SCHOOL OF COMPUTATION, INFORMATION AND TECHNOLOGY — INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Bachelor's Thesis in Information Systems

Exploring the boundaries of Visual, Auditory and Haptic Effects for Immersive Learning VR Expieriences

Petr Ershov

SCHOOL OF COMPUTATION, INFORMATION AND TECHNOLOGY — INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Bachelor's Thesis in Information Systems

Exploring the boundaries of Visual, Auditory and Haptic Effects for Immersive Learning VR Expierences

Untersuchung der Grenzen visueller, auditiver und haptischer Effekte für immersive VR-Lernumgebungen

Author:	Petr Ershov
Examiner:	Prof. Dr. Enkelejda Kasneci
Supervisor:	Prof. Dr. Enkelejda Kasneci
Submission Date:	07.11.2024

Contents

Declaration of Academic Integrity

I confirm that this bachelor's thesis is my own work and I have documented all sources and material used. I hereby declare that the thesis submitted is my own unaided work. All direct or indirect sources used are acknowledged as references. I am aware that the thesis in digital form can be examined for the use of unauthorized aid and in order to determine whether the thesis as a whole or parts incorporated in it may be deemed as plagiarism. For the comparison of my work with existing sources I agree that it shall be entered in a database where it shall also remain after examination, to enable comparison with future theses submitted. Further rights of reproduction and usage, however, are not granted here. This paper was not previously presented to another examination board and has not been published. Throughout the writing process of this thesis, I have used ChatGPT and Grammarly to spellcheck and improve the grammatical structure of my writing. After using these tools, I thoroughly reviewed and edited the content as needed. I take full responsibility for the final outcome and any and all consequences that may arise from my usage of these tools

Munich, 07.11.2024 Petr Ershov

Acknowledgments

In this part I want to express my gratitude and reverence to the Chair for Human-Centered Technologies for Learning for giving me an opportunity to do both practice course and a bachelor thesis with all the support, assistance and sponsorship. This help allowed me to gather valuable information, find survey participants and use cutting-edge technologies. I also want to thank Mengdi Wang for providing great support in doing the research. I am grateful to all the people who agreed to participate in the research, do the survey and share their honest opinions and feelings. Most of all, I am thankful to my entire family, especially my mother, Irena Ershova, and my father, Andrew Ershov, for providing heartfelt support, help, and encouragement during my studies at the Technical University of Munich.

Abstract

This study focuses on exploring the possibilities and boundaries of interactions in educative world of virtual reality serious games. Using three different human perception senses: Vision, Hearing and Touch(Tactile), the study explores what existing approaches there are and how those can be used in increasing the information comprehension. This paper also includes a survey game project of 40 participants, who were each assigned to groups, based on the hints provided: control, visual, auditory or haptic groups. Each group played a space exploration based VR game which included two tasks: one is to assemble a puzzle with four parts out of nine lying on the ground. Second task included understanding gravity variations between Mars, Venus and Moon. In contrast to control group, treated(Visual, Auditory and Haptic) users received hints based on their assigned sense perception group. The study then further research the effects of those hints and tries to prove if users under treatment achieve better results and using eye-tracking technologies the study compares the puzzle stages and grab durations to time focused at specific objects to understand cognitive processes hidden behind every treatment group. Other then that the study also includes the design implementation of the educational game and future work that can be done with it.

Using the survey results of user's scores, NASA TLX Ratings and eye-tracking data it was possible to prove the main hypothesis with p-value of 5 percent, suggesting that the Control group performs worse than at least one other group, indicating that hints significantly aid task performance. Study furthermore showed a significant statistical connection that the best performing group(Visual) also has the best cognitive load values.

Keywords: virtual reality, puzzle, perception senses, education, interactive, eye tracking

1 Introduction

1.1 Motivation and Background

Virtual reality classes are no longer just a myth, and many museums are already integrating Virtual Reality (VR) or Augmented Reality (AR) exhibits. An animal, spacecraft, or historical object can immediately be viewed virtually on a table in one's room by conducting a simple online search. This capability is enabled by AR technologies that have already established a place in everyday life.

A primary motivation for this study is that Virtual Reality can enhance learning in environments that are otherwise inaccessible within real-world educational settings. VR can also increase engagement through immersion and the enjoyment of solving gamified tasks. Astronomy science was selected as the focus for this paper, given its prevalence in educational contexts and the requirement for imagination to fully understand many of its concepts. With VR technologies, many components of astronomy can be simulated, as organizations like European Space Agency (ESA) and The National Aeronautics and Space Administration (NASA) are already utilizing VR to train astronauts and scientists. [NAS24]

Concurrently, one compelling motivation is that VR technology can be seamlessly integrated into museums, exhibitions, and schools as part of educational programs, with headsets becoming lighter and easier to configure. Simple VR games already offer users a more accessible understanding of complex topics while fostering interest through gamification, allowing for active engagement in learning experiences rather than passive observation.

1.2 Structure

This paper examines an often-overlooked aspect of many exhibitions and experiences: Interactions. These interactions enhance comprehension and understanding of a topic while providing enjoyment and engaging new user groups. This area of study is closely linked to human senses, which are crucial for perceiving and learning from the surrounding world. There are five primary senses in humans: sight, hearing, touch, smell, and taste.[Wik24b][SS20][TWL22][Boy23] [San+21] Sight is the most frequently employed sense in VR exhibitions and learning experiences. [SS20]

This research therefore focuses on three major senses: vision, hearing and touch. All three senses are analyzed by comparing their impact on learning in a VR game, quality of experience and cognitive load during the playtime.

The study includes a survey of 40 participants, divided into four groups: Visual, Auditory, Haptic, and Control, according to the hints. [SS20], [TWL22], [Boy23] Each group consisted of 10 users, trying to maintain diversity in age, education and previous VR knowledge within each group. Each group completed the same task. Following the task, participants answered a 15 question long multiple-choice questionnaire. During game play, data on eye-tracking, decision steps, and grab interactions was collected for use in the evaluation. Chapter 2 discusses current approaches to VR or AR learning and examples of its application in astronomy education. Chapter 3 outlines the design and implementation of the VR game. Chapter 4 outlines implementation principles, used tools and faced challenges during this step. Chapter 5 and Chapter 6 provide further details on the survey methodology. Chapter 6 explores further potential work to enhance understanding of interaction effects and run further statistical research.

1.3 Aims and Objectives/Hypotheses

This section outlines the key hypotheses, which are used across the whole study, with focus on the impact of different hints on participants' task performance and cognitive load. The primary objective is to test whether the Control group performs equally well or worse than the other groups, and whether hints (such as visual, auditory, or haptic) significantly enhance performance(user's score).

1.3.1 Main Hypothesis

- **H0**: The Control group achieves results that are equal to or better than each other group individually (Auditory, Haptic, and Visual), suggesting that hints do not significantly help solve the tasks.
- **H1**: The Control group performs worse than at least one other group, indicating that sensory hints significantly enhance performance(user's score).

Separate tests were conducted for each group listed above to determine whether the difference from the Control group is statistically significant.

1.3.2 Cognitive Load and Dominance Hypothesis

- **H0**: There is no difference in cognitive load between the best-performing group and each other group.
- **H1**: The best-performing group experiences lower cognitive load compared to other groups.

1.3.3 Visual Sensory Dominance Hypothesis

- **H0**: There is no significant advantage for the Visual group in task performance compared to other groups.
- **H1**: The Visual group performs better in tasks due to reliance on visual dominance.

1.3.4 Multimedia Learning Hypothesis

It's furthermore investigated on other connected theories, like Multimedia Learning Theory [May24] that could conclude following:

- **H0**: There is no significant difference in performance between the combined sensory information group and the control groups.
- **H1**: The combination of different sensory information in combined groups leads to improved performance compared to the control groups.

This theory can not be proven statistically in this paper due to the way the survey is structured(by observing individual group-specific results) and no option to run independent t-tests on combined groups. It will still be contemplated on possible outcomes for the theory and observed for common improvements of multiple sense group combinations.

2 Related Work

2.1 Virtual Reality in Education

VR is increasingly popular within schools and university courses. A notable example is the University of Canterbury [Sec24], where VR was used to support early childhood teacher training programs for students without access to working with real infants, by simulating infant behavior and providing feedback based on students' actions.

Previous studies have demonstrated that VR can be at least as effective (Null Hypothesis) as traditional educational methods [CTK16] [JE21].

Systematic literature reviews from the last 15-20 years indicate that VR can enhance higher education by fostering active learning, information retention, and engagement in the subject matter [Rad+20] [Luo+21].

From a commercial perspective, VR is recognized for its potential to transform and improve educational experiences. According to an Hewlett-Packard (HP) article [Inc21], VR contributes to a 25 percent increase in student engagement and a 20 percent boost in learning retention. Therefore, one of the most significant benefits of VR and AR in education is their ability to convert traditional study methods into motivating and engaging activities without substantial disruption. Nevertheless, both the HP article and [Rad+20] highlight challenges in the VR field, including high equipment costs and the requirement for technical experts near educational facilities. Although these factors are becoming less significant as prices decrease and headset usability improves, they remain obstacles for further research and full-scale integration.

2.2 Virtual Reality in Expos and Museums

VR programs are becoming increasingly employed in museums, expos, and galleries. According to Museum Next [Mus21], visitor engagement with a single exhibit (e.g., painting, game, or film) increases by 25 percent when VR technology is utilized. Moreover, a 90 percent engagement increase is observed among younger audiences, attributed to the gamification and immersive qualities of VR, which means even more significant growth in the next years.

Additionally, VR enhances accessibility, reaching a broader global audience. VR tours and exhibits allow certain individuals to experience these displays online, even from home. Overall, visitor's engagement has grown significantly, as VR provides immersive, hands-on experiences. Visitors of such programs can walk through historical events and interact with virtual versions of artifacts that would otherwise be inaccessible. it's also crucial that it opens a wide range of educational possibilities for people with disabilities and people, who might not afford travelling to museums around the world.

2.3 Virtual Reality for Professionals

For almost a decade, VR and AR have been used by highly skilled professionals on the International Space Station. According to NASA [NAS24], VR and AR technologies assist astronauts by providing real-time instructions, enhancing perception, and supplying data overlays. As stated in the article, AR aids in navigating repair processes. Additionally, VR is used for training by simulating missions and flights. This immersive training, combined with, for example, physics simulations (e.g., overload), provides a unique approach to astronaut preparation.

2.4 Interactions in Virtual Reality

A study by Yemon Zhong [YM24] investigates various user interaction techniques in VR environments, focusing on 3D object manipulation. The research compares high-fidelity (virtual hand) and mid-fidelity controls (controller) and demonstrates that hand manipulation results in faster task completion times and higher user performance.

The study concludes that in VR and AR, controls are more effective when using one-to-one mappings between physical actions and virtual responses, without intermediaries. In such cases, responses (e.g., haptic feedback) are also more intuitive.

High- and mid-fidelity techniques are further examined in another study by Robert McMahan [McM+11], which suggests that high-fidelity techniques, which offer direct and realistic mappings between users' physical actions and the virtual environment (e.g., a virtual hand closely aligned with real hand movements), significantly improve user performance in tasks such as object manipulation and visual search. These studies indicate that with advancing technologies, especially eye and hand tracking, immersion and learning potential could evolve significantly.

Furthermore, neuroscience field investigated Cognitive Load theories for years, trying to evaluate which sense stronger affect the perception of a human. According to [Xu+17] the Visual group may be taking the strongest jump in the recent years due to changes in worklife order(e.g. Home-office work).

3 VR Game Design

3.1 General Design and Structure

This chapter presents the design principles and structure of the environments used in the survey, alongside the learning curves observed throughout the play time. It also provides an overview of the existing interfaces and grabbable objects(the objects that can be picked up and used, referred to as grabbable due to Interface naming), as well as the calculations utilized to determine puzzle hints and gravity effects. Additionally, it covers the User Interface (UI) and data tracking processes used for a data evaluation in Chapter 6.

3.1.1 Idea

The core concept of the game is to assess how players with different hints are able to solve tasks and perceive gravitational effects. The chosen game environment centers around space exploration, selected for its potential to provide an immersive experience for users. Moreover, it offers a substantial database of scientific knowledge, aligning with the principles of serious games that aim to educate while entertaining.

3.1.2 Simplicity through Grabbing

The primary mechanic of the game is grabbing, one of the first and most basic abilities humans develop. It provides the game with a simple mechanism, that is easy to learn in short period of time. All interactions and effects within the game are based on this mechanic. By using grabbing and utilizing ability to rotate objects, users are encouraged to investigate even the smallest parts of objects, while the design allows for various hints, as described later in this chapter 3.2.2. 3.5

3.1.3 Visual Comfort and Reality-Based Assets

To maintain a realistic educational experience and improve user's interest, real assets from NASA and ESA were utilized. As the project focuses on learning, achieving visual accuracy was required.

Another key requirement for objects within the game was to ensure visual quality and stability across different game play sections to minimize cybersickness and maintain consistency in the hints provided. This consistency allows users to focus and immerse themselves fully in the game without experiencing discomfort. For that requirement the stable framerate was guaranteed and lightweight assets were used.

