Dissertation presented to the Instituto Tecnológico de Aeronáutica, in partial fulfillment of the requirements for the degree of Master of Science in the Graduate Program of Aeronautical and Mechanical Engineering, Field of Automation.

Ivan de Souza Rehder

VIRTUAL REALITY FOR THE HUMAN-CENTRED DESIGN OF ASSISTIVE DEVICES

Dissertation approved in its final version by signatories below:

Prof^{a}. Dr^{a}. Emilia Villani
Advisor

Prof. Dr. Edmar Thomaz da Silva
Co-advisor

Prof^{a}. Dr^{a}. Emilia Villani
Dean of Graduate Studies

Campo Montenegro
São José dos Campos, SP - Brazil
2022

Cataloging-in Publication Data
Documentation and Information Division

Rehder, Ivan de Souza
Virtual reality for the human-centred design of assistive devices / Ivan de Souza Rehder.
São José dos Campos, 2022.
126f.

Dissertation of Master of Science – Course of Aeronautical and Mechanical Engineering. Area of Automation – Instituto Tecnológico de Aeronáutica, 2022. Advisor: Profª. Drª. Emilia Villani. Co-advisor: Prof. Dr. Edmar Thomaz da Silva.

1. Virtual Reality. 2. Human Factors. 3. Collaborative Design. I. Instituto Tecnológico de Aeronáutica. II. Title.

BIBLIOGRAPHIC REFERENCE

REHDER, Ivan de Souza. **Virtual reality for the human-centred design of assistive devices**. 2022. 126f. Dissertation of Master of Science – Instituto Tecnológico de Aeronáutica, São José dos Campos.

CESSION OF RIGHTS

AUTHOR'S NAME: Ivan de Souza Rehder
PUBLICATION TITLE: Virtual reality for the human-centred design of assistive devices.
PUBLICATION KIND/YEAR: Dissertation / 2022

Ivan de Souza Rehder
Av. Francisco José Longo, 633
12.245-906 – São José dos Campos–SP

VIRTUAL REALITY FOR THE HUMAN-CENTRED DESIGN OF ASSISTIVE DEVICES

Ivan de Souza Rehder

Thesis Committee Composition:

Prof. Dr.	Luís Gonzaga Trabasso	Chairperson	-	ITA
Prof^{a}. Dr^{a}.	Emilia Villani	Advisor	-	ITA
Prof. Dr.	Edmar Thomaz da Silva	Co-advisor	-	ITA
Prof. Dr.	Christopher Shneider Cerqueira	Internal member	-	ITA
Prof. Dr.	Petter Krus	External member	-	Linköping University

ITA

“Not being the best at something, doesn’t mean that you are not capable of doing it.”

Acknowledgments

I would like to thank my advisor Profª Emilia Vilani and co-advisor Dr. Edmar Thomaz da Silva for their guidance and trust throughout this project.

To thank my parents, for all of the support that was given to me all of these years and that I will always be thankful for having.

To thank Sidney, that guided me in this universe of blind and visually impaired users. Without him, the devices and the results would be very different, and probably less efficient, than the final ones.

To thank Gabriel "Boi", Santiago, André, Wellington and every one of the Competence Center in Manufacturing (and all my other friends that I could not mention here) that shared experiences and knowledge of the daily life of a master's student and also helped me on some of the solutions used in the work.

To thank all of the participants that gave a little bit of their time and patience to provide me with data for this work.

"If we believe most people can't be trusted, that's how we'll treat each other, to everyone's detriment. Few ideas have as much power to shape the world as our view of other people."

— Rutger Bregman

Abstract

Society has developed technology to create autonomous vehicles and to connect different devices and machinery to exchange data and optimize production efficiency. With this technology, soon, it will be possible to achieve better methods to guide blind and visually impaired (BVI) users in their daily activities. The available products in the market have several limitations and do not satisfy BVI users. We believe that one of the reasons behind this problem is that they are not members of the development team or are not consulted by these.

The lack of an efficient solution for BVI users' navigation became even more significant with the SARS-CoV2 pandemic, in which people had to avoid contact with one another and not touch another surface.

The purpose of this paper is to use virtual reality (VR) to test and evaluate different designs of BVI products. Also to verify if BVI and non-BVI users have the same mental demand and situation awareness when using assistive products. The idea is to use VR as a testing ground where a BVI user can try different assistive solutions in different scenarios. By doing so, the user becomes part of the product design and evaluation, resulting in better and more user-friendly products. The proposed method includes not only the setup of the virtual environment but also the use of physiological sensors and subjective tests to assess the mental workload and situational awareness in different situations.

To illustrate the proposed method, a case study is proposed, in which the navigation of BVI users inside a medical clinic is studied. This case study is chosen due to the current undergoing SARS-CoV-2 pandemic and the impact on BVI people, so the simulated clinic is also applying COVID health protocols.

The scenes were made using Unity3D, a widely used development platform for virtual reality applications. The VR device was the Tobii Eye Tracking VR, a head-mounted display for virtual reality developed using the HTC VIVE. This VR device is used for defining the user position and orientation inside the virtual environment. Based on the current situation in the virtual environment, inputs are provided to the user using aural commands and haptics devices. To assess the mental workload, physiological sensors, from TEA Captiv T-Sens, are used. Among them, are an electrocardiogram sensor (ECG), to

gather heart-rate and heart-rate variance data, and a galvanic skin reaction sensor (GSR), to collect skin conductance. Besides these sensors, the users are also expected to answer mental workload assessment tests and situation awareness questionnaires.

Among the proposed method's expected benefits are the flexibility and agility to create different scenarios, and also the possibility to test all of them in the same physical room. The method could not only speed the design of new solutions but also improve the overall quality of the products and verify the need of a BVI user in the development team of an assistive product.

List of Figures

List of Tables

List of Abbreviations and Acronyms

AR	Augmented Reality
AV	Augmented Virtuality
BVI	Blind and Visually Impaired
CCM	Competence Center of Manufacturing
DLR	German Aerospace Center (*Deutsch Zentrum für Luft- Raumfahrt*)
ECG	Electrocardiogram
EF	Effort
EDA	Electrodermal Activity
FR	Frustation
GSR	Galvanic Skin Reaction
HR	Heart Rate
HRV	Heart Rate Variance
IEA	International Ergonomics Association
MR	Mixed Reality
MD	Mental Demand
MWL	Mental Workload
NASA-TLX	NASA Task Load Index
PCB	Printed circuit board
PD	Physical Demand
PE	Performance
RE	Real Environment
SA	Situation Awareness
SAGAT	Situation Awareness Global Assessment Technique
SWAT	Subjective Workload Assessment Technique
TD	Temporal Demand
UN	United Nations
VE	Virtual Environment
VR	Virtual Reality
WHO	World Health Organization

XR	Extended Reality

Contents

1 Introduction

According to the World Health Organisation (WHO), there are at least 2.2 billion people with some visual impairment degree (WORLD HEALTH ORGANIZATION AND OTHERS, 2019). Among them, 43,3 million are classified as blind and 295 million have moderate or severe vision impairment. In order to be fully integrated into our society, they rely on assistive devices, such as canes, braille speakers, among others (BOURNE *et al.*, 2021).

Although a range of products has already been proposed, incorporating different features, they do not entirely fulfill their aim. Among the problems, of the solutions available in the market, are the lack of practicality and portability, invasive and requiring too much effort to learn (LOZANO *et al.*, 2009).

The difficulty of using or learn how to use a device could be avoided if concepts from human factors, or ergonomics, were analysed during the product's development, using appropriate methods. The early application of these methods and tests could be a gamechanger for the success of the product's user experience (WOLF *et al.*, 2019).

Motivated by the dissatisfaction of blind people with the currently available products, this dissertation starts from the hypothesis that a human-factors-centred design of assistive devices for blind and visually impaired people (BVIs) requires the involvement of BVIs in the design process in order to evaluate the product under design. The user has to test the product under development to provide feedback for the design team to improve the product.

In order to approach this problem, this work proposes using virtual reality (VR) as a tool for creating virtual environments, where proof of concepts or prototypes of assistive devices could be easily tested by BVIs. VR can be used to create specific, immersive and interactive situations that could help the user to learn and train (FARRELL, 2018), and the the developers to create more user-friendly products.

In a virtual environment, as long as the BVI is wearing a locating system, he/she can navigate the environment. Any information about the scenario, such as the position of objects and their distances to the user, is known and could be extracted from the virtual platform. As a consequence the designer can test different ways of translating this information into inputs before actually implementing a prototype of the assistive device, providing a flexible, safe and easy way to have it evaluated by different users.

As a second motivation, this dissertation considers the COVID-19 pandemic scenario that dominated the world during the last two years. In order to try to slow the rate at which the virus spread, WHO recommended strategies such as wearing face masks, washing hands regularly, social distancing, and avoiding touching surfaces that have not been disinfected (WORLD HEALTH ORGANIZATION, 2020). The recommendations bring additional difficulties for BVI people as the touch is one of the senses they rely on to compensate for the visual impairment. The BVI depends on others to do daily activities (JONDANI, 2021). The development of solutions that can guide a BVI in an environment respecting social distance and other recommendations is also considered in this work.

1.1 Objectives

This dissertation proposes the use of virtual reality as a tool for evaluating proofs of concept of assistive devices for blind and visually impaired people from a human-factors perspective. The purpose is to provide a flexible and easily configured way of testing different concepts of assistive devices in order to support an agile and user-centered development.

This goal is related to the following research questions, which are investigated in this work:

- Is it possible to evaluate and compare concepts of assistive devices from a human factors perspective in a virtual environment? What are the main limitations of the use of a virtual reality environment?
- Do non-BVI users, when deprived of their vision, similarly evaluate assistive devices as BVI users?

To investigate these research questions and achieve the proposed goals, the following specific objectives are defined:

- Select a scenario for testing assistive devices and develop it in a virtual environment;
- Develop three concepts of assistive devices that use different senses to provide input to the BVI;

- Propose a set of methods for BVI to evaluate assistive devices from human factors perspective;
- Design and execute an experiment to evaluate the concepts of assistive devices in the virtual environment using the proposed methods.

1.2 Resources and methods

This work adopts an experimental approach to evaluate the proposal of this dissertation and to investigate the questions stated in Section 1.1. The work is organized in the following steps, illustrated in Figure 1.1:

Step 1 – Literature review

It is composed of two parts. The first is to review the fundamental concepts related to the topics covered in this work: human factors and virtual reality. The second part aims at contextualizing the dissertation's proposal. It reviews recently published works on the development and evaluation of assistive devices for BVI people.

Step 2 – Specification of examples of the virtual environment and assistive devices

This step consists of specifying one example of a virtual environment and a few examples of assistive devices to test the proposed approach of using virtual reality for evaluating purposes. Considering the above-mentioned motivation related to the covid-19 pandemic, the chosen virtual environment is the reception of a health clinic. The assistive devices used as examples are: an audio system, a haptic belt and a virtual cane, which could be used as stand-alone devices or combined.

Step 3 – Development of the specified virtual environment

The virtual environment of a health clinic reception is developed in the Unit3D environment. The HTC VIVE VR Head Mounted Device (HMD) is used as a localizing system to define the user's position inside the virtual environment.

Step 4 – Development of proofs of concept of the specified assistive devices

The three examples of devices are developed using low-cost and available laboratory equipment. The audio guide is developed using the audio system of HTV VIVE HMD, while the virtual cane is developed using the HTC VIVE VR hand controller. Finally, the haptic belt is developed using an ESP32 microcontroller, eight vibrating motors 1027 and 3D printed pieces.

Step 5 – Design and execution of the experiment

The proposed experiment is based on the best practices and principles of the Design of Experiment (DoE) discipline. The following techniques and tools are used for evaluating human factors:

a) Questionnaires adapted from the literature, such as NASA-TLX and SAGAT, or explicitly proposed for this work;

b) Physiological sensors, such as GSR and ECG, to capture the body's response.

Step 6 – Analysis of results

The results of the experiment are graphically and statistically analysed to estimate the user's mental workload and situation awareness. In the statistical analysis, the outputs are verified for normality distribution, then pairwise compared using the Student's T-Test, and their variance inside the group is verified using an ANOVA. Finally, when needed, a Fisher's Least Squared Difference Test is done to verify similarities between pairs.

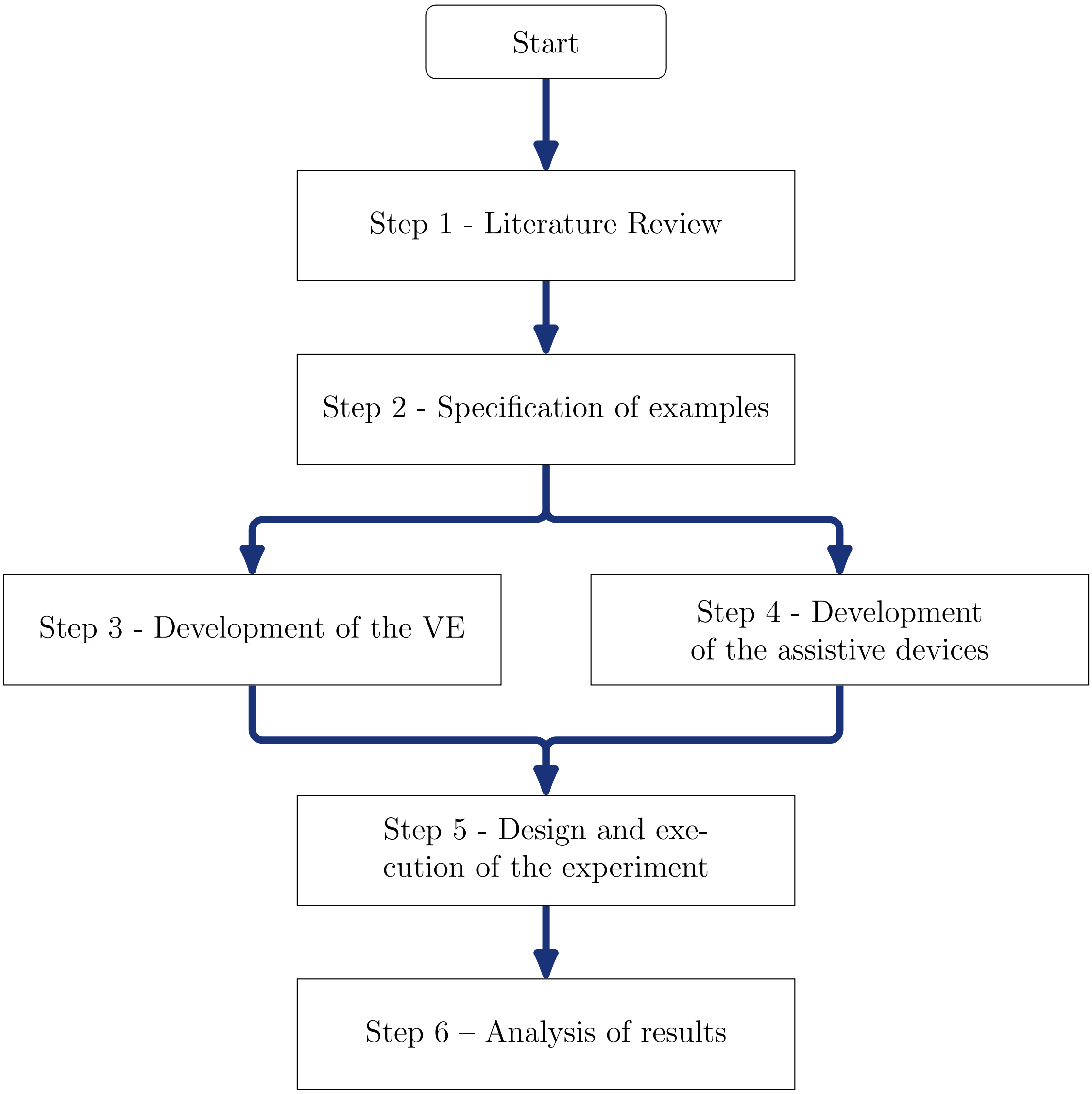


FIGURE 1.1 – Steps of this work

1.3 Research boundaries

The concepts of assistive devices presented as part of this work are used only as examples for investigating the research questions presented in Section 1.1. The challenges related to their full development up to high Technology Readiness Levels (TRLs), as well as their feasibility as commercial products, are out of the scope of this work.

1.4 Structure of the text

This dissertation is organized into seven additional chapters as follows.

Chapter 2 introduces the concepts and techniques that are used in this work. It starts with a review of human factors, emphasizing mental workload and situational awareness, and introduces some human factors' evaluation tools and techniques. Then, it presents the definitions of virtual reality (VR) and Extended Reality (XR) and, to conclude, discusses the concept of co-design.

Chapter 3 is dedicated to the state of the art. It brings a review of the literature and discusses published research works that are related to this dissertation. It covers the proposal and evaluation of BVI devices with emphasis on human factors analysis or virtual reality.

Chapter 4 details the proposal of this dissertation describing how virtual reality could be used to integrate BVI users into the design process of assistive design. It illustrates the proposed method by applying it to evaluate three different assistive devices (audio guide, virtual cane and haptic belt), as well as their mixed-use, in the environment of a hospital reception.

Chapter 5 describes the experiment designed to evaluate the dissertation's proposal and analyses the results in order to investigate the research questions of Section 1.1

Finally, Chapter 6 summarizes the main conclusions of this work and discusses future work.

2 Fundamentals

The proposal of this work combines the concepts of co-design and virtual reality with techniques from of Human Factors.

In order to facilitate the understanding of this dissertation, this chapter introduces these concepts and techniques. It starts with the definition of Human Factors, also known as Ergonomics. It describes mental workload and situation awareness, as well as the corresponding assessment methods used in this work. It then presents the concept of extended reality and defines virtual reality. Finally, it introduces the principles of co-design.

2.1 Human Factor or Ergonomics

Studies in the area of Human Factors started during the Second World War, motivated by performance shortfalls and failures related to the operation of equipments used by humans. The studies showed that these problems could diminish when, other than engineering, psychology and physiology were also considered when designing systems that would be handled by human beings (SANDOM; HARVEY, 2004).

This study area was named "Human Factors" in the United States and "Ergonomics" in Europe. Despite this difference in the names, today they are considered the same field of study. The International Ergonomics Association (IEA) defines Human Factors, therefore Ergonomics, as the following:

> *"Ergonomics (or human factors) is the scientific discipline concerned with the understanding of interactions among humans and other elements of a system, and the profession that applies theory, principles, data and methods to design in order to optimize human well-being and overall system performance. Human Factors professionals contribute to the design and evaluation of tasks, jobs, products, environments and systems in order to make them compatible with the needs, abilities and limitations of people"* (KARWOWSKI, 2012).

This definition shows that humans and their interaction with systems and devices should be considered during the design process (SANDOM; HARVEY, 2004; SANDERS; MCCORMICK, 1998; DUL; WEERDMEESTER, 2003). This need resulted in the proposal of an ISO Standard: BS EN ISO 13407 ”Human-centred design processes for interactive systems”. It is essencial to highlight that human-centered design does not mean that the product is designed specifically for an individual. The design has to be suited to everyone, i.e., anyone that may interact with the system (DUL; WEERDMEESTER, 2003).

The interaction between humans and machines can be abstracted as illustrated in Figure 2.1. The machine receives inputs from its environment and provides information to the human operator through displays and other monitoring devices. The operator perceives the available information, process it and decides on his/her control actions. Based on the environment's inputs and operator's commands, the machine defines its outputs to the environment.

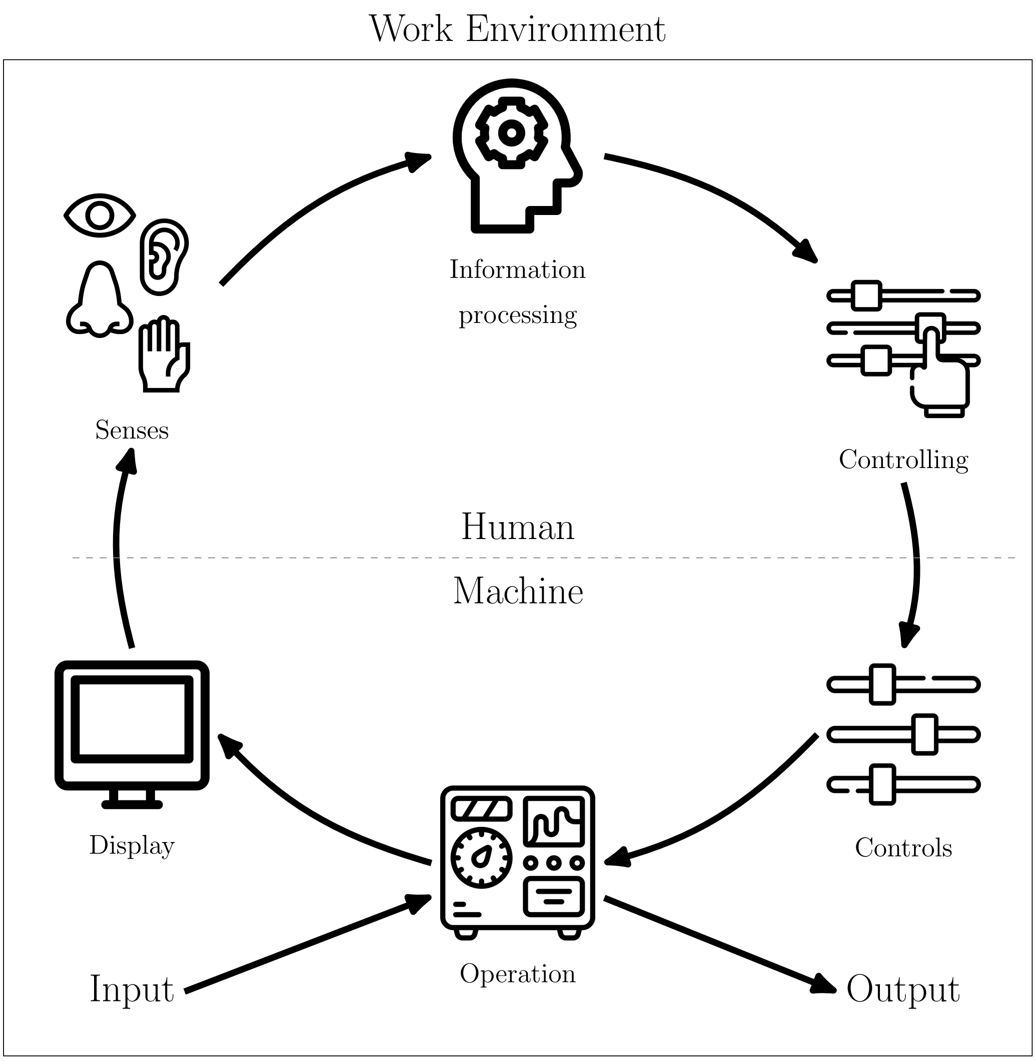


FIGURE 2.1 – Human-Machine system representation. Adapted from (SANDERS; MCCORMICK, 1998).

Humans handle devices, machines and equipment during their daily activities. All of these manipulations are susceptible to accidents or failures that can happen because of the interaction between operator, equipment and environment. Each interface with the operator can be a factor. For example:

- The operator's body position during an activity: the position can impact the operator's comfort and concentration throughout the activity, therefore, impacting the success rate or the chance of some accident happening (SANDERS; MCCORMICK, 1998).
- The environment's lighting: illumination can make details more noticeable without provoking discomfort or distraction and even increase productivity (SANDERS; MCCORMICK, 1998).
- The information displayed and manipulation of the device: the way a information is displayed on a screen, figure or text impacts how efficiently it will be understood by the operator. If this takes too long, it can draw the operator's attention for too long and compromise his/her reaction time.

Among the various human factors related to human-machine interaction, this work considers mental workload and situation awareness, which are explained in detail in the following sections.

2.1.1 Mental Workload (MWL)

Mental workload is one of the main concepts studied in Human Factors (STANTON *et al.*, 2004).

In order to explain it, Stanton *et al.* (2004) propose an analogy with the concept of physical workload. When an athlete must lift a dumbbell (one of those gym weights bars), the strength demand from the athlete is proportional to the dumbbell's mass being lifted. If the dumbbell is lighter than the athlete's capability, it is easy enough for him to lift it. If the athlete is strong enough to carry the dumbbell, he does not feel a physical demand bigger than his capabilities. In this case, the physical workload of this activity is appropriately fitted for this athlete. Two things can happen if the dumbbell is heavier than the athlete's capability. Either the athlete adapts to lift the dumbbell using tools (adjust the strategy), or the dumbbell is not lifted completely (performance degrades). This situation corresponds to the case of a operator executing a task, which is not fitted for his capabilities.

The mental workload is similar to the physical workload but refers to the mental capacity necessary to perform a task. Each human being has a finite mental capacity.

When the mental demand is higher than the operator's capacity, the person needs to adapt to finish the task, or the overall performance of the task is compromised. Otherwise, if the mental workload is too low, the operator may get bored and easily distracted and could also fail or not process the task's information.

It is important to say that mental workload is unique within each individual. It is influenced by the operator perception and also by other factors outside the task itself. These factor can be more related to the operator (like its skill, age, education, training) or the environment (like noise, heat and toxicity) (CAIN, 2007; FALLAHI *et al.*, 2016; CARDOSO; GONTIJO, 2012).

The mental workload is not a quantitative resource or something that one can directly measure, but several different methods have been proposed in the literature to infer it. Figure 2.2 illustrates three different classes of methods used to evaluate mental workload: methods based on task performance, methods based on physiological measures and methods based on subjective questionnaires.

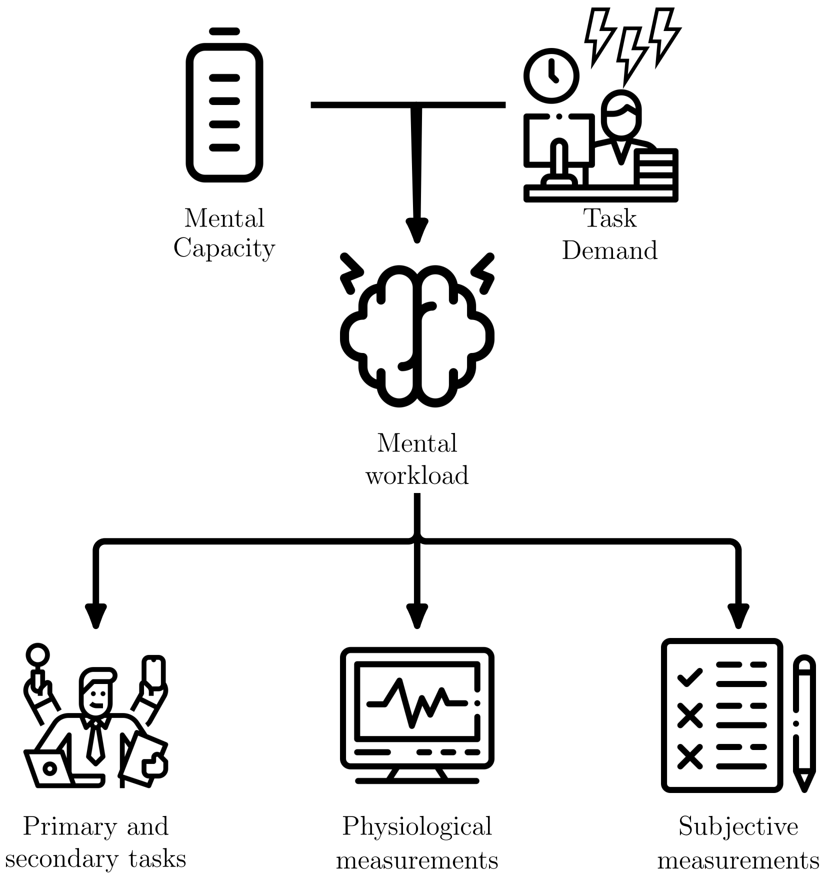


FIGURE 2.2 – A overview of mental workload and the methods to infer it.

Methods based on task performance

If the mental workload influences the task performance, it would be possible to infer it using the performance's variation of a task. Because there are cases where the user's mental capacity is too high for only one task, a common approach is to add a secondary task. In this case, the user is asked to maintain a good performance level and still try to execute both tasks. Both tasks should use the same skills (STANTON *et al.*, 2004; SANDERS; MCCORMICK, 1998).

For example, an experiment to assess mental workload in a flight simulator may use two tasks. The primary task is to fly the aircraft while maintaining a performance level. The second is something simple, like mentally summing two random numbers that appear on the screen. If the numbers' sum is odd, the pilot should press the left arrow key on the keyboard, otherwise should press the right arrow key. If the pilot's performance in the secondary task is too low, it means that the mental demand from the first task is too high to pay attention to the second task (MOHANAVELU *et al.*, 2020).

Methods based on physiological measurements

Many physiological measurements can be used to assess mental workload. The most common ones are heart and brain activity (CHAKLADAR *et al.*, 2020; ORLANDI; BROOKS, 2018), skin conductance, eye movement and pupillary contraction (STANTON *et al.*, 2004; RODRÍGUEZ *et al.*, 2015). These measurements are considered an unbiased assessment method (FALLAHI *et al.*, 2016). It is recommended to evaluate them alongside another method, as they can be influenced by unknown variables and external factors.

This work uses heart activity, obtained from an electrocardiogram (ECG) sensor, and electrodermal activity, obtained from a galvanic skin response (GSR) sensor, as physiological measurements to assess mental workload.

The electrocardiogram (ECG) is a recording of the heart's electrical activity. From it, it is possible to determine the intervals between heartbeats and the corresponding frequency (heart rate, HR). Another common variable is the heart rate standard deviation (heart rate variability, HRV) (CAIN, 2007). The heart activity is controlled by the sympathetic and parasympathetic nervous systems (STANTON *et al.*, 2004). During a task, the heart activity changes with the mental demand of the task. The heart rate is expected to increase with the mental workload, while the heart rate variability is expected to decrease. These are consequences of two reactions in our system when in a mental demand situation (STANTON *et al.*, 2004): a decrease in the parasympathetic nervous system activity and an increase in sympathetic nervous system activity. The ECG is a simple and non-invasive method used in many experiments to evaluate mental workload and other human factors' (MOHANAVELU *et al.*, 2020; MANSIKKA *et al.*, 2016; ZHANG *et al.*, 2014).

The skin electrodermal activity is affected by the person's sweating and the level of moisture in the environment. It can be used to reveal changes in our sympathetic system (NOURBAKHSH *et al.*, 2012; SHI *et al.*, 2007). It has been used in the literature as an assessment method for stress and arousal (NOURBAKHSH *et al.*, 2012; STANTON *et al.*, 2004; SHI *et al.*, 2007), the usability of human-computer systems (SHI *et al.*, 2007) and also mental workload (ZHANG *et al.*, 2014; BORGHINI *et al.*, 2014).

Methods based on subjective questionnaires

The use of subjective questionnaires to assess mental workload has been extensively discussed in the literature (SANDERS; MCCORMICK, 1998; STANTON *et al.*, 2004). They are sensitive to perceived difficulty, automation, concurrent activities and demand for multiple resources. The questionnaires can be unidimensional, which are more straighforward but has only a general workload score.

It is discussed if one should only use subjective measures to measure MWL (SANDERS; MCCORMICK, 1998; STANTON *et al.*, 2004). They are sensitive to perceived difficulty, automation, concurrent activities and demand for multiple resources. These tests can be unidimensional, which are more straighforward but has only a general workload score (STANTON *et al.*, 2004),or multidimensional. Examples of multidimensional questionnaires are the Subjective Workload Assessment Technique (SWAT) and the NASA Task Load Index (NASA-TLX). SWAT decomposes the mental in three dimensions: time load, mental effort load, and psychological stress. NASA-TLX, a questionnaire created by Hart e Staveland (1988), uses six dimensions, as described in Table 2.1. These questionnaires were proposed to evaluate only one task/activity. If the user has performed two tasks (a primary and secondary task), he/she should be oriented to answer about the primary task, not a combination of both (SANDERS; MCCORMICK, 1998).

Finally, it is essential to highlight that, to have a comprehensive evaluation of the mental workload, not to choose only one method, but to combine methods from the three classes (task performance, physiological measurements and subjective questionnaires). Mental workload is multidimensional and can reflect partially or differently in each method (SANDERS; MCCORMICK, 1998).

2.1.2 Situation Awareness (SA)

Situation awareness can be defined as “the perception of the elements within a volume of time and space (Level 1), the comprehension of their meaning (Level 2), and the

TABLE 2.1 – NASA-TLX dimensions and the description of each dimension. (STANTON *et al.*, 2004).

Dimension	**Explanation**
Mental demand (MD)	The mental and perceptive activity demanded by the task (chose, decide, think, calculate, search, etc.).
Physical demand (PD)	The physical activity demanded by the task (pull, lift, spin, drag, etc.).
Temporal demand (TD)	The time pressure felt by the user. A rating the leverages the time available and the time necessary to completed the task.
Performance (PE)	The user's satisfaction with its performance or result the task.
Effort (EF)	A rating of the effort necessary to achieve that performance felt by the user.
Frustration (FR)	A rating of stress, annoy or irritation felt by the user throughout the task.

projection of their status in the near future (Level 3)" as illustrated in Figure 2.3. One example is when an air traffic controller looks at a radar display (Level 1). He/she seeks to understand the aircraft's position and speed (Level 2) and then predict its position in the near future, 5, 10 or 15 minutes after (Level 3) (SANDERS; MCCORMICK, 1998). Similarly, when a pilot reads the cockpit panel (Level 1) and understands their data (Level 2) then he/she can predict the next reading of that same instrument or some other status of the aircraft after a couple of minutes (Level 3).

The term "situation awareness" was first proposed for the Aeronautics domain and today is considered a key factor for designing complex and dynamic systems from other domains, such as automotive, medical and nuclear (ENDSLEY, 1995). It is an essencial factor to make sure that the user will be capable to make important decisions correctly and achieve high-performance (ENDSLEY, 1988; ENDSLEY, 2018).

