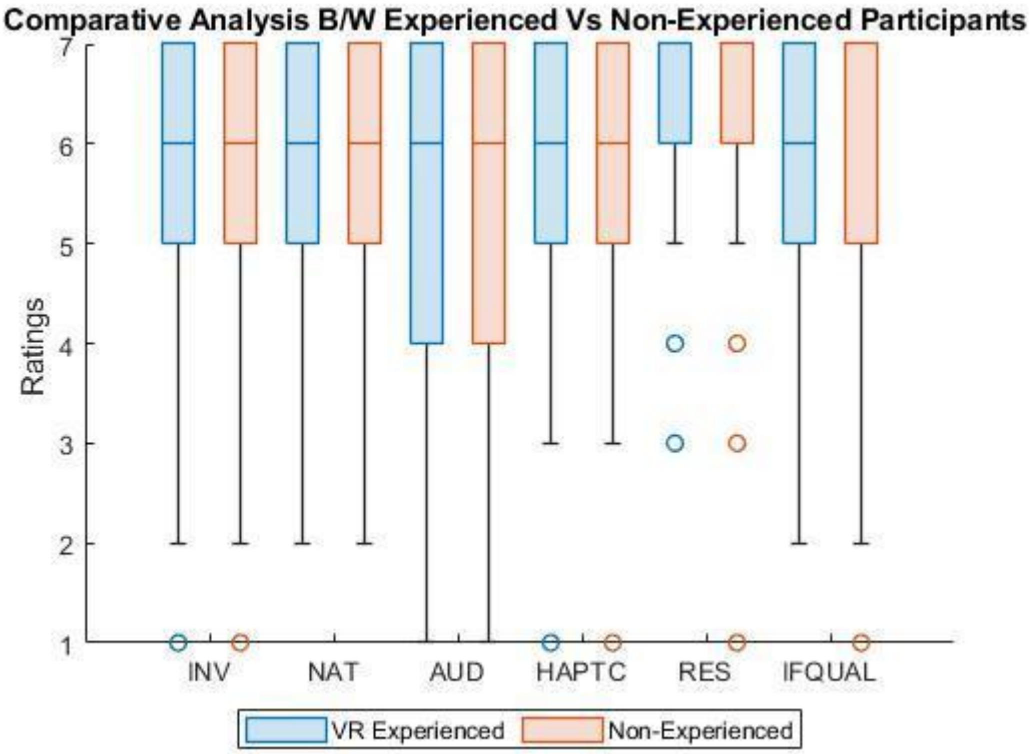


Figure 26. Comparative Analysis B/W Participants Based on Experience.

The table below shows the summary that represents that how experienced and non-experienced participants controlled or interacted with the 3D crane in IVR environment.

Table 13. Subscale Comparison Based on Participant's Experience.

Subscale	**Min** **Experienced/Non-Experienced**	**Q1** **Experienced/Non-Experienced**	**Median** **Experienced/Non-Experienced**	**Q3** **Experienced/Non-Experienced**	**Max** **Experienced/Non-Experienced**
INV/C	2 2	5 5	6 6	7 7	7 7
NAT	2 2	5 5	6 6	7 7	7 7
AUD	1 1	4 4	6 6	7 7	7 7
HAPTC	3 3	5 5	6 6	7 7	7 7
RES	5 5	6 6	6 6	7 7	7 7
IFQUAL	2 2	5 5	6 6	7 7	7 7

This table represents that the results are almost completely identical for both experienced and non-experienced participants, there is only little difference in experience of resolution and interface quality one extra outlier lies at minimum value one for non-experienced participants, which indicates that one participant from non-experienced user didn't feel good quality of resolution or interface at all.

6.4.1 Mean and SD of Different Variables

The following tables represents the mean and standard deviation (SD) of different variables. To calculate these, we write a code in MATLAB on our imported data.

Standard deviation and mean of four main variables are depicted in the following table.

Table 14. SD and Mean of Different Variables.

Variables	Standard Deviation	Mean
Normal Scale	1.65	5.4853
Big Scale	1.59	5.7178
Sound Off	1.54	5.47
Sound On	1.45	5.7331

Following table depicts the SD and mean of different variables with possible combinations.

Table 15. SD and Mean of Variables with Combination.