3.2 Controls

3.2.1 Movement

To prevent cybersickness and simplify movement controls for users, teleportation logic was chosen as the primary movement mechanic. According to [Bax+24], teleportation has become the preferred navigation method in many VR experiences due to its ability to allow users to move freely within a virtual space while minimizing the risk of inducing VR sickness. This method facilitates ease of use without causing unwanted effects.

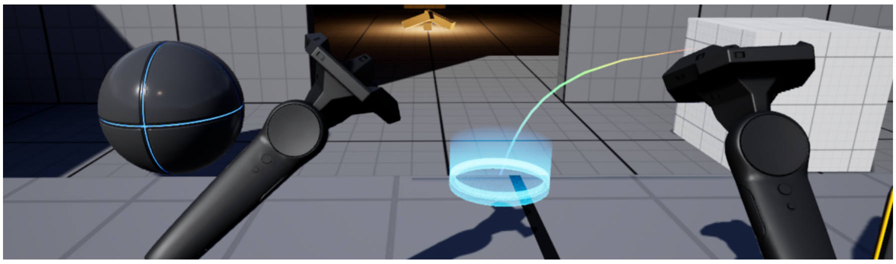


Figure 3.1: Teleportation

3.2.2 Grabbing

In this project, grabbing is implemented via a simple mechanism, called "hand-based" where the user's hand collides with any part of an object." Simple hand-based grabbing techniques, such as the Virtual Hand method, provide an intuitive way for users to interact with objects in VR. As described in [WZL20], the direct mapping of hand movements to virtual hand actions makes these techniques easily understandable, especially for short-distance interactions. This method supports basic manipulation tasks, such as picking up, rotating, and moving objects in the virtual environment, without requiring extensive training.

3.2.3 UI

For UI controls, a simple beam mechanism was used, allowing users to interact with provided buttons by pointing and selecting them via a controller. According to [Kra24], one advantage of using a beam from a finger for 2D component interaction in VR is that it mimics natural pointing gestures, providing an intuitive user experience. This approach offers precise selection and feedback without complex hand movements or controller usage, ensuring easy-to-learn interactions with virtual 2D elements. Example menu usage from Unreal Engine's Documentation [Gam24b] can be found on Figure 3.2

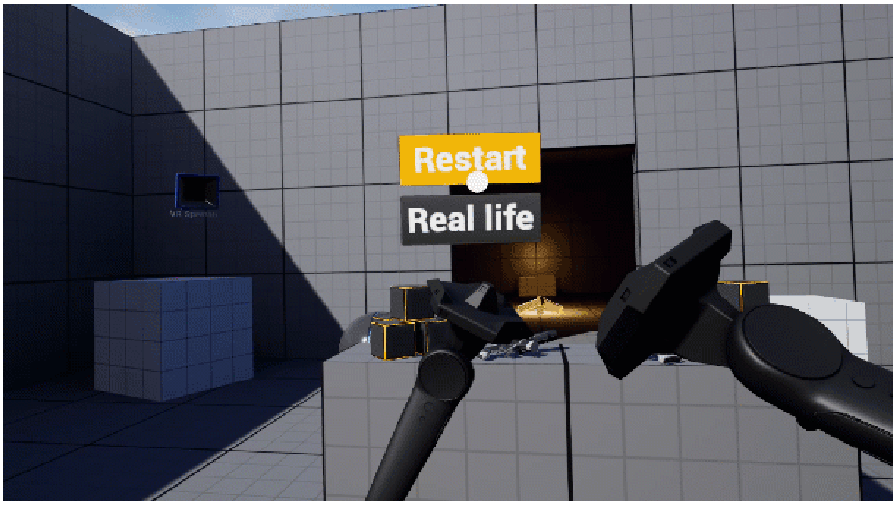
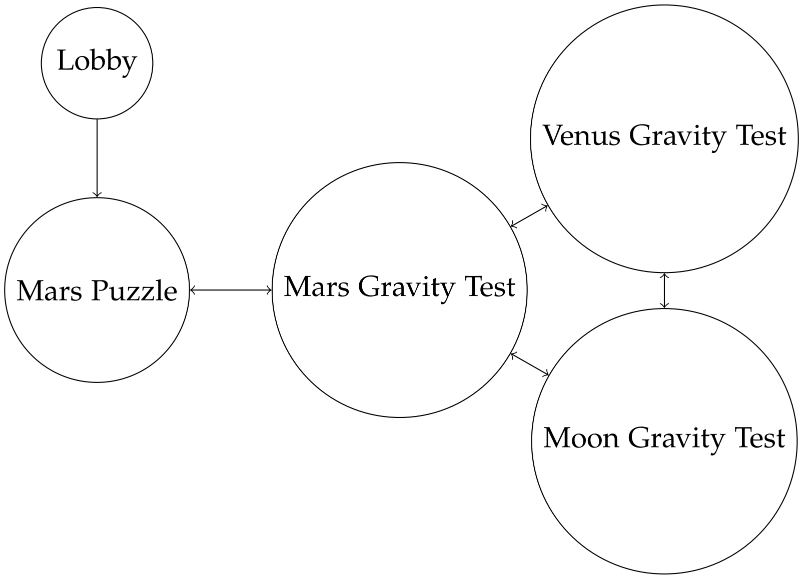


Figure 3.2: UI Interaction

3.3 Environments

This section provides an in-depth description of the game environments and explains the tasks users are required to accomplish. It also outlines the game flows and decision-making processes, which are analyzed later in 6, where decision tracking data is examined. The connection between environments is facilitated through red buttons. Whenever user presses a button with an according text label, the level starts. The timestamps of button presses is also stored in the steps data, which is analyzed further.



The graph above illustrates the interconnection between environments. Notably, to progress from the Mars Puzzle environment to the Mars Gravity Test environment, the puzzle must first be completed. Once completed, unrestricted travel is enabled.

Each environment is examined in more detail below.

3.3.1 Menu

Upon starting the game, users are presented with introductory text describing their challenge and the three possible hint groups. This text serves to explain the puzzle's logic.

Once users press the "Understood" button, they are prompted to select their assigned group, determining the hints they will receive during the game.

The group is managed by a tag attached to the player, storing all relevant information. Once the group is selected, the user is spawned in the lobby.

3.3.2 Lobby

The primary objective of the lobby is to allow users to familiarize themselves with the game's mechanics and controls. Various objects are placed throughout the level for the player to grab, throw, and release. Users can teleport around the environment, interact with different objects to understand their hints, and listen to an audio file providing information about the puzzle and gravity tasks. Additionally, a button initiates the experience, starting data collection and the timer. On average, users spent

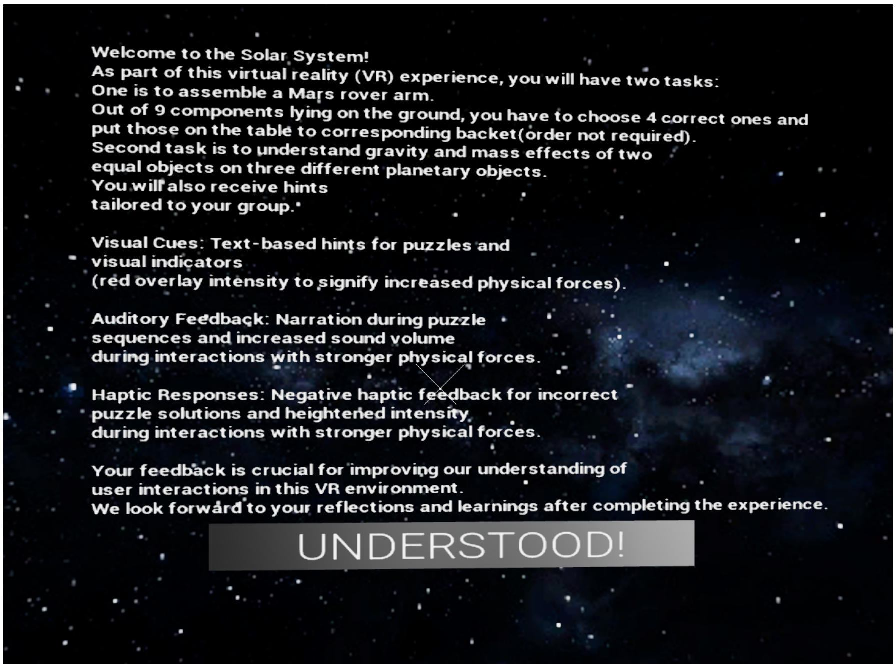


Figure 3.3: UI Description

approximately 8 minutes in this environment, learning the controls and listening to the introductory sound.

Once the button is pressed, each player receives a sound hint explaining the tasks that must be completed. The hint describes the process of collecting 4 puzzle pieces from a total of 9 objects scattered on the ground. This process is required as understanding the puzzle structure was not part of the hypothesis and does not influence the overall performance of participants. Users are given unlimited time to prepare themselves for the challenge, and data collection begins once the start button is pressed.

3.3.3 Assemble Puzzle

This environment was created using NASA textures and background images captured by the Curiosity rover. The ground textures were generated using the Reality Capture engine and images taken by the Perseverance rover. Mars was chosen as the environment due to its status as one of the most extensively studied planets in space exploration. Overall good knowledge, popularity and existing research base are making it a reliable subject for educational purposes. The following puzzle options were considered for this study:

Figure 3.4: UI Menu to Choose Assigned Group

- Assemble the rover parts by grabbing.
- Combine a chemistry puzzle using substances found on Mars (e.g., the Mars Oxygen In-Situ Resource Utilization Experiment (MOXIE) device used to produce oxygen on Mars).
- Complete a step-guided puzzle where users press buttons and observe different functionalities.
- Solve a balance puzzle by adjusting the power sources of the rover.

Compared to other puzzle types, the assembling and combining puzzles allowed for the use of various hints, which was essential for hypothesis testing in this study. The assembly puzzle was ultimately chosen as it offers a more engaging and comprehensible task, even for participants unfamiliar with the technicalities of oxygen production.

To ensure the puzzle could be completed within a 10-minute time frame, only the assembly of the rover's robotic arm, used for environmental observation and guidance, was included.

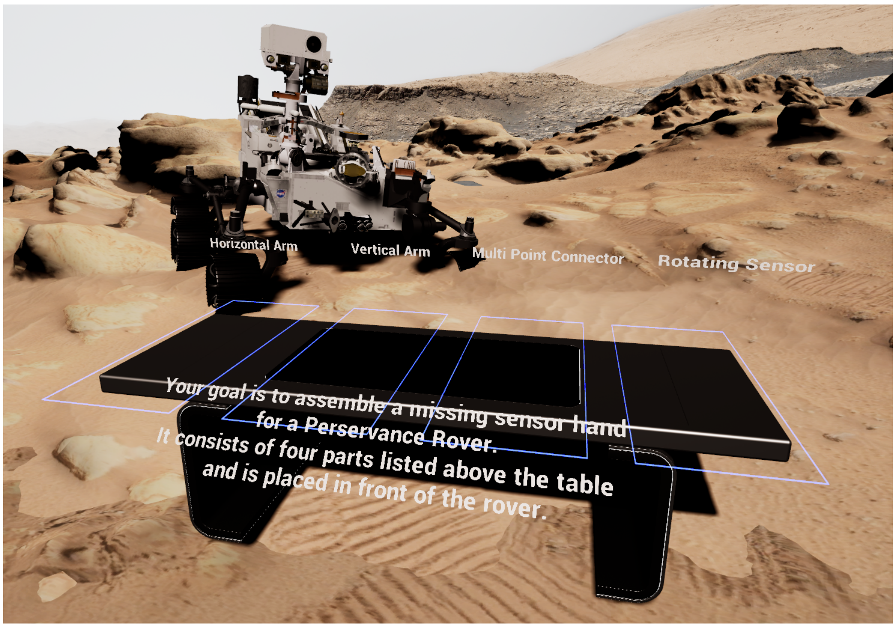


Figure 3.5: Puzzle Buckets and Rover

The primary goal of the assembly puzzle is to find and correctly place 4 parts out of the 9 available into the appropriate buckets on the table (as shown in Figure 3.5 for the Visual Group). Each group has corresponding labels for the buckets: Horizontal Arm, Vertical Arm, Multi-Point Connector, and Rotating Sensor. Each part has a unique function, and incorrect objects with similar functions are included to increase the challenge. For example, incorrect parts include a cable for the Horizontal Arm, a pivot arm for the Vertical Arm, and so on.

When a user attempts to place an object in the wrong bucket, the object is thrown away with an accompanying sound to indicate the error. If the correct part is placed in the appropriate bucket, an assembly animation plays to visually reward the user and focus their attention on the rover.

Upon successful assembly of all required parts, the remaining objects disappear, and a button to start the gravity perception task appears. Users can freely travel between environments once this button is available, though only two participants chose to return to the puzzle later.

The analysis of the steps data gathered during the game, as well as the evaluation of the puzzle-solving process, can be found in Chapter 6.

3.3.4 Gravity Perception Levels

The primary objective of the gravity perception levels is for users to experience the effects of different gravitational forces across various planetary environments. Each level contains two objects, which users can grab and throw or drop. The environments are designed such that the physics and atmosphere remain consistent for all groups. For example, one stone weighs 20 kg, while another weighs 200 kg, and users can observe the differences in their movement by throwing them. Control group participants rely solely on these observations, while other groups receive gravity force-based hints. Below, the design of each environment and its consistency across the three planetary environments is shown.