As it is for the mental workload, situation awareness is not a quantitative subject. The most common way to measure it is using subjective methods, among which one of the most famous is the Situation Awareness Global Assessment Technique (SAGAT). It was proposed by (ENDSLEY, 1988) and is based on how the information is processed inside the user's mind. The test application is made by freezing the operator activity, usually made in a simulation environment, and then asking the user some questions that were previously defined based on the user's activity. These questions should be as similar as possible to how the person thinks when reasoning about the situation to avoid extra effort in understanding it (STANTON *et al.*, 2004). Although freezing the activity may sound troublesome, empirical work has shown that it does not interfere with the user performance

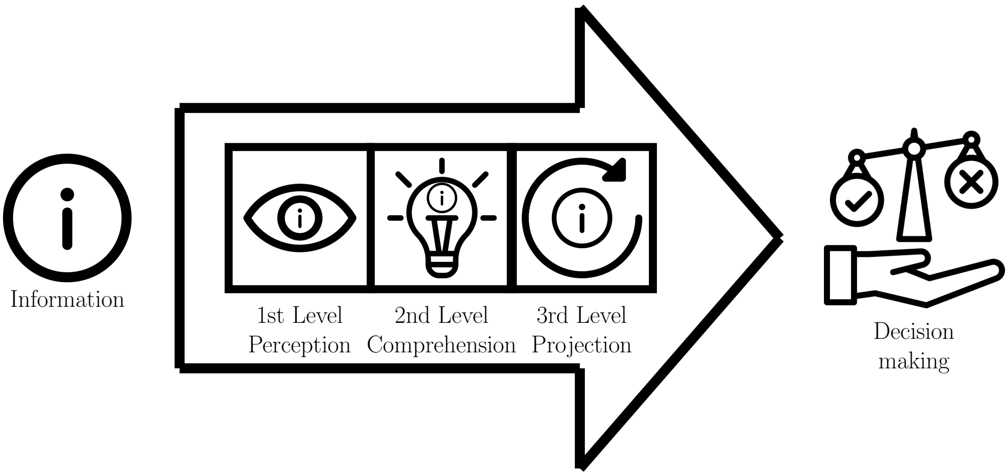


FIGURE 2.3 – An overview of situation awareness and the SAGAT.

and the user memory can withstand a break for as long as 5 to 6 min (ENDSLEY, 1988).

2.2 Extended Reality (XR)

Extended reality is a broad term that refers to all different ways of combining virtual and real entities in a human-machine interface system. It is usually decomposed into four classes (augmented reality, augmented virtuality, mixed reality, and virtual reality) that differ on the level of reality and virtuality involved in the interface system.

Milgram e Kishino (1994) organized these classes and created the concept of the “virtuality continuum”, as illustrated in Figure, as illustrated in Figure 2.4. On the left, the real environment represents the cases where the stimuli are not produced by a computer or any other digital system. Along the path to the right, the environment starts to incorporate digital elements until it reaches the far right, where all the elements in the environment are virtual and have a digital origin (NIJHOLT; TRAUM, 2005; DOOLANI *et al.*, 2020).

The extreme left means full reality, where the stimuli is not produced by any computer or any other digital system. Along the path to the right, the environment starts to have some digital elements until it reaches the far right, where all the environmental elements have a digital origin. The first step from the ”Real Environment” to ”Virtual Reality” is the augmented reality.

In the augmented reality system, the user sees some digital elements that are laid over the real environment. without making the user lose his sense of presence in the real world. These elements can be text, images, video, etc. Augmented reality can be used to assist workers in manufacturing and assembly tasks, as well as training (DOOLANI *et al.*, 2020; FARRELL, 2018; MA; CHOI, 2007).

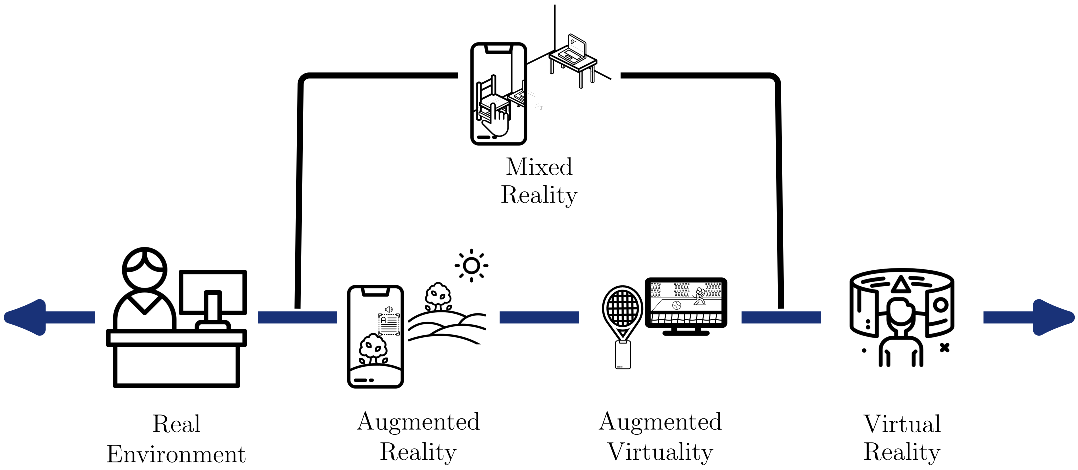


FIGURE 2.4 – The Virtuality Continuum concept. Adapted from (MILGRAM; KISHINO, 1994)

While the augmented reality brings digital elements to the real environment, the augmented virtuality creates an environment that could only exist digitally, such as a fantasy world from games or movies. This scenario is the background of some other activity that is being done by the user in the real environment. An example is to train a pilot in a virtual environment but with an accurate mock-up of the cockpit, which provides physical buttons and inceptors for the pilot to touch and hold (FARSHID *et al.*, 2018). Another example is to play sports, such as tennis, golf or baseball, in a complete digital arena but using the actual equipment with a tracker.

The mixed reality stays in between the real and virtual environments. Unlike augmented reality and augmented virtuality, in a mixed reality system the user can manipulate digital elements as if they were inside the real world (DOOLANI *et al.*, 2020). One example is when a client from a furniture store uses mixed reality not only to see how the furniture fits inside his room, but he can also move it and change its color, size and shape before buying or even going to the shop.

On the far right of the virtuality continuum, the virtual reality is when the user is the only non-digital element, everything else is digital, immersing the user in a virtual environment, but, of course, inside the physical limits of the real environment (MA; CHOI, 2007). If the feeling of presence inside that environment is well tailored, the user can momentarily forget about the real environment and act and react accordingly to the virtual environment (FARRELL, 2018).

Virtual reality is a powerful tool that allows a user to be transported to a tridimensional environment that could be out of reach or that does not exist but is needed for testing or training reasons (MUJBER *et al.*, 2004). Inside this virtual environment, the user can walk, look around and feel as if the environment was real (SALAH *et al.*, 2019).

Figure 2.5 shows the representations of each of these Extended Reality classes.

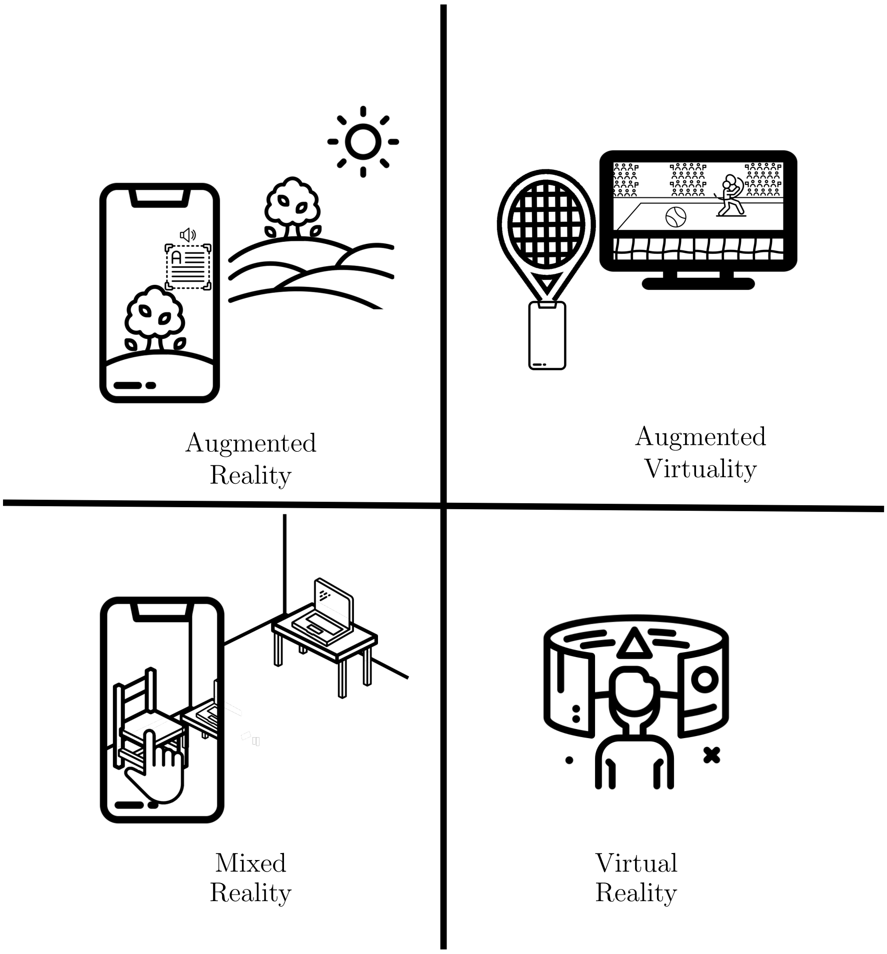


FIGURE 2.5 – A represantion of the differences between AR, AV, MR and VR

2.3 Co-Design

Co-design, or collaborative design, refers to a design process in which individuals of the design team have different backgrounds or bring different experiences, which can be essencial for the product under design. It is based on good communication and information sharing among the team (CHIU, 2002).

Kleinsmann (2006) provides the following definition:.

> *"Collaborative design is the process in which actors from different disciplines share their knowledge about both the design process and the design content. They do that to create a shared understanding of both aspects, to be able to integrate and explore their knowledge and to achieve the larger common objective: the new product to be designed."* (KLEINSMANN, 2006).

This definition emphasizes two critical aspects of co-design: knowledge sharing and integration. According to Kleinsmann (2006) knowledge is the data after the receiver's understanding or translating process, in a state that is possible to record or register, so that the person can remember and use it later. During the collaborative design, ideas, facts or concepts are exchanged between the actors. This exchange is a fundamental part of the co-design method since it is responsible for the growth of each individual's knowledge. Once the knowledge is shared among the actors, they can use it when performing their tasks, resulting in knowledge integration (KLEINSMANN, 2006).

2.4 Final remarks

This chapter is dedicated to introducing the fundamental concepts, which are necessary for understanding the proposal of this Master's dissertation. It starts by describing the importance of Human Factors in improving human-machine interaction. Two important concepts of Human Factors are presented: mental workload and situation awareness. For each of them, the main assessment methods are described.

Next, the concepts of extended and virtual reality are discussed, as well as other variations such as augmented reality and augmented virtuality. This work's idea is to explore the concept of using a virtual environment as a method of quickly evaluating prototypes of assistive devices.

Finally, the approach of co-design is proposed as a way of integrating different stakeholders into the design process.

The next chapter discusses recently published work that are related to the proposal of this work.

3 Literature review

This chapter discusses a set of selected works from the literature related to different aspects of this work. Their selection was performed in the Scopus and Web of Science databases, using keywords such as “human factors”, “virtual reality” and “blindness”. From an initial set of 344 papers, a set of seven were selected as more relevant to this work and are detailed in the following sections.

3.1 Virtual reality for BVI users

Motivated by the popularization of virtual reality technology, Siu *et al.* (2020) developed a white cane to be used by BVI users in a virtual environment. Their purpose was to make virtual reality applications available for BVI users.

The traditional white cane transmits three sources of information to the user: detection of obstacles, surface topography and foot placement preview. In their work, these sources of information were transmitted through sounds or haptics (SIU *et al.*, 2020), which would be defined based on the cane position in the virtual environment. For obstacle detection, the cane was built with a three-degree-of-freedom brake mechanism that would stop the movement when the cane hit an obstacle. A coil actuator was used to simulate surface properties. Lastly, a wave-based acoustic simulation was used to render geometry-aware sound effects in other to give the user a sense of the surroundings (echo localization).

In order to evaluate their proposal, the authors performed an experiment where the participants had to play a “scavenger hunt” using an HTC Vive system. During the experiment, each participant had two tasks: collect targets along the way (primary task) and avoid virtual obstacles and walls (secondary task). The targets appeared, one at a time, once the previous target was collected, and they emitted a sound that acted like an audio beacon for the participant. The obstacles did not emit any sound as a beacon, but the participant could detect it by the shape and the noise it emitted when in contact with the cane. The experiment was performed with 8 blind users (4 female, 4 male) from 25 to 70 years old. All of them did a training section where the virtual environment was presented.

Among the relevant findings of Siu *et al.* (2020) is that not all the participants reacted the same to a particular stimulus. The vibration of the cane was considered confusing by some participants, while others were familiar with it. This difference affected the performance of the participants. The ones that had already used vibrating devices performed better. It shows that user's previous experiences can impact their performance in the virtual environment.

Another interesting observation was that, similar to what happens in the real world, it was easier for the participants to navigate in larger areas than in tight spaces. Moreover, the authors observed that the participants focused their attention on the primary task, without freely exploring the environment, which might have impacted the low time to achieve the goal and the low number of obstacle hits.

Among the limitations pointed out by the authors is the lack of feedback possibilities for situations such as when the obstacle contacts a point along with the cane, not the tip of it, and the fact that the brake system did not stop the participant when he/she walked forwards toward a wall.

Comparing the work of Siu *et al.* (2020) to this work, Siu *et al.* (2020) were focused on providing mechanisms for a BVI user to navigate inside virtual environments. In this work, the purpose is to use the virtual environment to collect data about how the BVI user would navigate in a real environment. Another difference is in the functioning of the virtual cane, which in this work is limited to vibration, with no brake system, as the BVI user does not need to touch the environment with it. One common observation of both works is the sound importance for the BVI guidance and the need to use high-quality spatialized audio to increase the realism of the virtual environment.

3.2 Feeling of presence in virtual reality

The second work discussed in this literature review is an evaluation of what affects the user's feeling of presence in virtual reality, i.e. when the user feels drawn into the virtual environment and starts to occupy it instead of the real one (CUMMINGS; BAILENSON, 2016).

One of the many feelings that flourish during the use of a VR is the feeling of presence. This feeling, inside the virtuality context, is when someone feels drawn into a VE and starts to occupy the VE instead of the real one (CUMMINGS; BAILENSON, 2016).

Jicol *et al.* (2021) aim to correlate the feeling of presence with one's agency (which is the self-perception that the user is in control of a situation or some actions (FARRER; FRITH, 2002) and emotion. For this purpose, the authors created two virtual environment, one that would trigger happy emotions, and other that would trigger fear. For each, two variations were provided: one that the user could interact with (with agency) and another that it could not (without agency).

Following, they performed an experiment where 121 participants were randomly assigned to one of the four virtual environments. The purpose was to evaluate three hypotheses: 1) The intensity of the dominant emotion correlates positively with the presence; 2) Presence is significantly higher in environments where participants have agency; and 3) Agency moderates the effect of the emotion on the presence.

The experiment's results confirmed the first hypothesis: no matter if the feeling was positive (happiness) or negative (fear), the users did feel a more substancial presence when the positive or negative feelings were more intense. The second hypothesis was only partially confirmed. In the virtual environment that induced fear, agency did make a difference and induced a higher feeling of presence, while in the environment that induced happiness, agency did not affect the presence. The same could be said about the third hypothesis.

Although the study of Jicol *et al.* (2021) was limited to sighted people, it highlights the importance of the feeling of presence in virtual reality. It provides important inputs to this work, such as the need of including mechanisms for the user to interact with the virtual environment in order to increase the feeling of presence.

3.3 Information for BVI navigation

Bradley and Dunlop published two works (2002, 2005) about how BVI navigates and how much it is similar or different to how a sighted person navigates.

The first work of Bradley and Dunlop was published in 2002 and discussed which type of information BVI uses to navigate in an environment and how it compares to sighted people. The data were collected during structured interviews where the participant had to explain how to arrive at two different locations as if they were talking to someone with the same vision condition (BRADLEY; DUNLOP, 2002).

Based on the answers, the authors defined 11 categories of information: 1) directional (e.g. left/right, north/south); 2) structural (e.g. road, monument, church); 3) environmental (e.g. hill, river, tree); 4) textual-structural (e.g. name of shops, places, restaurants); 5) textual-area/street-based (e.g. name of street, neighbourhoods, squares); 6) numerical (e.g. first, second, 100m); 7) descriptive (e.g. steep, tall); 8) temporal/distance based (e.g. ”walk until you reach...” or ”before you get to”); 9) sensory (e.g. the sound of engines, the smell of bread from a bakery); 10) motion (e.g. cars passing by, doors opening); 11) social contact (e.g. asking people or using a guide dog for help) (BRADLEY; DUNLOP, 2002).

As an output from the interviews, the authors provided the average number which each category was used by each group and is reproduced in Figure 3.1. From the results, the researchers observed that BVI participants used less text-based information than the sighted participants. However BVI participants used more words to describe a path than the sighted participants. Another essential result was that visually impaired people used, on average, 9 to 10 categories to describe a route, while sighted people used around 6 categories.

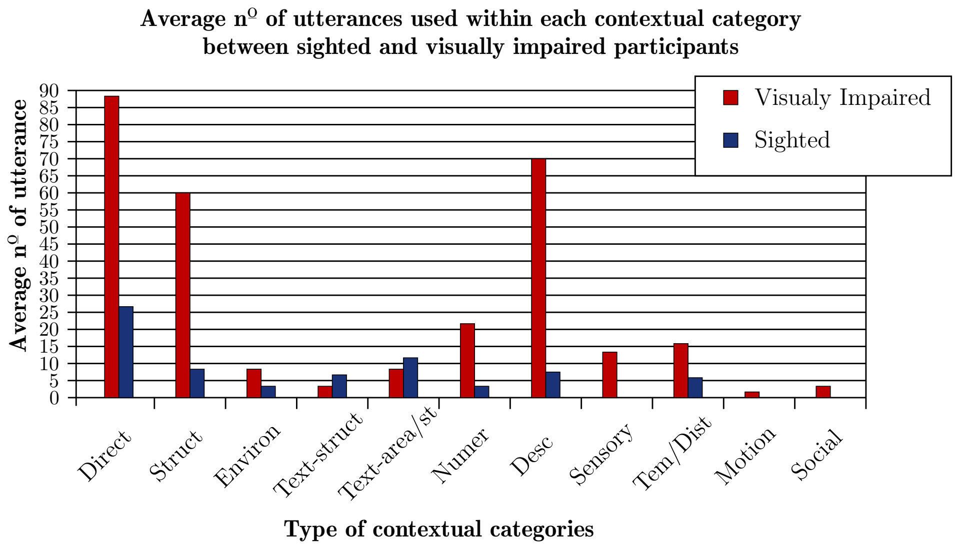


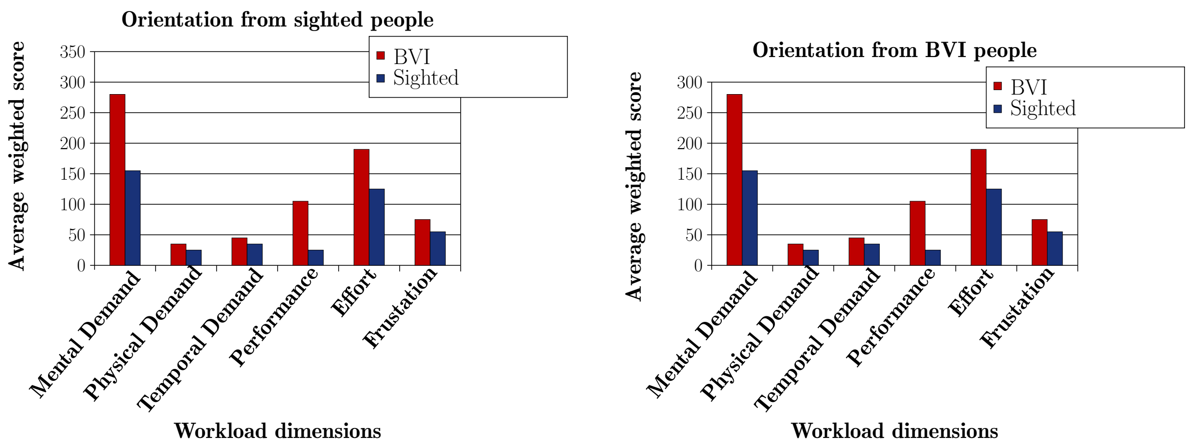
FIGURE 3.1 – Comparison between sighted participants with BVI participants (Adapted from Bradley e Dunlop (2002)).

Among the comments provided by BVI participants, a common one was about the limitations of available navigation methods, such as white canes and guide dogs. They also emphasize that, when navigating, using different senses is essential for confirming one piece of information.

To extend the findings of their previous work, Bradley and Dunlop designed an experiment to investigate if there is a difference between the perceived workload of BVI participants and sighted participants when they navigate using user-tailored information created with the results of the previous experiments (BRADLEY; DUNLOP, 2005).

The experiment was performed with 16 participants, 8 sighted and 8 BVI, who were recruited to walk to four pre-determined landmarks in the centre of Glasgow. They followed the orientations recorded during the interviews from their previous work. For each participant, orientations for 2 of the 4 landmarks were made using sighted users' interviews, while the other 2 used data from BVI interviews. The results showed that BVI users reached landmarks significantly quicker when given the information made for that group, but still longer than sighted users.

Another issue analysed during the experiment was the perceived workload. After each landmark, the participant was asked to complete the NASA-TLX questionnaire. The average score for each dimension of the NASA-TLX is reproduced in Figure 3.2. As expected, it shows that BVI participants systematically have a higher workload than sighted participants. It also confirms that BVI did have a higher workload when guided by orientations provided by sighted people, as well as the sighted participants did with orientations from BVI. Another essential piece of information that stands out is the high frustration score given by the BVI users when they were guided by the orientations of sighted people.



(a) Condition 1. (b) Condition 2.

FIGURE 3.2 – Comparison of the NASA-TLX between the participants (Adapted from Bradley e Dunlop (2005)).

The work of Bradley e Dunlop (2002) brings some relevant information for developing this work. Firstly, it shows the differences between the way sighted and BVI people navigate, highlighting the importance of including BVI in the design process of assistive technologies. It confirms the limitations of the current solutions. It brings essential insights on what type of information to include in developing audio systems, which is one of the assistive devices evaluated in this work. Finally, it shows the importance of using different workload assessment methods when evaluating assistive technologies.

3.4 Audio navigation for BVI

Despite the several existing navigation systems for BVI users, their limitations have been pointed out in many works. Yang *et al.* (2014) explored the effect of two factors when BVIs use a standard GPS navigation system. The first factor is the amount of detail of the provided information (information completeness). The second one is the distance between the information reproduction and the object referred by that same information (broadcasting timing).

In order to evaluate the impact of these two factors, Yang *et al.* (2014) experimented with BVI users where each factor had two levels. The completeness of the information could be "complete" and "simple" and the broadcast timing could be 5m and 7m. As outputs the authors evaluated the participants' performance by their precision and time in finding a goal, and evaluated their perceived workload with NASA-TLX.

The independent variables were analysed by a two-way ANOVA hypothesis test. They found that the precision in finding the goal was only influenced by the broadcasting timing. The time in finding the goal was influenced by both variables. The task's workload was influenced by the broadcasting timing and the interaction between it and the information completeness.

The work of Yang *et al.* (2014) shows the importance of synchronizing the information provided by the audio system with the current position of the BVI user – a point to be taken into account when developing the audio solution used in this work. However, concerning the lack of influence of information completeness, it is relevant to observe that, in a certain way, this result contradicts the conclusions of Bradley and Dunlop of 2002, 2005 and, therefore, should be considered with caution. It may be due to the difference between the two levels (complete and simple) adopted in the experiment. Finally, Yang *et al.* (2014) confirms the NASA-TLX as a feasible tool to evaluate workload in experiments with BVI participants.

3.5 Comparison of assistive devices

Marston *et al.* (2006) compare assistive devices for BVI. Two guidance displays were evaluated, one based on haptics and another based on sound. They were tested in two scenarios: a busy street block with a variety of street furniture, parked bicycles and people, and a park, with paths made of concrete, crushed gravel and paver blocks.

The experiment was performed with 8 BVI participants. As output, the authors collected the time to reach a set of waypoints, the errors made by the participants, the travelled distance and the percentage of the total time that the users accessed the guidance device. All participants were able to complete the task with both devices. However, the configuration (audio x haptic, street x park) that resulted in the best performance varied among the participants. One relevant consideration is that the haptic device caused strain on the participants' arm and was considered less acceptable when compared to the sound device, which required no use of the arms.

Similar to the study of Marston *et al.* (2006), this work also compares different devices based on haptics and sound. But, complementary to Marston *et al.* (2006), this work also evaluates the combined use of devices, as well as a guidance system currently familiar by the BVI participant.

3.6 Virtual reality in the design process

The use of virtual reality for design purposes is not new. It has been studied and evaluated in several areas, including Aeronautics. Moerland-Masic *et al.* (2021) investigated using virtual reality during the aircraft cabin's design process with the purpose of facilitating the communication between the design team and the client.

The cabin design process (Figure 3.3) is often said to be complex because it involves several stakeholders, each with his/her own set of preferences and requirements. According to the authors, the time needed to satisfy the multiple demands tends to be long, and the process usually requires building multiple mock-ups and attending many meetings with the stakeholders.

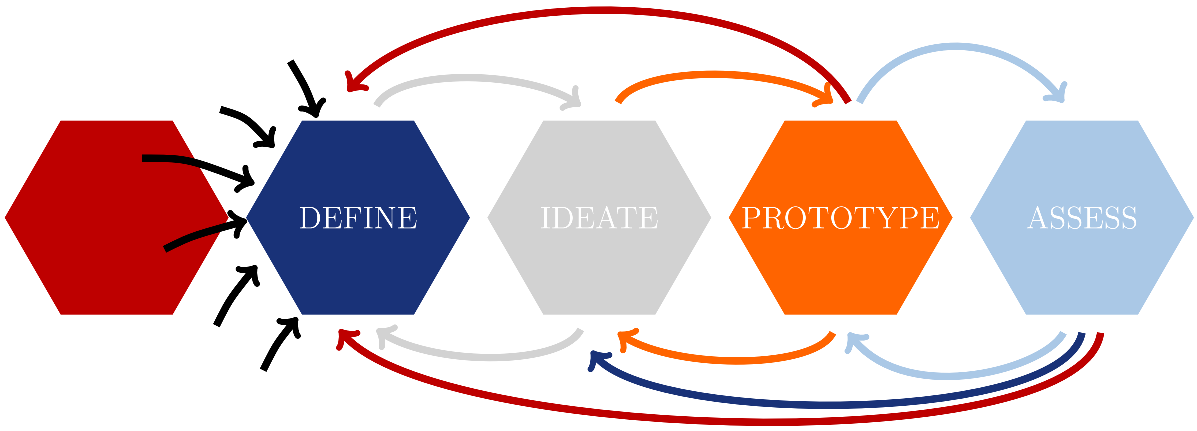


FIGURE 3.3 – Simplified cabin design process (Adapted from Moerland-Masic *et al.* (2021)).

Moerland-Masic *et al.* (2021) proposed to anticipate the involvement of the final users based on co-design. In their proposal, the users can influence the product's development from the beginning, as shown in Figure 3.4. However, for the involvement to happen, a communication channel needed to be established, and it was done using virtual reality.

The authors described the application of the proposed approach to a use case. Three different designers initiated a cabin design. In the traditional method, the results were illustrated in a sketch, which could only present a glance of what the cabin would be, and in a 3D model, which had more details. However, any modification required a new rendering session and this could take hours, or even days. The same solution was also

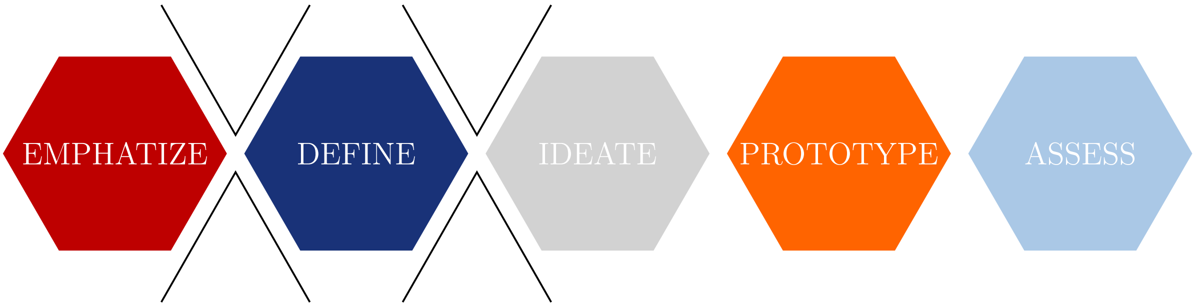


FIGURE 3.4 – Best moments for user involvement (Adapted from Moerland-Masic *et al.* (2021)).

illustrated in a virtual reality environment. The sketch was inside the aircraft cabin, where the client or the stakeholder could draw and give their opinions from the beginning of the design process. The 3D models could be imported to increase the sketch's level of detail.

The use case showed some benefits and disadvantages of using virtual reality. The virtual reality helped to bring the client closer to the design team, allowing them to draw quick sketches in brainstorming gatherings. It was associated with a steep learning curve for the designers. Among the disadvantages, it was considered a high-cost tool, and its use for a long time was associated with nausea.

The work of Moerland-Masic *et al.* (2021) is an example of how virtual reality can be used to bring the user into the design process. Similarly, in this work, virtual reality is explored to create a test environment where BVI users can try out the device under development, contributing to improving its usability. Differently from the work of Moerland-Masic *et al.* (2021), in this work, users are not expected to feel sick, as it is usually associated with discrepancies between the motion of the image in the virtual environment and the motion perceived by the vestibular system of the user.

3.7 Final Remarks

This chapter discussed published works that are related to this dissertation. These works combine at least two of the following concepts: "human factors", "virtual reality" and "blindness".

Some of the works discuss the design of navigation tools, investigating different aspects of their solutions, while others focus on providing a virtual reality experiment to BVI users. Among the assessment methods used in the literature, the NASA-TLX is a recurrent option, as well as the proposal of performance metrics.

Generally, all the surveyed works bring exciting conclusions about the designing of BVI devices and contributed in some way to the proposal of next chapter.

4 The proposed method

This chapter describes the method proposed in this work for evaluating assistive devices using virtual reality. The method is organized into 5 phases, further decomposed into 10 steps, as illustrated in Figure 4.1.

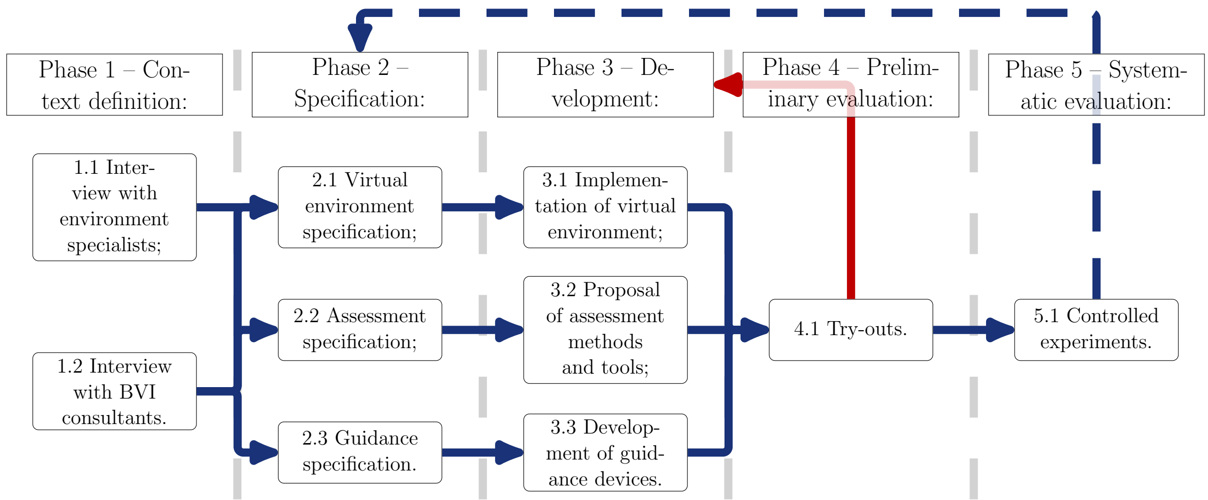


FIGURE 4.1 – Methodology's diagram

The first phase is the context definition. It consists of defining the main features of the environment in which the assistive device will be used, based on interviews from specialists. It also includes a step for understanding the limitations of the current assistive devices and defining the main features of assistive devices to be designed. This last step is based on interviews with BVI users.

The second phase is devoted to specification and is composed of three steps. It includes the specification of a virtual environment that represents the real environment where the device will be used. In parallel, the relevant human factors for the device evaluation are defined. Finally, the assistive devices are also specified.

The third phase is dedicated to developing the virtual environment, the evaluation tools and methods, and the first proof of concept of the assistive devices, which should be integrated into the virtual environment for testing.

The fourth phase provides a preliminary assessment of the devices through its unstructured experimentation by BVI consults. This preliminary assessment provides feedback for improving the device concept. The cycle of "try-out and improve device concept" can be repeated until the device concept is considered mature to be tested through a systematic set of controlled experiments.

The fifth phase consists of executing a campaign of controlled experiments, following the best practices of the DoE (Design of Experiments) discipline, and analysing the results. Concluded this phase, the results should provide information for the design team to decide between proceeding to the detailed design of the assistive devices or performing a new evaluation cycle.

In the last case, the experiment should provide feedback on improving the devices' specification, refine the assessment methods, expand the virtual environment to include new tasks and/or situations, or even abandon a device concept. Each evaluation cycle should increase the maturity of the assistive devices under development.