Variables	Standard Deviation	Mean
Sound On Small Scale	1.47	5.6098
Sound On Big Scale	1.42	5.8481
Sound Off Small Scale	1.79	5.3691
Sound Off Big Scale	1.74	5.578

The following table represents the SD and mean based on participant's distribution.

Table 16. SD and Mean Based on Participants Distribution.

Variables	Standard Deviation	Mean
Male Participants	1.63	5.62
Female Participants	1.63	5.38
VR Experienced Participants	1.59	5.53
Non- Experienced	1.66	5.67
Age 15-20	1.82	5.57
Age 21-25	1.70	5.50
Age 26-30	1.60	5.71
Age 31-35	1.46	5.37
Age 41-45	1.62	6.01
Above 50	1.37	5.61

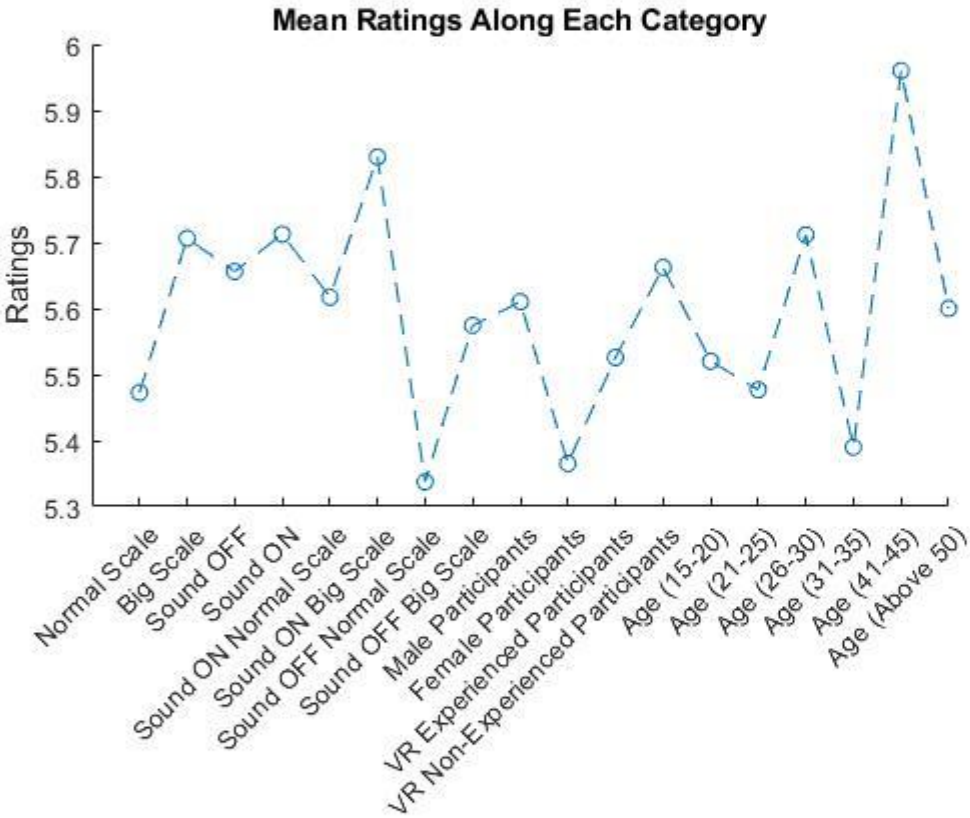


Figure 27. Mean Ratings of Each Category.

The average response of a normal scale is less than the average response of a big scale. The average response of sound off is less than the average response of sound on. Similarly, the average response of on sound when the scale was small is less than when sound was on, and scale was big. When sound was off on a small scale, the average score was lesser than the average score when sound was on, and the scale was big. This represents those participants felt more immersed or present in the IVR environment when sound was on or the scale was big. Participants between the age of 41-45 were more immersed and present in the IVR environment than those in other age groups, as their mean value is highest compared to their age groups. We didn't have any participants in our experiments who were between the ages of 36-40, thus the mean value for this age group is missing.

7 Conclusions

Virtual reality is growing in importance and dominance in the everyday lives of a growing number of people on a daily basis. Different aspects of the VR environment are crucial for the quality of the experience.