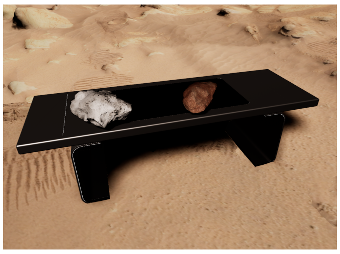


Figure 3.6: Mars

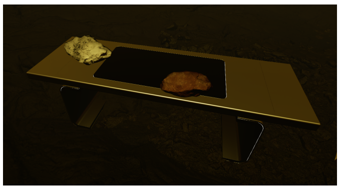


Figure 3.7: Venus

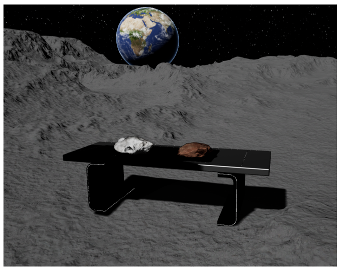


Figure 3.8: Moon

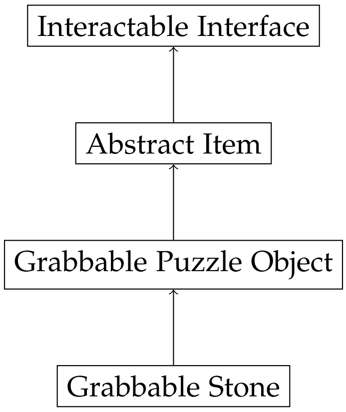
To avoid bias due to mass and size, the objects remain consistent across all environments, despite the fact that the appearance of red stones on the Moon violates realistic design principles. This decision was necessary to ensure the effects of hints were observable and to maintain meaningful survey results. Gravity is measured in gravitational force (e.g., 9.8 m/s^2 on Earth), with each environment influenced by equal forces: gravitational Z force and atmospheric thickness. All of this is simulated using Unreal Engine's integrated physics.

As depicted in Figure 3.3, each environment contains buttons that allow travel between levels. The Mars Gravity level includes an additional button to restart the puzzle, though when the puzzle is completed, the button appears at the beginning of the level rather than at the end, as it would if the puzzle had not been completed.

3.4 Grabbing Objects

To better understand the structure and hierarchy of objects, this section provides a brief explanation of the implementation, though a detailed description can be found in Chapter 4. All grabbable objects in the game adhere to object-oriented programming principles, meaning that each object has a common parent abstract class (which defines shared features) and an interface (which defines shared functionalities). Every grabbable object possesses interaction functionality(by implementing Interactable interface 4, which can trigger different functionality based on the group assigned to the participant.

The following graph depicts the object hierarchy:



While it may seem counterintuitive that stones are classified as children of puzzle objects, this decision was made for future research purposes. Each grabbable stone (used for gravity perception) must have puzzle features, allowing them to be used in both contexts (e.g., for eye-tracking behavior research in cases involving multiple activated human senses). For this particular study, it was essential to maintain a clear separation between tasks.

3.4.1 Interaction Interfaces and Abstract Item

There is only one interface in the game, responsible for triggering the interaction logic. This interface manages level transitions, grab sound effects, visual hints, and haptic effects. It also tracks whether users are attempting to grab a grabbable object.

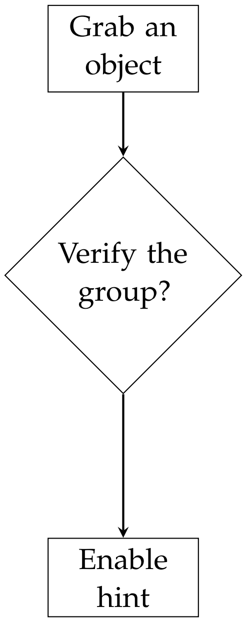
The Abstract Item class implements this interface, with its default functionality designed to start levels and store data. This functionality is precisely what buttons are required to perform, making them direct children of the Abstract Item class.

The functionality of the two groups of grabbable objects is described below.

3.4.2 Puzzle Objects

Puzzle objects contain three essential components: the Niagara component (which handles visual effects), a Text component (which displays text above the object), and a Grab component (which triggers interaction).

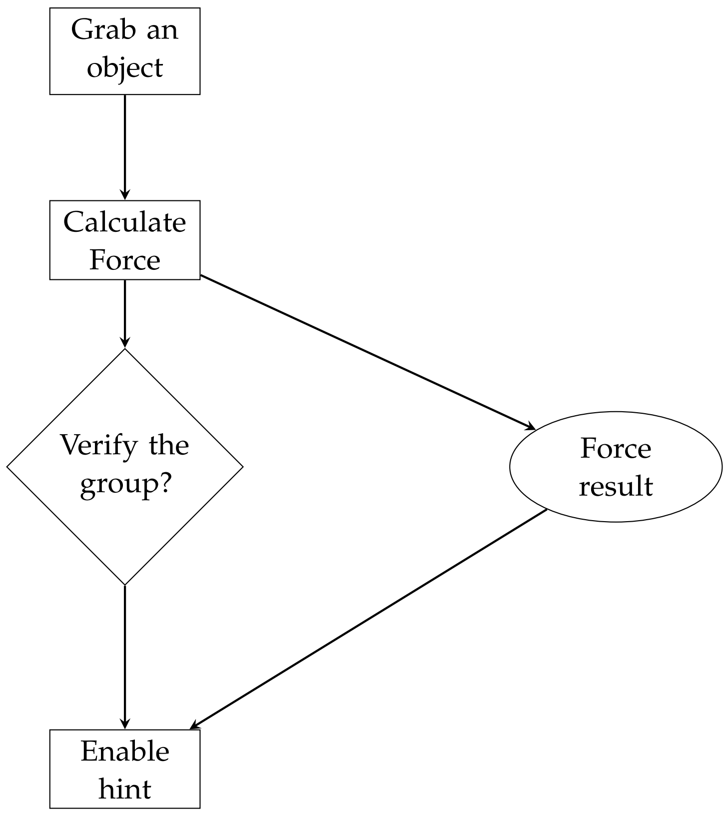
The Niagara component is responsible for the disappearing effect when an object is placed in the correct bucket, while the text renderer displays explanations for participants in the Visual group. Notably, text hints are only visible to participants in the Visual group.



As shown in the diagram above, the functionality is straightforward, requiring only inputs that are predefined for each object: Group and Hint Effect (Text, Sound, or Haptic effect).

3.4.3 Stones

The primary function of grabbable stones is to trigger gravity-based effects based on the current environment, while providing hints according to the player's assigned group. Hints in the gravity perception levels are based on the gravitational force applied to the object. These hints are dynamic, updated every frame, and further calculations are described later in the chapter.



This process resembles that of the puzzle object, with the key distinction being that before the group is verified, the current gravitational force is calculated. This value influences the strength of the hint effects. The force is calculated every frame.

The hint effects are examined in more detail below.

3.5 Hint Effects

The visual and haptic effects are similar, in that the text displayed or spoken is the same for both. The following text cues correspond to each object:

- Vertical Arm (Correct) - This two-pivot arm can carry substantial weight and may be helpful in carrying heavy objects and withstanding strong winds.
- Horizontal Arm (Correct) - This arm is designed to support heavy objects and could be used for tasks requiring stability on a flat surface.
- Cable - Hmm, is this a cable?

- Pivot Arm - This arm looks like a pivot point between other solid objects.
- Camera - This component rotates and has a camera attached.
- Rotating Sensor (Correct) - Interesting... This can rotate and has six cameras. What could it be?
- Multi-Point Connector (Correct) - This connector seems to connect to an arm and possibly rotates.
- Single-Point Drag Connector - This part appears to be used for dragging objects.
- Name Plate - This plate has a logo and some strange numbers of Explorers on it. What could it be?

3.5.1 Visual

Puzzle

Figure 3.9: Visual Hint on a Rotating Sensor

As previously mentioned, the visual hint for the puzzle appears as text above the object. As shown in Figure 3.9, the text is displayed whenever a specific object is picked up, offering users a hint about its function without explicitly revealing its identity.

Moreover, the text is always aligned with the user's camera to ensure ease of reading. Regardless of how the object is rotated, the text remains oriented towards the user's view. This design ensures users can easily read and comprehend the hints provided.

Gravity

For the gravity perception task, three visual hint options were considered:

- Display only the text indicating the object's weight.
- Display both text and visual force effects.
- Display only visual force effects.

As other groups received gravity force-based hints, both the force effect and the text were included in the visual group to maintain consistency across the hinting system. While the force effect is not entirely realistic, it was included to help users interpret the gravity effects accurately. The visual force effect is a red border, created using Niagara effects, with intensity values ranging from 1 to 50 (where 50 represents the most intense red, and 1 represents no effect).

The visual force effect is applied whenever an object hits the ground and lasts for two seconds to allow users to observe the effect.

Figure 3.10 illustrates both the weight text and the red borders of the visual gravity hint.

3.5.2 Auditory

Puzzle

Before beginning the puzzle, users in the auditory group hear an introductory sound describing the puzzle's structure and requirements.

Each puzzle part also triggers a sound cue when picked up. This sound is played through the user's headphones or the VR headset speakers and is designed to provide hints regarding the object's identity.

When a user releases an object and does not pick it up within the next 2 seconds, the sound fades out. The same applies when a user places an object in the correct bucket. If a user grabs an object again within 2 seconds or with another hand, the sound continues playing, ensuring a consistent auditory experience without unnecessary repetition.

If a second object is picked up simultaneously, the sound for the second object is added to a queue and plays after the first sound concludes.

The sound cues align with the hints described in 3.5. The voice of Callum from ElevenLabs was used [Ele24], ensuring that the audio was slow and easily understandable. The Otter tool [Ott24] was employed to reverse-transcribe the sound cues, ensuring a transcription consistency of 95

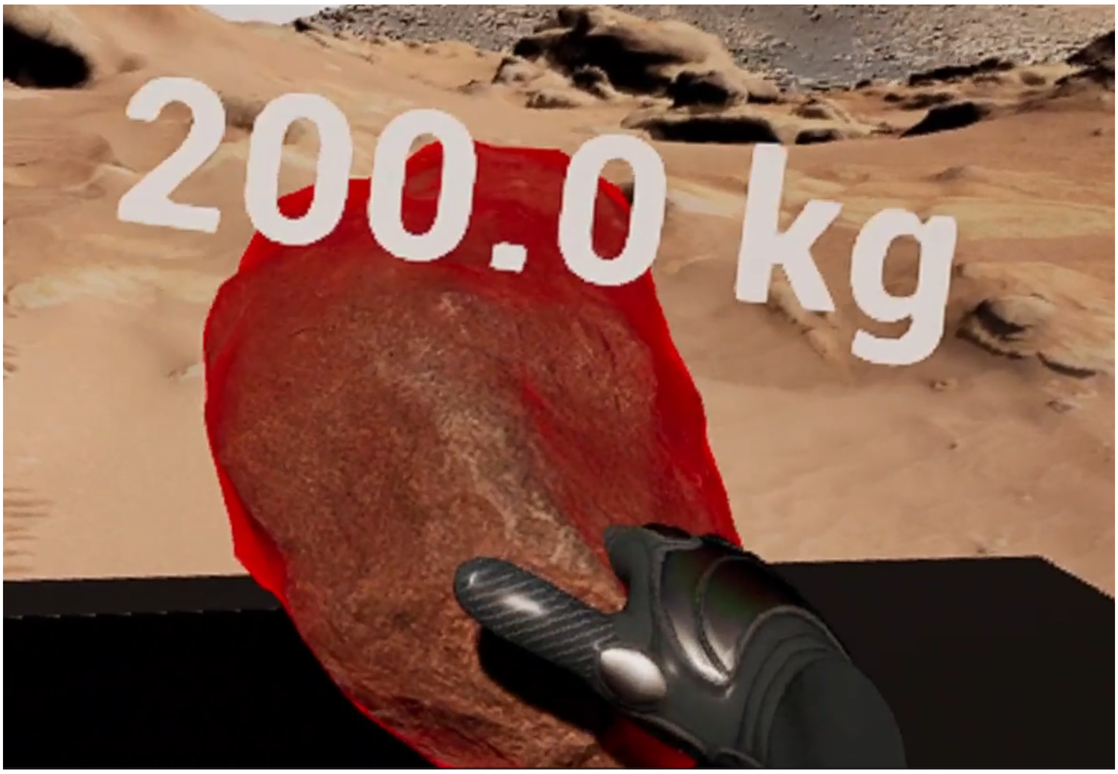


Figure 3.10: Visual Gravity Hint

Additionally, each sound file underwent Root Mean Square (RMS) Volume Consistency analysis, ensuring that RMS values remained between 30 and 50, with a slightly louder volume for the introductory sound to ensure clarity. Figure 3.13 displays the RMS values for each object.

Gravity

For the gravity perception task, the auditory hints function similarly. Each time an object is grabbed and moved, the force value is updated, and a corresponding sound is played. The sound's volume is proportional to the gravitational force. When an object hits the ground, a falling sound is triggered, with the volume based on the distance fallen and the force of the impact. This provides users with an additional auditory cue regarding the strength of gravity in the environment.