In order to illustrate it, the proposed method is applied to the evaluation of assistive devices concepts that should be able to guide BVI users in a hospital or a medical clinic during the covid-19 pandemic.

4.1 Phase 1 – Context definition

As previously stated, the context in which the assistive devices are evaluated is the BVI navigation through a hospital or medical clinic in a covid-19 pandemic situation.

Step 1.1 Interview with environment specialists

In order to gather information about this context, the first step of Phase 1 consists of interviewing specialists from hospitals to understand the usual hospital procedures and the standard features of this environment.

For this work, employees from two hospitals of São José dos Campos were interviewed (Vivalle Hospital and the Municipal Hospital). They were questioned about the layout of rooms and spaces and the processes that a new patient follows through the hospital, from the check-in until he/she gets in the doctor's office.

According to the interviews, the procedures adopted by patients in both hospitals are similar and are usually composed of the following steps:

1. Enter the hospital;
2. Use the sanitizer to clean their hands;

3. Take a queue number and wait for the call of the receptionist;
4. Go to the receptionist and check in;
5. Sit in the waiting area and wait until the patient's name is called;
6. Leave the waiting area toward the doctor's office.

Step 1.2 Interview with BVI consultants

One of the motivations of this work is that BVI users are not completely satisfied with the current guidance products. We hypothesize that BVI users are not usually consulted from the early stages of the products' development.

In this work, two BVI acted as consultants for the proposal of guiding methods and the design of a virtual environment that would be familiar to their reality. They had different visual impairments. One of them became blind at the age of 13 years old, while the other was diagnosed with Usher's disease.

The interviews with BVI consultants were critical to understanding how they perceive a medical clinic as they walk in and how they interact with the environment. Among the inputs provided by BVI interviews' are the importance of including in the virtual environment background noise typical of a reception, such as a keyboard typing and a telephone ringing, as well as a few physical expected objects to interact with (chairs and desk).

4.2 Phase 2 - Specification

In Phase 2, the information collected through the interviews of Phase 1 is used to make critical decisions about the virtual environment where the evaluation of the assistive devices will be carried out. It is also used to define which human factors should be assessed. Finally, it contributes to define the guidance methods that should be implemented in the assistive devices.

Step 2.1 Virtual environment specification

Based on the data collected during the interviews, the scenario was defined as the hospital or medical clinic reception. The BVI patient should be able to navigate through the reception hall and reach a reception desk. After he/she is registered, the BVI patient should wait until he/she is called and then goes to his/her doctor's office.

At this step, a preliminary layout of the virtual environment was sketched, combining the elements that were common to the reception of the two hospitals, such as a waiting area with chairs (some of which should remain empty due to covid distancing), a reception desk, an alcohol totem, among others.

One limitation that had to be taken into account in the specification of the virtual environment is the corresponding physical area (real empty room) that should be available so that the BVI could walk through the real world. while he/she navigates in the virtual world. Typical dimensions of a reception are around 15x20m. Due to the limited space available at the laboratory, the reception was scaled down to fit into the available physical space, with an area of 7x10m but maintaining the same elements (reception desk, waiting area, sanitizer, among others)

Step 2.2 Assessment specification

Based on the review of the literature and the BVI interviews, the assessment method should include the following dimensions, which were commonly evaluated in previous works:

A) Performance evaluation;

B) Mental workload evaluation.

Additionally, a third dimension was also included, based on lessons learned from the Aeronautics area:

C) Situation awareness evaluation.

This dimension was included to verify if the BVI user can build a mental map of the environment based on the inputs provided by the devices.

Finally, a fourth dimension is included for evaluating the user preferences when comparing the assistive devices.

D) Guidance method evaluation.

Step 2.3 Guidance specification

Most current solutions for supporting BVI navigation use sound, vibration or both to communicate relevant information about the environment to the user.

Based on the interviews with BVI consults, the three options were selected to be tested in this work (audio, vibration and audio/vibration combination).

Moreover, an exciting property was also evaluated: the effect of information being transmitted with or without the user's command. This evaluation was applied to the vibration method and resulted in it being split in two variants: one that works around the user and another that works with where he/she decides.

In summarizing, four guidance methods were chosen to be evaluated:

A) Audio guidance;

B) Vibration guidance with command;

C) Vibration guidance without command;

D) Mixture of audio and vibration guidance.

It was also pointed out that the assessment should compare these four methods and the familiar guidance method of the BVI participants among them (e.g. white cane) with each other.

4.3 Phase 3 - Development

With the specifications from the previous phase, it is possible to start the development of the virtual environment, the guidance methods and the human factors assessment tools.

Step 3.1 Implementation of the virtual environment

The virtual reality environment was developed using the Unity3D platform, a well-known tool for virtual reality applications and development of games. Unity 3D has some built-in tools, but it is also possible to customize functions for more specific use(WANG *et al.*, 2010).

The implementation of the virtual environment should be followed by preparing the corresponding physical space to perform the test campaign. During this step, additional simplifications were introduced. In the real world, the person's position is tracked by two stations of the virtual reality system. According to the specification of the system, the maximum distance between the stations should be 5 m to guarantee tracking quality. This limitation led to simplify further the virtual environment limiting it to a floor area of 4x4 m, which, in the real world corresponded to the CCM entry hall.

Following the recommendations from the BVI interviews, typical reception furniture was placed in the virtual environment: a reception desk and a waiting area, composed of 2-3 chairs (two standard seats and one marked with an ”X” due to maintaining a minimum distance due covid-19), as illustrated in Figure 4.2. Were added a telephone and a laptop to the scene, and they were programmed to emit sounds at random moments. The purpose was to increase the feeling of immersion and indicate to the BVI participant where the reception desk was located. The participant navigation through the reception was composed of 4 tasks, also illustrated in Figure 4.2.

1. Clean the hands at the sanitizer totem (COVID-19 procedures);
2. Go to the reception desk to receive a queue number;
3. Go to the waiting area and wait for the number calling;
4. Leave the room when called.

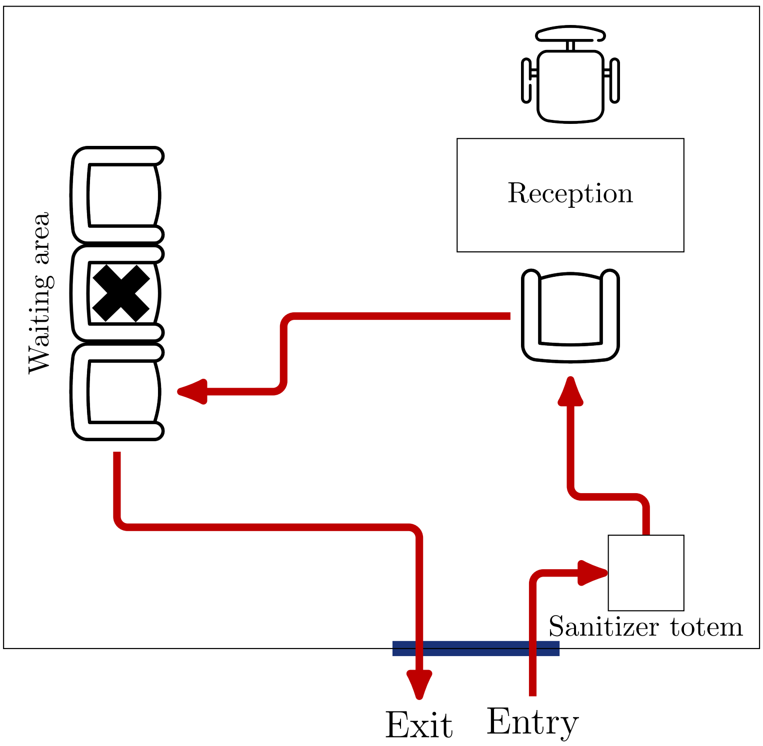


FIGURE 4.2 – Scheduled task of the experiment and their order.

These tasks should engage the user and make it navigate through the room. The purpose is to verify if he/she can draw a mental map of the scene and use the available information about obstacles to avoid them when needed. These sources of distraction were added in order to increase the immersion and to be a distraction as well. Otherwise, the virtual scene would not replicate the reality of a reception scenario.

Figure 4.3 shows the virtual environment created in Unity3D and the corresponding real environment assembled at the CCM entrance hall.

(a) Virtual environment screenshot (b) Real environment photo

FIGURE 4.3 – Environment comparisson

Step 3.2 Proposal of assessment methods and tools

As previously defined, the assessment should evaluate performance, workload and situation awareness. Following, we detail the selected methods and tools for each case.

A) Performance;

In the hospital reception scenario, the proposed measurement related to BVI performance is the number of times the BVI user hits the furniture in the virtual environment during the tasks' execution.

B) Workload;

Following the recommendations from the review of literature, the workload is estimated using two approaches:

- Physiological measures obtained from an ECG sensor and a GSR sensor;
- NASA-TLX subjective questionnaire (Appendix A.2).

C) Situation awareness;

In order to evaluate the BVI situation awareness, a modified version of the SAGAT questionnaire is proposed. This questionnaire was based on the proposed idea of Endsley

(1988) and is presented in Appendix A.3. The actor acting as the receptionist questioned when the user sitted at the reception desk.

As the original idea, the proposed version is based on 3 levels of situation awareness:

Level 1 – Perception

It aims to evaluate if the user can perceive the environment surrounding him/her. It is not expected that the user details about all the objects in the environment, only about some key objects. Example: "Is there an object around you?"

Level 2 – Comprehension

After the user answer about an detected object, he/she is asked to point to where the object is located. The same question is made when the user moves from the reception desk to the waiting chairs. During this movement, the user needed to find out where are the waiting chairs, and that could make them lose their sense of direction.

Level 3 – Projection

This level is measured after every question that asks the location of an object. He/she is then required to answer how far he/she supposes that this object is

The questions written for this questionnaire was made with the support of a blind consultant. He was explained about the concept of situation awareness and then suggested questions. The result analysis was based on the number of corrected answers. This adpated SAGAT is on Appendix A.3.

D) Guidance method evaluation;

Finally, a questionnaire is proposed for evaluating the guidance methods. This questionnaire was also guided by the same blind consultant from the Adapted SAGAT questionnaire. The consultant was asked to make questions about each method that he considered necessary for an assistive device to have and he focused on the comfort, the sense of safety, the sense of confusion and on the precision that the manipulation of the device caused. This questionnaire is on Appendix A.4.

Step 3.3 Development of guidance devices

As previously stated, four guidance methods were proposed to be evaluated in this work.

A) Audio guidance;

The first method is audio guidance. Basically, in the course of the experiment, the participant could give two different voice commands:

- "What is around me?";

The answer to this command was a quick description of the closest furniture around the user.

- "Where is (something)?".

The answer to this command was the direction and distance of something asked by the user.

Although an automatic audio guidance system that recognizes the two voice commands and answer them could be easily developed, for the proof of concept of the audio guidance system, this was done with the interference of a member of the design team.

B) Vibration guidance with command – virtual cane;

The vibration guidance with command was implemented in a device named "virtual cane", as it was inspired on the long cane. When using a white cane, the user points it to check nearby obstacles in a specific direction. The virtual cane has a similar way of functioning, but instead of connecting the user to the nearby object through the cane, it vibrates when it detects an obstacle in the direction the user pointed it. A virtual reality hand-control was used to implement it.

The virtual cane algorithm runs in the virtual environment. When requested by the user, it identifies the nearest object in the direction pointed by the user and calculates its distance. If the nearest object is within a specific range, it then calculates the vibration intensity based on the distance between the object and the user, and sends the corresponding command to the hand-control device.

C) Vibration guidance without command – haptic belt

A haptic belt was developed as a device that uses vibration guidance without command. The belt has appended 8 vibration units that vibrate accordingly to the direction and distance of the closest object around the user.

The main differences between the virtual cane and the haptic belt is that the haptic belt checks 360° around the user. When objects are within a certain limit, it vibrates indicating to the user the direction of the closest object.

The project of the haptic belt was inspired by a haptic compass (KYLECORRY31, 2020). However, instead of having the input being defined by a magnetometer, it uses the information available in the Unity3D environment.

The haptic belt was developed using an Arduino Mega 2560 and the following materials: an ESP32 DevKit v1, a printed circuit board (PCB), a leather belt, 8 coin vibrators 1027, and a 3D printed case (Figure 4.4). The PCB was designed in the EasyEDA web platform (Figure 4.5).

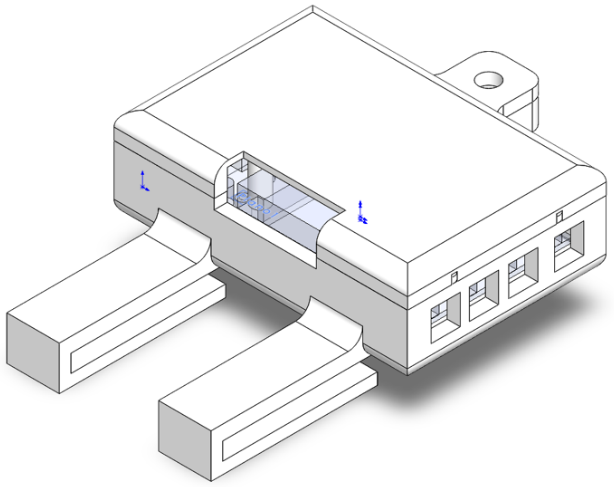


FIGURE 4.4 – CAD model of the designed case

The haptic belt is illustrated in the Figure 4.6 and communicates with the Unity3D environment using a Bluetooth connection.

The haptic belt algorithm is divided into two modules: one implemented in the virtual environment and one implemented in the ESP32 kit. In the virtual environment, it determines the direction of the nearest object to define the vibration motor(s) that must be activated or deactivated and the corresponding intensity. It then sends a message to the ESP32 kit with the corresponding command. The algorithm implemented in the ESP32 kit receives the command from the virtual environment and activates/deactivates the corresponding motors.

D) Mixture of audio and vibration guidance

This option is implemented making the three options available to the user: audio guidance, haptic belt and virtual cane.

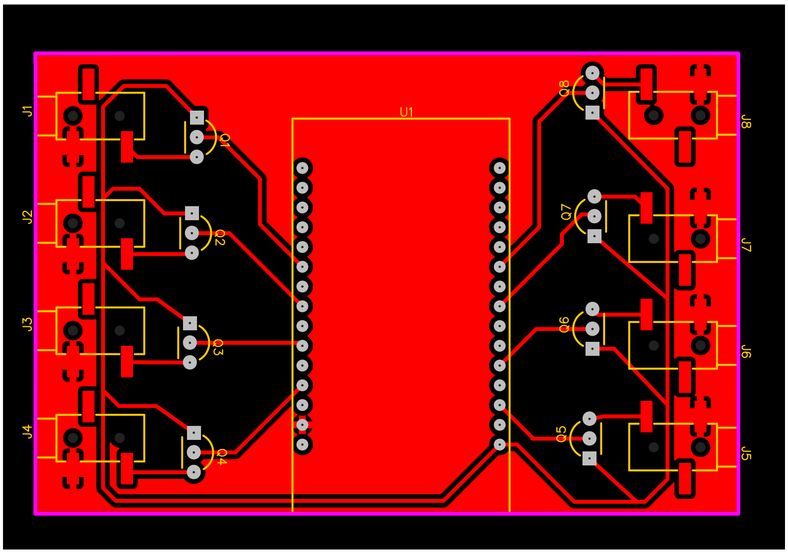


FIGURE 4.5 – Printed circuit board wiring

FIGURE 4.6 – The haptic belt

4.4 Phase 4 – Preliminary evaluation

Step 4.1 Try-out

During the preliminary evaluation phase, the concepts of assistive devices were tested by two BVI users and feedback was provided to improve both the devices and the virtual environment.

Regarding the devices, the main feedback provided by the BVI users was about the virtual cane. Initially, the virtual cane algorithm emitted a different sound when the virtual cane was pointed to the floor, walls or ceiling. This sound was found annoying by the BVI users and was removed from the algorithm.

Regarding the virtual environment, BVI users pointed out the need to add more sources of noise in order to improve the immersion. The suggestions were both for internal and external sounds.

Regarding internal sounds, they point to the lack of people chattering and the noise that came from a TV. Both were included in the virtual environment. In order to simulate people chattering, dialogues between two people, collected from videos or series available on the internet, were added to the virtual environment. The TV noise was collected from famous Brazilian TV programs. Another missing artifact noticed by the team was the queue machine, calling different names at random intervals.

Regarding external sounds, the BVI observed that usually they use external sounds to locate the exit of a room. As a suggestion, the following sequence of sound was created and added to the virtual environment to run at random moments:

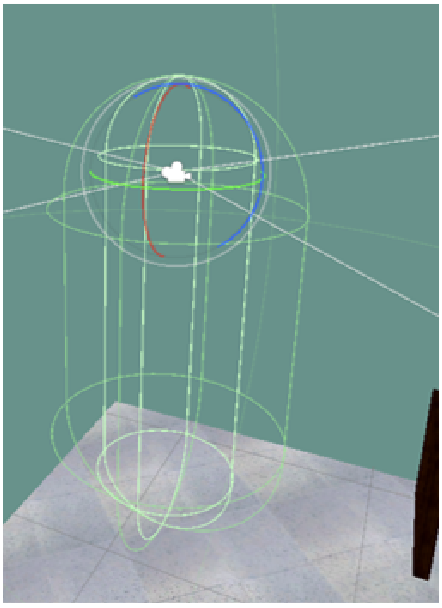
1. Sound of a door opening;
2. Noise from an exterior space (like people walking, cars passing by, horns, etc.);
3. Sound of a door closing.

The sequence was associated with a sound-emitting point in the virtual environment, running at random moments.

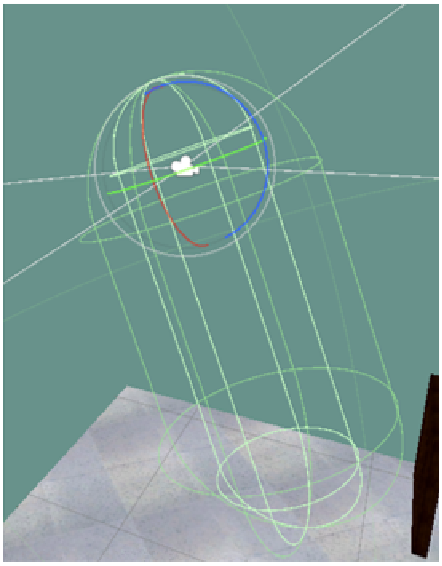
Once the preliminary evaluation of the environment was concluded, different versions of the reception were created, modifying the position of the objects, in order to make possible the execution of multiple tests with the same person without the cumulative effect of learning.

Another contribution that emerged from the try-out regards detection of the user's collision. A routine was implemented in the virtual environment to detect and register the collisions as a performance metric. The tests with BVI users showed that the routine for automatic collision detection in the virtual world does not work correctly because it does single monitor parts of the human body, such as legs, hands and arms.

The only information provided as input is the position and orientation of the head (captured by the HTC VIVE HMD). Based on this information, the virtual platform approximates the volume occupied by the human body to a vertical capsule, as illustrated in Figure 4.7a. If the user tilts his/her head down, as if facing the ground, the capsule rotates about the HMD point, making the virtual body of the user occupy a different space from user's body, as illustrated in Figure 4.7b. This approximation leads to several errors, both related to detecting collision that did not happened and not detecting collisions that happened.



(a) The user's capsule while the participant is straight and looking forward.



(b) The user's capsule while the participant is straight but looking down.

FIGURE 4.7 – Two different capsule positions based on the user's head position.

The solution to the collision detection problem would be the acquisition of additional sensors to monitor the user's hands, arms and legs. However, this solution was not feasible within the time frame available for this work. As a result, the performance was removed from the list of factors to be evaluated. Developing an automatic solution for detecting collision in the virtual world is recommended as future work.

4.5 Phase 5 – Systematic evaluation

During the systematic evaluation phase, the evaluation experiment is designed and executed. It consists of inviting BVI volunteers to use the assistive devices to navigate inside the virtual environment and perform the proposed tasks. The assessment techniques are applied during and after the use of each device.

In the case of this work, the proposed experiment consists of asking the participant to use five guidance methods: the four methods presented in Section 4.2 (audio guidance, haptic belt, virtual cane, mixed) and, additionally, the guidance method used daily by the BVI (e.g., white cane). Moreover, each participant should use each guidance method twice ("first visit" and "return visit"), in order to provide some information about how the guidance method performs in new and known environments.

In order to avoid the learning effect from one device to the another, five versions of the reception scene are developed, changing the position of the objects - one version to be used with each guidance method. The scene order is randomized for each participant.

Moreover, particularly in the case of this work, an additional round of tests is added to the experiment to investigate the differences between the evaluation performed by BVI users and sighted (non-BVI) users. For this purpose, the same experiment is repeated with a set of non-BVI users. The purpose is to investigate whether or not performing the analysis with non-BVI users could lead to different conclusions.

The experiment is organized in the following way:

- Briefing:
  - The experiment's purpose is explained to the participant, followed by the signature or recording of the free and informed consent form.
  - An explanation about the physiological sensor is provided and the participant is invited to wear it.
  - The assistive devices are introduced to the participant.
  - The participant is invited to wear the VIVE HMD and start the experiment.
- Execution of the experiment - the following sequence is repeated for each guidance method:
  - The participant receives a brief introduction about the guidance method and is given some time to familiarize and train with it. During this step, no data is gathered.
  - First visit: the participant performs the task and is asked the SAGAT questionnaire during the task.
  - The participant answers the NASA-TLX for the first visit.
  - Return visit: the participant performs the task for the second time in the same scene. However, a few things are changed, such as the presence, or not, of

a television or random people talking. The participant answers the SAGAT questionnaire again.

  - The participant answers the NASA-TLX for the return visit.
  - After both visits, the participant answers the questionnaire about the guidance method.

- Conclusion
  - The physiological sensors are removed from the participant and the experiment is concluded.

As a result of the experiment, the following data are collected:

- Answers to the NASA-TLX questionnaire (Appendix A.2);
- ECG and GSR signals;
- Answers to the SAGAT questionnaire (Appendix A.3);
- Answers to the guidance method questionnaire (Appendix A.4).

The data analysis is discussed in the next chapter.

4.6 Final Remarks

This chapter describes the method proposed in this work to evaluate early concepts of assistive devices using virtual reality. It also discusses and proposes a set of assessment techniques.

The next chapter shows the results obtained from the execution of an experimental campaign.

5 Results' analysis and discussion

The purpose of the experiment discussed in this chapter is to investigate the two research questions proposed for this work:

- Is it possible to evaluate and compare concepts of an assistive device from a human factors' perspective in a virtual environment? What are the main limitations of the use of a virtual reality environment?
- Do non-BVI users, when deprived of their vision, similarly evaluate assistive device as BVI users?

For this purpose, the experiment described in Section 4 was performed with the following groups:

- Blind group: composed of 4 participants with ages varying from 26 to 56, all male, three of them graduated and one with ongoing graduation.
- Sighted group: composed of 4 participants with ages varying from 22 to 31, three males and one woman, all graduated.

In order to answer the two research questions, this chapter is organized in the following way. Section 5.1 is dedicated to the first question and brings an analysis performed only with data from blind participants. Then, Section 5.2 repeats the same analysis now with data from sighted participants and compares the results with those obtained from blind participants in order to answer the second research question.

In both sections, the data analysis follows the following sequence:

- Analysis of subjective questionnaires:
  - NASA-TLX: it aims at assessing the workload perceived by the user in six dimensions, including 'mental demand'. It is expected a decrease in the mental workload between the 'first' to the 'return' round. It is also expected that some guidance methods would differ regarding the required mental workload.

  - Adapted SAGAT: it aims at assessing the situation awareness and the user's mental map. It is expected that the SAGAT score would increase from the 'first' to the 'return' round. It is also expected that some guidance methods would differ regarding the required situation awareness provided to the user.
  - Guidance method's questionnaire: It assess the user experience with each method. It is also expected that some guidance methods would differ regarding the score received in this questionnaire.

- Analysis of physiological sensors:
  - ECG: it aims at assessing the user workload. Two features are extracted from the ECG signal, heartrate (BPM) and heartrate variance (SDNN). The heartrate is expected to decrease slightly from the 'first' to the 'return' round, while the heartrate variance is expected to increase slightly.
  - GSR: it aims at assessing the user workload and stress. It is expected that the GSR average would increase at every 'first' round and then a slight decrease in the 'return' round.

Particularly in the case of this work, an additional round of tests is added to the experiment to investigate the differences between the evaluation performed by BVI users and sighted (non-BVI) users. For this purpose, the experiment is repeated with a set of non-BVI users and the same data is collected. The purpose is to investigate whether or not performing the analysis with non-BVI users could lead to different conclusions.

5.1 Evaluation of assistive device from a human factors' perspective in a virtual environment

5.1.1 Subjective data

5.1.1.1 NASA-TLX

The NASA-TLX provides two relevant pieces of information to the workload analysis. The first is the score attributed to the "mental demand" dimension and the second is the average obtained from NASA-TLX's six dimensions. The two analyses are presented in the next subsections.

5.1.1.1.1 Analysis of the mental demand scale

Table 5.1 presents the "mental demand" score of each blind participant to each guidance method. The base method refers to the guidance method that the person uses in his/her daily life (e.g., white cane).

TABLE 5.1 – Score of NASA-TLX mental demand for the blind participants.

Participant	Round	Base	Audio	Haptic Belt	Virtual Cane	Mixture
001C	First	3	1	14	3	6
	Return	1	1	10	2	6
002C	First	5	1	1	10	12
	Return	1	1	1	10	3
003C	First	5	5	5	8	1
	Return	3	1	1	2	1
004C	First	9	10	15	10	10
	Return	7	10	14	8	10

The mean value obtained for each guidance method is illustrated in Figure 5.1. It shows a systematic reduction in the perceived mental workload between the rounds for all methods, confirming that the participants get familiar with the devices after the first use. It also shows that although the haptic belt obtained the most considerable mean, it also had the most significant variation, showing that the effort required from the user may vary significantly.

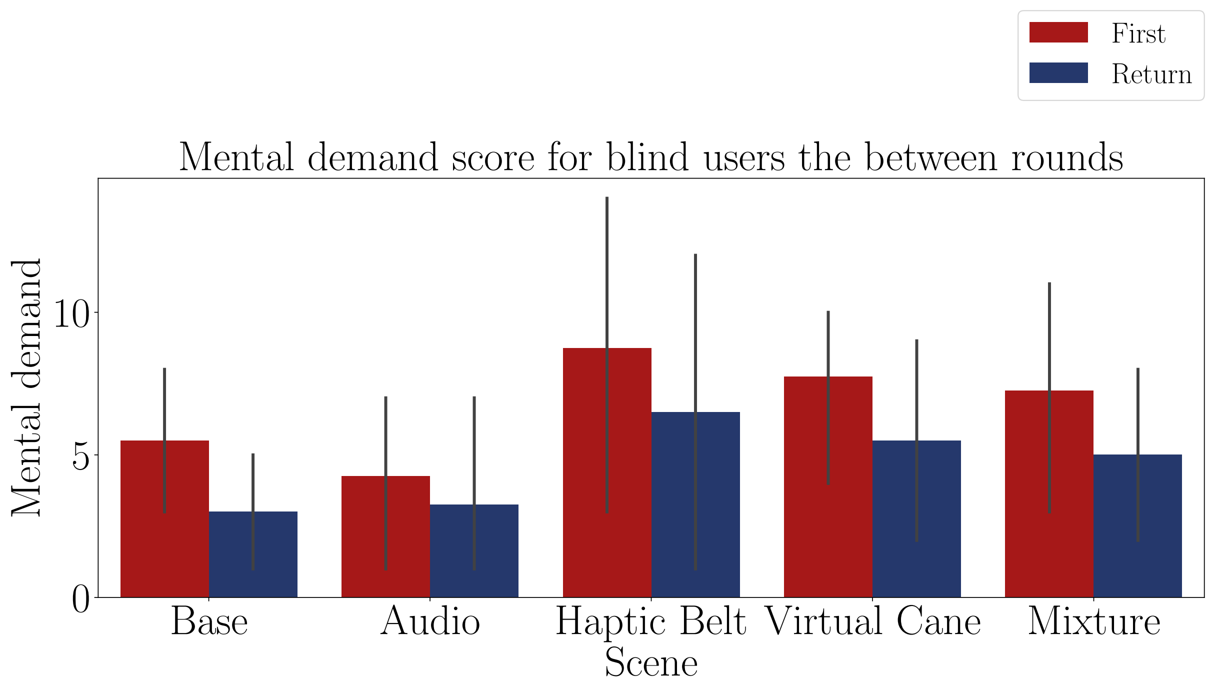


FIGURE 5.1 – Mean and standard deviation of mental demand of blind participants for each method.

Figure 5.2 presents a boxplot of the mental demand score grouped by the methods.

This figure shows that there may be two groups: one associated with lower demand, composed of base and audio, and another with higher demand, composed of haptic belt, virtual cane and mixture. It indicates that maybe a guidance method that uses vibration as input is not intuitive. Figure 5.3 presents a boxplot of the mental demand grouped by the rounds, confirming the general tendency to reduce the required "mental demand".

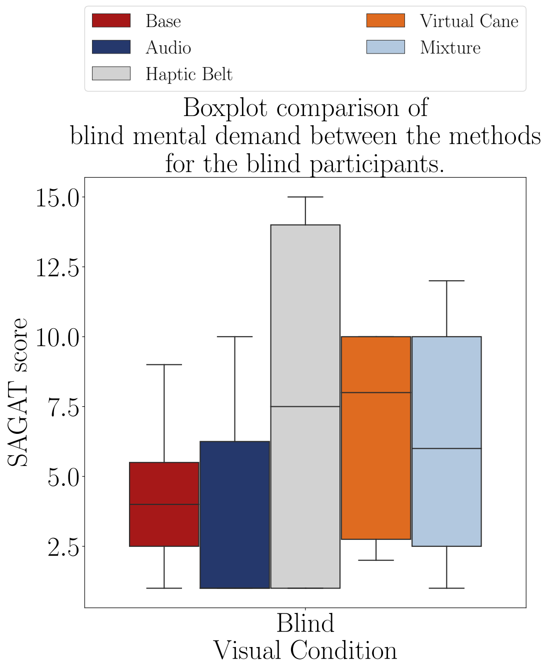


FIGURE 5.2 – Boxplot of the mental demand of the blind participants grouped by the methods.

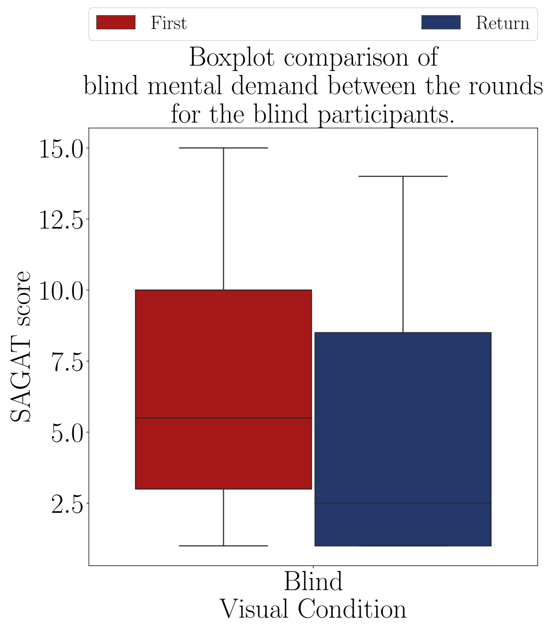


FIGURE 5.3 – Boxplot of the mental demand of the blind participants grouped by the rounds.

In order to support the statistical analysis, Figures 5.4 and 5.5 presents the QQ-plot and the residual plot of the "mental demand" data, confirming that the data follow a normal distribution and the residues are homogenous.

Figures 5.4 and 5.5 show the distribution and variance of Table 5.1. These figures show that the data are normally distributed and that the methods have a similar variance. Table 5.2 shows the ANOVA test p-values of the mental demand of the "blind" sample between the guidance methods. The methods' and the rounds' p-values indicate that there is no influence from them in the mental demand. The interaction between the methods and the round also does not influence the mental demand.

Following, the statistical model of Equation 5.1 is used for the analysis of variance (ANOVA):

$$y_{ijk} = \mu + \tau_i + \beta_j + \omega_k + (\tau\beta_{ij}) + e \tag{5.1}$$

where:

- y_{ij} - output variable for method i, round j and participant k;

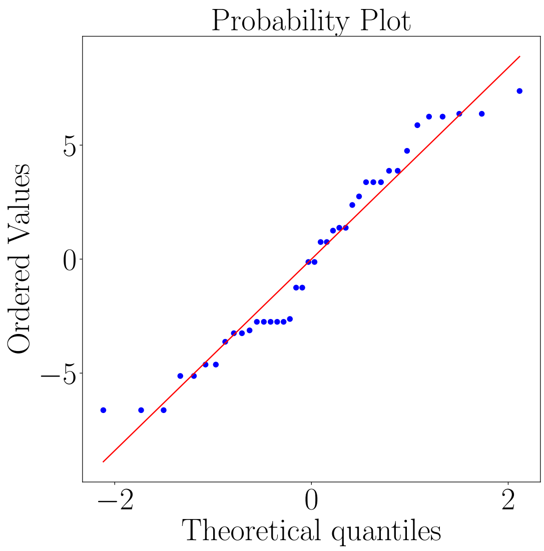


FIGURE 5.4 – QQ plot of the mental demand of the blind participants on each method.

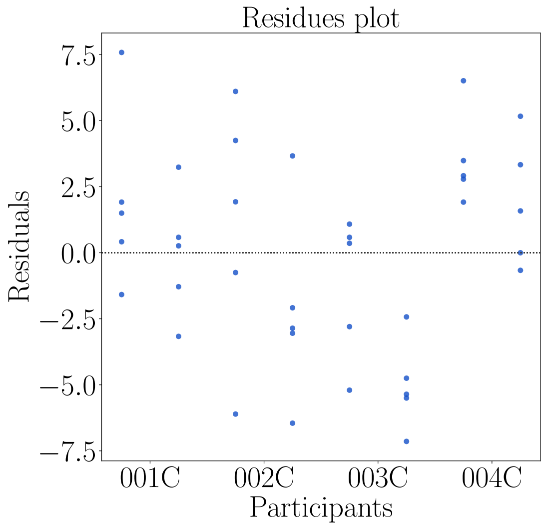


FIGURE 5.5 – Residual plot of the mental demand score the blind participants on each method.