In this thesis a system and method for measuring audio visual elements in IVR is suggested. To measure the impact of these two elements on subjects in IVR, a digital twin of a 3D crane was used and manipulated in such a way that it can measure audio visual elements. The manipulation or upgradation of DT of 3D includes addition of sound and varied scale. The data was collected by 58 participants and each participant did two experiments with variation of sound and scale. In this way participants interacted with a 3D crane in an IVR environment with sound, without sound, laboratory-scale (Normal Scale) and bigger scale 3D crane. To measure these elements a standrized presence questionnaire was used.

The results from the feedback of participants indicates that sound and visual elements play a vital role in the IVR environment. If there was no sound the level of presence was decreased compared to when there was sound on. Also participants were more involved and had better experience when the scale of the 3D crane was big compared to the normal scale. From this, several important conclusions can be drawn. First, it is apparent that sound plays an important role in the process of immersion: up until this point the digital twin of the laboratory-scale 3D crane did not have audio, and it is now apparent that sound should be properly implemented. Second, the greater the scale of the crane, the more believable the simulation becomes. The reason for this is most likely that the lab-scale model of the crane does not appear in the experience of an average person; on the other hand, true-to-scale cranes are in abundance everywhere and hence can be related to more easily.

The percentage of the mean value of all subscales when sound was on is 81.6% and sound off percentage 78.1% which shows that the overall mean percentage was increased by 3.5% when sound was on. As far as visual elements are concerned, it also shows that big scale was more immersive as compared to normal scale (laboratory scale). The percentage of the mean value of all subscales when the scale was big is 81.68%, and at normal scale, the percentage is 78.36% which reflects that there was an increase of 3.32% in the mean value of all subscales, or we can say that average response was increased by 3.32% when the scale was big in IVR environment.

The four parameters that we used in our data collection process were sound, scale, eye gate, and hand controller. These factors had an effect on our experiments as well. Our investigation was limited to two variables: sound and scale. More study can be done in the future using a combination of these four elements for future advancements.

References

[1] Smiith, S. (2019). The Effect of Increased Immersion and its Relationship with Presence on Episodic Memory Performance within Virtual Reality Environments. https://doi.org/10.17615/qy9n-5374

[2] C. Machover and S. E. Tice, "- Virtual Reality - Virtual Reality," *IEEE Comput. Graph. Appl.*, vol. 1, no. January, pp. 15–16, 1997.

[3] "Ueber das Organ der Seele. von Soemmerring, Samuel Thomas v.: (1796) | Antiq. F.-D. Söhn - Medicusbooks.Com." https://www.zvab.com/Ueber-Organ-Seele-Soemmerring-Samuel-Thomas/861487385/bd.

[4] N. C. Burbules, "Rethinking the Virtual Repensando o virtual," pp. 89–107, 2004.

[5] K. S. Hale and K. M. Stanney, "Handbook of virtual environments : design, implementation, and applications.," p. 1418.

[6] A. Fuller, Z. Fan, C. Day, and C. Barlow, "Digital Twin: Enabling Technologies, Challenges and Open Research," *IEEE Access*, vol. 8, pp. 108952–108971, 2020, doi: 10.1109/ACCESS.2020.2998358.

[7] Negri, E., Fumagalli, L., & Macchi, M. (n.d.). *NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/). Peer-review under responsibility of the scientific committee of the 27th International Conference on Flexible Automation and Intelligent Manufacturing ScienceDirect A review of the roles of Digital Twin in CPS-based production systems-review under responsibility of the scientific committee of the 27th International Conference on Flexible Automation and Intelligent Manufacturing*. https://doi.org/10.1016/j.promfg.2017.07.198

[8] S. Jersov and A. Tepljakov, "Digital Twins in Extended Reality for Control System Applications," 2020 43rd International Conference on Telecommunications and Signal

Processing (TSP), 2020.

[9] T. Gabor, L. Belzner, M. Kiermeier, M. T. Beck, and A. Neitz, “A simulation-based architecture for smart cyber-physical systems,” *Proc. - 2016 IEEE Int. Conf. Auton. Comput. ICAC 2016*, pp. 374–379, 2016, doi: 10.1109/ICAC.2016.29.

[10] M. Bajaj, D. Zwemer, and B. Cole, “Architecture to geometry - Integrating system models with mechanical design,” *AIAA Sp. Astronaut. Forum Expo. Sp. 2016*, 2016, doi: 10.2514/6.2016-5470.

[11] V. Kuts, T. Otto, T. Tähemaa, and Y. Bondarenko, “Digital twin based synchronised control and simulation of the industrial robotic cell using virtual reality,” *J. Mach. Eng.*, vol. 19, no. 1, pp. 128–144, 2019, doi: 10.5604/01.3001.0013.0464.