It is important to note that the sounds used for puzzle interaction and gravity perception are not tied to the same sound component, ensuring that they do not interfere with one another. This is important if in further research it's decided that

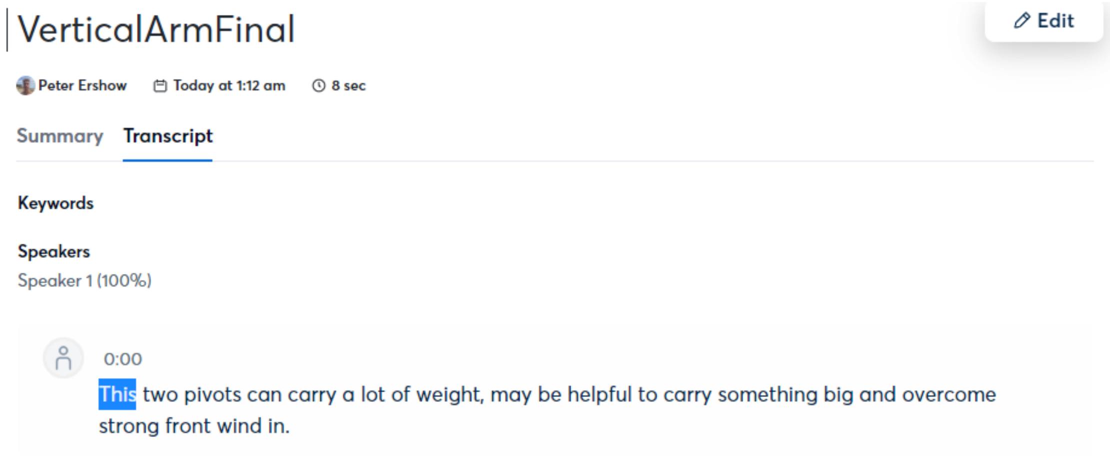


Figure 3.11: Audio Transcript 1

grabbable objects may involve both functionalities simultaneously.

3.5.3 Haptics

The haptic group hints provide a unique and immersive experience.

Puzzle

For the puzzle challenge, the following options were considered:

- Provide assignment assistance by increasing intensity when closer to the correct bucket.
- Provide consistent error feedback for incorrect objects.
- Provide varying error haptic effects depending on the bucket.

The first two options were ultimately rejected, as they provided excessive guidance without encouraging users to engage with the puzzle-solving process. This led to a reliance on haptics, undermining the purpose of the task.

The final implementation provides consistent error feedback when an incorrect object is picked up, allowing users to determine which parts are correct and complete the puzzle independently. This design allowed for meaningful data collection, further detailed in 6.

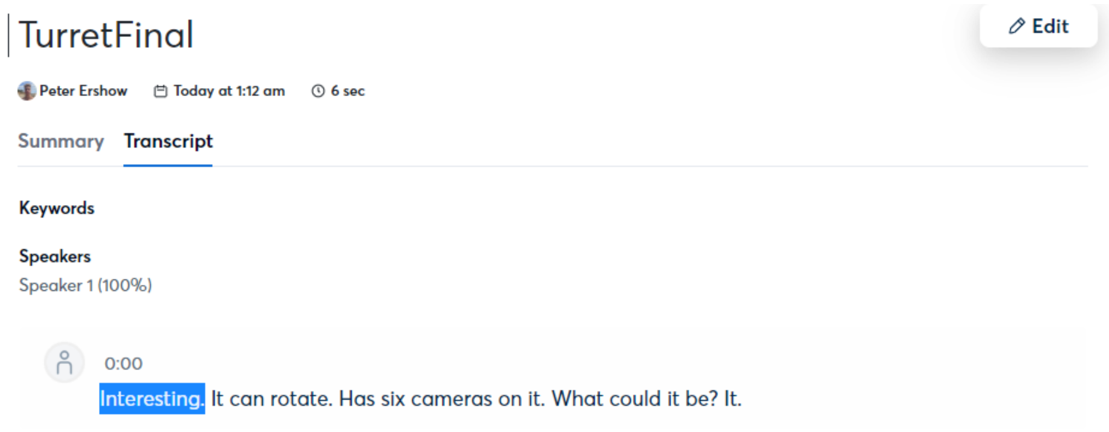


Figure 3.12: Audio Transcript 2

Gravity

For the gravity perception task, the haptic hints offer a highly realistic experience. The intensity of the haptic feedback is determined by the current force applied to the object, calculated as described earlier.

One challenge was the limitations of the Meta Quest controllers, which have a restricted range of haptic feedback. If users shook an object too vigorously, the feedback could become too strong, and in cases where the object's mass was too small, the effect could be too weak. Therefore, the intensity values were clamped within a range to ensure consistency across the experience.

3.6 Gravity Perception and Calculations

This section details the gravity calculations, which apply to all four groups and provide the baseline hint for all users.

3.6.1 Engine Physics

According to the documentation [Epi24], Unreal Engine simulates gravity using Newtonian physics, where gravity is applied as a constant acceleration, typically defined as $9.8\,\mathrm{m/s^2}$. The force due to gravity is given by:

$$F = m \times g$$

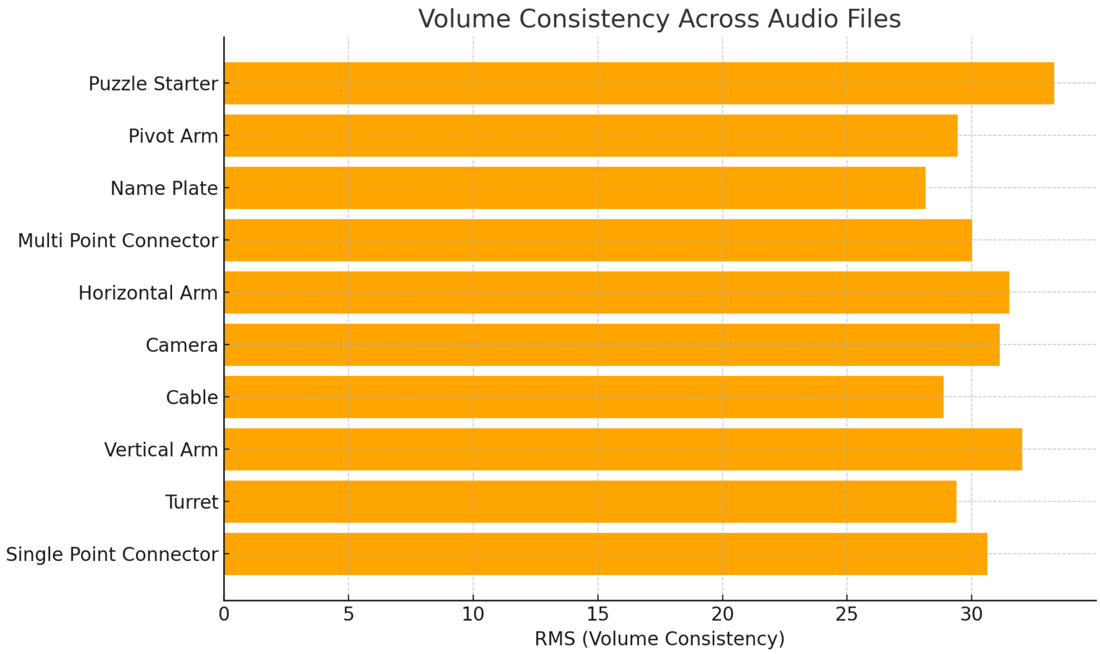


Figure 3.13: Audio Consistency Analysis

where F is the gravitational force, m is the object's mass, and g is the gravitational acceleration. The velocity v of an object after time t is:

$$v = g \times t$$

and its displacement d after time t is:

$$d = \frac{1}{2} g \times t^2$$

This calculation is used as the basis for all environments. For Mars, the acceleration is defined as 3.71 m/s^2, for Venus as 8.87 m/s^2, and for the Moon as 1.62 m/s^2.

As users primarily experience gravity effects when throwing objects, the grab function's interaction with object physics is examined below:

When a user grabs an object, the object's physics and collision detection are disabled, and it is attached to the user's hand, moving with the hand as long as the button is pressed. Upon release, the object's motion is determined by its initial velocity (based on the user's hand movement) and the gravitational forces acting on it.

Initial Velocity Based on Hand Movement The initial velocity v_0 of the object is determined by the movement of the user's hand just before release. The velocity is calculated as the change in hand position over time:

$$v_0 = \frac{\Delta p}{\Delta t}$$

where:

- v_0 is the object's initial velocity.
- Δp is the change in the hand's position between two time points.
- Δt is the time interval between those two points.

This velocity determines the direction and speed of the throw.

Object's Velocity Over Time Once the object is released, gravity acts on it, and the velocity at any time t after the throw is updated as follows:

$$v(t) = v_0 + g \times t$$

where:

- $v(t)$ is the object's velocity at time t.
- v_0 is the initial velocity from the throw.
- g is the gravitational acceleration (which varies by environment).
- t is the time since the object was released.

The graph in Figure 3.14 shows the parabolic paths of objects in different environments.

A limitation of VR controls is that the initial velocity cannot be controlled meaningfully without restricting hand movement speed (as is done in some entertainment games). This can result in excessively powerful throws, distorting the physics calculations. Therefore, before the experiment began, participants were instructed to throw stones with equal, minimal force.

3.6.2 Hint Calculations

The intensity of hint effects is calculated through the following steps:

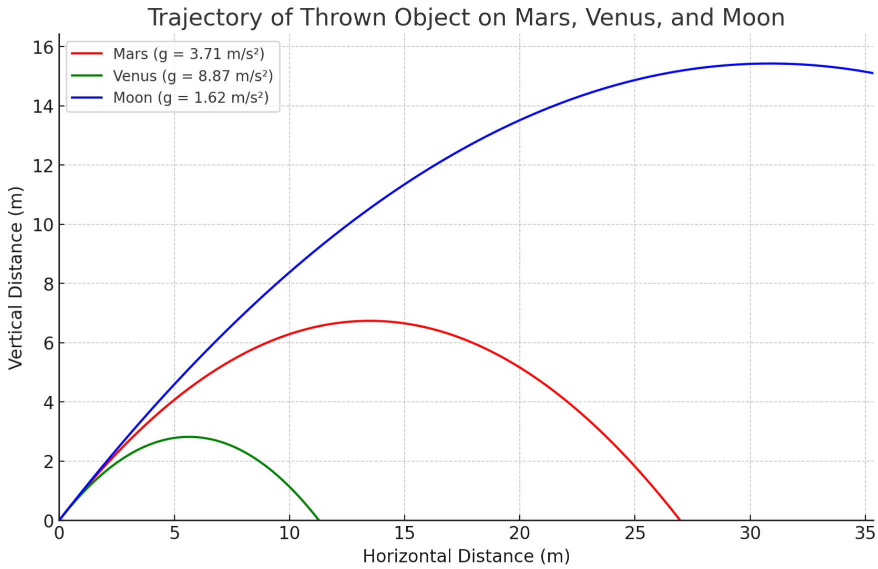


Figure 3.14: Throw Acceleration in Three Environments

Step 1: Multiply Speed by Gravity and Weight First, the force is calculated by multiplying the current grab speed (distance the object traveled over the last 6 frames divided by time) by gravitational acceleration (g) and weight (W):

$$F = \text{GrabSpeed} \times g \times W$$

Step 2: Divide by Constant Value Next, the force value is divided by a constant value of 75,000 to bring it into a range suitable for VR controllers and other hint effects:

$$F_{\text{new}} = F/75000$$

Step 3: Take Absolute Value The absolute value of the force is then calculated:

$$F_{\text{abs}} = |F_{\text{new}}|$$

Step 4: Clamp the Result Finally, the result is clamped between a minimum of 0 and a maximum of 1.0 to ensure it remains within a suitable range for intensity:

$$F_{\text{clamped}} = \text{Clamp}(F_{\text{abs}}, 0, 1.0)$$

In conclusion, the final intensity calculation is as follows:

$$\text{Vibration Intensity} = \text{Clamp}\left((\text{GrabSpeed} \times g \times W/75000), 0, 1.0\right)$$

3.7 Data Collection Methodologies

The data collection functionality is implemented using Comma-Separated Values (CSV) parsing of entries recorded every frame. The following data was collected during the game, named according to filenames stored by system:

- GazeData: Gathered by Meta Quest Pro every frame, including gaze direction, fixation point vector location, confidence value and traced object.
- LookData: Summarizing the frames during which the player looked at specific objects, including an object name, time looked at the object, and a timestamp when the eyes left the object collision box.
- GrabData: Recording the times when specific objects were grabbed, including object name and timestamp.
- StepsData: Logging puzzle attempts, both wrong and correct, with object names, level starts, and puzzle completion timestamps.

For consistency, object name can be used as primary key for LookData and GrabData or as foreign key in Steps data for possible analysis. The primary analysis focuses on the Look, Grab, and Steps data, with GazeData used to verify the consistency and accuracy of other collected information.