- μ - mean of all the observations;
- τ_i - variance from method i;
- β_j - variance from round j;
- $\varpi_{\bar{k}}$ - variance from participant k, which is treated as a block;
- $\tau\beta_{ij}$ - combined variance from the interaction between method i and round j;
- e - residual error.

The results of ANOVA are presented in Table 5.2. ANOVA tests the hypothesis that the means of independent data groups are equal or not. In the literature, a p-value of 0.05 is commonly adopted as a threshold to confirm the hypothesis. A p-value < 0.05 indicates that the means of the groups are statistically different with 95% of confidence. According to this criterion, neither method or round have a significant influence on the mental demand.

However, due to the low number of participants, the threshold of 0.1 could also be considered. In this case, it indicates, with 90% confidence, that the mean of the first and return rounds are different. For the guidance method, the p-value of 0.170 is close to the threshold but slightly higher, suggesting that the means may be different. However, this hypothesis is not statistically confirmed with the current data.

In order to conclude the analysis of the NASA-TLX mental demand, Table 5.3 brings the average difference between the mental demand of the first and return rounds. Unexpectedly, it shows that the most significant variation is obtained to the base, i.e., the

TABLE 5.2 – ANOVA p-value for mental demand – blind participants

Source	P-Value
Methods	0.170
Rounds	0.075
Interaction	0.993

guidance method the participant uses and, therefore, should not present a significant variation. The methods with the lower variation was audio, probably because it already had a shallow score in the first round.

TABLE 5.3 – Mental demand variation grouped by participant and visual condition

Visual Condition	Base	Audio	Haptic Belt	Virtual Cane	Mixture
Blind	-2.5	-1.0	-2.2	-2.2	-2.2

5.1.1.1.2 Analysis of the NASA-TLX score

This section repeats the analysis steps of the previous section but now considers the mean value of all dimensions of NASA-TLX, referred to in this text as the global score. Table 5.4 presents the global score of each blind participant.

TABLE 5.4 – NASA-TLX score felled by the blinded participants.

Participant	Round	Base	Audio	Haptic Belt	Virtual Cane	Mixture
001C	First	4.833	4.000	8.833	5.167	6.333
	Return	4.167	4.000	6.667	4.500	6.167
002C	First	6.333	4.833	4.833	9.000	7.000
	Return	4.500	4.833	4.833	7.000	5.167
003C	First	4.000	4.000	5.333	6.667	3.500
	Return	4.000	3.833	3.667	3.500	3.500
004C	First	9.833	10.000	12.667	9.667	11.000
	Return	8.667	9.167	11.667	9.333	10.833

Figure 5.6 brings the corresponding barplot with the mean value and standard deviation for each guidance method and each round. In a qualitative comparison with Figure 5.1, the differences between the methods are confirmed but softened. It is possible to notice that the mean score of audio and base are still lower than that of the other methods.

The differences between first and return rounds are also reduced. However, the standard deviation is also considerably reduced for all methods, and especially for the haptic belt.

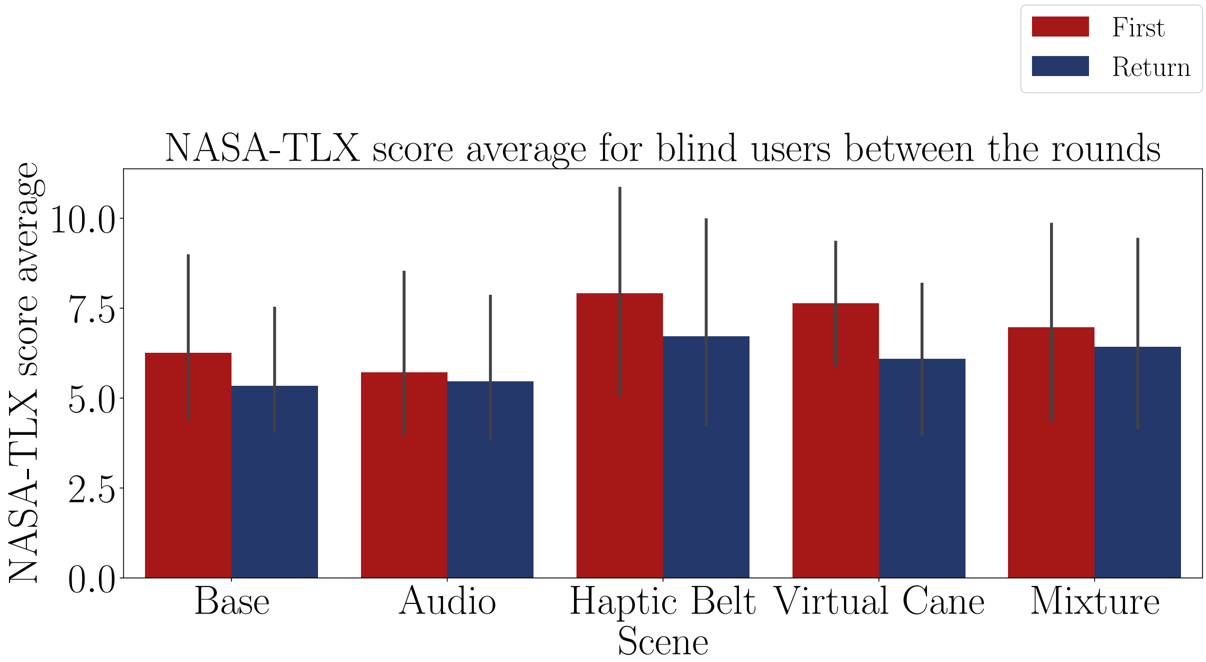


FIGURE 5.6 – Barplot of the average NASA-TLX score of the blind participants on each method.

Figure 5.7 presents the boxplot with the NASA-TLX global score grouped by the methods. Similar to what happened for the "mental demand", it is possible to split the methods into two different groups: base and audio, which require a lower level of workload, and another group, which requires a higher level. Figure 5.8 presents a boxplot with the NASA-TLX global score grouped by the rounds, showing that the two groups are still different.

Figures 5.9 and 5.10 presents the QQ plot and residual distribution of the NASA-TLX global score, showing that the data are normally distributed. However, the residuals are not so homogeneous as in the previous case, showing that the participants have different variability among them.

Table 5.5 brings the p-value resulting from ANOVA. In this case, both the methods and the rounds were appointed as significant variables that influence the mean value of the NASA-TLX global score.

TABLE 5.5 – Anova p-value for the NASA-TLX score on each method for blinded users.

Source	P-Value
Methods	0.029**
Rounds	0.022**
Interaction	0.814

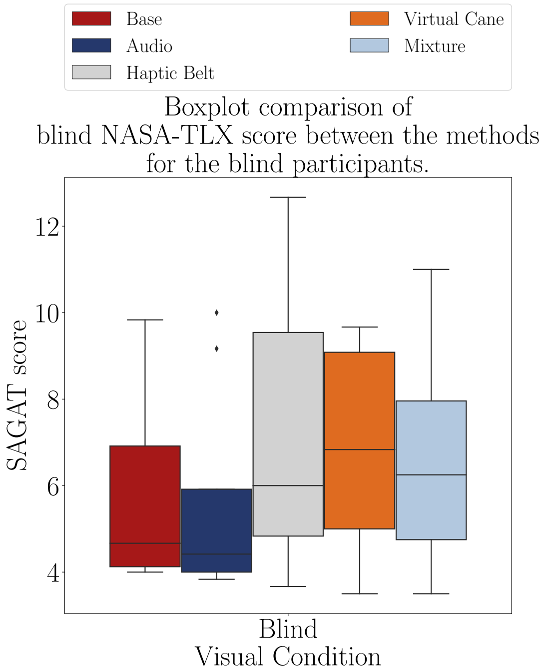


FIGURE 5.7 – Boxplot of the NASA-TLX of the blind participants grouped by the methods.

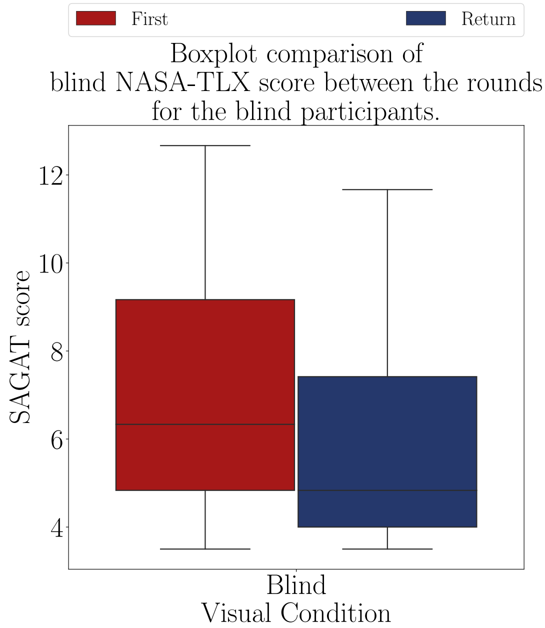


FIGURE 5.8 – Boxplot of the NASA-TLX demand of the blind participants grouped by the rounds.

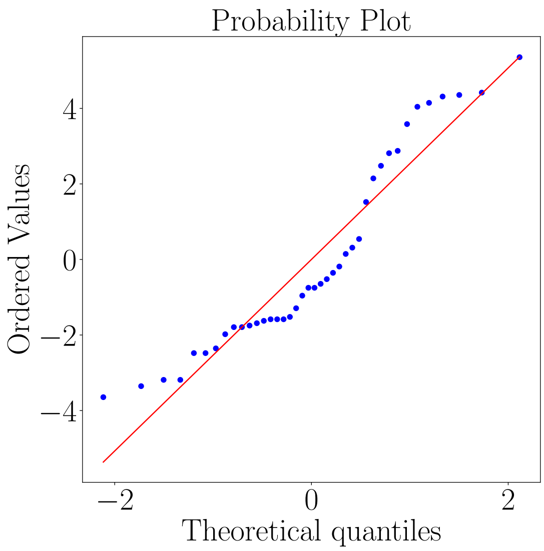


FIGURE 5.9 – QQ plot of the NASA-TLX score of the blind participants on each method.

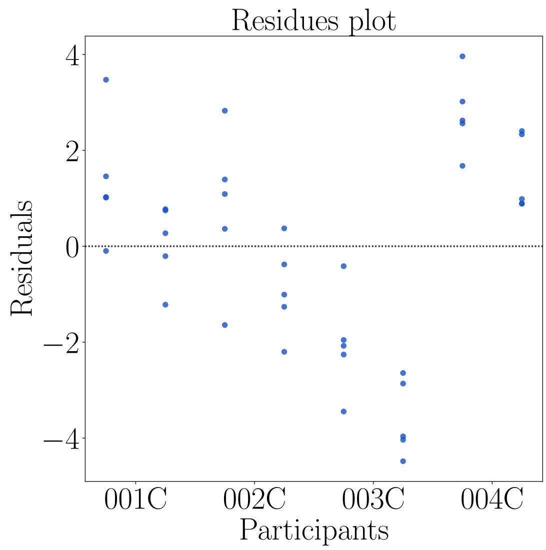


FIGURE 5.10 – Residual plot of the NASA-TLX score the blind participants on each method.

Finally, Table 5.6 presents the results of a pairwise Fisher LSD test comparing each pair of guidance methods. The results show that only audio is similar base. All the other methods are different from each other.

Table 5.7 shows the difference in the NASA-TLX global score between the first and return rounds. It shows that the audio difference is the lowest among all methods, while the highest difference is for the virtual cane.

TABLE 5.6 – Cross validation p-value for the NASA-TLX score on each method for blinded users.

Method			Analysis	
Base	X	Audio	$H_0 : \mu_{Base} = \mu_{Audio}$	
Base	X	Haptic Belt	$H_1 : \mu_{Base} \neq \mu_{HapticBelt}$	**
Base	X	Virtual Cane	$H_1 : \mu_{Base} \neq \mu_{VirtualCane}$	**
Base	X	Mixture	$H_1 : \mu_{Base} \neq \mu_{Mixture}$	**
Audio	X	Haptic Belt	$H_1 : \mu_{Audio} \neq \mu_{HapticBelt}$	**
Audio	X	Virtual Cane	$H_1 : \mu_{Audio} \neq \mu_{VirtualCane}$	**
Audio	X	Mixture	$H_1 : \mu_{Audio} \neq \mu_{Mixture}$	**
Haptic Belt	X	Virtual Cane	$H_1 : \mu_{HapticBelt} \neq \mu_{VirtualCane}$	**
Haptic Belt	X	Mixture	$H_1 : \mu_{HapticBelt} \neq \mu_{Mixture}$	**
Virtual Cane	X	Mixture	$H_0 : \mu_{VirtualCane} = \mu_{Mixture}$	

TABLE 5.7 – NASA-TLX score grouped by participant and visual Condition.

	Base	Audio	Haptic Belt	Virtual Cane	Mixture
Visual Condition					
Blind	-0.92	-0.25	-1.21	-1.54	-0.54

5.1.1.2 Adapted SAGAT

This section discusses the results of the adapted SAGAT questionnaire, which aims at assessing the participant's situation awareness and their mental map of the environment.

For each question of the SAGAT questionnaire, the participant could score 1 point or a fraction of it. The total score achieved by each blind participant is presented in Table 5.8. Figure 5.11 illustrates the corresponding bar plot, indicating the mean and standard deviation for each guidance method and each round. This figure shows clearly that the participants improved their situation awareness in the return round, when they already had some information about the environment. Also, it is possible to observe that the worst situation awareness is obtained in the first round for the virtual cane. However, on the return round, the SAGAT mean score becomes equivalent to the audio method.

Figure 5.12 brings the boxplot of the SAGAT score grouped by the guidance methods. It shows that the methods can be divided into two groups. The first one is composed of base, haptic belt and the mixture. This group received scores higher than the second group, composed of audio and virtual cane. Figure 5.13 shows the boxplot of the data grouped by round and confirms the general improvement of situation awareness from the first to the return round.

Proceeding to the statistical analysis of the data, Figures 5.14 and 5.15 present the QQ plot and the residual distribution, which confirms the normal distribution assumption

TABLE 5.8 – SAGAT global score felled by the blinded participants.

Participant	Round	Base	Audio	Haptic Belt	Virtual Cane	Mixture
001C	First	6.25	5.50	5.33	5.83	3.500
	Return	6.25	6.50	8.50	5.50	5.500
002C	First	6.75	4.50	3.99	4.50	6.250
	Return	5.25	5.00	4.00	6.50	8.500
003C	First	7.25	7.50	7.49	4.66	9.000
	Return	10.00	10.00	8.50	9.00	9.000
004C	First	7.50	6.00	7.66	4.99	6.500
	Return	9.00	6.00	9.25	7.25	9.000

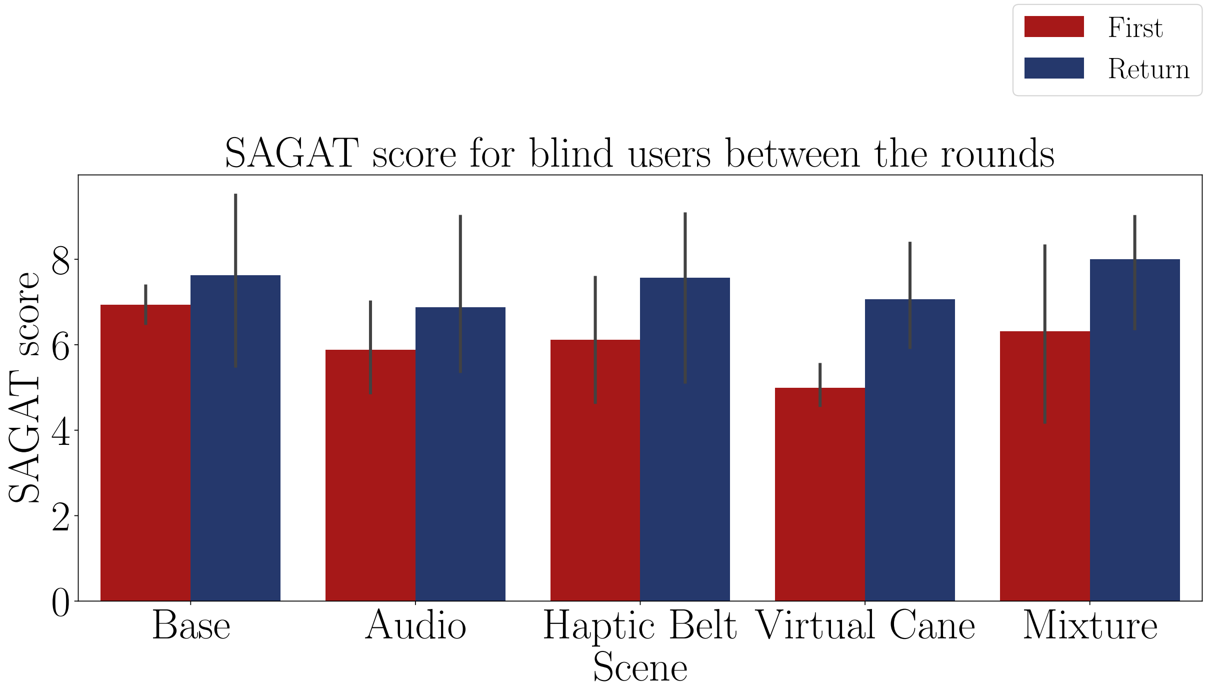


FIGURE 5.11 – Barplot of the average SAGAT score of the blind participants on each method.

and the homogeneity of variances.

Table 5.9 shows the ANOVA test p-value of the SAGAT score. It indicates that the round is a significant variable that influences the value of the SAGAT score. The same cannot be said for the method, which has no significant influence.

TABLE 5.9 – Anova p-value for the SAGAT score on each method for blinded users.

Source	P-Value
Methods	0.277
Rounds	0.002**
Interaction	0.834

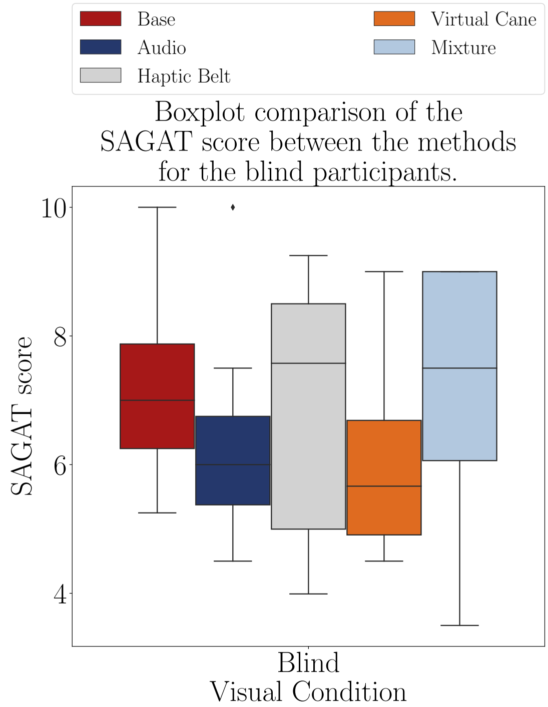


FIGURE 5.12 – Boxplot of the SAGAT score of the blind participants grouped by the methods.

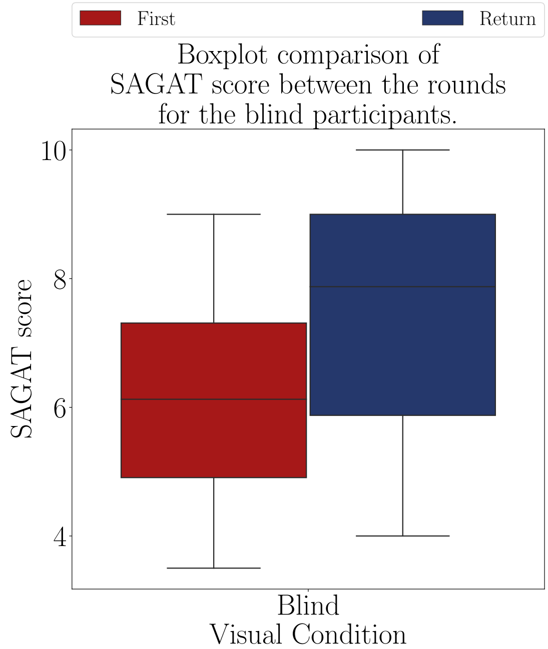


FIGURE 5.13 – Boxplot of the SAGAT score of the blind participants grouped by the rounds.

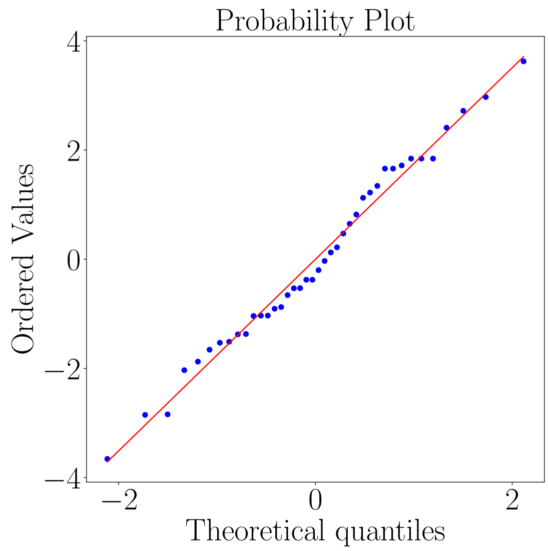


FIGURE 5.14 – QQ plot of the SAGAT score of the blind participants on each method.

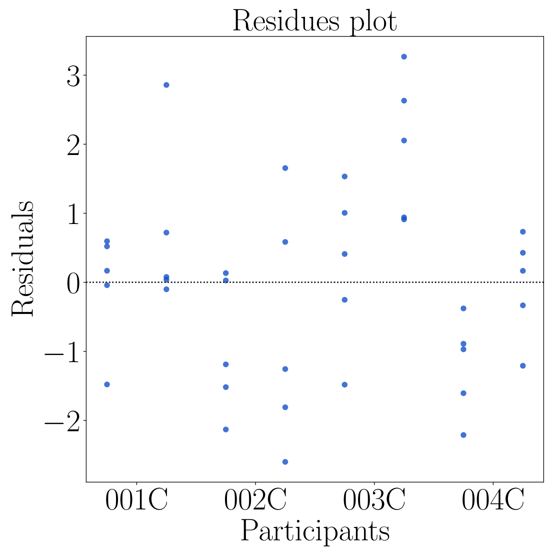


FIGURE 5.15 – Residual plot of the SAGAT score the blind participants on each method.

Finally, Table 5.10 presents the mean difference in the SAGAT score between the first and return rounds for each guidance method. It shows that the base and audio methods have the lowest difference, while the highest was obtained for the virtual cane.

TABLE 5.10 – Adapted Sagat global score variation grouped by participant and visual Condition

Visual Condition	Base	Audio	Haptic Belt	Virtual Cane	Mixture
Blind	8.93	15.66	23.49	44.30	32.90

5.1.1.3 Guidance method's questionnaire.

The data from the questionnaire for evaluating the user experience with each guidance method is also analysed. The higher the score, the more satisfied the user is with the method. It is essential to observe that this analysis does not include the base method as the questions are specific about each method and the base may vary among the participants. Also, there is no distinction between first and return rounds. Each questionnaire is answered only once for each method.

Table 5.11 presents the score attributed to each method by each participant. The mean values are plotted in Figure 5.16 and show a dissatisfaction with the methods that only use vibration for communicating with the participant, i.e., the haptic belt and the virtual cane.

TABLE 5.11 – Guidance method questionnaire score felled by the blinded participants.

Participant	Audio	Haptic Belt	Virtual Cane	Mixture
001C	0.774	0.543	0.629	0.865
002C	0.857	0.743	0.543	0.935
003C	0.929	0.571	0.543	0.745
004C	0.881	0.486	0.400	0.730

Figure 5.17 brings the questionnaire boxplot, which clearly shows the difference between two groups: haptic belt and virtual cane, and audio and mixture.

Figures 5.18 and 5.19 show that the data follows a normal distribution. However, the residual variance is not strictly homogenous among the participants.

The results of ANOVA are presented in Table 5.12 and it shows that the method, with a p-value of 0.001, is indeed a significant variable that affects the user's satisfaction.

TABLE 5.12 – Anova p-value for the questionnaire score on each method for blinded users.

Source	P-Value
Method	0.001**

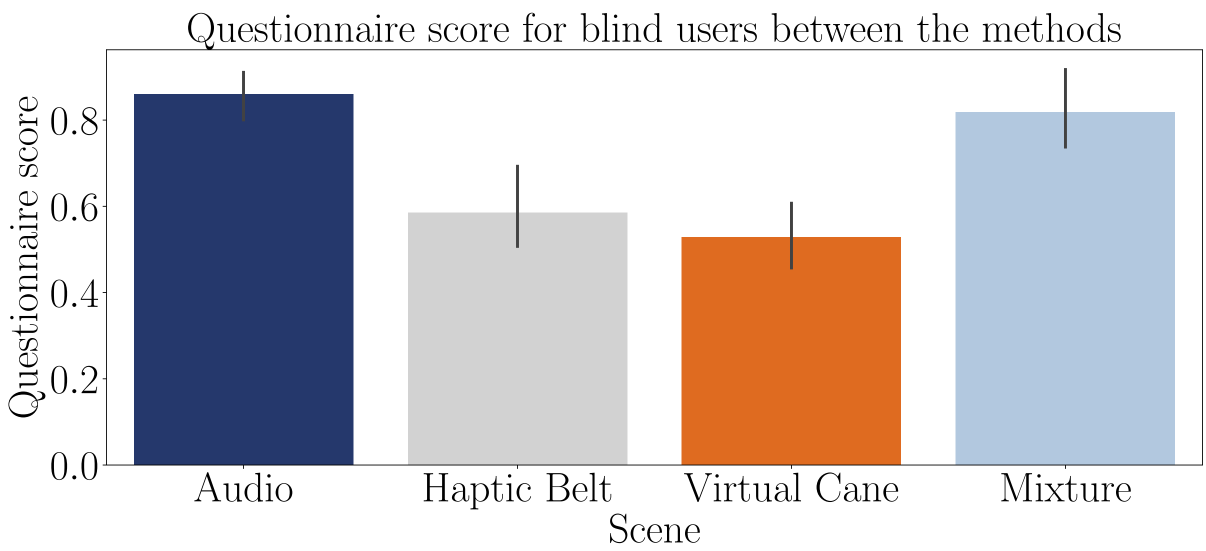


FIGURE 5.16 – Barplot of the average questionaire score of the blind participants on each method.

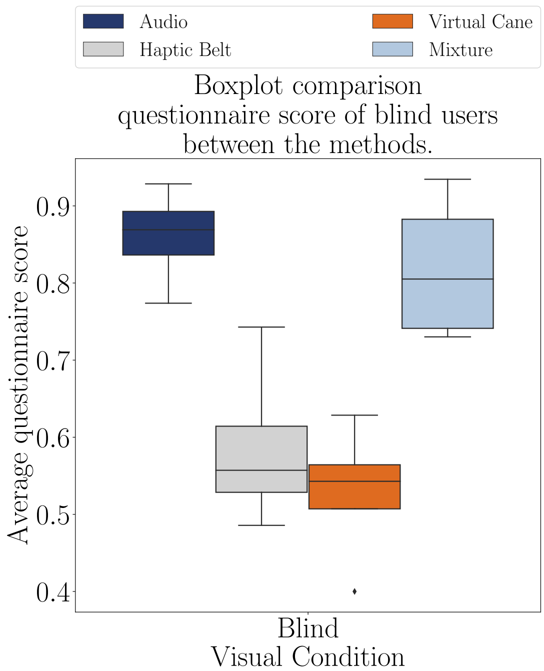


FIGURE 5.17 – Boxplot of the questionaire score of the blind participants grouped by the methods.

In order to complement the ANOVA analysis, the pairwise comparison of the methods obtained from the Fisher LSD test is presented in Table 5.13. The results show that audio and mixture are equivalent from the perspective of user satisfaction. All the other comparisons indicate there is a difference between the methods.

Additional to the scores, the participants also expressed their dissatisfaction with the answers to the open questions of the questionnaire, where they commented that the haptic

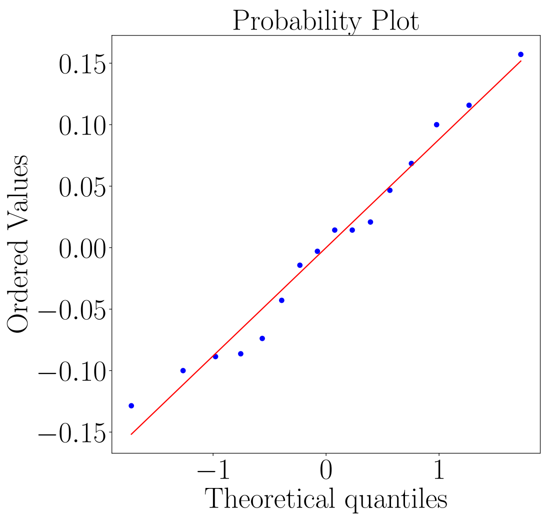


FIGURE 5.18 – QQ plot of the questionnaire score of the blind participants on each method.

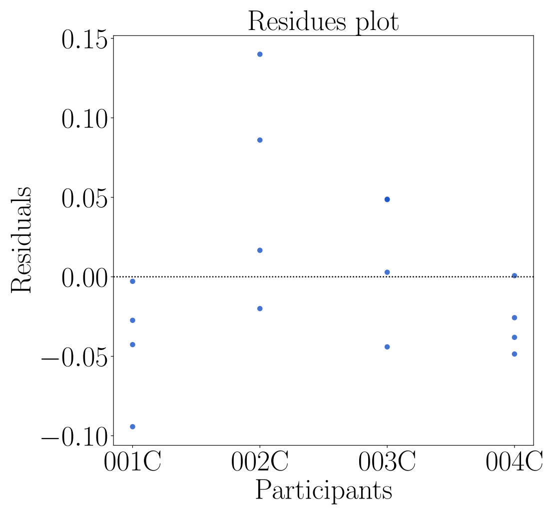


FIGURE 5.19 – Residual plot of the questionnaire score the blind participants on each method.

TABLE 5.13 – Cross validation p-value for the questionnaire score on each method for blinded users.

Method			Analysis	
Audio	X	Haptic Belt	$H_1 : \mu_{Audio} \neq \mu_{HapticBelt}$	**
Audio	X	Virtual Cane	$H_1 : \mu_{Audio} \neq \mu_{VirtualCane}$	**
Audio	X	Mixture	$H_0 : \mu_{Audio} = \mu_{Mixture}$	
Haptic Belt	X	Virtual Cane	$H_1 : \mu_{HapticBelt} \neq \mu_{VirtualCane}$	**
Haptic Belt	X	Mixture	$H_1 : \mu_{HapticBelt} \neq \mu_{Mixture}$	**
Virtual Cane	X	Mixture	$H_1 : \mu_{VirtualCane} \neq \mu_{Mixture}$	**

belt and the virtual cane are confusing, are not precise enough, and are very different from what they are used to.

5.1.2 Physiological data

During the experiment, data from two physiological sensors were captured: ECG and GSR. As commonly found in the literature, these data are used to assess mental workload. The corresponding analysis is presented in this section.

5.1.2.1 Electrocardiogram (ECG) data

As previously stated, the ECG analysis is based on two variables: the heartrate (BPM – beats per minute) and heartrate variance (SDNN – standard deviation of NN intervals).

After the experiment, the ECG signal processing is organized in the following steps (Figure 5.20):

- Filtering and removing outliers. Since the participants moved during the whole experience, the sensors also captured some noise data.
- Normalization between -1 and 1;
- Peak detection and evaluation – if the results were not of good quality, the peak detection method's parameters were adjusted to improve it;
- Calculation of BPM using Kubius HRV Standard;
- Calculation of SDNN using Kubius HRV Standard.

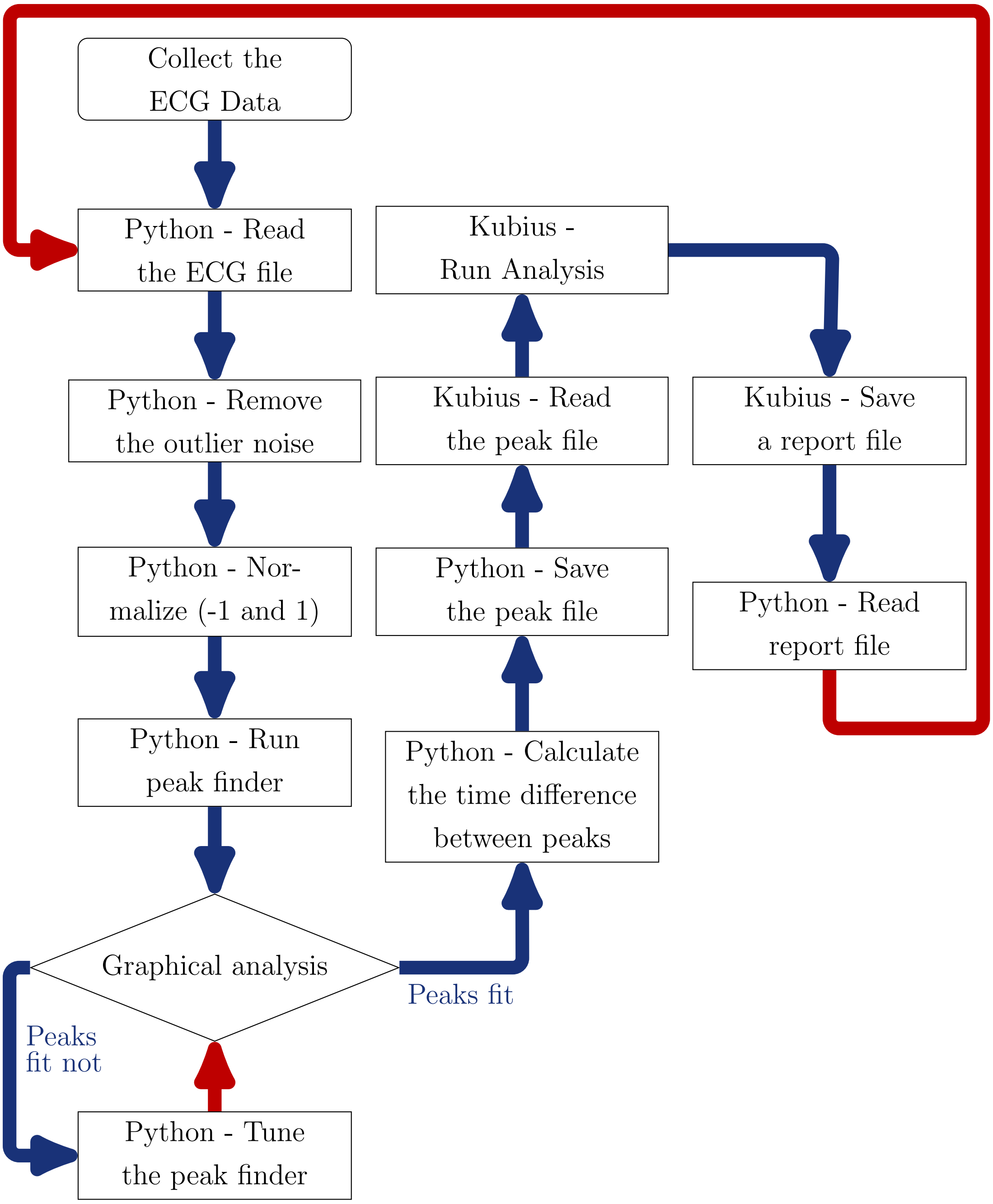


FIGURE 5.20 – ECG's data treatment algorithim.