[12] V. V. Makarov, Y. B. Frolov, I. S. Parshina, and M. V. Ushakova, “The design concept of digital twin,” *Proc. 2019 12th Int. Conf. "Management Large-Scale Syst. Dev. MLSD 2019*, pp. 1–4, 2019, doi: 10.1109/MLSD.2019.8911091.

[13] K. Rogers, G. Ribeiro, R. R. Wehbe, M. Weber, and L. E. Nacke, “Vanishing importance: Studying immersive effects of game audio perception on player experiences in virtual reality,” *Conf. Hum. Factors Comput. Syst. - Proc.*, vol. 2018-April, Apr. 2018, doi: 10.1145/3173574.3173902.

[14] E. Brown and P. Cairns, “A grounded investigation of game immersion,” *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 1297–1300, 2004, doi: 10.1145/985921.986048.

[15] C. I. Jennett, “Is game immersion just another form of selective attention? An empirical investigation of real world dissociation in computer game immersion,” *PhD Thesis UCL*, no. July, p. 243, 2010, [Online]. Available: http://eprints.ucl.ac.uk/20225/.

[16] C. Jennett *et al.*, “Measuring and Defining the Experience of Immersion in Games. International Journal of Human Computer Studies 66(9),” *Int. J. Human_Computer Stud.*, vol. 66, no. 9, pp. 641–661, 2008, [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.157.4129&rep=rep1&type=pdf.

[17] W. R. Sherman and A. B. Craig, “Understanding virtual reality : interface, application,

and design," p. 582, 2003.

[18] Y. Ali, School of Information Technologies "Investigation of the Stroop Effect in a Virtual Reality Environment" TALLINN 2020.

[19] K. Das Thoondee and A. Oikonomou, "Using virtual reality to reduce stress at work," *Proc. Comput. Conf. 2017*, vol. 2018-Janua, no. July, pp. 492–499, 2018, doi: 10.1109/SAI.2017.8252142.

[20] J. Tham, A. H. Duin, L. Gee, N. Ernst, B. Abdelqader, and M. McGrath, "Understanding Virtual Reality: Presence, Embodiment, and Professional Practice," *IEEE Trans. Prof. Commun.*, vol. 61, no. 2, pp. 178–195, Jun. 2018, doi: 10.1109/TPC.2018.2804238.

[21] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino, "<title>Augmented reality: a class of displays on the reality-virtuality continuum</title>," *Telemanipulator Telepresence Technol.*, vol. 2351, no. January, pp. 282–292, 1995, doi: 10.1117/12.197321.

[22] "Virtual reality definition | Strate, School of Design." https://www.strate.education/gallery/news/vr-definition (accessed Oct. 31, 2021).

[23] F. Ruotolo *et al.*, "Immersive virtual reality and environmental noise assessment: An innovative audio-visual approach," *Environ. Impact Assess. Rev.*, vol. 41, pp. 10–20, 2013, doi: 10.1016/j.eiar.2013.01.007.

[24] G. M. Echevarria Sanchez, T. Van Renterghem, K. Sun, B. De Coensel, and D. Botteldooren, "Using Virtual Reality for assessing the role of noise in the audio-visual design of an urban public space," *Landsc. Urban Plan.*, vol. 167, no. June, pp. 98–107, 2017, doi: 10.1016/j.landurbplan.2017.05.018.

[25] J. Y. Hong and J. Y. Jeon, "The effects of audio-visual factors on perceptions of environmental noise barrier performance," *Landsc. Urban Plan.*, vol. 125, pp. 28–37, 2014, doi: 10.1016/j.landurbplan.2014.02.001.

[26] L. Maffei *et al.*, "The effects of vision-related aspects on noise perception of wind turbines in quiet areas," *International Journal of Environmental Research and Public Health*, vol. 10, no. 5. pp. 1681–1697, 2013, doi: 10.3390/ijerph10051681.

[27] J. Y. Hong, J. He, B. Lam, R. Gupta, and W. S. Gan, "Spatial audio for soundscape design: Recording and reproduction," *Appl. Sci.*, vol. 7, no. 6, 2017, doi: 10.3390/app7060627.