3.7.1 Eye Tracking

ObjectName	TimeLookedAt	ElapsedTime
head	0.32	96.83
arm	0.02	96.88
SinglePointConnect	0.04	97.21
arm	0.11	99.23
Arm2	0.32	115.02
SinglePointConnect	0	115.87
arm	0.14	128.41
Turret	0.06	129.57
Connector	0.02	129.68
Turret	0.09	129.84
Names	0.04	130.04
Connector	0	130.21
head	0.32	130.71
Names	0	130.74
Turret	0.21	138.05
Connector	0	138.06
Names	0.03	138.23
Turret	0.03	140.47

The figure to the left provides an example of Look Data for one user. This table includes object name entries with the TimeLookedAt and Elapsed time (total time since the start). By analyzing this data, it's possible to extract typical eye-tracking behaviors. In such cases, calculate the difference between subsequent timestamps, subtract the TimeLookedAt, and consider the remaining time as the period when the user was not looking at any specific object.

3.7.2 Grab Tracking

ObjectName	ElapsedTime
Cable	165.46
Cable	170.66
Turret	221.12
Turret	251.39
Connector	269.74
Connector	297.24
arm	354.97
arm	368.73
SinglePointConnect	395.52
Turret	423.62

The figure to the left shows an example of Grab Data, which logs the grabbed object and the timestamp when it was grabbed. Using this data, it's possible can calculate the probability of a user from a specific group attempting to place an object in the wrong or correct bucket.

3.7.3 Decision Tracking

Step	ElapsedTime
StarterLevelMenu	0.21
L_Mars_VisionPuzzl	88.56
Wronged Connector	243.5
Wronged arm with T	253.4
Wronged Turret with	286.94
Solved Connector	306.11
Solved arm	387.04
Solved Turret	435.04
Wronged Arm2 with	490.59
Solved Arm2	500.96
PUZZLE SOLVED	507.21
L_Mars_Vision	521.55
L_Titan	602.65
L_Mars_Vision	662.98
L_Venus_GravityTes	674.38
L_Mars_Vision	740.05

The figure to the left provides an example of Steps Data for one of the users. In this table, the 'Step' column describes the decision made. Entries beginning with 'L' represent level starts, while 'Wronged' or 'Solved' entries indicate puzzle attempts, including the object name and bucket attempted. In the case of a wrong entry, the correct item for that bucket is also logged. Additionally, the 'Puzzle Solved' entry logs the timestamp when the last item was inserted. Using this data, it's possible to calculate the puzzle-solving time by subtracting the timestamp of the level start (e.g., LMarsVisionPuzzle) from the PuzzleSolved timestamp.

4 Implementation

In this chapter, the principles, stack, challenges, and limitations encountered during implementation are described.

4.1 Introduction

As the main framework for the game, Unreal Engine 5 (UE5) 5.3.2 [Gam24b] was used. The code is implemented in C++ Programming Language (C++) for major upper hierarchy classes, which are inherited by lightweight Blueprint instances. The Blueprint Visual Scripting system in UE5 is a visual programming language that uses a node-based interface to create gameplay elements. This node-based workflow provides designers with a wide range of scripting concepts and tools generally only accessible to programmers. Additionally, Blueprint-specific markup available in UE5's C++ implementation provides programmers with a way to create baseline systems that designers can extend [Gam24a].

4.2 Engine

The UE5 is one of the most powerful and scalable engines in the game development world, used in fields such as gaming, architectural visualization, physics calculations, software simulation, and, most relevant to this paper, educational serious gaming.

Using UE5 templates, the fundamental framework for the game was established. The VR Template, which includes essential VR-specific interaction frameworks such as key bindings and interfaces for UI interactions, movement (teleportation), and object manipulation, was utilized. This template accelerates the development process by automatically configuring necessary plugins for Open Extended Reality (OpenXR) [Gro24] and Meta Quest [Dev24] libraries.

Furthermore, UE5 provides tools specifically for optimizing VR performance, which is crucial due to the high frame rate and responsiveness required for a smooth and reliable gameplay experience.

4.2.1 CI/CD

To communicate changes, improve observability, and guarantee file consistency for project participants, the Perforce Helix Core server was used [Sof24]. Helix Core serves as a centralized version control system (Version Control System (VCS)) allowing the use of large-scale environments hosted locally, without the typical Git-related challenges of handling large data volumes and increased transfer costs. The technology also supports dynamic file locking, restricting access to files currently being modified by another contributor. This feature facilitated a smooth, revert-free implementation cycle.

4.3 C++

C++ code in the project served as a base for more than 15 actors and components in the game. The structural behavior and communication between instances are implemented in C++ following the Observer Pattern principles [Gur24]. The following logic is presented: abstract items (e.g., grabbing interactions such as sound interface, gravity calculations, and puzzle events); puzzle subsystem (activating correct or incorrect effects, triggering assembly logic for the actor, and tracking solution steps data); player controller and pawn logic (playing sounds from three different sound components, including music, grab effects, and narration, haptic, visual, and audio effects on the headset and controllers, grabbing logic, fading in and out of levels, storing eye-tracking and grab data); and game subsystem (storing persistent variables, extracting data to CSV, level connections).

Using C++ helped maintain code structure and improved performance due to the limitations of Blueprints in this regard.

4.4 Hardware

For game testing, the Meta Quest Pro [Pla24] headset was used, providing excellent performance and tracking capabilities, such as eye-tracking. It also offered ergonomic ease for each survey participant.

Using Meta Quest Link, the game was rendered and run on the Personal Computer (PC) hardware to which the headset was connected. This setup increased performance and allowed data to be stored directly on the computer and later transferred to the cloud. Table 4.1 lists the PC's technical specifications.

Component	Specification
Graphics Card	GeForce 3060 Ti
Processor	Ryzen 3600x
Memory	64 GB RAM
VR Headset	Meta Quest Pro
VR Link Cable	Meta Quest Link
VR Runtime	OpenXR runtime

Table 4.1: Hardware Specifications

4.5 Statistical Tools

4.5.1 R

For the purpose of statistical tests, like t-tests, Shapiro-Wilk and ANOVA test in Chapter 6 the RStudio based on R-language version R 4.4.2 was used. Example console input for doing boxplot of distribution of scores by group 6.5

```
boxplot(Score_Numeric ~ SessionType, data=data,

main="Distribution of Scores by Group", xlab="Group", ylab="Score")
```

The following lines were used to run a t-test between the user scores of Control and Auditory groups.

```
control_auditory<-subset(data, SessionType %in% c("Auditory", "Control"))
t.test(Score_Numeric~SessionType, data=control_auditory)
```

4.5.2 Python Panda

For the analysis of big loads of data, especially eye-tracking data, Python pandas library was used. Example for Cognitive load heatmap console command6.6:

```
plt.figure(figsize=(10, 8))
sns.heatmap(mean_time_looked_at_sorted, cmap="YlGnBu", annot=True)
plt.title("Cognitive Load Heatmap (Sorted Objects)")
plt.xlabel("Session Type")
plt.ylabel("Object")
plt.show()
```

4.6 Challenges

This section describes the main challenges encountered during development.

4.6.1 Collision Detection

The main challenge associated with game physics, which could influence game play and potentially bias some gravity effects, was collision detection for components that could be grabbed.

The grabbing mechanic functions through collision detection between the user's hand and an object. When an object overlaps with a hand and the grab button is pressed, the object attaches to the corresponding controller. To ensure functionality, it is necessary to disable further collisions and physics for the object until the hold button is released. As a result, due to how the engine's physics work, the object may overlap with other objects, like buttons or other grabbables, while being held.

Therefore, if the user releases an object in such an overlap, collision re-enables immediately upon release, possibly triggering an overload of force on the object and pushing it away without accounting for weight or gravity.

This could affect users' perception of weight and gravity. Consequently, buttons were configured to have disabled collisions with grabbed objects to prevent unintended level changes when users did not intend them.

In conclusion, it is essential to note that in VR environments with grabbing mechanics, developers should be aware of the freedom users have when physics are enabled on all objects. This issue, as part of users' overall constraints, remains crucial to consider when anticipating the immersive quality of VR games that aim for realistic physics simulation.

4.6.2 Physical Controllers and Anti-Physics

Another challenge was to achieve sufficient immersion through VR controllers. Due to the nature of gravity and physics, the weight, mass, and gravity of an object influence the force applied to the person holding it. However, it is not possible to apply such weight to VR controllers, leading to two potential issues:

- Users do not feel the weight of an object, making the experience less immersive. The closest alternative is haptic effects, which may vary across different manufacturers.
- When an object is thrown by a user, the engine does not account for the counterforce, and therefore cannot simulate it realistically. This made it possible for some users, who threw objects quickly enough, to not notice differences in gravity.

One possible solution for these issues is to manually restrict the speed of objects based on their weight, though this would reduce immersion and require significant development time. Alternatively, as how it was done in this study, users can be informed of these limitations, with control measures applied during each game session(Controlling the behavior of each user).This solution may not be applicable to every survey.

These issues highlight the constraints of VR in real-world physics simulations and warrant further analysis.

5 Survey Setup and Structure

In this chapter, the design and structure of the survey supporting the experiment are discussed, along with a full list of questions.

5.1 Survey Overview

This chapter outlines the design, structure, and implementation of the survey used in the study. It focuses on how participant responses were gathered, the flow of the questionnaire, and the structure of questions used to assess participants' understanding of gravitational effects and their performance across different session types. Additionally, it describes how various response types were processed to extract meaningful data for further analysis. Please not, that the word "performance" in the whole experimental setup and experimental evaluation 6 parts, is used for user's questionnaire scores.

5.1.1 Survey Goals

The main goal of the survey was to measure participants' understanding of puzzle-solving accomplishments and gravity perception in various planetary environments. Additionally, the survey collected demographic information, educational background, and prior knowledge of virtual reality (VR) and augmented reality (AR) technologies. These factors were considered in analyzing performance to minimize potential biases, such as prior knowledge of the topic.

5.1.2 Survey Structure

The survey was divided into several sections to capture various aspects of the participant experience:

Pre-Questionnaire Participants completed a pre-questionnaire before launching the experience to provide data necessary for identifying and eliminating potential biases. Questions for this part can be found in Table 5.1.

- **Demographic Information:** This section collected basic data such as age, gender, field of study, and educational level. These factors were essential for understanding the diversity of participants and how their backgrounds might influence their perception of puzzles and gravity in a virtual environment.
- **VR Knowledge Assessment and Prior Knowledge of the Topic:** Participants were asked to self-assess their prior knowledge of VR technologies. This enabled analysis of the correlation between their familiarity with VR technologies and their performance. In the last question of this part, participants were asked to rank Mars, Venus, and the Moon based on gravitational strength. This question allowed for the exclusion of six participants who already knew the correct ranking, ensuring they were not counted in the gravity performance statistics.

Demographic Information and Previous Knowledge
How would you rate your prior knowledge of Virtual Reality (VR) technologies?
How would you rate your prior knowledge of Augmented Reality (AR) technologies?
How many times have you used a VR headset or other VR technology before?
How would you describe your overall experience with VR?
Have you had any formal training or education in VR or AR? (School/University/-Courses) Please specify if yes.
Have you had any expo/museum VR/AR experience? Please specify if yes.

Table 5.1: Demographic Information and Previous Knowledge

Post-Questionnaire The second part of the survey included a post-questionnaire, divided into two main sections: puzzle-related questions and gravity-related questions. Each section also included user experience questions and an assessment of cognitive load using the NASA Task Load Index (NASA TLX) [NAS86].

- **Experience and Cognitive Load Rating using NASA TLX:** At the beginning of the post-questionnaire, participants were asked to rate their enjoyment and comfort level with the VR experience and provide feedback on the task design. This section provided qualitative data on their subjective experience, including any challenges they encountered during the tasks. In addition, NASA TLX analysis was conducted to assess the mental, physical, and temporal demands of the task, along with perceived success, effort, and frustration levels. Questions for this section are listed in Table 5.2. According to existing research papers, NASA TLX ratings analysis can be helpful in assessing the cognitive load and therefore be a

proof for
1.3.2 Cognitive Load Hypothesis if aligned with user's performance data. [HS88]

- **Task-Related Puzzle Questions:** This section consisted of questions designed to assess participants' knowledge gained during the puzzle-solving experience. The questions addressed both attention to detail (e.g., "What function does an object with the text 'Mars 2020' have?") and knowledge of the topic (e.g., "Which part is required for precise adjustments in the arm assembly?"). For ease of reference, images of all objects were provided. Questions are listed in Table 5.3.
- **Task-Related Gravity Questions:** This section included questions aimed at assessing participants' attention to detail and understanding of the gravity-based tasks they completed. Questions ranged from identifying gravitational forces on different planets to ranking planets based on gravitational strength. This section sought to evaluate participants' level of comprehension and task performance. Questions can be found on Table 5.4.

NASA TLX (Cognitive Load)
How mentally demanding was the task?
How physically demanding was the task?
How hurried or rushed was the pace of the task?
How successful were you in accomplishing what you were asked to do?
How insecure, discouraged, irritated, stressed, and annoyed were you?

Table 5.2: NASA TLX (Cognitive Load)

5.1.3 Participant Sessions

Participants were assigned to different session types, or "assigned groups": Auditory, Haptic, Visual, and Control. The survey structure enabled comparisons of performance and understanding across these groups. Each participant's session type was recorded and analyzed in relation to their survey responses, facilitating study of how various sensory cues influenced decision-making processes and task outcomes. Verification was conducted to ensure each row in the data stored by the game included the session type chosen at the beginning 3.3.1.