At the beginning of each experiment, a baseline was collected to establish a comparison between the relaxed state of the participant and the scenes' induced state. However, the results were not consistent. During the experiment, it was expected that the heartrate would increase compared to the baseline because the participants were at rest. However, for most of the participants, it decreased, indicating a systematic problem may have occurred. Due to this fact, the analysis is based only on absolute values.

5.1.2.1.1 Analysis of the heartbeat frequency (BPM)

Table 5.14 presents the heart rate of each blind participant for each guidance method. It is possible to observe that there is no systematic difference between the methods. Also, there are significant differences among the participants, with some presenting values significantly lower than others.

TABLE 5.14 – Average BPM felled by the blinded participants [BPM].

		Base	Audio	Haptic Belt	Virtual Cane	Mixture
Participant	Round					
001C	First	75.75	60.71	71.17	59.07	68.24
	Return	71.05	58.61	66.22	64.20	70.76
002C	First	48.69	38.67	48.74	46.89	52.23
	Return	52.46	47.58	58.97	56.75	58.25
003C	First	68.37	69.89	70.95	69.41	66.94
	Return	67.34	67.44	69.68	68.82	67.37
004C	First	75.09	73.55	73.70	71.94	74.03
	Return	74.74	74.79	74.02	72.69	67.34

Figure 5.21 presents the mean heart rate. It shows a slight increase in the heartrate between the rounds, except for the base method, indicating that the participants felt the return round harder.

Figures 5.22 and 5.23 brings the corresponding boxplot, grouped by method and round. In both cases, it is not possible to observe significant differences among the methods or rounds.

Figures 5.24 and 5.25 bring the QQ Plot and residual distribution. The last one shows that the participants do not have a similar variance, which may jeopardize the results of ANOVA. Considering this limitation, Table 5.14 brings the p-value obtained by ANOVA, which confirmed the previous analysis, as it does not indicate a significant influence of either the guidance methods or the rounds in the participants' heartrate.

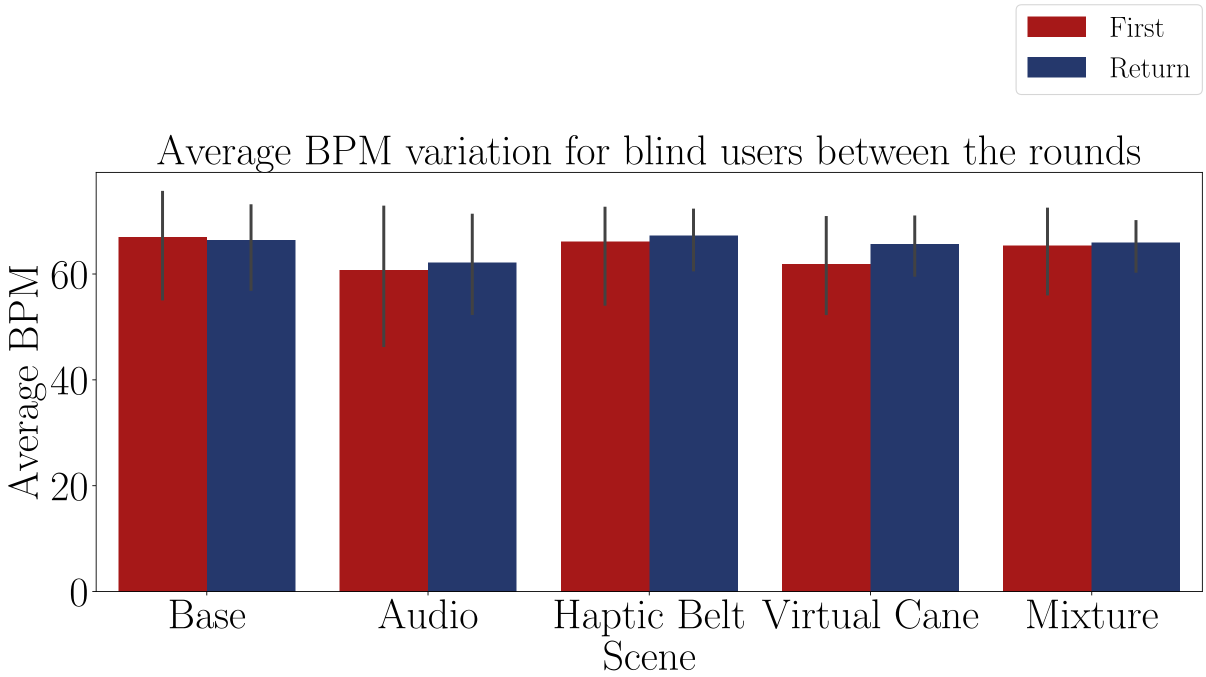


FIGURE 5.21 – Barplot of the average BPM of the blind participants on each method.

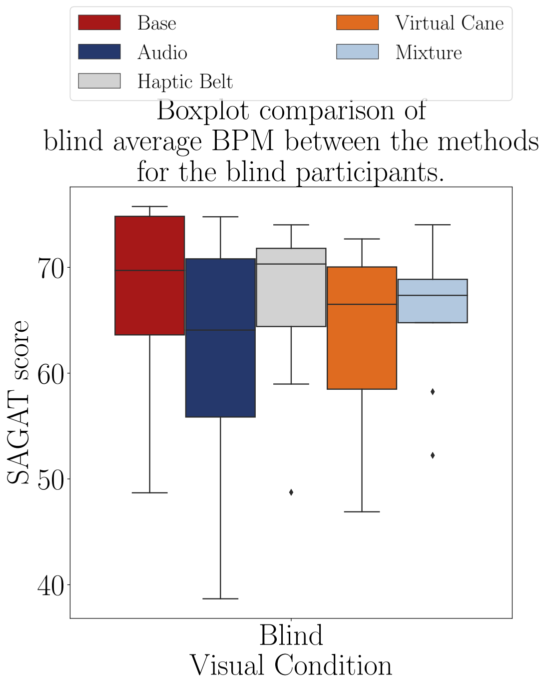


FIGURE 5.22 – Boxplot of the BPM of the blind participants grouped by the methods.

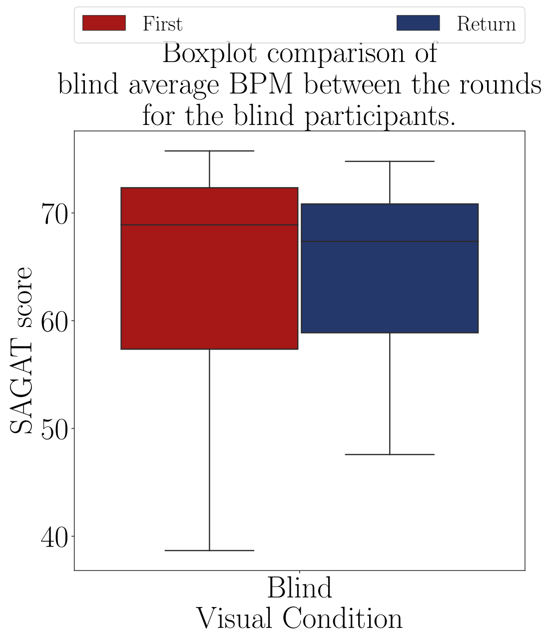


FIGURE 5.23 – Boxplot of the BPM of the blind participants grouped by the rounds.

TABLE 5.15 – Anova p-value for the BPM on each method for blinded users.

Source	P-Value
Methods	0.100
Rounds	0.371
Interaction	0.894

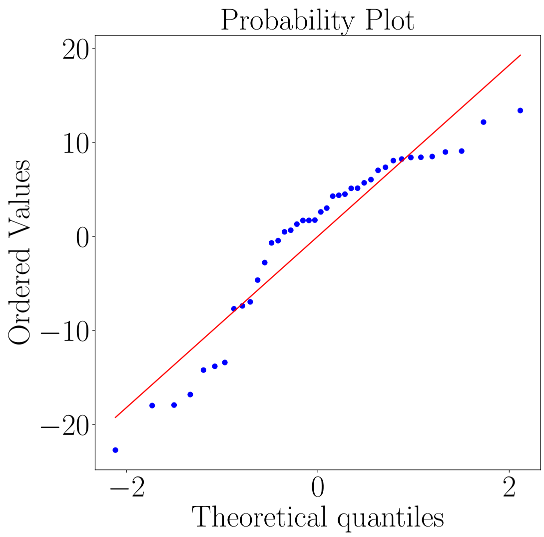


FIGURE 5.24 – QQ plot of the BPM of the blind participants on each method.

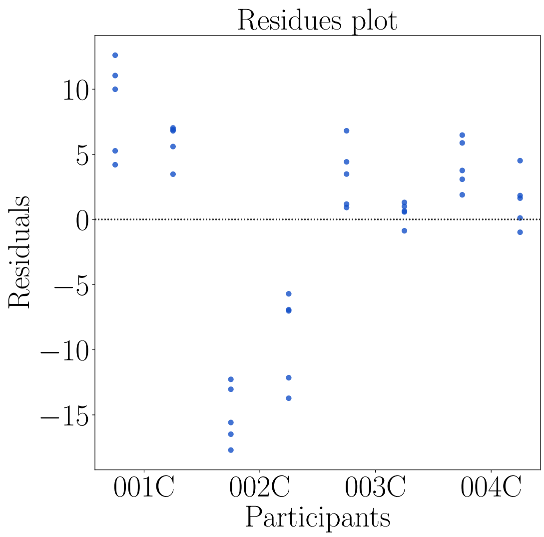


FIGURE 5.25 – Residual plot of the BPM score the blind participants on each method.

5.1.2.1.2 Analysis of the heartbeat variance (SDNN)

Table 5.16 brings the value of the second variable extracted from the ECG: SDNN, the standard deviation of the interbeat interval. Similar to what is observed for the BPM, it is not possible to draw a pattern from this data. The participants had increased or decreased their heartrate with different methods.

TABLE 5.16 – Average SDNN of the blind participants during the each round and method.

Participant	Round	Base	Audio	Haptic Belt	Virtual Cane	Mixture
001C	First	81.292	107.061	124.737	163.968	129.054
	Return	120.719	130.885	131.590	157.589	124.786
002C	First	73.761	98.863	81.140	33.977	79.289
	Return	108.940	49.627	42.815	114.057	107.545
003C	First	36.870	38.325	35.101	42.392	43.692
	Return	52.750	41.196	44.256	42.602	46.145
004C	First	70.728	86.827	62.560	85.900	70.472
	Return	71.950	74.895	70.017	66.089	104.040

The barplot of Figure 5.26 shows the average SDNN for each method. It is possible to notice that some methods are associated with an increase in the SDNN between the rounds, while others present a slight decrease.

Figure 5.27 and Figure 5.28 bring the SDNN barplot grouped by the methods and the rounds. There is a slight tendency among the participants to increase the heartbeat in the return round.

Figures 5.29 and 5.30 bring the QQ Plot and residual distribution. In this case,

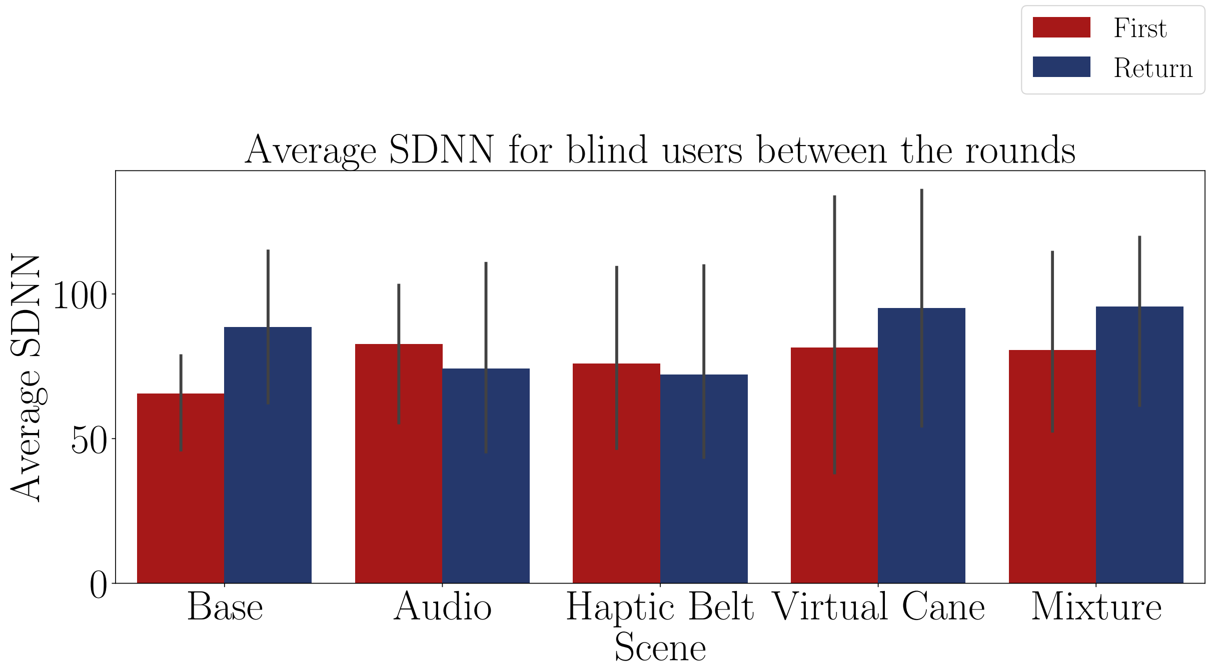


FIGURE 5.26 – Barplot of the average SDNN of the blind participants on each method.

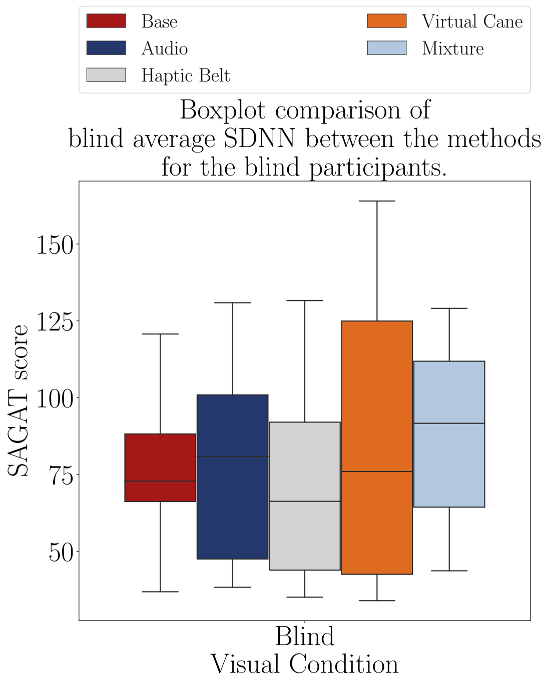


FIGURE 5.27 – Boxplot of the SDNN of the blind participants grouped by the methods.

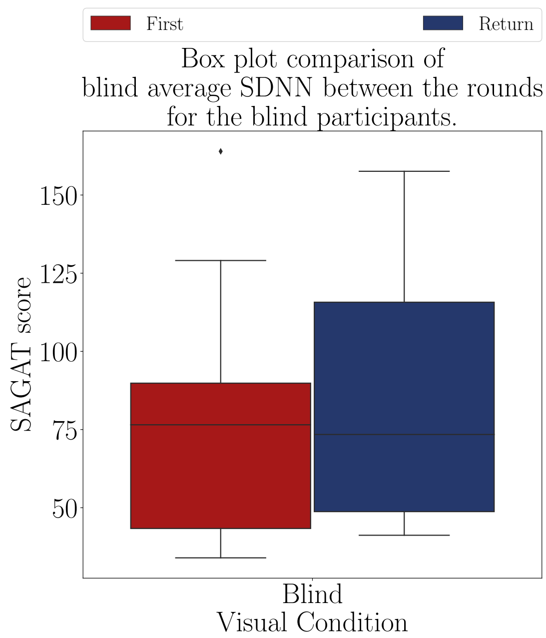


FIGURE 5.28 – Boxplot of the SDNN of the blind participants grouped by the rounds.

the residual distribution is more uniform than in Figure 5.25. The ANOVA results are presented in Table 5.17 and do not confirm any influence of the methods nor the rounds on the ECG heartrate variance.

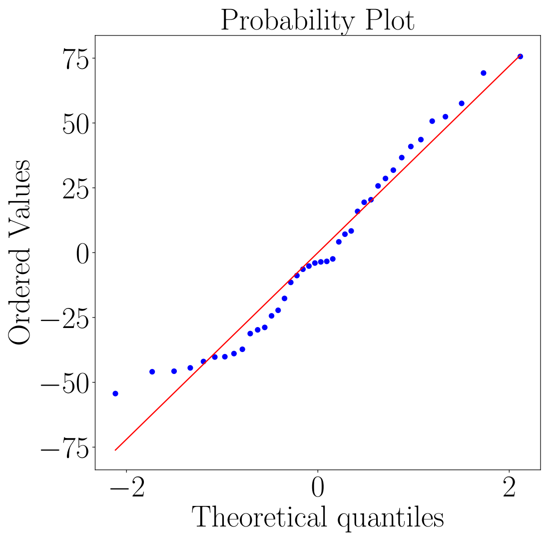


FIGURE 5.29 – QQ plot of the SDNN of the blind participants on each method.

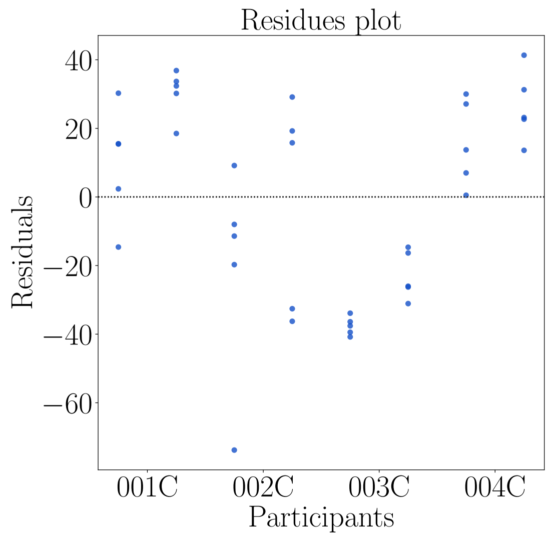


FIGURE 5.30 – Residual plot of the SDNN of the blind participants on each method.

TABLE 5.17 – Anova p-value for the average SDNN on each method for blinded users.

Source	P-Value
Methods	0.486
Rounds	0.223
Interaction	0.473

5.1.2.2 Galvanic skin reaction and temperature data;

The GSR analysis is based on the signal's average level. Each experiment's round is compared to the participant baseline collected before the experiment. The GSR sensor was worn on the left hand for right-handed participant and on the right hand for left-handed participants. One of the blind participants had the GSR sensor removed during the experiment because it was not appropriately fixed.

Table 5.18 presents the GSR average values for the three remaining participants. For all the participants, the baseline was smaller than the values obtained during the experiment, as expected. Moreover, in most cases, the skin conductance has risen from the first to the return, indicating an increase in the mental workload.

Table 5.19 brings the percentual increase in the GSR average compared to the baseline value. Figure 5.31 shows the corresponding barplot. The presence of a haptic device causes an increase in the skin conductance, hence its mental workload. Also, it is possible to observe the increase in GSR average between the two rounds, except for the haptic belt.

Figure 5.32 presents the boxplot of the percentual variation in the skin conductance for each method. The base method has the lowest variation among all methods. Also, the introduction of vibration increases the method variance. Figure 5.33 presents the GSR

TABLE 5.18 – Average GSR felled by the blind participants [μS].

Participant	Round	Baseline	Base	Audio	Haptic Belt	Virtual Cane	Mixture
001C	First	0.37	0.48	1.03	3.14	3.79	3.90
	Return		0.83	1.58	2.81	4.04	4.57
003C	First	0.30	0.56	0.56	0.62	0.85	1.09
	Return		0.62	0.63	0.65	0.92	1.06
004C	First	1.24	2.34	3.07	3.49	2.28	2.23
	Return		2.57	2.95	3.20	2.21	2.24

TABLE 5.19 – Average GSR variation in relation to the baseline in each round of the blind participants [μS].

Participant	Round	Base	Audio	Haptic Belt	Virtual Cane	Mixture
001C	First	30.58%	176.54%	746.10%	920.72%	951.71%
	Return	125.29%	327.42%	656.99%	988.93%	1132.39%
003C	First	85.36%	84.23%	104.19%	182.35%	258.80%
	Return	105.34%	109.23%	112.95%	202.35%	249.72%
004C	First	89.62%	148.53%	182.84%	84.33%	80.69%
	Return	108.22%	138.64%	159.00%	78.73%	81.61%

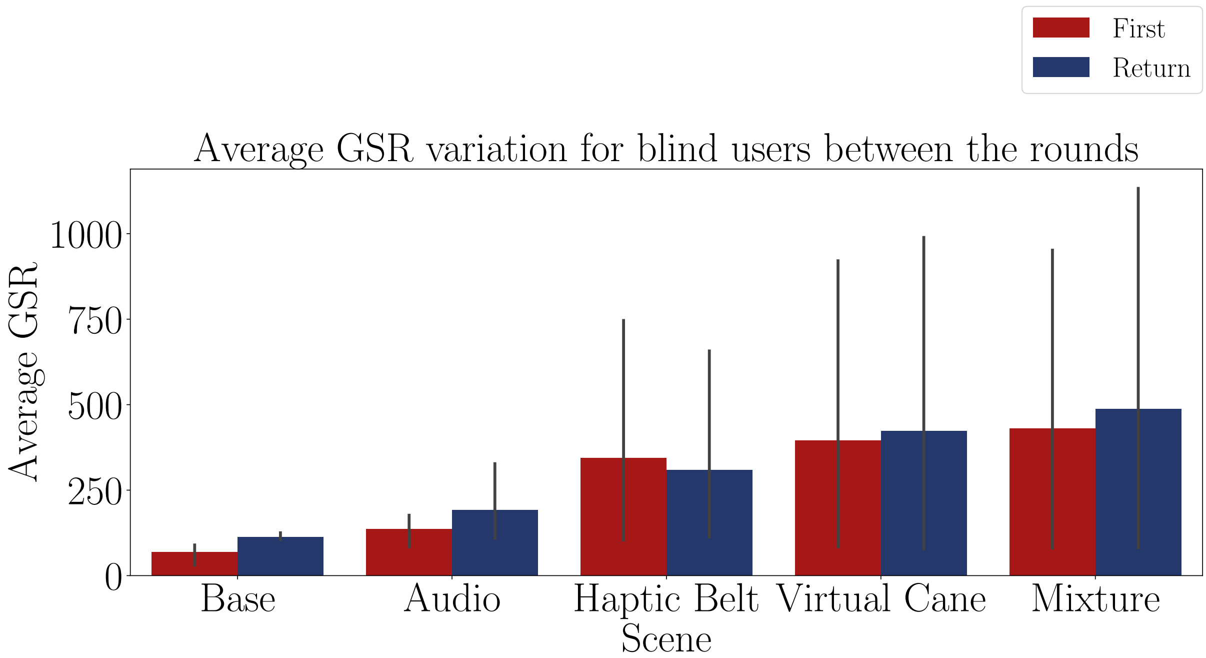


FIGURE 5.31 – Barplot of the average SDNN of the blind participants on each method.

grouped by the rounds. In this case, there is no apparent difference between the rounds.

Figures 5.34 and 5.35 shows the QQ plot and the residual distribution. Table 5.20

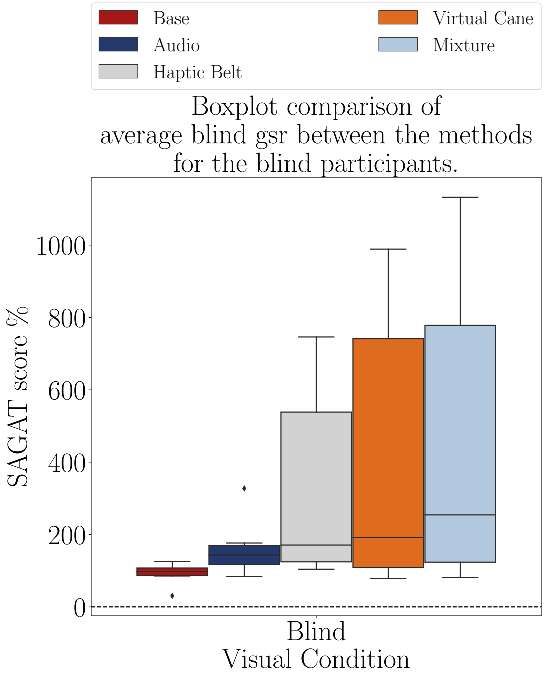


FIGURE 5.32 – Boxplot of the GSR of the blind participants grouped by the methods.

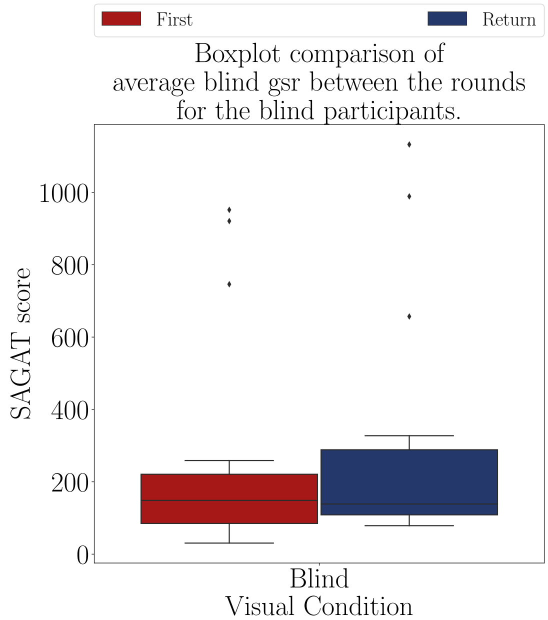


FIGURE 5.33 – Boxplot of the GSR of the blind participants grouped by the rounds.

shows the ANOVA test p-value for the GSR percentual variance. Although the p-value for the method is not below the threshold of 0.05, it is close to it, indicating that probably the GSR is affected by it.

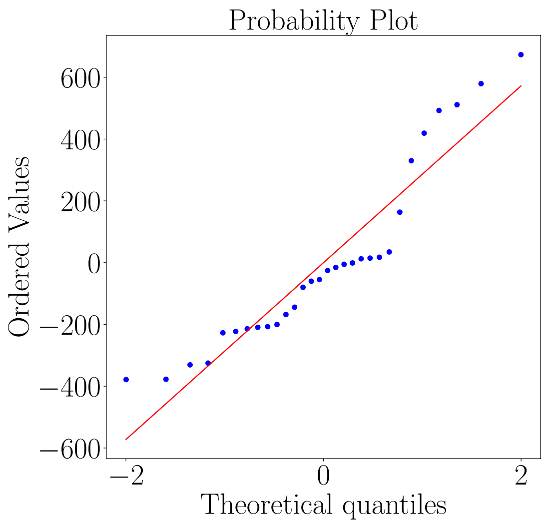


FIGURE 5.34 – QQ plot of the SDNN of the blind participants on each method.

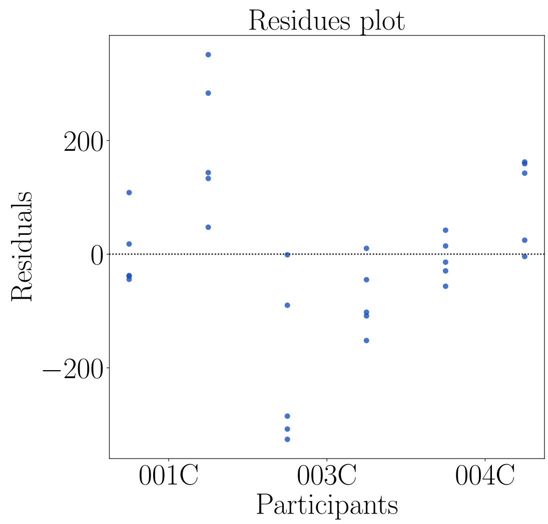


FIGURE 5.35 – Residual plot of the SDNN of the blind participants on each method.

TABLE 5.20 – Anova p-value for the mental demand average on each method for blinded users.

Source	P-Value
Methods	0.051
Rounds	0.722
Interaction	0.996

5.1.3 Final Remarks

To summarize the conclusion obtained from the analysis of the data from blind participants, the audio method showed a lower score both for NASA-TLX mental demand and NASA-TLX global score. In contrast, the methods that include vibration achieved higher scores. This probably happened because the participants are already used to using sound to guide themselves, especially environmental sounds. The environment sounds used in the scenes were always the same (telephone ringing, laptop keyboard sounds, exterior noise, door opening and closing). The participants likely felt more relaxed when they only had to focus on the sounds around him/her. This is reinforced by the fact that, during the experiment with the audio method, half of the participants did not ask for any information, or the audio command option was used only a few times.

The fact that the haptic devices caused a higher workload is probably due to the fact that the users had to learn and get used to them. Besides, for being just conceptual, their precision was not as good as they were expecting. That explains why their results were not as good as the base or audio methods. The NASA-TLX results are correctly related to the satisfaction questionnaires, which scored them as the unsatisfied devices.

As expected, most of the variables from subjective questionnaires (NASA-TLX and SAGAT) show some influence of the rounds. On the other hand, the results from the physiological sensors did not show a clear tendency.

The statistical analysis based on ANOVA tests confirmed some of the observations from the bar and box plots. However, in many cases, the residual distributions were not homogenous and the statistical analysis was affected by the small number of samples.

All the blind participants showed great enthusiasm before, during and after the experiment. They also made several recommendations for both the virtual environment and the devices, such as:

- The speakers of the HMD are not good enough to give them the precise location of the sound origin
- The HMD is too large and covers half of the participant's face. It gives them a strange sensation, since some of them use the air or the wind feeling on the face to

give them hints about the location of walls or other high obstacles;

- The precision of the vibration for both the haptic belt and the virtual cane needs to be improved. It is not enough for them to use the devices. This problem is related to how the HMD sets the position of the user in the virtual environment.

- The vibration from the haptic belt was not intense enough.

5.2 Comparison between BVI users and sighted users

This section investigates the second research question of this work: "do non-BVI users, when deprived of their vision, similarly evaluate assistive devices as BVI users?".

To do so, the analysis performed in the previous section is now repeated with the data obtained from sighted participants. However, the data corresponding to the "base" method is omitted, as the daily method used by sighted people is based on their vision.

5.2.1 Subjective data

5.2.1.1 NASA-TLX

5.2.1.1.1 Analysis of the mental demand scale

Table 5.21 presents the mental demand score of all participants, while the corresponding barplot is presented in Figure 5.36. It is interesting to observe that sighted people gave a higher score to audio, as they are not so familiar with using sounds as source of guidance.

Figures 5.37 and 5.38 presents the box plot for both groups, organized by the methods and the rounds. The mental demand is systematically higher for sighted people, which is expected. However, while blind participants considered the audio method less demanding, sighted participants prefered to the virtual cane. For both groups, we observe a decrease in the mental demand.

Figures 5.39 and 5.40 show the QQ plot and residual distribution for the sighted data, confirming that the data is normally distributed and participants have similar variance. Table 5.22 brings the results of ANOVA. Unlike the blind participants, in the case of sighted ones, the p-value for the methods is below the threshold of 0.05, confirming it as a significant variable for the mental demand. In the case of the rounds, the data from

TABLE 5.21 – Mental demand felled by the participants.

Participant	Visual Condition	Round	Audio	Haptic Belt	Virtual Cane	Mixture
001	Sight	First	12	11	5	9
		Return	13	13	5	10
001C	Blind	First	1	14	3	6
		Return	1	10	2	6
002C	Blind	First	1	1	10	12
		Return	1	1	10	3
003	Sight	First	18	18	16	10
		Return	12	15	11	8
003C	Blind	First	5	5	8	1
		Return	1	1	2	1
004	Sight	First	17	20	12	20
		Return	12	15	10	15
004C	Blind	First	10	15	10	10
		Return	10	14	8	10
005	Sight	First	4	12	10	13
		Return	6	10	6	12

both sighted and blind participants resulted in the exact p-value of 0.075, which is close to the traditional threshold of 0.05 but slightly higher.