[28] M. Naef, O. Staadt, and M. Gross, "Spatialized audio rendering for immersive virtual environments," *ACM Symp. Virtual Real. Softw. Technol. Proceedings, VRST*, no. October, pp. 65–72, 2002, doi: 10.1145/585740.585752.

[29] L. Maffei, M. Masullo, F. Aletta, and M. Di Gabriele, "The influence of visual characteristics of barriers on railway noise perception," *Sci. Total Environ.*, vol. 445–446, pp. 41–47, 2013, doi: 10.1016/j.scitotenv.2012.12.025.

[30] W. Zhang, P. N. Samarasinghe, H. Chen, and T. D. Abhayapala, "Surround by sound: A review of spatial audio recording and reproduction," *Appl. Sci.*, vol. 7, no. 5, pp. 1–19, 2017, doi: 10.3390/app7050532.

[31] Z. Yan, J. Wang, and Z. Li, "A multi-criteria subjective evaluation method for binaural audio rendering techniques in virtual reality applications," *Proc. - 2019 IEEE Int. Conf. Multimed. Expo Work. ICMEW 2019*, pp. 402–407, 2019, doi: 10.1109/ICMEW.2019.00075.

[32] C. Xu, T. Oberman, F. Aletta, H. Tong, and J. Kang, "Ecological Validity of Immersive Virtual Reality (IVR) Techniques for the Perception of Urban Sound Environments," *Acoustics*, vol. 3, no. 1, pp. 11–24, 2020, doi: 10.3390/acoustics3010003.

[33] J. Y. Jeon and H. I. Jo, "Effects of audio-visual interactions on soundscape and landscape perception and their influence on satisfaction with the urban environment," *Build. Environ.*, vol. 169, no. August 2019, p. 106544, 2020, doi: 10.1016/j.buildenv.2019.106544.

[34] F. Pallavicini, A. Pepe, and M. E. Minissi, "Gaming in Virtual Reality: What Changes in Terms of Usability, Emotional Response and Sense of Presence Compared to Non-Immersive Video Games?," *Simul. Gaming*, vol. 50, no. 2, pp. 136–159, Apr. 2019, doi: 10.1177/1046878119831420.

[35] J. M. Normand, E. Giannopoulos, B. Spanlang, and M. Slater, "Multisensory stimulation

can induce an illusion of larger belly size in immersive virtual reality," *PLoS One*, vol. 6, no. 1, 2011, doi: 10.1371/journal.pone.0016128.

[36] J.-Y. Lai, J.-L. Doong, J.-R. Hwang, M.-C. Jeng, and C.-Y. Gau, "Development of a Virtual Simulation System for Crane-Operating Training," *Proc. ASME Turbo Expo*, Dec. 2014, doi: 10.1115/97-AA-045.

[37] H. Song, T. Kim, J. Kim, D. Ahn, and Y. Kang, "Effectiveness of VR crane training with head-mounted display: Double mediation of presence and perceived usefulness," *Autom. Constr.*, vol. 122, no. June 2020, p. 103506, 2021, doi: 10.1016/j.autcon.2020.103506.

[38] B. E. Riecke, A. Väljamäe, and J. Schulte-Pelkum, "Moving sounds enhance the visually-induced self-motion illusion (circular vection) in virtual reality," *ACM Trans. Appl. Percept.*, vol. 6, no. 2, 2009, doi: 10.1145/1498700.1498701.

[39] R. Chinta, "Measurement of Game Immersion through Subjective Approach," no. June, 2017.

[40] S. Adam Smith. (2019). THE EFFECT OF INCREASED IMMERSION AND ITS RELATIONSHIP WITH PRESENCE ON EPISODIC MEMORY PERFORMANCE WITHIN VIRTUAL REALITY ENVIRONMENTS. *A Dissertation Submitted to the Faculty at the University of North Carolina at Chapel Hil*, iii–26.

[41] I. Hupont, J. Gracia, L. Sanagustin, and M. A. Gracia, "How do new visual immersive systems influence gaming QoE? A use case of serious gaming with Oculus Rift," *2015 7th Int. Work. Qual. Multimed. Exp. QoMEX 2015*, 2015, doi: 10.1109/QoMEX.2015.7148110.

[42] "3D Crane | INTECO." https://www.inteco.com.pl/products/3d-crane/ (accessed Oct. 31, 2021).