5.1.4 Data Collection Methodology

Responses were collected through an online form and stored in Comma-separated values CSV [con23] format for further analysis. The gathered data included both

Puzzle Perception
How much did you enjoy the puzzle-solving experience?
How effective was the information provided in helping you solve the puzzles?
How intuitive was the information provided in guiding you through the puzzles?
Which component is designed to maintain lateral stability and support weight across a flat surface?
Which piece is essential for allowing full directional movement and precise adjustments within the assembly?
Which component acts as rotating support for a sensor?
Which component is specifically designed to bear weight and overcome strong front wind pressure?
Which part is not suitable for bearing significant weight or contributing to the structural integrity of the hand assembly?
Which piece is crucial for terrain observation and data collection but does not contribute to structural support of the sensor hand?
How many parts does the front sensor subsystem consist of?
What function does an object with the text 'Mars 2020' have?
What function does an object with the text 'Perseverance' have?

Table 5.3: Puzzle Perception

objective answers (e.g., scores on gravity-related questions) and subjective ratings (e.g., enjoyment of the experience). These were systematically organized by participant and session type to facilitate cross-group analysis in subsequent sections of this paper.

Further details on the analysis of this data, including statistical tests and comparisons across groups, are provided in the Evaluation Chapter (6).

Gravity Perception
How much did you enjoy the gravity perception experience?
Did you find the gravity effects enhanced your overall experience of the game?
How realistic did the gravity of the stones feel to you?
How well did the provided information help you understand the gravity effects of the stones?
How can you tell the difference in gravity between Mars and Venus when an object hits the ground?
If you throw an object on Mars, which one is expected to fall faster?
The gravitational acceleration on Mars is approximately 3.7 m/s^2. What is the gravitational acceleration on Venus?
The gravitational acceleration on Mars is approximately 3.7 m/s^2. What is the gravitational acceleration on the Moon?
Considering gravity, which planet would be more challenging to launch a spacecraft from, Mars or Venus?
Rank these space objects based on their gravity strength (from low to high).

Table 5.4: Gravity Perception

6 Experimental Evaluation

6.1 Pre-treatment Analysis

6.1.1 Previous Knowledge and Potential Biases

Study Field

This section examines participants' prior knowledge of relevant concepts and the potential biases they may carry into the experience. To assess the distribution of knowledge and demographics, three main aspects were evaluated: field of study, age, and previous VR AR knowledge. These factors provide insight into how different backgrounds could affect the perception of the game and its outcomes.

For the analysis of potential biases related to age and study field, an ANOVA test [Scr24a] was conducted for each bias against the user score in the content survey. The null hypothesis assumed that these biases did not have a significant impact on the results. The analysis was performed with significance thresholds of 5 and 10 percent, meaning that if the p-value is less than 0.05 or 0.1, the null hypothesis can be rejected.

Figure 6.2 displays the distribution of participants' study fields across four major groups: STEM [Dic24], Social Sciences, Humanities, and Health Sciences. The majority of participants, 58.5 percent, come from a STEM background, suggesting familiarity with scientific and technological concepts that could influence engagement with the space exploration theme of the game. This group may also demonstrate a higher baseline knowledge of gravity and physics, potentially reducing their learning curve during game-play.

To assess whether participants' study fields had a statistically significant impact on survey scores, a one-way ANOVA test was conducted, comparing the mean scores across the groups (STEM, Social Sciences, Humanities, Health Sciences). The results were as follows:

- **F-statistic**: 1.96
- **p-value**: 0.137

Since the p-value is greater than the significance level of 0.05, the null hypothesis is not rejected. This indicates no significant evidence that the field of study affects

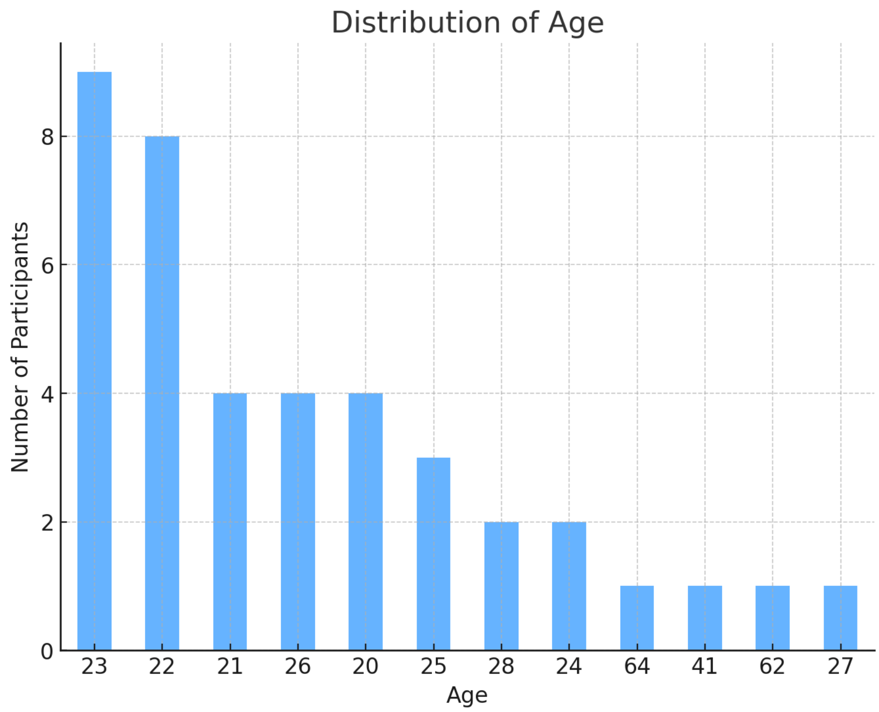


Figure 6.1: Age Distribution of Participants

participants' survey scores, suggesting that study field does not have a statistically significant impact on performance.

Age

Additionally, age was tested as a potential bias. The age distribution for all participants is shown in Figure 6.1.

An ANOVA test was conducted to assess whether participants' age significantly impacted their performance. The analysis compared survey scores across different ages, yielding the following results:

- **F-statistic**: 0.64
- **p-value**: 0.783

Since the p-value is greater than 0.05, there is no statistically significant difference in survey scores across age groups, indicating that age did not introduce any measurable bias in participants' performance on the survey.

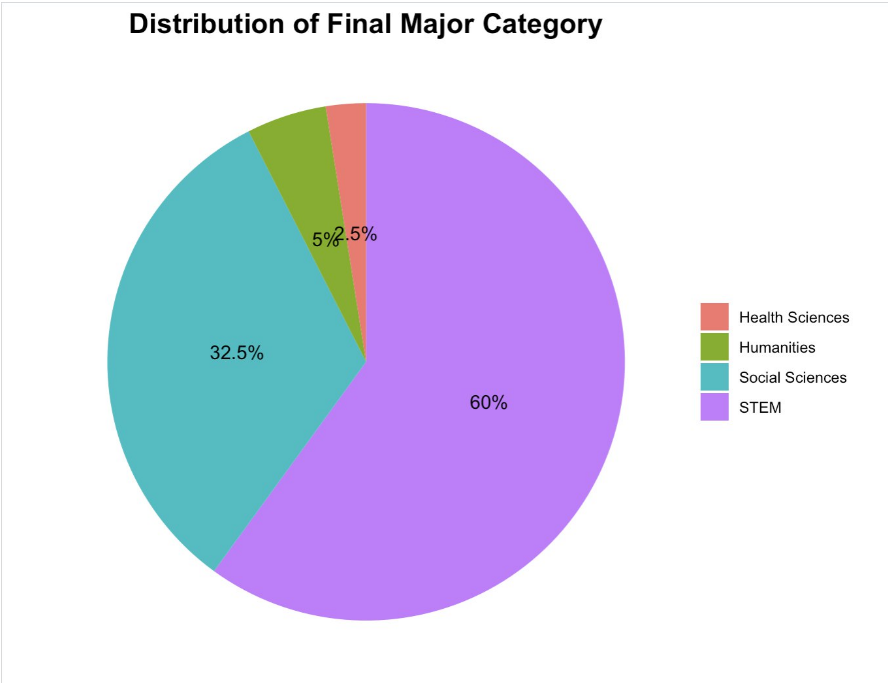


Figure 6.2: Distribution of Study Fields

Previous VR Knowledge

To assess VR knowledge and experience, the survey included questions regarding:

- Prior knowledge of AR technologies
- Prior knowledge of VR technologies
- Frequency of VR headset or VR technology use
- Overall VR experience

Responses were combined to calculate a mean rating from 1 to 4 for each participant's previous VR knowledge.

Table 6.1 shows an example calculation for the first five entries.

For this analysis, as the values in the Final VR Knowledge variable are continuous, a Pearson correlation test [Scr24b] was conducted between the **Final VR Knowledge Rating** and the survey score. The test yielded the following results:

- **Correlation coefficient**: 0.078

AR Knowledge	VR Knowledge	VR Usage	VR Experience	Final VR Knowledge
3	1	1	1	1.875
2	4	2	2	3.125
2	2	2	2	2.500
3	2	3	2	3.125
1	1	2	2	1.875

Table 6.1: VR Knowledge Ratings for Participants

- **p-value**: 0.634

Since the p-value is greater than 0.05, there is no statistically significant correlation between previous VR knowledge and survey scores. This indicates that participants' prior experience with VR did not introduce measurable bias in their performance scores.

6.2 NASA TLX Index

6.2.1 Setup

In the first part of the post-questionnaire, the NASA Task Load Index (NASA-TLX) [NAS86] assessed perceived workload across six dimensions. Below are the questions used to evaluate each dimension:

1. **Mental Demand**: How mentally demanding was the task?
2. **Physical Demand**: How physically demanding was the task?
3. **Temporal Demand**: How hurried or rushed was the pace of the task?
4. **Performance**: How successful were you in accomplishing what you were asked to do?
5. **Effort**: How hard did you have to work to achieve your level of performance?
6. **Frustration**: How insecure, discouraged, irritated, stressed, and annoyed were you during the task?

6.2.2 Interpretation

In Figure 6.3, the weighted NASA TLX ratings are displayed for all groups in the survey. Observing the figure, values approach the center of the bar chart, approximately 2.5. To statistically confirm this, an ANOVA test was conducted across the four groups to examine independence.

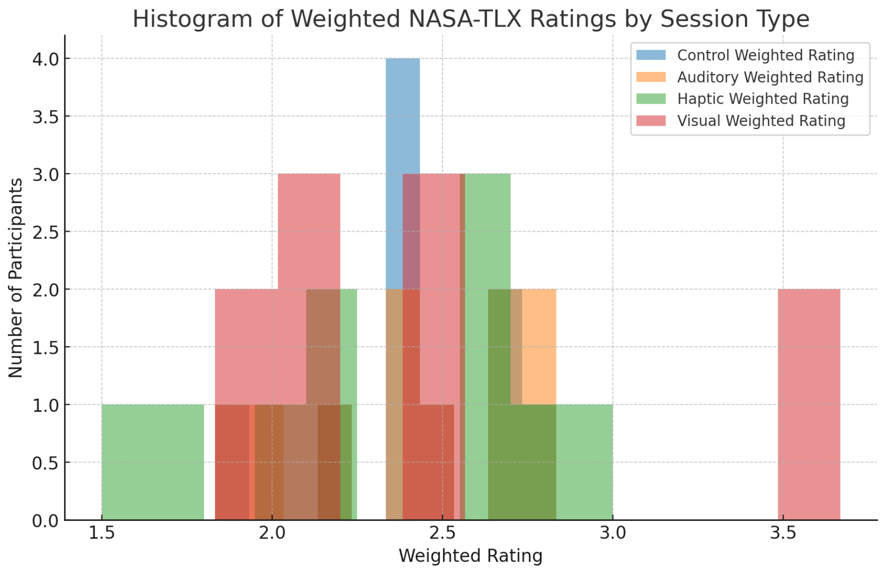


Figure 6.3: NASA TLX Weighted Scores Among Groups

The analysis results are as follows:

- **F-statistic**: 0.225
- **p-value**: 0.879

Since the p-value is much greater than 0.05, there is no statistically significant difference in NASA-TLX scores across different session types. This suggests that perceived workload is similar across the Control, Auditory, Haptic, and Visual sessions. This is also visible in Figure 6.4, where average scores per group appear similar. Slightly higher scores in the Visual group might indicate greater engagement with the task, potentially explained by cognitive load theory [con24a], which posits that visual presentations are often more efficient than auditory or tactile ones, allowing quicker processing and potentially enhancing performance.