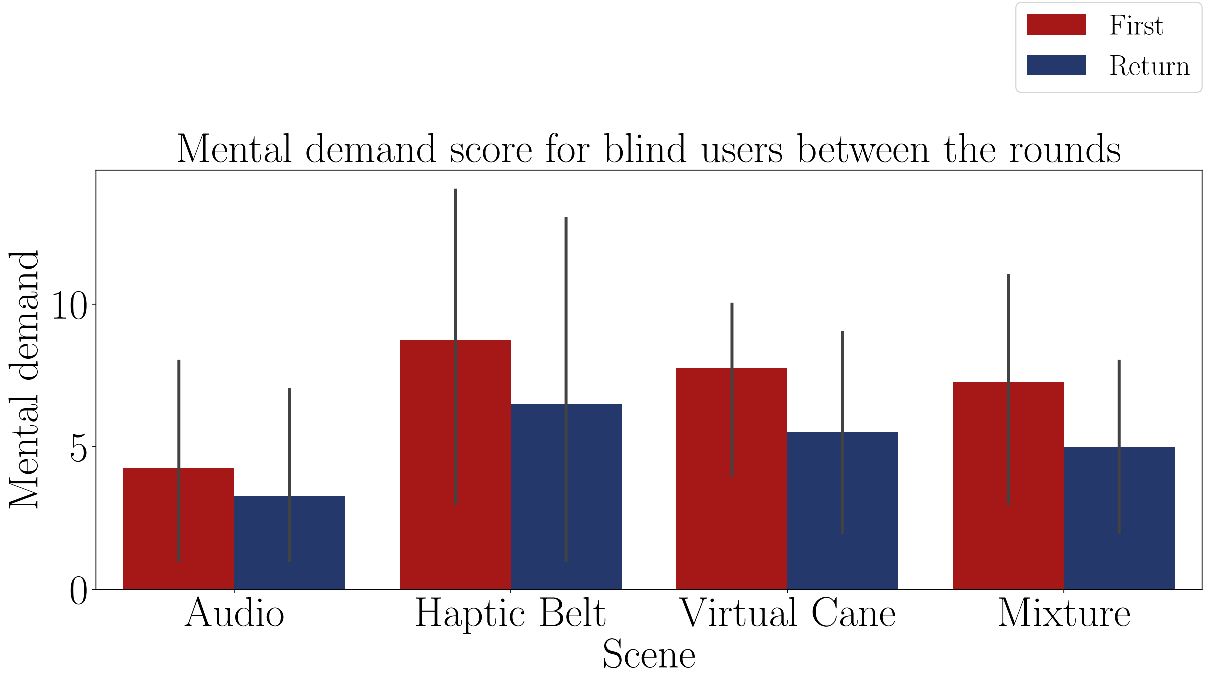
TABLE 5.22 – Anova p-value for the mental demand average on each method'

(a) Blind participants

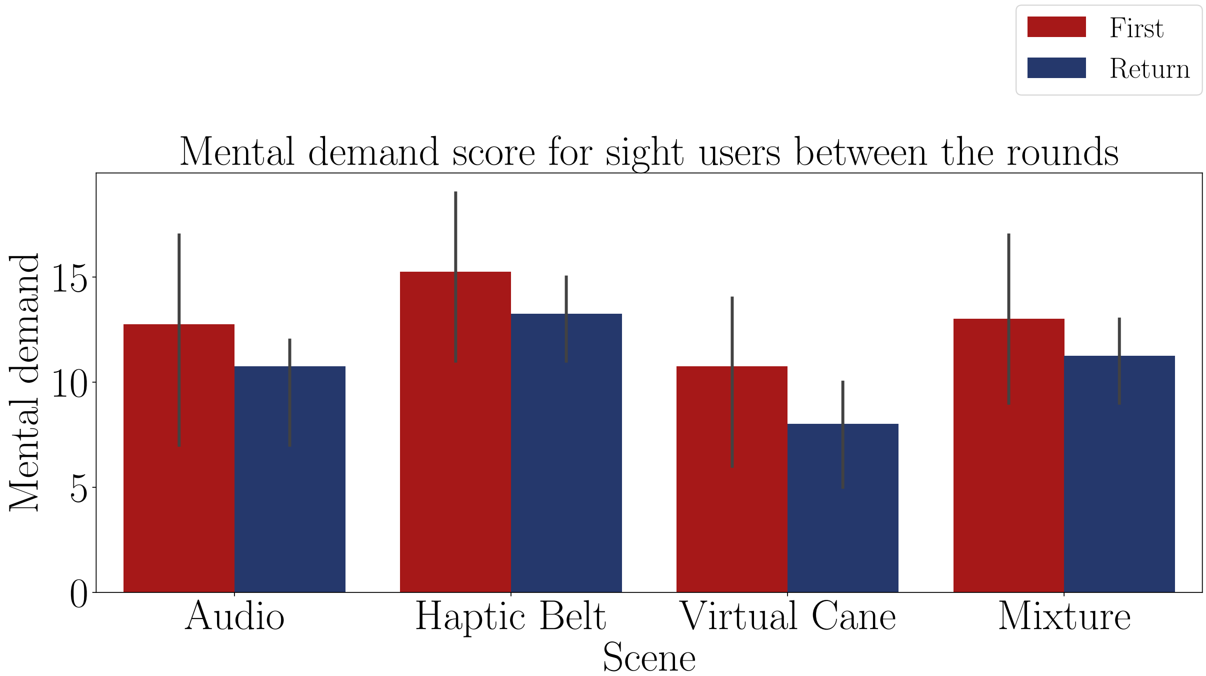
Source	P-Value
Methods	0.170
Rounds	0.075
Interaction	0.993

(b) Sight participants

Source	P-Value
Methods	0.049**
Rounds	0.075
Interaction	0.990



(a) Blind participants



(b) Sight participants

FIGURE 5.36 – Barplot of the average mental demand on each method and each round.

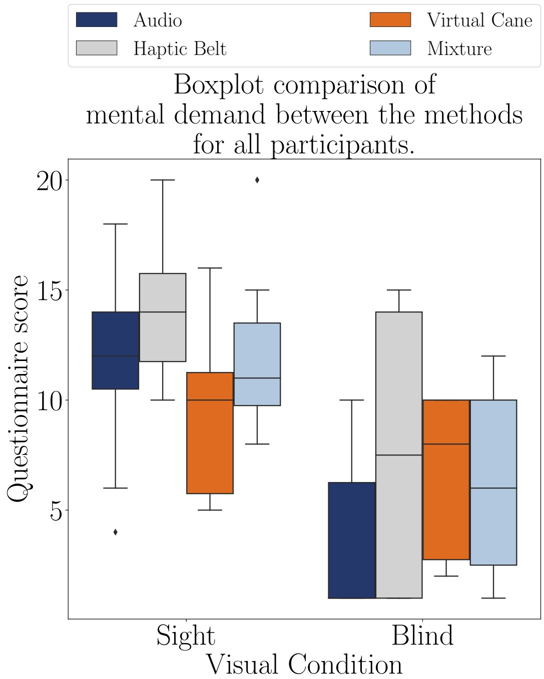


FIGURE 5.37 – Boxplot of the mental demand of the participants grouped by the methods.

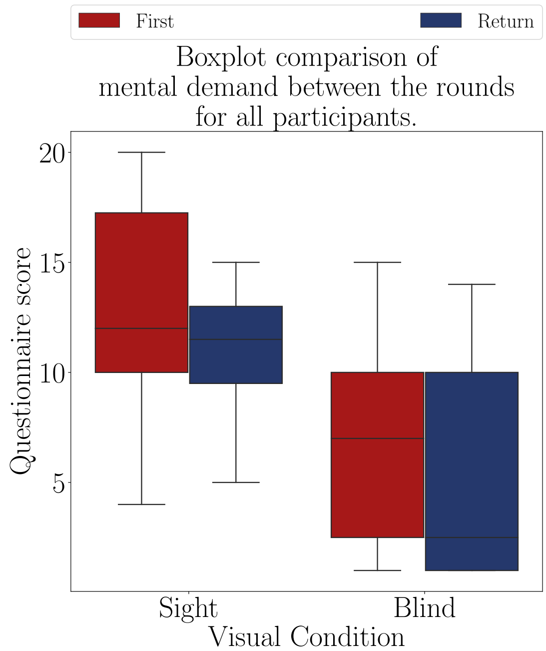


FIGURE 5.38 – Boxplot of the mental demand of the participants grouped by the rounds.

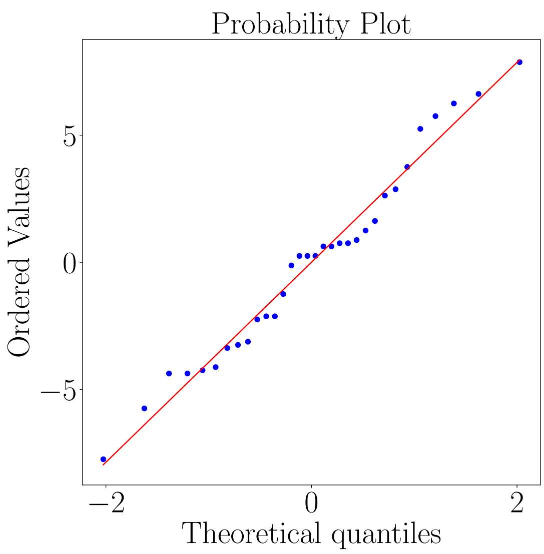


FIGURE 5.39 – QQ plot of the mental demand of the sight participants on each method.

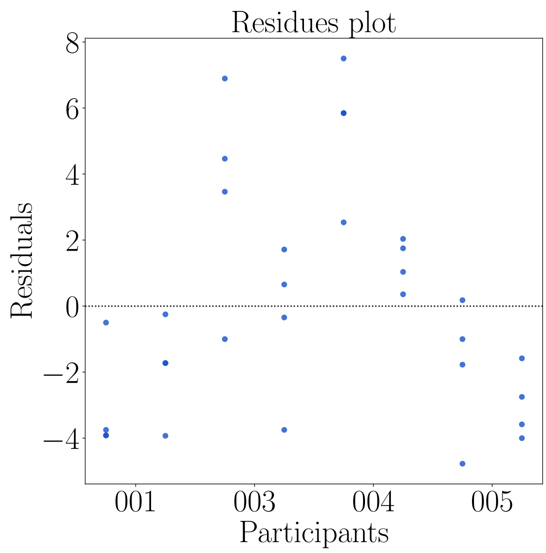


FIGURE 5.40 – Residual plot of the mental demand score the sighted participants on each method.

5.2.1.1.2 Analysis of the NASA-TLX score

Table 5.23 brings the NASA-TLX global score of all participants, while the corresponding barplot is presented in Figure 5.41.

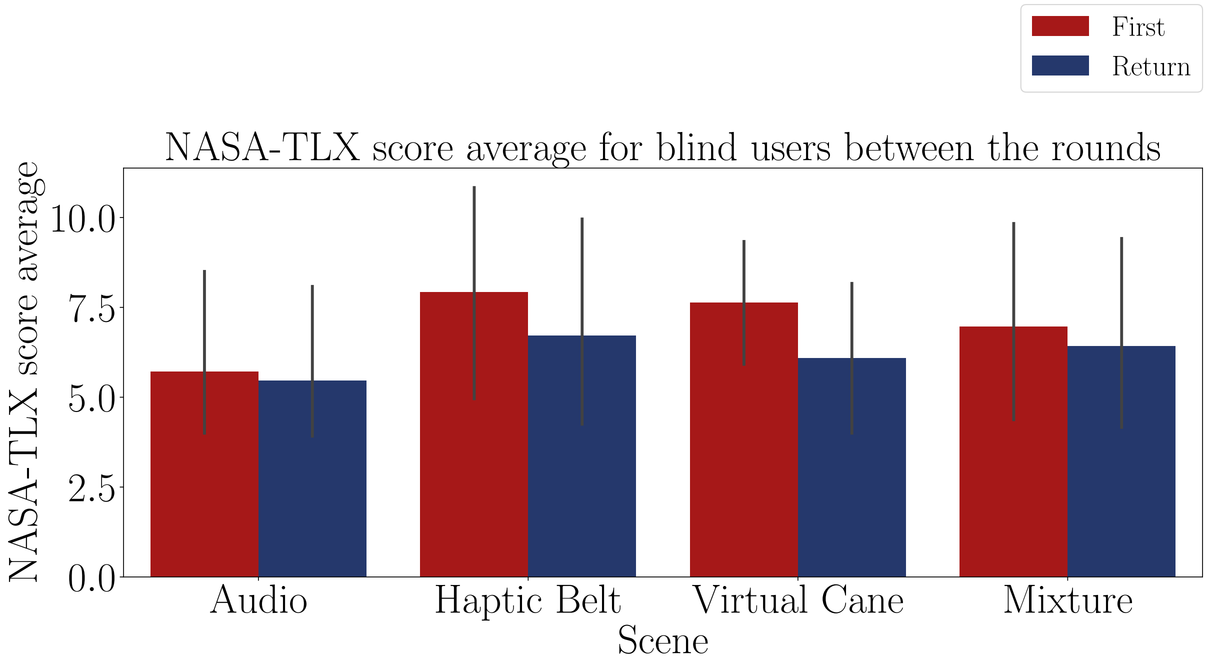
TABLE 5.23 – NASA-TLX score felled by the participants.

Participant	Visual Condition	Round	Audio	Haptic Belt	Virtual Cane	Mixture
001	Sight	First	10.167	9.833	7.000	9.000
		Return	11.000	10.833	6.167	9.333
001C	Blind	First	4.000	8.833	5.167	6.333
		Return	4.000	6.667	4.500	6.167
002C	Blind	First	4.833	4.833	9.000	7.000
		Return	4.833	4.833	7.000	5.167
003	Sight	First	9.833	10.167	9.500	6.500
		Return	6.667	9.667	7.833	4.833
003C	Blind	First	4.000	5.333	6.667	3.500
		Return	3.833	3.667	3.500	3.500
004	Sight	First	14.833	13.667	11.500	15.833
		Return	11.833	11.833	10.833	12.167
004C	Blind	First	10.000	12.667	9.667	11.000
		Return	9.167	11.667	9.333	10.833
005	Sight	First	7.667	9.000	8.000	9.667
		Return	7.667	8.667	7.667	6.000

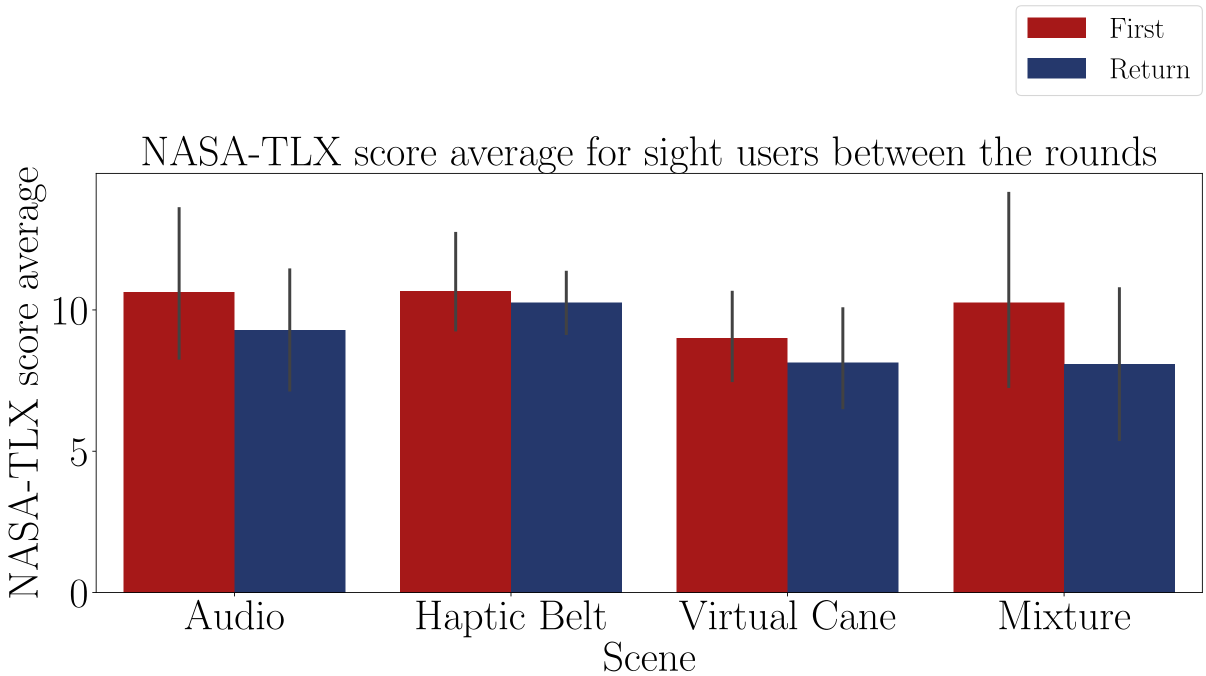
From Figure 5.41, it is possible to see that, similar to blind participants, sighted participants consider that the workload of the return round was lower than that of the first round. However, similar to what happened for the mental demand, sighted participants considered virtual cane as the methods with the lowest workload, while, for blind participants, it was the audio.

Figures 5.42 and 5.43 present the boxplots of the NASA-TLX global score. Again, it is possible to see that sighted people usually give higher workload scores than blind ones. The influence of the round is approximately the same. However, the order of preference of the methods is different.

Figures 5.44 and 5.45 bring the QQ plot and residual distribution of the data from sighted participants, showing that ANOVA can be used. The p-values for both groups are presented in Table 5.24. It confirms the influence of the round for both sighted and blind people. In the case of the methods, the p-value of blind is lower than the threshold of 0.5, while that of sighted is slightly higher.



(a) Blind participants.



(b) Sight participants.

FIGURE 5.41 – Barplot of the NASA-TLX score on each method and each round.

TABLE 5.24 – Anova p-value for the NASA-TLX score on each method

(a) Blind participants

Source	P-Value
Methods	0.029**
Rounds	0.022**
Interaction	0.814

(b) Sight participants

Source	P-Value
Methods	0.086
Rounds	0.034**
Interaction	0.688

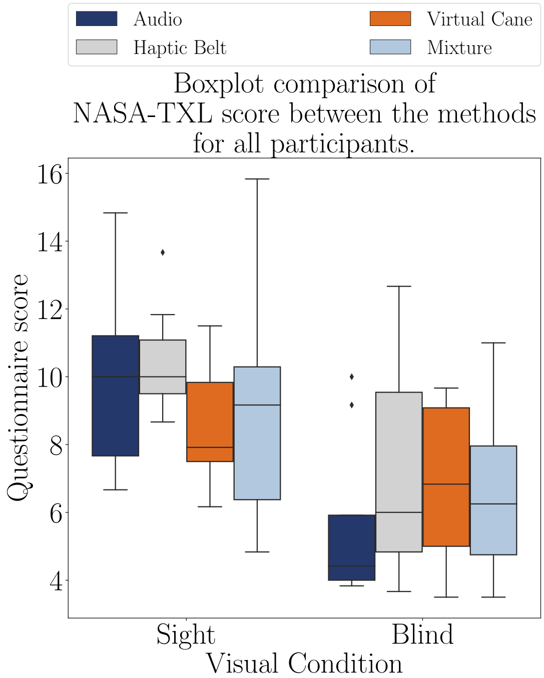


FIGURE 5.42 – Boxplot of the NASA-TLX score of the participants grouped by the methods.

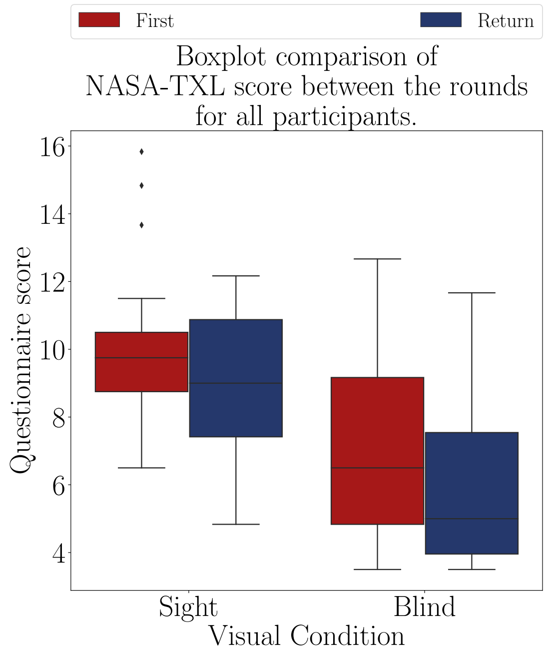


FIGURE 5.43 – Boxplot of the NASA-TLX score of the participants grouped by the rounds.

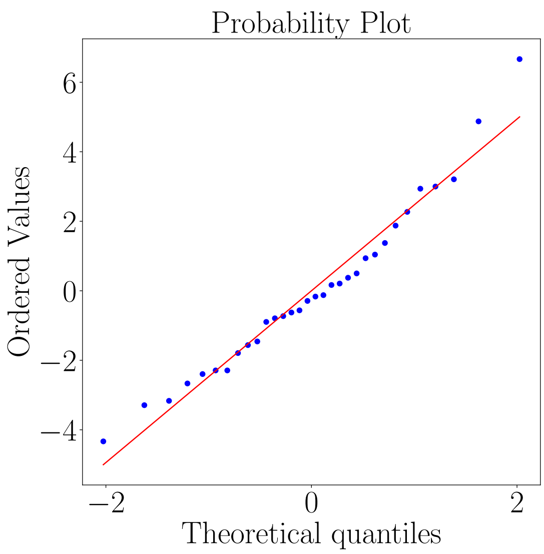


FIGURE 5.44 – QQ plot of the NASA-TLX score of the sight participants on each method.

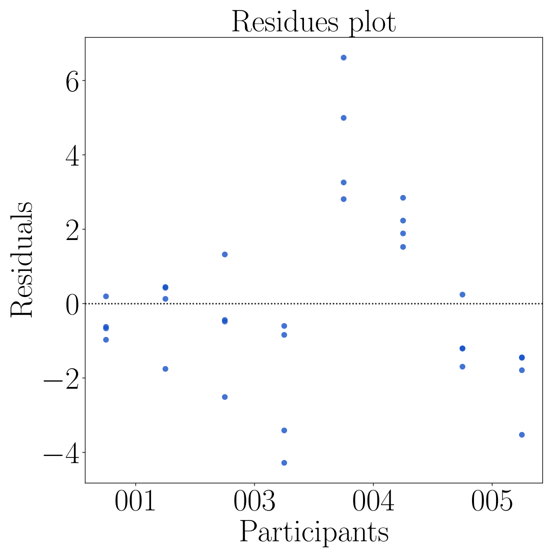


FIGURE 5.45 – Residual plot of the NASA-TLX score the sight participants on each method.

5.2.1.2 Adapted SAGAT

Table 5.25 presents the SAGAT score of all participants. The corresponding barplot is presented in Figure 5.46.

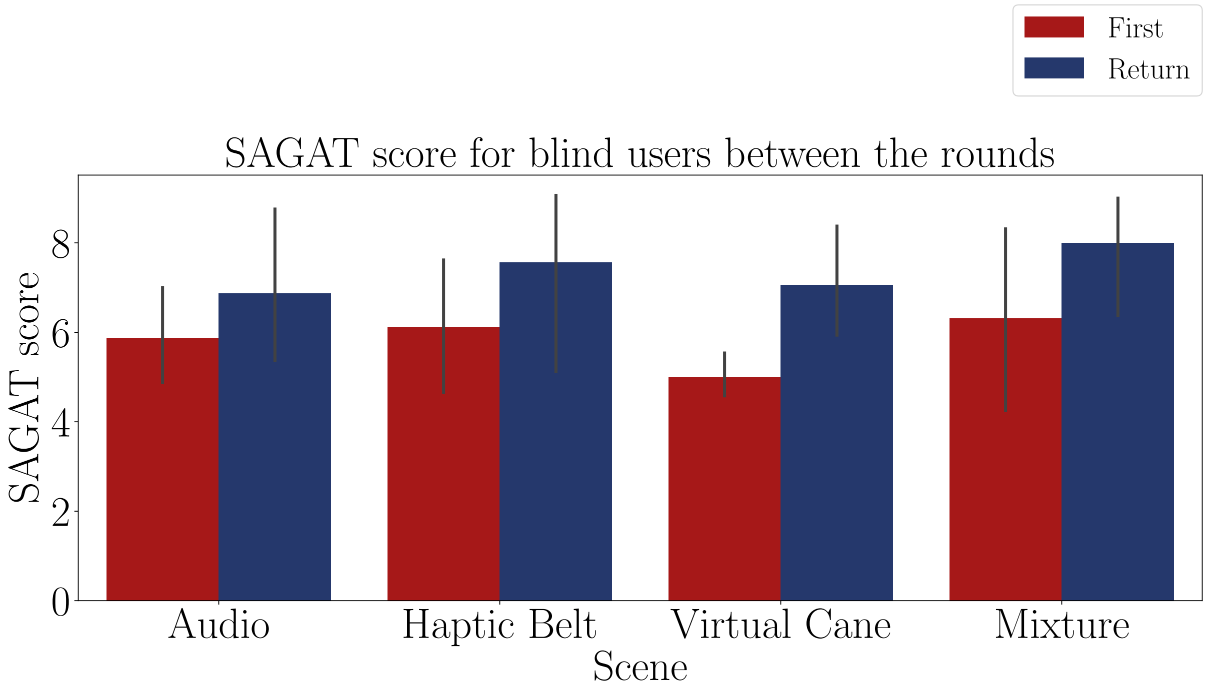
TABLE 5.25 – SAGAT global score felled by the participants.

Participant	Visual Condition	Round	Audio	Haptic Belt	Virtual Cane	Mixture
001	Sight	First	4.500	4.330	2.660	6.500
		Return	6.000	5.000	5.000	4.500
001C	Blind	First	5.500	5.330	5.830	3.500
		Return	6.500	8.500	5.500	5.500
002C	Blind	First	4.500	3.990	4.500	6.250
		Return	5.000	4.000	6.500	8.500
003	Sight	First	6.750	5.990	3.990	6.750
		Return	6.000	7.250	6.250	7.500
003C	Blind	First	7.500	7.490	4.660	9.000
		Return	10.000	8.500	9.000	9.000
004	Sight	First	7.250	7.990	5.990	8.250
		Return	7.750	9.500	8.250	7.000
004C	Blind	First	6.000	7.660	4.990	6.500
		Return	6.000	9.250	7.250	9.000
005	Sight	First	3.000	3.160	3.990	4.000
		Return	3.750	3.000	2.000	6.000

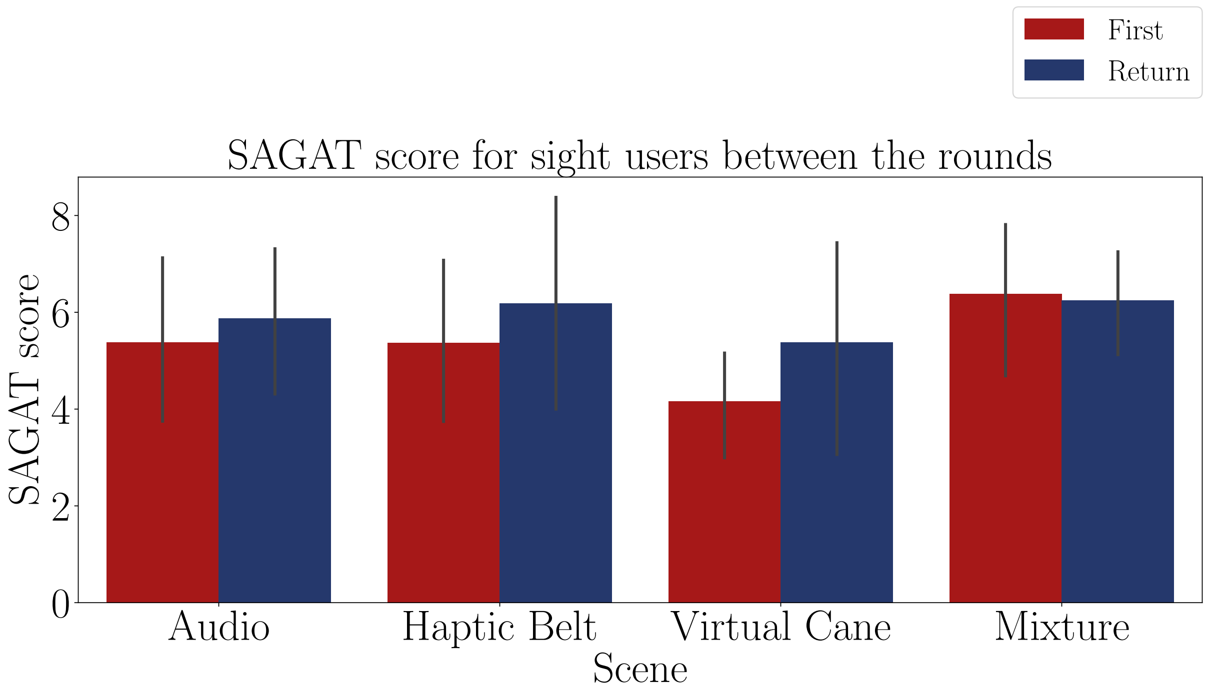
Figure 5.46. shows that the SAGAT score for sighted participants is, on average lower than that of blind participants, which is expected as they are not used to navigating without vision. Also, the increase in situation awareness from the first to the return round is lower. In the case of the mixture method, the SAGAT score did not improve at all. For both groups, the virtual cane was the method with the lowest score in the first round.

Figures 5.47 and 5.48 bring the boxplots. According to Figure 5.47, both groups presented a higher situation awareness with 'mixture' and 'haptic'. On the other hand, Figure 5.48 confirms that the difference between the rounds is more significant for blind participants.

Figures 5.49 and 5.50 brings the QQ plot and residual distribution. The variance of the residuals is not equal among the participants. Table 5.26 brings the p-value from ANOVA. While for the blind participants, the rounds are a significant factor and the methods are not, for the sighted participants the result is the opposite, showing a significant influence of the methods and not of the rounds.



(a) Blind participants.



(b) Sight participants.

FIGURE 5.46 – Barplot of the SAGAT score on each method and each round.

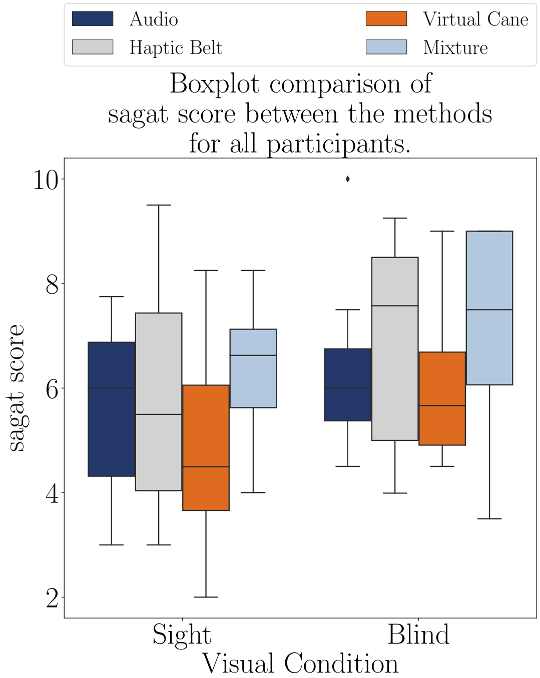


FIGURE 5.47 – Boxplot of the Sagat score of the participants grouped by the methods.

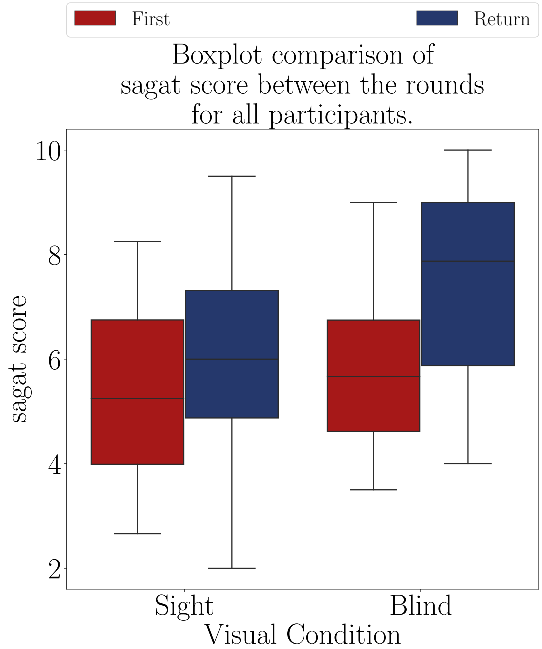


FIGURE 5.48 – Boxplot of the Sagat score of the participants grouped by the rounds.

TABLE 5.26 – Anova p-value for the SAGAT score on each method

(a) Blind participants

Source	P-Value
Methods	0.277
Rounds	0.002**
Interaction	0.834

(b) Sight participants

Source	P-Value
Methods	0.035**
Rounds	0.095
Interaction	0.578

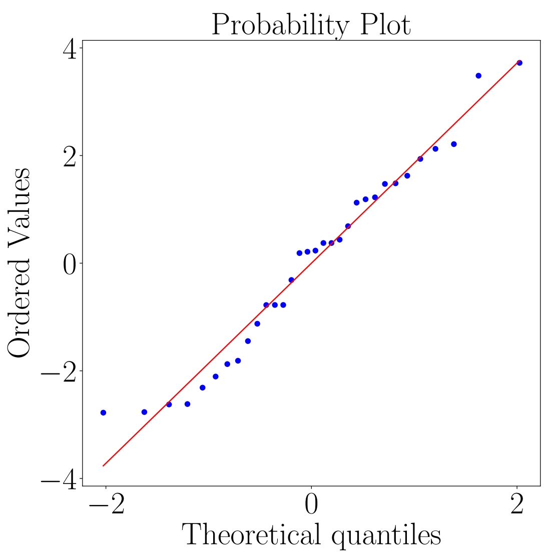


FIGURE 5.49 – QQ plot of the mental demand of the sight participants on each method.