[43] S. Alsaleh, A. Tepljakov, M. Tamre and E. Petlenkov, "Towards Artificial Intelligence Driven Immersive Environments in Virtual Reality for Industrial Applications," *2021 44th International Conference on Telecommunications and Signal Processing (TSP),*

2021, pp. 340-345, doi: 10.1109/TSP52935.2021.9522676.

[44] Zamana, O., & Tallinna Tehnikaülikool. (2021). *Personal identification in Virtual and Augmented Reality: Control Systems Perspective*. Tallinna Tehnikaülikooli Kirjastus. https://digikogu.taltech.ee/et/Download/ca0ea7c6-c4a6-4b40-837d-7c12a31b6c9f

[45] Witmer, B. G., & Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. In *Presence* (Vol. 7, Issue 3).

[46] Alhazza, K., Masoud, Z., and Alotaibi, N. (2016) A Smooth Wave-Form Shaped Command with Flexible Maneuvering Time: Analysis and Experiments. *Asian Journal of Control*, 18: 1376– 1384. doi: 10.1002/asjc.1204.

[47] Turbosquid, shutterstack. (2021). *rtg crane 3d model*. (n.d.). Retrieved from https://www.turbosquid.com/3d-models/rtg-crane-3d-model/883586

[48]https://www.zoom.co.jp/sites/default/files/styles/product_feature/public/products/product-features/h1_sides.png?itok=cw8s91mB

[49] Abel, M., & Tallinna Tehnikaülikool. (2019). *Towards a Framework for Analysis of Human Behaviour in VR-based Applications*. https://digikogu.taltech.ee/en/Download/0ca0563c-43d3-4d0b-81d9-87d4eb724cc5

[50] Schuemie MJ, van der Straaten P, Krijn M, van der Mast CA. Research on *presence in virtual reality: a survey*. *Cyberpsychol Behav*. 2001 Apr;4(2):183-201. doi: 10.1089/109493101300117884. PMID: 11710246.

[51] Slater, M., & Usoh, M. (1993). Representations Systems, Perceptual Position, and Presence in Immersive Virtual Environments. *Presence: Teleoperators and Virtual Environments*, *2*(3), 221–233. https://doi.org/10.1162/pres.1993.2.3.221

[52] Sheridan, T. B. (1992). Musings on Telepresence and Virtual Presence. *Presence: Teleoperators and Virtual Environments*, *1*(1), 120–126. https://doi.org/10.1162/pres.1992.1.1.120

[53] Steuer, J., & Reeves, B. (1992). Defining Virtual Reality: Dimensions Determining Telepresence. In *Journal of Communication* (Vol. 42, Issue 4). http://www,cyborganic.com/

[54] *Understanding Boxplots. | by Michael Galarnyk | Towards Data Science*. (n.d.). Retrieved from https://towardsdatascience.com/understanding-boxplots-5e2df7bcbd51

[55] V. Schwind, P. Knierim, N. Haas, and N. Henze, "Using presence questionnaires in virtual reality," *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 1–12, 2019, doi: 10.1145/3290605.3300590.

[56] McLeod, S. A. (2019, July 19). What does a box plot tell you? Simply psychology: https://www.simplypsychology.org/boxplots.html

Appendix 1 – Non-exclusive licence for reproduction and publication of a graduation thesis[1]

I , Daud Suleman

1. Grant Tallinn University of Technology free licence (non-exclusive licence) for my thesis "Measuring Audio Visual Elements in Immersive Virtual Reality" supervised by Saleh Al-saleh and Aleksei Tepljakov.
   1.1. to be reproduced for the purposes of preservation and electronic publication of the graduation thesis, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright;
   1.2. to be published via the web of Tallinn University of Technology, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright.
2. I am aware that the author also retains the rights specified in clause 1 of the non-exclusive licence.
3. I confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights, the rights arising from the Personal Data Protection Act or rights arising from other legislation.

03.01.2022

1 The non-exclusive licence is not valid during the validity of access restriction indicated in the student's application for restriction on access to the graduation thesis that has been signed by the school's dean, except in case of the university's right to reproduce the thesis for preservation purposes only. If a graduation thesis is based on the joint creative activity of two or more persons and the co-author(s) has/have not granted, by the set deadline, the student defending his/her graduation thesis consent to reproduce and publish the graduation thesis in compliance with clauses 1.1 and 1.2 of the non-exclusive licence, the non-exclusive license shall not be valid for the period.