Another possible explanation is Sensory Dominance Theory, which suggests that individuals often rely on their dominant sensory modality for information processing

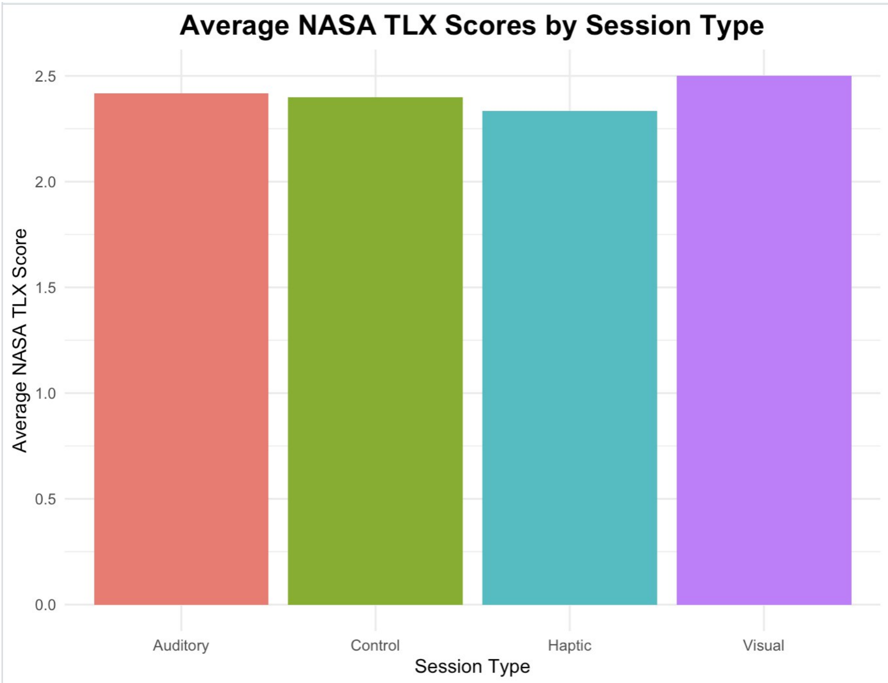


Figure 6.4: Average NASA TLX Score Among Groups

[con24b]. Modern lifestyles, with increased digitalization and remote work, may further emphasize the dominance of visual sensory processing. Different articles already contemplate on this, especially that several studies demonstrated for visual stimuli to be influencing awareness and behavior. [Xu+17] Research suggests that the specific direction of sensory dominance depends on the level of processing, with vision dominating at earlier stages and audition at later stages. It can be seen on the graph 6.4 that Visual and Auditory groups indeed provide higher average TLX Score.

6.3 Main Hypothesis Analysis

In this section, common results across groups are analyzed, and the main hypothesis 1.3.1 of comparison between Control and other groups is tested.

6.3.1 Normality Tests

To ensure that each group follows a normal distribution, normality tests were conducted, as parametric tests (e.g., t-tests or ANOVA) require approximate normality. Shapiro-Wilk [Bui24] tests were performed, where W-statistic is a measure of how well a given sample approximates normal distribution, and a p-value above 0.05 confirms that the data is approximately normal. Results for each group are as follows:

Auditory Group

$$W = 0.90,\ p = 0.203$$

The p-value indicates that the Auditory group data does not significantly deviate from normality.

Haptic Group

$$W = 0.88,\ p = 0.140$$

No significant deviation from normality is also not observed for the Haptic group.

Visual Group

$$W = 0.93,\ p = 0.403$$

The Visual group data is approximately normal.

Control Group

$$W = 0.972,\ p = 0.906$$

The Control group data is also does not deviate from normality.

None of the groups show significant deviation from normality at the 0.05 level, indicating approximate normality for each group. According to that, it's possible to perform parametric tests. 6.3.1

Figure 6.5 presents a box plot visualization of the mean results for each group.

6.3.2 T-Test Analysis

T-tests were conducted to compare Control group performance against the Auditory, Haptic, and Visual groups each to determine if the hints provided significantly aided performance compared to no hints. Similar approach was used in [MJL18] to compare user satisfaction and perceived realism scores between different simulator designs.

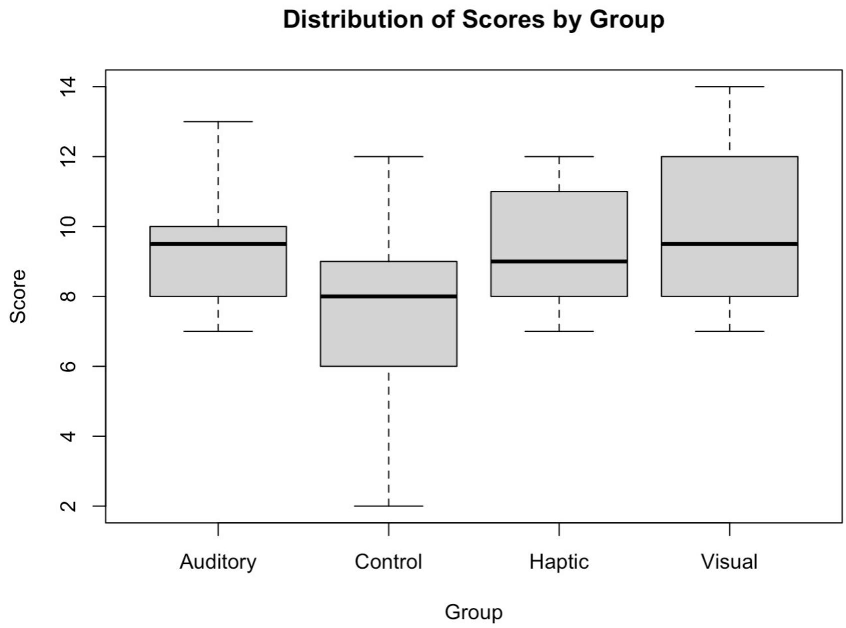


Figure 6.5: Distribution of Scores by Group

Auditory vs. Control

The t-test between the Auditory and Control groups was conducted to assess whether auditory hints improved task performance.

$$\mathrm{df} = 15.197, \mathrm{t}(N) = -1.62, \mathrm{p} = 0.1247$$

With a p-value greater than 0.05, no statistically significant difference is observed between the Control and Auditory groups.

Haptic vs. Control

A t-test was conducted between the Haptic and Control groups to evaluate performance differences.

$$\mathrm{df} = 14.918, \mathrm{t}(-N) = -1.74, \mathrm{p} = 0.103$$

As the p-value exceeds 0.05 or 0.1, no significant performance difference is found.

Visual vs. Control

The t-test between the Visual and Control groups was conducted to determine if visual hints aided performance.

$$\mathrm{df} = 17.8, \mathrm{t}(N) = -1.93, \mathrm{p} = 0.06919$$

Though close to significance at 5 percent, the p-value is greater than 0.05 but less than 0.1. Therefore, the null hypothesis is not rejected at the 5 percent level but is at the 10 percent level, suggesting visual hints improve performance.

6.3.3 Conclusion

The Visual hint group showed statistically significant improvement in performance compared to the Control group, while Haptic and Auditory results approached significance, indicating further investigation may be warranted with a larger sample size. Neuroscience research [CS13] already highlights how visual environments may reinforce the primacy of visual senses and improve the understanding on earlier stages of cognitive processing, which might be crucial for short-term experiences like serious Virtual Reality games. Further investigation on this topic is required.

6.4 Cognitive Load Analysis and Multimedia Learning Theory

1.3.2 In this section, eye-tracking data for each group is analyzed. It's possible to measure cognitive load by reflecting changes in eye behavior [SHV12]. Cognitive load is measured by filtering eye-tracking data for objects viewed longer than 100 ms, reducing noise and accidental glances. The resulting data is visualized in Figure 6.6.

Independent t-tests were conducted to assess differences in cognitive load between the Visual group and other groups. Results are as follows:

- **Visual vs Auditory in Cognitive Load**:
  - t-statistic: 0.84
  - p-value: 0.408

  The p-value indicates no statistically significant difference in cognitive load between the Visual and Auditory groups.

- **Visual vs Haptic in Cognitive Load**:
  - t-statistic: 2.48
  - p-value: 0.025

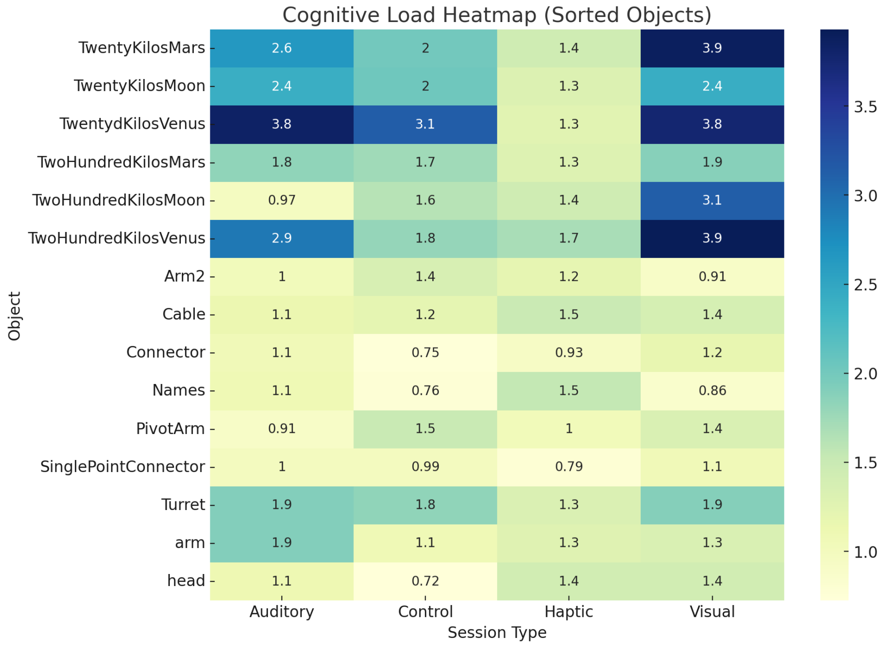


Figure 6.6: Heatmap of Cognitive Load (Mean Viewing Time)

With a p-value below 0.05, a significant difference in cognitive load is observed between the Visual and Haptic groups.

- **Visual vs Control in Cognitive Load**:
  - t-statistic: 1.57
  - p-value: 0.131

  The p-value indicates no statistically significant difference in cognitive load between the Visual and Control groups.

Results suggest the Visual group shows a significant cognitive load difference from the Haptic group, with close significance to the Control group. The Visual group's superior performance supports Visual Sensory Dominance theory and indicates potential for confirming Multimedia Learning Theory. Further analysis is warranted for the Auditory group, with results as follows:

- **Auditory vs Haptic in Cognitive Load**:
  - t-statistic: 1.77
  - p-value: 0.096

The p-value at the 10 percent level suggests a significant difference in cognitive load between the Auditory and Haptic groups.

- **Auditory vs Control in Cognitive Load**:
  - t-statistic: 0.75
  - p-value: 0.462

  The p-value indicates no statistically significant difference in cognitive load between the Auditory and Control groups.

The results suggest a cognitive load difference with the Haptic group, supporting Multimedia Learning Theory. However, further testing is required for comparisons with the Control group.

6.5 Evaluation

Overall, the data analysis confirms the main hypothesis 1.3.1 that hints enhance performance and perception in serious games. In terms of cognitive load 1.3.2 (both NASA TLX Index and eye-tracking analysis), the Visual group demonstrates dominance over others, supporting Visual Sensory Dominance theory. This trend could be influenced by modern lifestyles, where visual information processing plays a critical role in work, entertainment, and everyday activities. As mentioned before in implementation 4.6.2, Multimedia Learning Theory 1.3.4 can not be tested statistically without running survey on the groups with combinations of multiple hints enabled. Nevertheless, from the individual statistic results, it may be observed that auditory and visual groups are at most statistically significant and therefore warrant further research of those combined.

7 Further Work

This chapter discusses potential further work that can be conducted following this research.

7.1 Multimedia Learning Theory Hypothesis

Due to restrictions in this thesis regarding the number of users and an increased focus on examining the effects of research group treatments individually, it would also be valuable to explore whether combining multiple hint groups might lead to more statistically significant results. The following hypothesis could be used for future research:

H0: There is no significant difference in performance between the combined sensory information group and the control groups.

H1: The combination of different sensory information in combined groups leads to improved performance compared to the control groups.

In further research on this topic, attention should be given to The Cambridge Handbook of Multimedia Learning [May12] and other psychological research projects and publications by Richard E. Mayer, especially The Past, Present, and Future of the Cognitive Theory of Multimedia Learning [May24]. It will also be important to consider the potential challenges and biases inherent in Virtual or Augmented Reality environments.

7.2 Puzzle Variations for Different Environments

As outlined in the Design chapter 3, there are various puzzle types suitable for use in Virtual Reality games.

Beyond those used in this research, it is possible to investigate cognitive load and performance across other approaches to puzzle tasks in Serious Games [NS20] [Lee23].

7.3 Hardware Possibilities for Realistic Physics

Due to the challenges described in 4.6.2, achieving realistic physics in combination with grabbable mechanics and weight perception requires compatible hardware. Solutions could involve using devices with strong haptic feedback [Lim23] [Wan22] or employing muscle stimulation and arm and body resistance systems, possibly using exoskeletons and motion capture suits [SDB21] 7.4. Additionally, further investigation could be done on presenting weight through visual and auditory cues, as explored in previous VR projects 3.5.2 [Yen24].

7.4 Further Human Senses to Investigate

It may be worthwhile to examine cognitive load and user perception related to other sensory groups. Human senses may be divided into the visual system (sense of vision), auditory system (sense of hearing), somatosensory system (sense of touch), olfactory system (sense of smell), and gustatory system (sense of taste) [Wik24b].

In this research, only the first three were investigated. Future studies could focus on the olfactory and gustatory systems. Some research has already extended these boundaries in the context of Virtual Reality. For instance, May I Smell Your Attention: Exploration of Smell and Sound for Visuospatial Attention in Virtual Reality discusses the introduction of smell in VR games [San+21]. Similarly, Sensory VR: Smelling, Touching, and Eating Virtual Reality explores the inclusion of smell and taste in unique environments [Har18]. According to these studies, incorporating smell and taste interactions requires environmental integration, such as bringing users into a specific environment or utilizing real food. The effects of temperature, treated as an environmental sensory factor, could follow a similar approach, as shown in Hyper Reality [Wik24a].