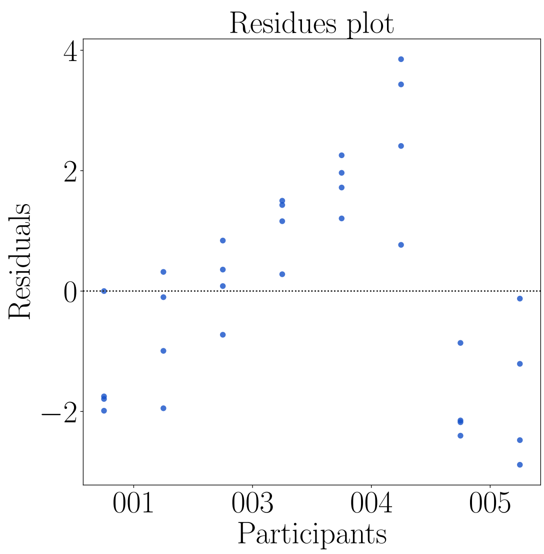


FIGURE 5.50 – Residual plot of the mental demand score the sight participants on each method.

5.2.1.3 Guidance method's questionnaire.

As for the blind users, the sighted user also answered the Guidance questionnaire to give their thoughts about their experience with the guidance methods. Table 5.27 shows the score of both groups. The corresponding barplots are presented in Figure 5.51. Both groups prefer audio and mixture methods. The difference lies in the preference between the haptic belt and virtual cane. The blind users tend to prefer the first one, while the sighted users tend to prefer the last.

TABLE 5.27 – Guidance method questionnaire score grouped by participant.

Participant	Audio	Haptic Belt	Virtual Cane	Mixture	Visual Condition
001	0.75	0.49	0.57	0.69	Sight
001C	0.77	0.54	0.63	0.87	Blind
002C	0.86	0.74	0.54	0.93	Blind
003	0.76	0.54	0.54	0.78	Sight
003C	0.93	0.57	0.54	0.74	Blind
004	0.86	0.60	0.79	0.76	Sight
004C	0.88	0.49	0.40	0.73	Blind
005	0.61	0.57	0.75	0.84	Sight

The Figure 5.52 presents the box plot with the distribution of the scores. It is possible to see that there is some similarity between the two groups, except for the virtual cane method, which has a broader distribution for the sighted users. Also, it seems that the audio and mixture have similar acceptance for sighted and blind users.

Figures 5.53 and 5.54 brings the QQ plot and residual distribution, which confirm that ANOVA can be applied. The result of ANOVA is presented in Table 5.28 and indicates that the method is an effective variable for the sighted participants, as it is for the blind ones.

TABLE 5.28 – Anova p-value for the questionnaire score on each method

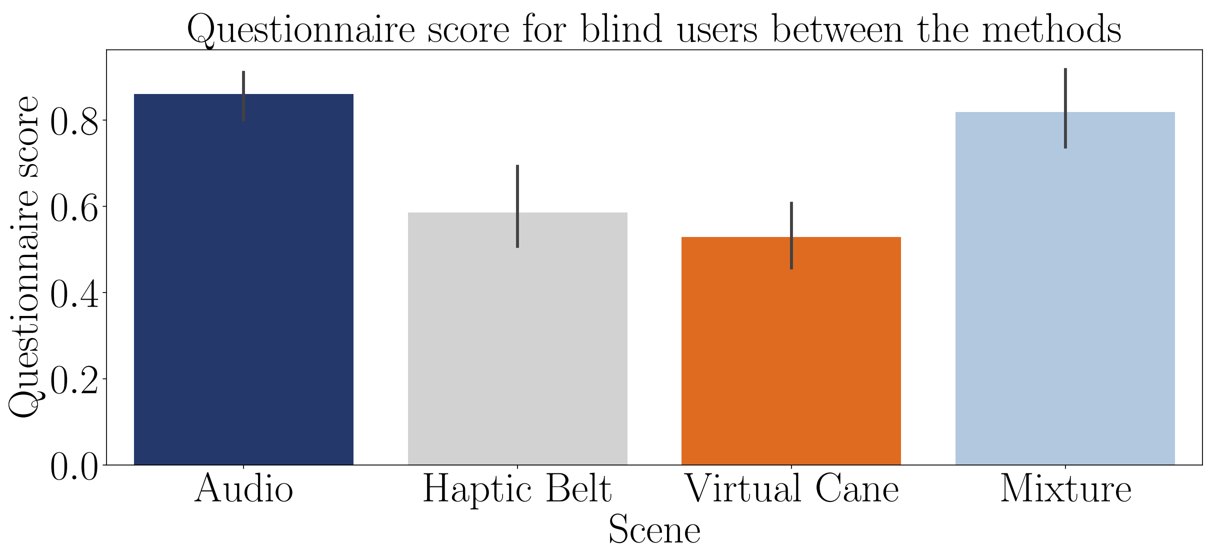
(a) Blind participants.

Source	P-Value
Method	0.001**

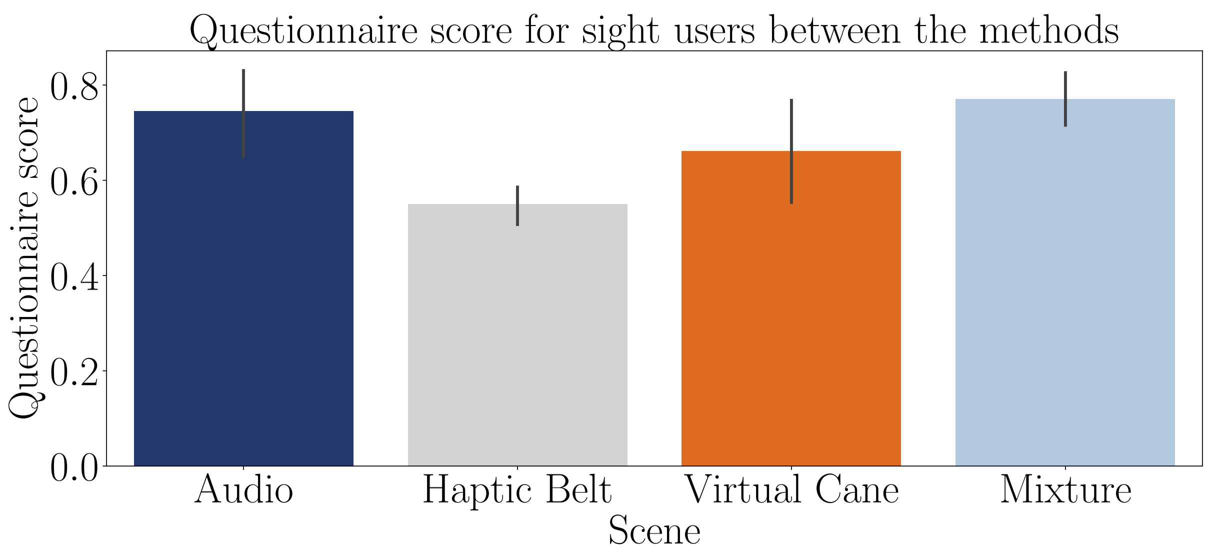
(b) Sight participants.

Source	P-Value
Method	0.016**

Table 5.29 presents the conclusion of a pairwise Fisher LSD test between all the guidance methods for both groups, showing that the results are coincident.



(a) Blind participants.



(b) Sighted participants.

FIGURE 5.51 – Barplot of the average questionaire score on each method.

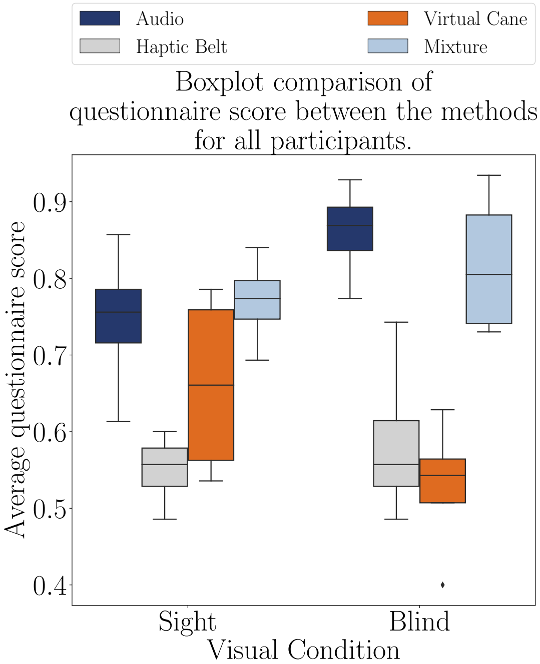


FIGURE 5.52 – Boxplot of the questionaire score of the the participants grouped by the methods.

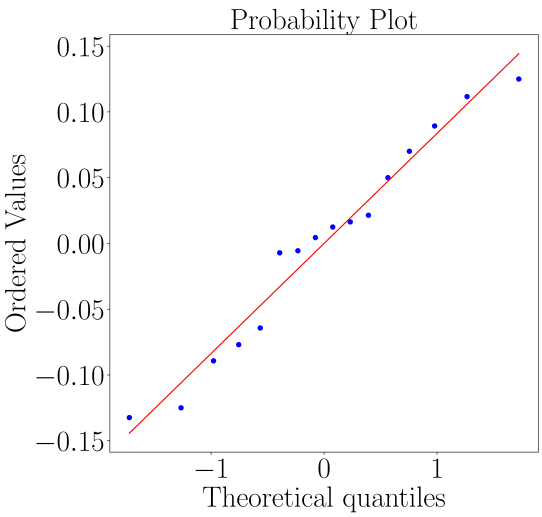


FIGURE 5.53 – QQ plot of the questionnaire score of the sighted participants on each method.

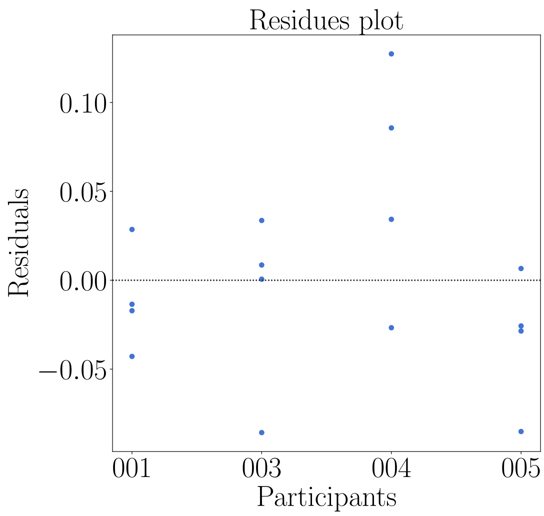


FIGURE 5.54 – Residual plot of the questionnaire score the sighted participants on each method.

TABLE 5.29 – Anova p-value for the mental demand average on each method'

(a) Blind participants.

Method			Analysis	
Audio	X	Haptic Belt	$H_1 : \mu_{Audio} \neq \mu_{HapticBelt}$	**
Audio	X	Virtual Cane	$H_1 : \mu_{Audio} \neq \mu_{VirtualCane}$	**
Audio	X	Mixture	$H_0 : \mu_{Audio} = \mu_{Mixture}$	
Haptic Belt	X	Virtual Cane	$H_1 : \mu_{HapticBelt} \neq \mu_{VirtualCane}$	**
Haptic Belt	X	Mixture	$H_1 : \mu_{HapticBelt} \neq \mu_{Mixture}$	**
Virtual Cane	X	Mixture	$H_1 : \mu_{VirtualCane} \neq \mu_{Mixture}$	**

(b) Sight participants.

Method			Analysis	
Audio	X	Haptic Belt	$H_1 : \mu_{Audio} \neq \mu_{HapticBelt}$	**
Audio	X	Virtual Cane	$H_1 : \mu_{Audio} \neq \mu_{VirtualCane}$	**
Audio	X	Mixture	$H_0 : \mu_{Audio} = \mu_{Mixture}$	
Haptic Belt	X	Virtual Cane	$H_1 : \mu_{HapticBelt} \neq \mu_{VirtualCane}$	**
Haptic Belt	X	Mixture	$H_1 : \mu_{HapticBelt} \neq \mu_{Mixture}$	**
Virtual Cane	X	Mixture	$H_1 : \mu_{VirtualCane} \neq \mu_{Mixture}$	**

5.2.2 Physiological data

5.2.2.1 Electrocardiogram (ECG) data

5.2.2.1.1 Analysis of the heartbeat frequency (BPM)

Table 5.30 presents the average heart rate for both sighted and blind groups. The barplots are presented in Figure 5.55. Comparing the two groups, the audio method is associated with a slightly lower heartrate for blind people, but the opposite happens for sighted participants. Moreover, data from blind participants have a significant variance. This significant variance can also be observed in the boxplot of Figures 5.56 and 5.57.

Figures 5.58 and 5.59 show the QQ plot and residual distributions for the sighted participants of Table 5.30. These figures show that the data are normally distributed and that the methods have a similar variance. Table 5.31 brings the results from ANOVA, which are similar for both sighted and blind participants.

TABLE 5.30 – ECG average BPM felled by the participants using the proposed methods [BPM].

Part.	Visual Condition	Round	Audio	Haptic Belt	Virtual Cane	Mixture
001	Sight	First	71.23	63.02	64.85	58.77
		Return	73.18	61.18	66.78	66.26
001C	Blind	First	60.71	71.17	59.07	68.24
		Return	58.61	66.22	64.20	70.76
002C	Blind	First	38.67	48.74	46.89	52.23
		Return	47.58	58.97	56.75	58.25
003	Sight	First	63.47	71.80	70.90	72.76
		Return	72.75	71.23	67.49	73.01
003C	Blind	First	69.89	70.95	69.41	66.94
		Return	67.44	69.68	68.82	67.37
004	Sight	First	66.85	62.45	65.94	67.86
		Return	69.48	65.65	64.58	71.86
004C	Blind	First	73.55	73.70	71.94	74.03
		Return	74.79	74.02	72.69	67.34
005	Sight	First	71.34	66.93	66.46	67.06
		Return	69.57	65.97	67.00	65.47

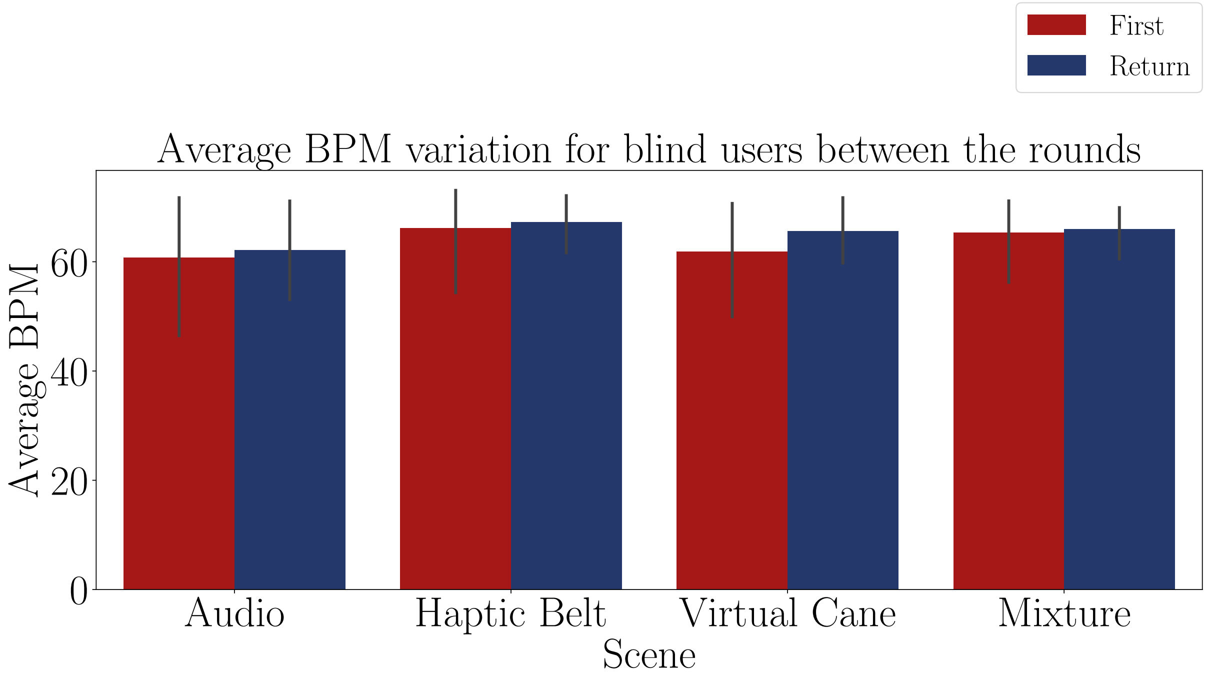
TABLE 5.31 – Anova p-value for the BPM on each method.

(a) Blind participants

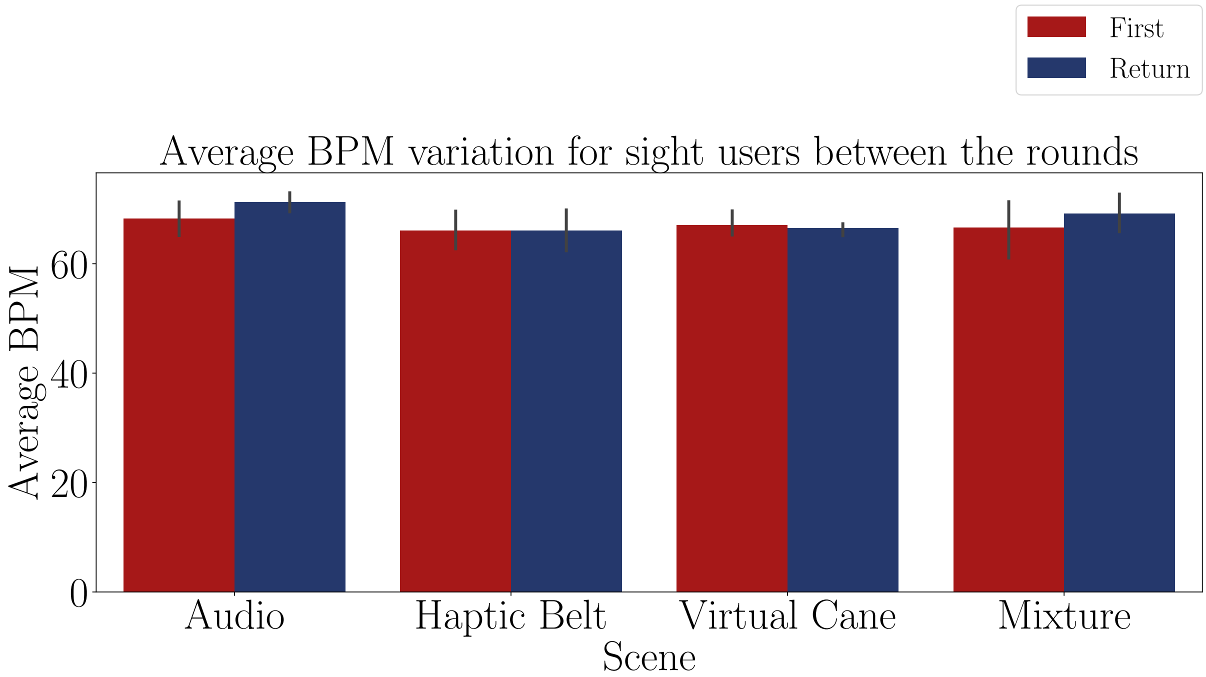
Source	P-Value
Methods	0.100
Rounds	0.371
Interaction	0.894

(b) Sight participants

Source	P-Value
Methods	0.166
Rounds	0.308
Interaction	0.631



(a) Blind participants.



(b) Sight participants.

FIGURE 5.55 – Barplot of the average BPM of the on each method and each round.

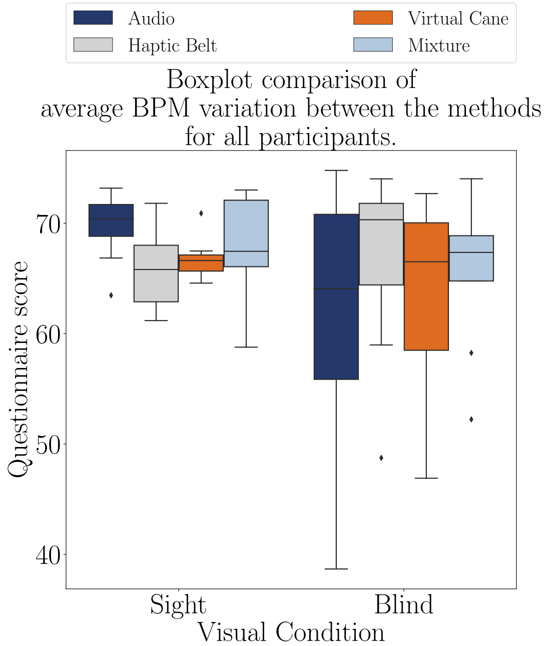


FIGURE 5.56 – Boxplot of the average BPM of the participants grouped by the methods.

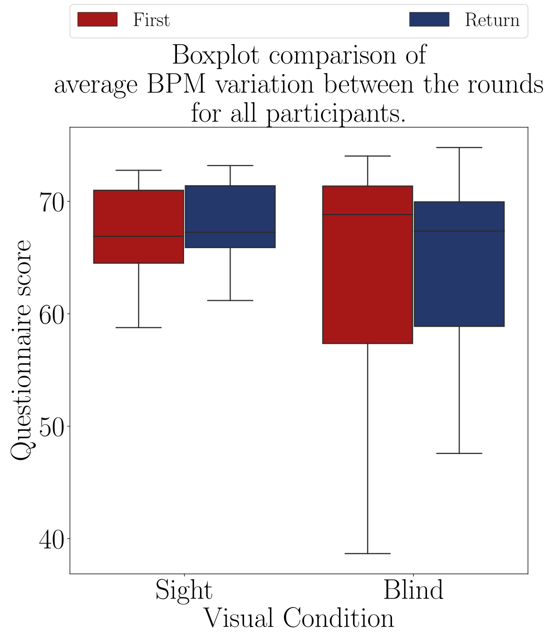


FIGURE 5.57 – Boxplot of the average BPM of the participants grouped by the rounds.

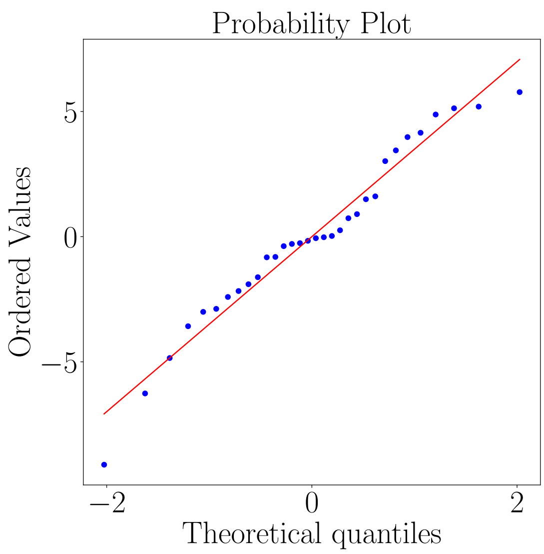


FIGURE 5.58 – QQ plot of the BPM of the sight participants on each method.

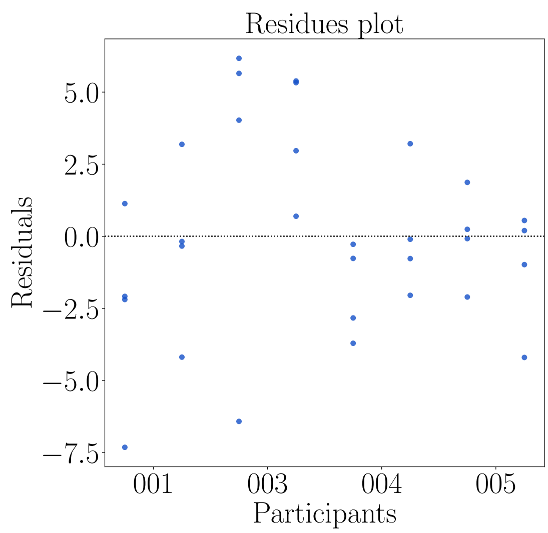


FIGURE 5.59 – Residual plot of the BPM score the sight participants on each method.

5.2.2.1.2 Analysis of the heartbeat variance (SDNN)

Table 5.32 presents the SDNN for both sighted and blind participants. The mean values are presented in the barplots of Figure 5.60.

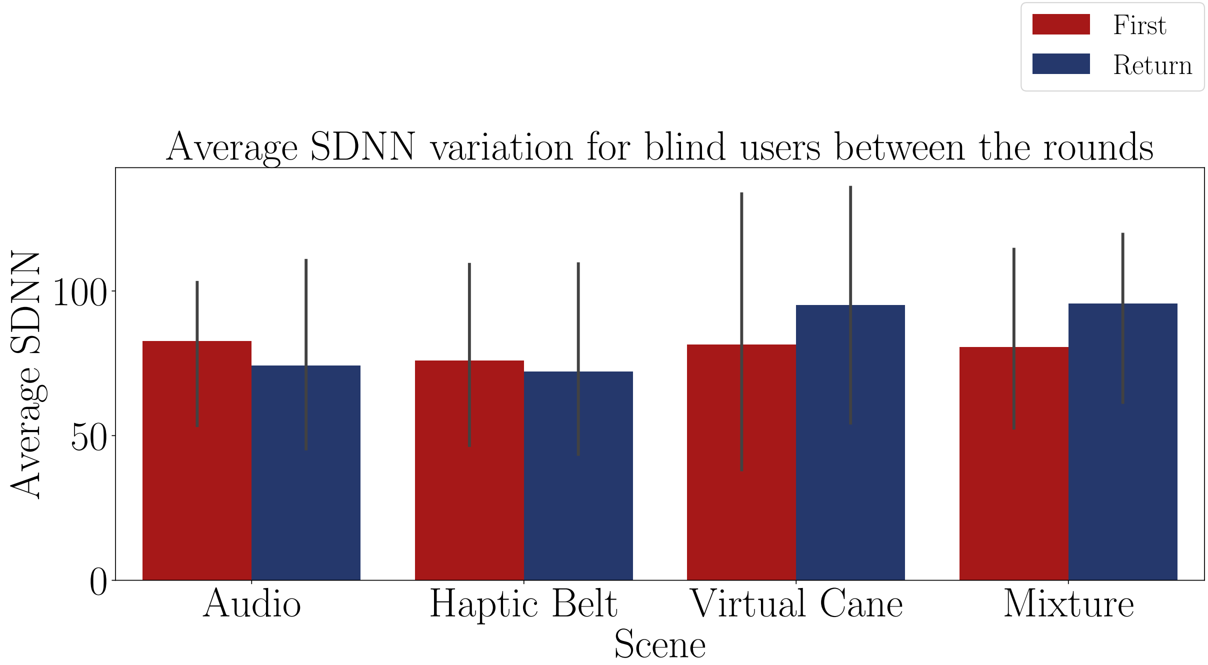
TABLE 5.32 – Average SDNN by the participants during the each round and method.

			Audio	Haptic Belt	Virtual Cane	Mixture
Part.	Visual Condition	Round				
001	Sight	First	82.185	134.530	134.773	225.408
		Return	69.479	318.747	116.003	136.507
001C	Blind	First	107.061	124.737	163.968	129.054
		Return	130.885	131.590	157.589	124.786
002C	Blind	First	98.863	81.140	33.977	79.289
		Return	49.627	42.815	114.057	107.545
003	Sight	First	79.600	51.782	68.676	60.842
		Return	45.709	40.927	66.323	47.823
003C	Blind	First	38.325	35.101	42.392	43.692
		Return	41.196	44.256	42.602	46.145
004	Sight	First	121.130	154.718	128.477	125.947
		Return	100.366	122.563	140.115	119.260
004C	Blind	First	86.827	62.560	85.900	70.472
		Return	74.895	70.017	66.089	104.040
005	Sight	First	87.686	120.522	88.591	102.796
		Return	93.207	122.839	141.305	96.035

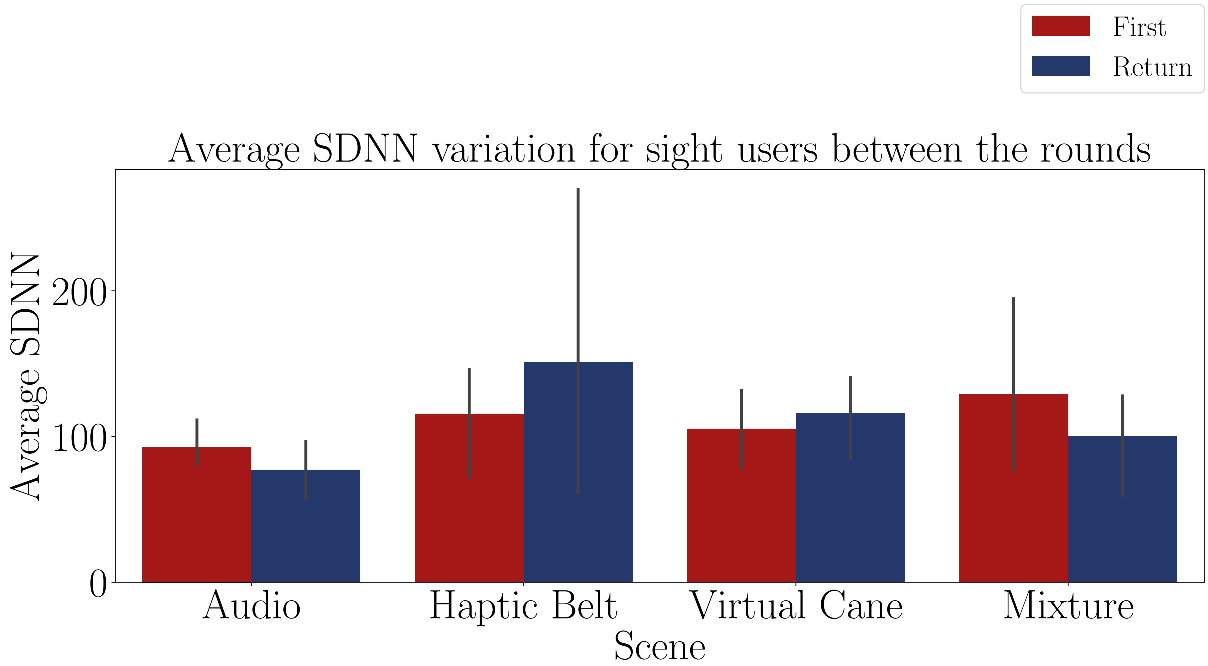
No clear pattern is evident from this figure. For some methods, the return round resulted in a decrease in the SDNN, while for others, it increased.

Figures 5.61 and 5.62 shows the boxplots for both groups. Both pictures show that the SDNN of the sighted users was higher than that of the blind users, indicating that sighted users had a lower mental workload than the blind users.

Figures 5.63 and 5.64 bring the QQ Plot and residual distribution. Figure 5.63 hints that the data from sighted users contain two outliers. Table 5.33 shows the ANOVA test p-values. For both groups, none of the factors have a significant influence on the SDNN value.



(a) Blind participants.



(b) Sight participants.

FIGURE 5.60 – Barplot of the average SDNN of the on each method and round.

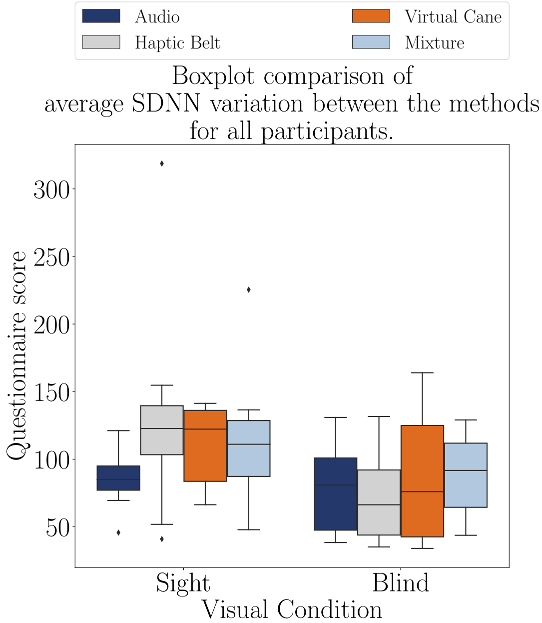


FIGURE 5.61 – Boxplot of the average SDNN of the participants grouped by the methods.

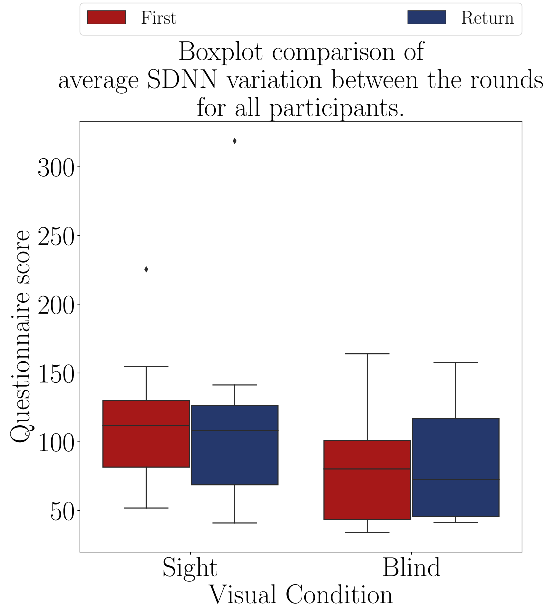


FIGURE 5.62 – Boxplot of the average SDNN of the participants grouped by the rounds.

TABLE 5.33 – Anova p-value for the average SDNN on each method.'