Sense	**Research Direction**
Olfactory (Smell)	Incorporate scent in VR to assess its impact on attention and immersion, possibly through environmental integration.
Gustatory (Taste)	Investigate taste simulation in VR, with a focus on environmental integration and real food usage.
Temperature	Assess temperature effects as an environmental factor to enhance realism in VR.
Balance (Equilibrioception)	Use motion capture to evaluate balance-related cognitive load and its impact on learning.
Body Position (Proprioception)	Study proprioception's role in VR interactions and its effects on cognitive load and performance.

Table 7.1: Sensory Groups and Research Directions

Additionally, other sensory groups may offer further opportunities for research and test the cognitive load and performance. Virtual Reality combined with advanced technologies, such as motion capture, could introduce new interactive environments. For example, Balance and Motion Simulation in Virtual Reality: Enhancing Immersive Experiences [LW20] suggests that motion capture could enhance learning accessibility and reduce cognitive load for human senses . Exploring the effects of body position (proprioception) interactions, as discussed in [DC23], may also be beneficial.

Abbreviations

VR Virtual Reality

AR Augmented Reality

UI User Interface

VCS Version Control System

NASA The National Aeronautics and Space Administration

HP Hewlett-Packard

ESA European Space Agency

RMS Root Mean Square

CSV Comma-Separated Values

MOXIE Mars Oxygen In-Situ Resource Utilization Experiment

RMS Root Mean Square

C++ C++ Programming Language

OpenXR Open Extended Reality

PC Personal Computer

UE5 Unreal Engine 5

List of Figures

List of Tables

Bibliography

[Bax+24] G. Baxter, J. Fox, J. Williamson, and A. Miller. “Performance and Navigation Behavior of using Teleportation in VR First-Person Shooter Games.” In: *Proceedings of the 30th ACM Conference on Computer-Human Interaction*. Association for Computing Machinery, 2024. DOI: 10.1145/3661133.

[Boy23] L. Boyd. “Designing Sensory Environments and Sensory Interactions in Virtual Reality.” In: *The Sensory Accommodation Framework for Technology* (2023). DOI: 10.1007/978-3-031-48843-6_4.

[Bui24] Builtin. *Shapiro-Wilk Test: A Complete Guide*. 2024. URL: https://builtin.com/data-science/shapiro-wilk-test.

[con23] W. contributors. *Comma-separated values*. 2023. URL: https://en.wikipedia.org/wiki/Comma-separated_values.

[con24a] W. contributors. *Cognitive load — Wikipedia, The Free Encyclopedia*. 2024. URL: https://en.wikipedia.org/wiki/Cognitive_load.

[con24b] W. contributors. *Colavita visual dominance effect — Wikipedia, The Free Encyclopedia*. 2024. URL: https://en.wikipedia.org/wiki/Colavita_visual_dominance_effect.

[CS13] Y.-C. Chen and C. Spence. “The Crossmodal Facilitation of Visual Object Representations by Sound: Evidence from Visual Recalibration of Sound Localization.” In: *The Journal of Neuroscience* 33.17 (2013), pp. 7109–7116. DOI: 10.1523/JNEUROSCI.5327-12.2013.

[CTK16] D. B. Clark, E. E. Tanner-Smith, and S. S. Killingsworth. “Digital games, design, and learning: a systematic review and meta-analysis.” In: *Review of Educational Research* 86.1 (2016), p. 122. DOI: 10.3102/0034654315582065.

[DC23] J. Doe and L. Chen. “The Cognitive Basis for Virtual Reality Rehabilitation of Upper-Extremity Motor Function after Neurotraumas.” In: *Journal of Virtual Rehabilitation* 29.1 (2023), pp. 45–61. DOI: 10.1007/s12193-023-00406-9.

[Dev24] M. Developers. *Horizon Documentation: Unreal Engine*. 2024. URL: https://developers.meta.com/horizon/documentation/unreal/unreal-engine/.

[Dic24] C. Dictionary. *STEM*. 2024. URL: https://dictionary.cambridge.org/dictionary/english/stem.

[Ele24] Eleven Labs. *Eleven Labs - Text to Speech Platform*. 2024. URL: https://elevenlabs.io/.

[Epi24] Epic Games. *Custom Gravity in Unreal Engine 5.4*. 2024. URL: https://dev.epicgames.com/community/learning/tutorials/w6l7/unreal-engine-custom-gravity-in-ue-5-4.

[Gam24a] E. Games. *Introduction to Blueprints Visual Scripting in Unreal Engine*. 2024. URL: https://dev.epicgames.com/documentation/en-us/unreal-engine/introduction-to-blueprints-visual-scripting-in-unreal-engine.

[Gam24b] E. Games. *Unreal Engine*. 2024. URL: https://www.unrealengine.com/en-US.

[Gro24] K. Group. *OpenXR: Cross-Platform, High-Performance Virtual Reality (VR) and Augmented Reality (AR)*. 2024. URL: https://www.khronos.org/openxr/.

[Gur24] R. Guru. *Observer Design Pattern*. 2024. URL: https://refactoring.guru/design-patterns/observer.

[Har18] P. Harley. "Sensory VR: Smelling, Touching, and Eating Virtual Reality." In: *Journal of Virtual Reality and Broadcasting* 15.2 (2018), pp. 123–136.

[HS88] S. G. Hart and L. E. Staveland. "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research." In: *Human Mental Workload* 52 (1988), pp. 139–183.

[Inc21] H. Inc. *Virtual Reality in Education: How VR is Changing the Future of Learning*. 2021. URL: https://www.hp.com/hk-en/shop/tech-takes/post/virtual-reality-in-education.

[JE21] J. S. M. Julian Abich IV Jason Parker and M. Eudy. "A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda." In: *Virtual Reality* 25 (2021), p. 160. DOI: 10.1007/s10055-021-00548-9.

[Kra24] N. Kramer. "VR Interaction Design: Best Practices and Principles." In: *Daily.dev* (2024).

[Lee23] L.-K. Lee. "A Systematic Review of the Design of Serious Games for Innovative Learning." In: *Electronics* 13.5 (2023), p. 890. DOI: 10.3390/electronics13050890.

[Lim23] W. N. Lim. "Weight Perception Simulation in Virtual Reality with Passive Force Feedback Device." In: *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces*. 2023. DOI: 10.1109/VR.2023.10031797.

[Luo+21] H. Luo, G. Li, Q. Feng, Y. Yang, and M. Zuo. "Virtual Reality in K-12 and Higher Education: A Systematic Review of the Literature from 2000 to 2019." In: *Journal of Computer Assisted Learning* 37.3 (2021), pp. 887–901. DOI: 10.1111/jcal.12538.

[LW20] J. Lee and K. Wong. "Balance and Motion Simulation in Virtual Reality: Enhancing Immersive Experiences." In: *Handbook of Virtual Reality Technology*. Ed. by A. Smith and E. Johnson. Springer, 2020, pp. 157–173. DOI: 10.1007/978-3-030-37336-4_13.

[May12] R. E. Mayer. *The Cambridge Handbook of Multimedia Learning*. Cambridge University Press, 2012. DOI: 10.1017/CBO9780511816819.

[May24] R. E. Mayer. "The Past, Present, and Future of the Cognitive Theory of Multimedia Learning." In: *Educational Psychology Review* 36.8 (2024). DOI: 10.1007/s10648-023-09842-1.

[McM+11] R. P. McMahan, D. A. Bowman, D. J. Zielinski, and R. B. Brady. "Interaction Fidelity: The Uncanny Valley of Virtual Reality Interactions." In: *Proceedings of the 2011 IEEE Virtual Reality Conference*. 2011, pp. 1–8. DOI: 10.1109/VR.2011.5759453.

[MJL18] A. Martinez, B. Johnson, and C. Lee. "Assessing the Effectiveness of Gamification Elements in Educational Software." In: *Journal of Educational Technology Research* 15 (2018), pp. 123–139.

[Mus21] MuseumNext. *How Museums are Using Virtual Reality*. 2021. URL: https://www.museumnext.com/article/how-museums-are-using-virtual-reality/.

[NAS24] NASA. *Nine Ways We Use AR and VR on the International Space Station*. 2024. URL: https://www.nasa.gov/missions/station/nine-ways-we-use-ar-and-vr-on-the-international-space-station/.

[NAS86] NASA. *NASA Task Load Index (NASA-TLX)*. Developed by NASA Ames Research Center Human Performance Research Group. 1986. URL: https://humansystems.arc.nasa.gov/groups/TLX/.

[NS20] T. Nyyssönen and J. Smed. "Exploring Virtual Reality Mechanics in Puzzle Design." In: *International Journal of Serious Games* 7.4 (2020), pp. 85–102. DOI: 10.1007/s40869-020-00120-6.

[Ott24] Otter.ai. *Otter: AI-Powered Speech to Text Transcription*. 2024. URL: https://otter.ai/.

[Pla24] M. Platforms. *Meta Quest Pro*. 2024. URL: https://www.meta.com/de/en/quest/quest-pro/.

[Rad+20] J. Radianti, T. A. Majchrzak, J. Fromm, and I. Wohlgenannt. "A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda." In: *Computers and Education* 147 (2020), p. 103778. DOI: 10.1016/j.compedu.2019.103778.

[San+21] P. M. dos Santos, L. Martens, J. Skopp, and A. K. Engel. "May I Smell Your Attention: Exploration of Smell and Sound for Visuospatial Attention in Virtual Reality." In: *Frontiers in Psychology* 12 (2021), p. 671470.

[Scr24a] Scribbr. *One-way ANOVA: Formula, Example and When to Use It*. 2024. URL: https://www.scribbr.com/statistics/one-way-anova/.

[Scr24b] Scribbr. *Pearson Correlation Coefficient: Formula, Example and Significance*. 2024. URL: https://www.scribbr.com/statistics/pearson-correlation-coefficient/.

[SDB21] A. Smith, J. Doe, and R. Brown. "Triggermuscle: Exploring Weight Perception for Virtual Reality Through Electrical Muscle Stimulation." In: *Frontiers in Virtual Reality* 2 (2021), p. 754511. DOI: 10.3389/frvir.2021.754511/full.

[Sec24] T. Sector. *University of Canterbury explores virtual reality in early childhood teacher training*. 2024. URL: https://thesector.com.au/2024/06/24/university-of-canterbury-explores-virtual-reality-in-early-childhood-teacher-training/.

[SHV12] H. Schmidtke, K. Huesing, and M. Vorlander. "Eye-tracking in complex work environments: New ways of using real-time monitoring to measure cognitive load." In: *Cognition, Technology and Work* 14.4 (2012), pp. 351–365. DOI: 10.1007/s10111-012-0234-7.

[Sof24] P. Software. *How to Use Unreal Engine 5 with Perforce*. 2024. URL: https://www.perforce.com/blog/vcs/how-use-unreal-engine-5-perforce.

[SS20] M. Slater and M. V. Sanchez-Vives. "Presence and Immersion in Virtual Reality." In: *Sptringer* (2020). DOI: 10.1007/978-3-319-08234-9_162-1.

[TWL22] J. Tang, M. Wei, and L. Liu. "The Effects of Sensory Cues on Immersive Experiences for Pro-Environmental Behavioral Intention." In: *Behavioral Sciences* 12.10 (2022), p. 361.

[Wan22] X. Wang. "VibroWeight: Simulating Weight and Center of Gravity Changes of Objects in Virtual Reality for Enhanced Realism." In: *arXiv preprint arXiv:2201.07078* (2022).

[Wik24a] Wikipedia contributors. *Hyper Reality Experience*. Accessed: 2024-11-03. 2024. URL: `https://en.wikipedia.org/wiki/Hyper_Reality_Experience`.

[Wik24b] Wikipedia contributors. *Sense*. Accessed: 2024-11-03. 2024. URL: `https://en.wikipedia.org/wiki/Sense#:~:text=Humans%20have%20various%20sensory%20organs,system%20(sense%20of%20taste).`.

[WZL20] M. Weise, R. Zender, and U. Lucke. "How Can I Grab That?: Solving Issues of Interaction in VR by Choosing Suitable Selection and Manipulation Techniques." In: *i-com: Journal of Interactive Media* 19.2 (2020), pp. 67–85. DOI: `10.1515/icom-2020-0011`.

[Xu+17] M. Xu, Y. Wang, W. Luo, and J. Qiu. "Visual Dominance in Information Processing." In: *Frontiers in Psychology* 8 (2017), p. 303. DOI: `10.3389/fpsyg.2017.00303`.

[Yen24] W. N. L. Y. L. K. M. Y. C. C. Yen. "Enhancing Weight Perception in Virtual Reality: An Analysis of Kinematic Features." In: *Virtual Reality* (2024). DOI: `10.1007/s10055-024-00948-7`.

[YM24] A. M. C. Yemon Lee and S. Marks. "Mixed interaction: evaluating user interactions for object manipulations in virtual space." In: *Journal on Multimodal User Interfaces* (2024). DOI: `10.1007/s12193-024-00431-2`.