(a) Blind participants

Source	P-Value
Methods	0.486
Rounds	0.223
Interaction	0.473

(b) Sight participants

Source	P-Value
Methods	0.189
Rounds	0.969
Interaction	0.455

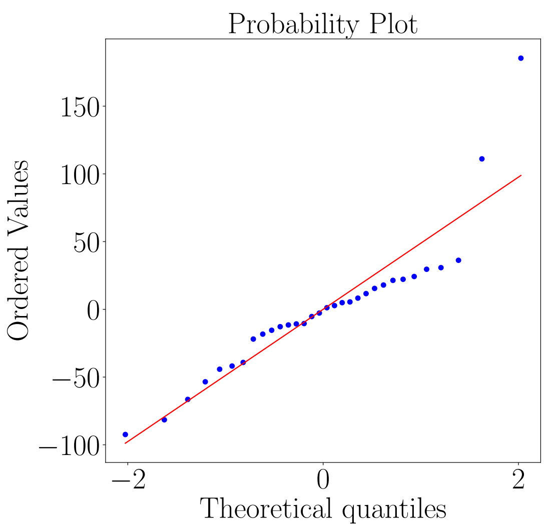


FIGURE 5.63 – QQ plot of the average SDNN of the sight participants on each method.

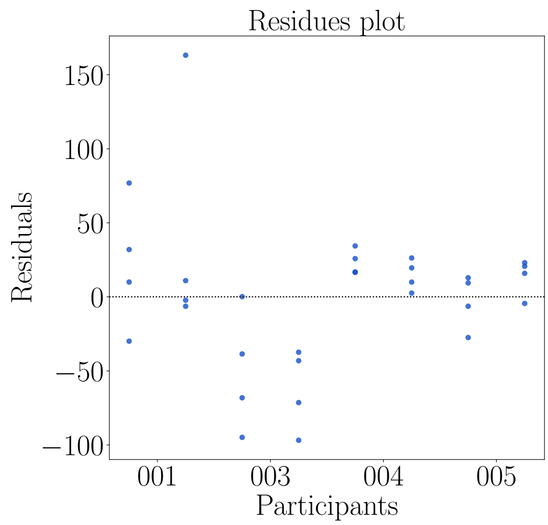


FIGURE 5.64 – Residual plot of the average SDNN score the sight participants on each method.

5.2.2.2 Galvanic skin reaction and temperature data;

Table 5.34 presents the average skin conductance for both groups, while the percentual variation related to the baseline is presented in Table 5.19.

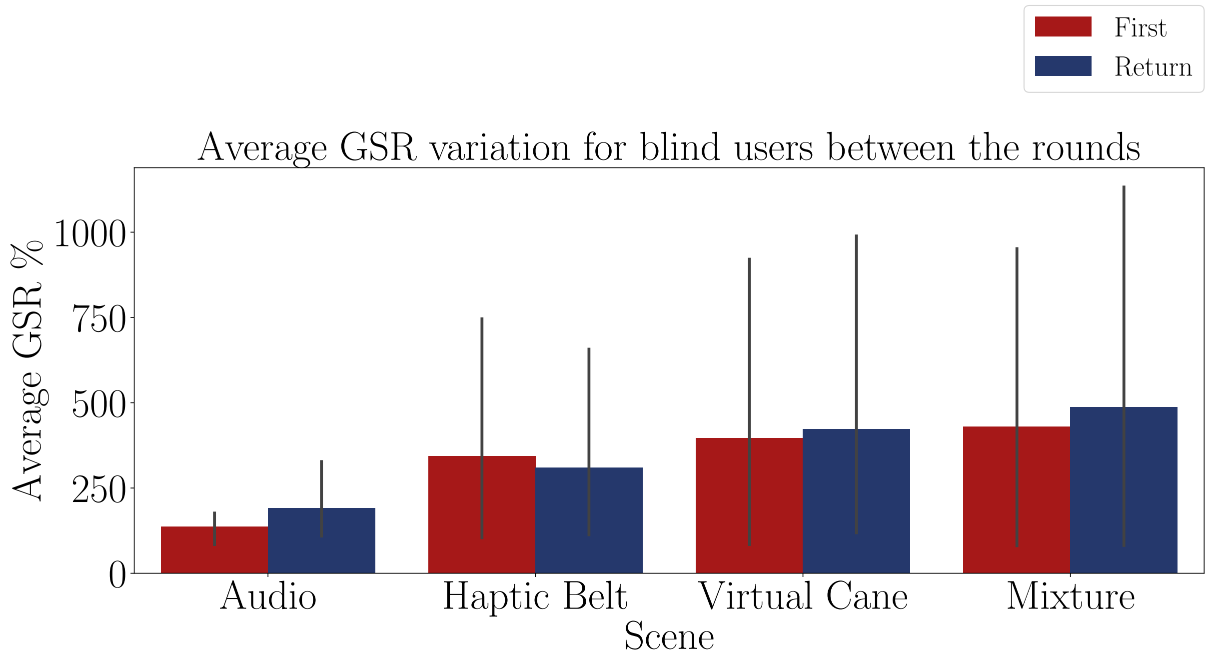
TABLE 5.34 – Average GSR felled by the participants [μS].

Participant	Visual Condition	Round	Baseline	Audio	Haptic Belt	Virtual Cane	Mixture
001	Sight	First	4.27	15.19	15.67	15.19	14.15
		Return		14.95	15.09	15.72	21.52
001C	Blind	First	0.37	1.03	3.14	3.79	3.90
		Return		1.58	2.81	4.04	4.57
003C	Blind	First	0.30	0.56	0.62	0.85	1.09
		Return		0.63	0.65	0.92	1.06
004	Sight	First	2.60	11.18	12.60	12.92	10.34
		Return		11.97	12.25	13.47	10.16
004C	Blind	First	1.24	3.07	3.49	2.28	2.23
		Return		2.95	3.20	2.21	2.24
005	Sight	First	0.47	1.58	1.44	1.37	1.33
		Return		1.53	1.47	1.49	1.33

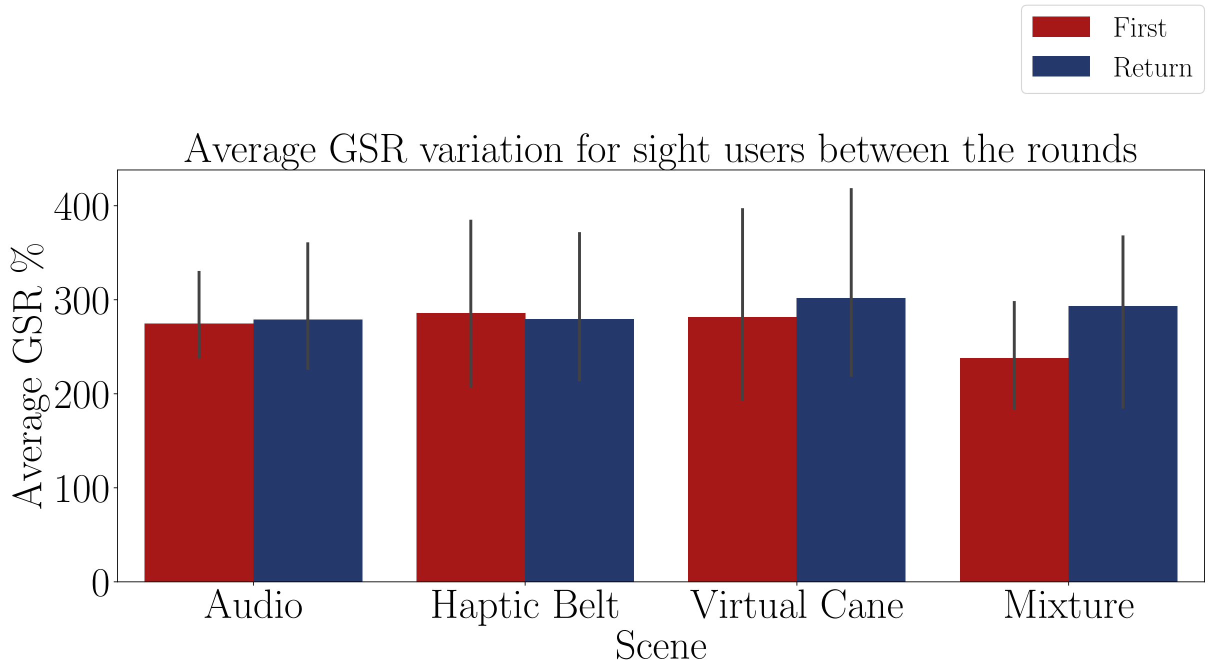
TABLE 5.35 – Average GSR variation in relation to the baseline in each round [μS].

Participant	Visual Condition	Round	Audio	Haptic Belt	Virtual Cane	Mixture
001	Sight	First	255.76%	266.93%	255.69%	231.52%
		Return	250.18%	253.32%	268.25%	403.90%
001C	Blind	First	176.54%	746.10%	920.72%	951.71%
		Return	327.42%	656.99%	988.93%	1132.39%
003C	Blind	First	84.23%	104.19%	182.35%	258.80%
		Return	109.23%	112.95%	202.35%	249.72%
004	Sight	First	329.08%	383.54%	395.83%	297.05%
		Return	359.53%	370.35%	417.17%	289.96%
004C	Blind	First	148.53%	182.84%	84.33%	80.69%
		Return	138.64%	159.00%	78.73%	81.61%
005	Sight	First	239.16%	207.74%	193.85%	184.71%
		Return	227.06%	214.91%	219.59%	185.86%

The barplots of the two groups are presented in Figure 5.65. While the GSR varied for the blind participants, increasing for methods with vibration, the same does not happen for sighted participants. Also, the variance of GSR data for blind participants is significantly higher than that of sighted ones. The same conclusion can be drawn from the boxplots in Figures 5.61 and 5.62.



(a) Blind participants.



(b) Sight participants.

FIGURE 5.65 – Barplot of the average GSR on each method and round.

Figures 5.68 and 5.69 bring the QQ Plot and residual distribution. The results from ANOVA are presented in Table 5.36. In the case of blind participants, the p-value for the method is just slightly over the threshold, indicating a possible influence of the method. The same does not happen with sighted participants, where the p-value of the method factor is the highest and well above the 0.05 threshold.

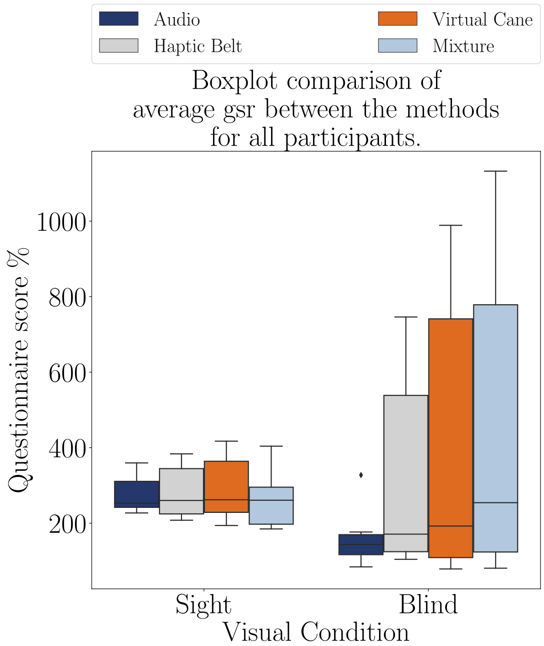


FIGURE 5.66 – Boxplot of the average GSR of the participants grouped by method.

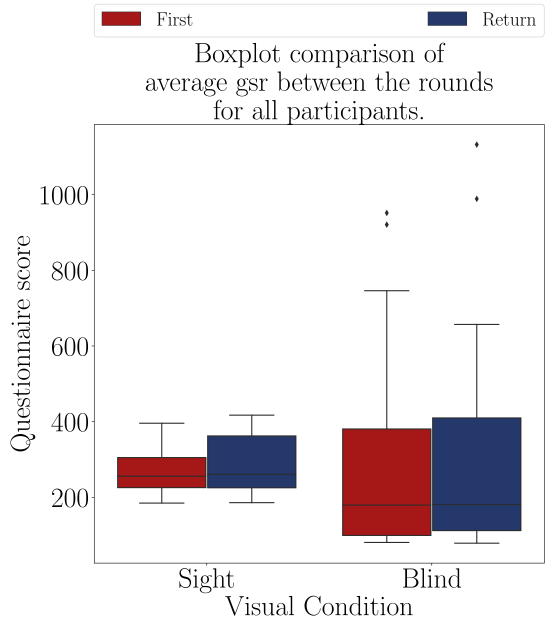


FIGURE 5.67 – Boxplot of the average GSR of the participants grouped by round.

TABLE 5.36 – Anova p-value for the skin conductance average on each method

(a) Blind participants

Source	P-Value
Methods	0.051
Rounds	0.722
Interaction	0.996

(b) Sight participants

Source	P-Value
Methods	0.802
Rounds	0.354
Interaction	0.686

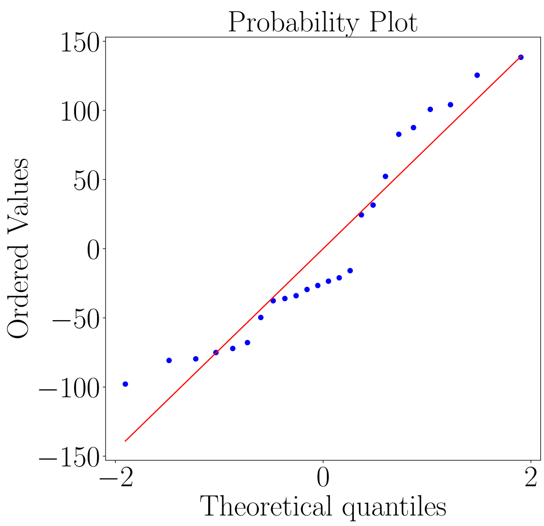


FIGURE 5.68 – QQ plot of the average skin conductance of the sight participants on each method.

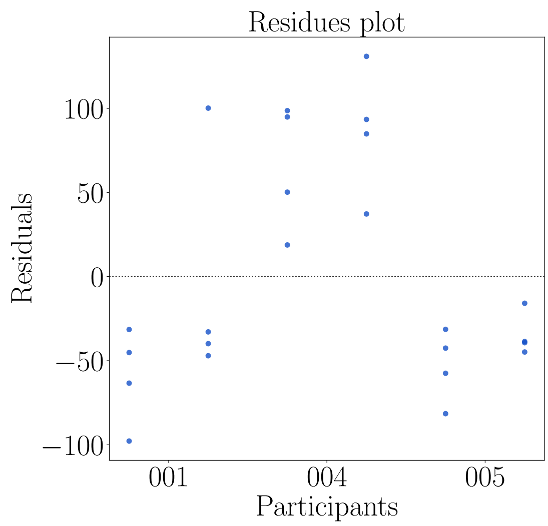


FIGURE 5.69 – Residual plot of the average skin conductance score the sight participants on each method.

5.2.3 Final Remarks

The comparison between the results from the blind participants and the sighted participants showed that there are significant differences in the evaluation performed by each group.

The sighted users evaluated the mental demand and other dimensions of NASA-TLX higher than blind ones. Also, blind participants were more familiar with audio methods and therefore gave a lower score to its mental demand. In the case of sighted participants, the method that received the lowest score was the virtual cane.

The adapted SAGAT questionnaire showed a more significant influence of the round factor for blind participants, which significantly improved their situation awareness on the return round. In the case of sighted users, the difference between the rounds was not so striking. Also, the score achieved by sighted participants was lower than that of blind users, which was expected.

Another difference is that, for blind participants, it was possible to observe a difference between the methods that use vibration and those that do not. This difference was not clear for sighted participants.

Besides these results, the sighted participants also gave feedback about the experiment. They felt considerably insecure when walking, even when hand-guided by another person. On the other hand, blind participants were already used to bumping their bodies when exploring new spaces. The sighted participants did not want that to happen and approached the furniture with caution. Similar to the blind participants, they also noticed the lack of precision of the haptic devices, but they did rely on them to navigate.

6 Conclusion

This work proposes using virtual reality to create an environment where concepts of assistive devices could be evaluated at the early stages of development by blind and visual impaired (BVI) people.

In order to systematize this proposal, this work presents a method composed of five phases that guide the development of the virtual environment in parallel with the design of the assistive devices and the proposal of assessment methods.

In order to illustrate the proposed method and investigate two research questions related to this work, it is described an application of the method for the evaluation of four different solutions of assistive devices in the environment of a hospital reception. For this example, it proposes as an assessment method the use of subjective questionnaires and physiological sensors.

In order to evaluate situation awareness, this work proposes an adapted version of the SAGAT questionnaire, which was initially introduced for evaluating the situation awareness of air traffic controllers.

Based on the results from the hospital reception example, the two research questions are discussed.

Is it possible to evaluate and compare concepts of assistive devices from a human factors' perspective in a virtual environment? What are the main limitations of the use of a virtual reality environment?

The example presented in this work showed that it is possible to evaluate both situation awareness and workload using experiments performed in a virtual environment. The tests performed in the virtual environment made possible the comparison of the assistive devices both qualitatively and quantitatively.

However, some limitations were identified during the development of this work, regarding both the virtual environment and the assessment techniques.

One of the most recurrent observations was the unsatisfactory quality of the sound system. According to blind participants, the headphone of the VIVE HMD does not provide sounds with a quality good enough for them to locate the source of a sound. A regular comment was, "I feel like the sound origin is inside my head". This limitation may be solved by placing a sound source in the real environment and use the HMD only for localizing the participant in the virtual environment.

Another limitation is the actual position of the furniture. After a first round, the furniture was not precisely aligned with its virtual model. A future solution for this problem would be to use a locator on each piece of furniture.

Among the main limitations identified during this work, it is worth mentioning the failure to detect collisions in the virtual environment, which could be solved by integrating sensors that monitor the position of each arm and leg of the user.

Regarding the assessment methods for evaluating human factors, the physiological sensors did not show any systematic difference among the methods under analysis. This result may be due the noise in the sensors' data, compromising its quality. Another problem is that the low number of participants may compromise the statistical analysis, due to the large variability among users.

Do non-BVI users, when deprived of their vision, similarly evaluate assistive devices as BVI users?

Comparing the results of the experiments performed with blind and sighted participants, some differences were observed. Among the most important, is the relative evaluation of audio and haptic devices. Due to their enhanced sensitivity to sounds, BVI users tend to evaluate audio solutions better than non-BVI. Also, the effect of repeating a task in the same environment, i.e., performing different rounds of the same experiment, may differ between sighted and blind users.

Generally, the results reinforce the importance of having BVI users involved in the design of assistive devices from the early stages of the specification of requirements.

6.1 Future works and suggestions

The following topics are of interest for future research:

- Perform a comparison between an evaluation campaign executed in a real environment and the same campaign executed in a virtual environment, to assess the main differences brought by the use of virtual reality;

- Further improve the virtual reality environment by providing better sound solutions, using different sources of sound in the real environment instead of using the sound from the HMD;
- Develop a solution for automatic collision detection in the virtual system to introduce performance metrics in the assessment.
- Repeat the experimental campaign with large sample of both BVI and non-BVI users, in order to improve the statistical analysis.
- Investigate the sources of noise of the physiological sensors and improve their data acquisition.

Bibliography

BORGHINI, G.; ASTOLFI, L.; VECCHIATO, G.; MATTIA, D.; BABILONI, F. Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. **Neuroscience & Biobehavioral Reviews**, Elsevier, v. 44, p. 58–75, 2014. 31

BOURNE, R.; STEINMETZ, J. D.; FLAXMAN, S.; BRIANT, P. S.; TAYLOR, H. R.; RESNIKOFF, S.; CASSON, R. J.; ABDOLI, A.; ABU-GHARBIEH, E.; AFSHIN, A. *et al.* Trends in prevalence of blindness and distance and near vision impairment over 30 years: an analysis for the global burden of disease study. **The Lancet global health**, Elsevier, v. 9, n. 2, p. e130–e143, 2021. 20

BRADLEY, N. A.; DUNLOP, M. D. Investigating context-aware clues to assist navigation for visually impaired people. In: **Proceedings of Workshop on Building Bridges: Interdisciplinary Context-Sensitive Computing, University of Glasgow**. [S.l.: s.n.], 2002. ix, 39, 40, 41, 42

BRADLEY, N. A.; DUNLOP, M. D. An experimental investigation into wayfinding directions for visually impaired people. **Personal and Ubiquitous Computing**, Springer, v. 9, n. 6, p. 395–403, 2005. ix, 39, 40, 41, 42

CAIN, B. A review of the mental workload literature. Defence research and development Toronto (Canada), 2007. 29, 30

CARDOSO, M. d. S.; GONTIJO, L. A. Evaluation of mental workload and performance measurement: Nasa tlx and swat. **Gestão &, Produção**, v. 19, p. 873–884, 2012. 29

CHAKLADAR, D. D.; DEY, S.; ROY, P. P.; DOGRA, D. P. Eeg-based mental workload estimation using deep blstm-lstm network and evolutionary algorithm. **Biomedical Signal Processing and Control**, Elsevier, v. 60, p. 101989, 2020. 30

CHIU, M.-L. An organizational view of design communication in design collaboration. **Design studies**, Elsevier, v. 23, n. 2, p. 187–210, 2002. 35

CUMMINGS, J. J.; BAILENSON, J. N. How immersive is enough? a meta-analysis of the effect of immersive technology on user presence. **Media psychology**, Taylor & Francis, v. 19, n. 2, p. 272–309, 2016. 38

DOOLANI, S.; WESSELS, C.; KANAL, V.; SEVASTOPOULOS, C.; JAISWAL, A.; NAMBIAPPAN, H.; MAKEDON, F. A review of extended reality (xr) technologies for manufacturing training. **Technologies**, Multidisciplinary Digital Publishing Institute, v. 8, n. 4, p. 77, 2020. 33, 34

DUL, J.; WEERDMEESTER, B. **Ergonomics for beginners: a quick reference guide**. [S.l.]: CRC press, 2003. 27

ENDSLEY, M. R. Design and evaluation for situation awareness enhancement. In: SAGE PUBLICATIONS SAGE CA: LOS ANGELES, CA. **Proceedings of the Human Factors Society annual meeting**. [S.l.], 1988. v. 32, n. 2, p. 97–101. 32, 33, 52

ENDSLEY, M. R. Measurement of situation awareness in dynamic systems. **Human factors**, SAGE Publications Sage CA: Los Angeles, CA, v. 37, n. 1, p. 65–84, 1995. 32

ENDSLEY, M. R. Automation and situation awareness. In: **Automation and human performance: Theory and applications**. [S.l.]: CRC Press, 2018. p. 163–181. 32

FALLAHI, M.; MOTAMEDZADE, M.; HEIDARIMOGHADAM, R.; SOLTANIAN, A. R.; MIYAKE, S. Effects of mental workload on physiological and subjective responses during traffic density monitoring: A field study. **Applied ergonomics**, Elsevier, v. 52, p. 95–103, 2016. 29, 30

FARRELL, W. A. Learning becomes doing: Applying augmented and virtual reality to improve performance. **Performance Improvement**, Wiley Online Library, v. 57, n. 4, p. 19–28, 2018. 20, 33, 34

FARRER, C.; FRITH, C. D. Experiencing oneself vs another person as being the cause of an action: the neural correlates of the experience of agency. **Neuroimage**, Elsevier, v. 15, n. 3, p. 596–603, 2002. 38

FARSHID, M.; PASCHEN, J.; ERIKSSON, T.; KIETZMANN, J. Go boldly!: Explore augmented reality (ar), virtual reality (vr), and mixed reality (mr) for business. **Business Horizons**, Elsevier, v. 61, n. 5, p. 657–663, 2018. 34

HART, S. G.; STAVELAND, L. E. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In: **Advances in psychology**. [S.l.]: Elsevier, 1988. v. 52, p. 139–183. 31

JICOL, C.; WAN, C. H.; DOLING, B.; ILLINGWORTH, C. H.; YOON, J.; HEADEY, C.; LUTTEROTH, C.; PROULX, M. J.; PETRINI, K.; O'NEILL, E. Effects of emotion and agency on presence in virtual reality. In: **Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems**. [S.l.: s.n.], 2021. p. 1–13. 38, 39

JONDANI, J. A. Strategies for addressing the special needs of people with visual impairments during the covid-19 pandemic. **Journal of Visual Impairment & Blindness**, SAGE Publications Sage CA: Los Angeles, CA, v. 115, n. 3, p. 263–267, 2021. 21

KARWOWSKI, W. The discipline of human factors and ergonomics. **Handbook of human factors and ergonomics**, Citeseer, v. 4, p. 3–37, 2012. 26

KLEINSMANN, M. S. Understanding collaborative design. 2006. 35, 36

KYLECORRY31. **Haptic compass belt**. 2020. Disponível em: <https://www.instructables.com/Haptic-Compass-Belt/>. 54

LOZANO, C. A.; KACZMAREK, K. A.; SANTELLO, M. Electrotactile stimulation on the tongue: Intensity perception, discrimination, and cross-modality estimation. **Somatosensory & motor research**, Taylor & Francis, v. 26, n. 2-3, p. 50–63, 2009. 20

MA, J. Y.; CHOI, J. S. The virtuality and reality of augmented reality. **J. Multim.**, Citeseer, v. 2, n. 1, p. 32–37, 2007. 33, 34

MANSIKKA, H.; SIMOLA, P.; VIRTANEN, K.; HARRIS, D.; OKSAMA, L. Fighter pilots' heart rate, heart rate variation and performance during instrument approaches. **Ergonomics**, Taylor & Francis, v. 59, n. 10, p. 1344–1352, 2016. 31

MARSTON, J. R.; LOOMIS, J. M.; KLATZKY, R. L.; GOLLEDGE, R. G.; SMITH, E. L. Evaluation of spatial displays for navigation without sight. **ACM Transactions on Applied Perception (TAP)**, ACM New York, NY, USA, v. 3, n. 2, p. 110–124, 2006. 42, 43

MILGRAM, P.; KISHINO, F. A taxonomy of mixed reality visual displays. **IEICE TRANSACTIONS on Information and Systems**, The Institute of Electronics, Information and Communication Engineers, v. 77, n. 12, p. 1321–1329, 1994. ix, 33, 34

MOERLAND-MASIC, I.; REIMER, F.; BOCK, T. M.; MELLER, F.; NAGEL, B. Application of vr technology in the aircraft cabin design process. **CEAS Aeronautical Journal**, Springer, p. 1–10, 2021. ix, 43, 44

MOHANAVELU, K.; POONGUZHALI, S.; RAVI, D.; SINGH, P. K.; MAHAJABIN, M.; RAMACHANDRAN, K.; SINGH, U. K.; JAYARAMAN, S. Cognitive workload analysis of fighter aircraft pilots in flight simulator environment. **Defence Science Journal**, v. 70, n. 2, 2020. 30, 31

MUJBER, T. S.; SZECSI, T.; HASHMI, M. S. Virtual reality applications in manufacturing process simulation. **Journal of materials processing technology**, Elsevier, v. 155, p. 1834–1838, 2004. 34

NIJHOLT, A.; TRAUM, D. The virtuality continuum revisited. In: **CHI'05 Extended Abstracts on Human Factors in Computing Systems**. [S.l.: s.n.], 2005. p. 2132–2133. 33

NOURBAKHSH, N.; WANG, Y.; CHEN, F.; CALVO, R. A. Using galvanic skin response for cognitive load measurement in arithmetic and reading tasks. In: **Proceedings of the 24th Australian Computer-Human Interaction Conference**. [S.l.: s.n.], 2012. p. 420–423. 31

ORLANDI, L.; BROOKS, B. Measuring mental workload and physiological reactions in marine pilots: Building bridges towards redlines of performance. **Applied ergonomics**, Elsevier, v. 69, p. 74–92, 2018. 30

RODRÍGUEZ, S.; SÁNCHEZ, L.; LÓPEZ, P.; CAÑAS, J. J. Pupillometry to assess air traffic controller workload through the mental workload model. In: **Proceedings of the 5th international conference on application and theory of automation in command and control systems**. [S.l.: s.n.], 2015. p. 95–104. 30

SALAH, B.; ABIDI, M. H.; MIAN, S. H.; KRID, M.; ALKHALEFAH, H.; ABDO, A. Virtual reality-based engineering education to enhance manufacturing sustainability in industry 4.0. **Sustainability**, Multidisciplinary Digital Publishing Institute, v. 11, n. 5, p. 1477, 2019. 34

SANDERS, M. S.; MCCORMICK, E. J. Human factors in engineering and design. **Industrial Robot: An International Journal**, Emerald Group Publishing Limited, 1998. ix, 27, 28, 30, 31, 32

SANDOM, C.; HARVEY, R. S. **Human factors for engineers**. [S.l.]: Iet, 2004. v. 2. 26, 27

SHI, Y.; RUIZ, N.; TAIB, R.; CHOI, E.; CHEN, F. Galvanic skin response (gsr) as an index of cognitive load. In: **CHI'07 extended abstracts on Human factors in computing systems**. [S.l.: s.n.], 2007. p. 2651–2656. 31

SIU, A. F.; SINCLAIR, M.; KOVACS, R.; OFEK, E.; HOLZ, C.; CUTRELL, E. Virtual reality without vision: A haptic and auditory white cane to navigate complex virtual worlds. In: **Proceedings of the 2020 CHI conference on human factors in computing systems**. [S.l.: s.n.], 2020. p. 1–13. 37, 38

STANTON, N. A.; HEDGE, A.; BROOKHUIS, K.; SALAS, E.; HENDRICK, H. W. **Handbook of human factors and ergonomics methods**. [S.l.]: CRC press, 2004. xiv, 28, 30, 31, 32

WANG, S.; MAO, Z.; ZENG, C.; GONG, H.; LI, S.; CHEN, B. A new method of virtual reality based on unity3d. In: IEEE. **2010 18th international conference on Geoinformatics**. [S.l.], 2010. p. 1–5. 49

WOLF, A.; BINDER, N.; MIEHLING, J.; WARTZACK, S. Towards virtual assessment of human factors: A concept for data driven prediction and analysis of physical user-product interactions. In: CAMBRIDGE UNIVERSITY PRESS. **Proceedings of the Design Society: International Conference on Engineering Design**. [S.l.], 2019. v. 1, n. 1, p. 4029–4038. 20

WORLD HEALTH ORGANIZATION. **Advice for the public on covid-19**. World Health Organization, 2020. Disponível em: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public>. Acesso em: 8 fev. 2022. 21

WORLD HEALTH ORGANIZATION AND OTHERS. **World report on vision**. [S.l.], 2019. 20

YANG, C.-H.; HWANG, S.-L.; WANG, J.-L. The design and evaluation of an auditory navigation system for blind and visually impaired. In: IEEE. **Proceedings of the 2014 IEEE 18th International Conference on Computer Supported Cooperative Work in Design (CSCWD)**. [S.l.], 2014. p. 342–345. 41, 42

ZHANG, H.; ZHU, Y.; MANIYERI, J.; GUAN, C. Detection of variations in cognitive workload using multi-modality physiological sensors and a large margin unbiased regression machine. In: IEEE. **2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society**. [S.l.], 2014. p. 2985–2988. 31

Appendix A - Questionnaires

A.1 BVI condition

Código de identificação do voluntário: ______________

Os dados deste questionário são anônimos e somente serão utilizados para fins acadêmicos e de pesquisa, ficando proibido seu manuseio ou uso sem consentimento do coordenador da pesquisa.

Parte 1 - Questionário sobre dados pessoais.

Idade: ______________ **Sexo:** ( ) Masculino ( ) Feminino

Como deficiente visual você se identifica como:

( ) Cego
( ) Surdo-Cego
( ) Baixa Visão
( ) Monocular
( ) Não sou deficiente visual

USO DO PESQUISADOR

Ordem dos cenários:

____

A.2 NASA-TLX

Código de identificação do voluntário: ____________

Os dados deste questionário são anônimos e somente serão utilizados para fins acadêmicos e de pesquisa, ficando proibido seu manuseio ou uso sem consentimento do coordenador da pesquisa.

Parte 2 - Questionário sobre carga mental (NASA-TLX).

TESTE ( ) BASE ( ) ÁUDIO ( ) CINTO HÁPTICO ( ) BENGALA ( ) MISTO

Avalie cada um dos itens abaixo em uma escala de 1 a 20.

Primeira visita

1) Demanda Mental — Baixo | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | Alto

2) Demanda Física — Baixo | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | Alto

3) Demanda Temporal — Baixo | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | Alto

4) Desempenho — Baixo | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | Alto

5) Esforço — Baixo | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | Alto

6) Frustração — Baixo | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | Alto

Retorno

1) Demanda Mental — Baixo | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | Alto

2) Demanda Física — Baixo | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | Alto

3) Demanda Temporal — Baixo | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | Alto

4) Desempenho — Baixo | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | Alto

5) Esforço — Baixo | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | Alto

6) Frustração — Baixo | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | Alto

A.3 Adapted SAGAT

Código de identificação do voluntário: ____________

Os dados deste questionário são anônimos e somente serão utilizados para fins acadêmicos e de pesquisa, ficando proibido seu manuseio ou uso sem consentimento do coordenador da pesquisa.

Parte 3 - Questionário sobre consciência situacional (SAGAT).

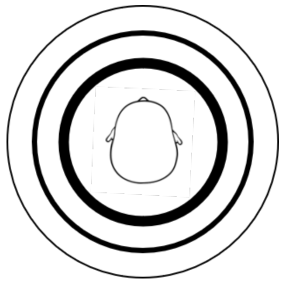
TESTE ( ) BASE ( ) ÁUDIO ( ) CINTO HÁPTICO ( ) BENGALA ( ) MISTO

Primeira visita

Nível 1 – Percepção

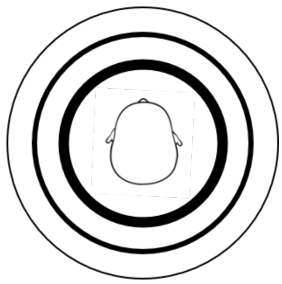
1) Existe algum objeto próximo de você?

2) Sinalize onde o objeto está:



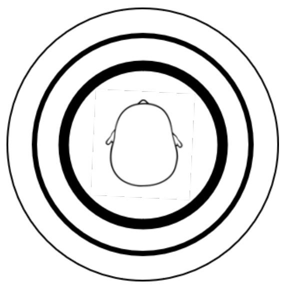
3) Existe alguém perto de você?

4) Sinalize onde a pessoa está:



5) Você percebeu alguma fonte de som característica do lugar onde você se encontra?

6) Sinalize de onde vem o som:



Código de identificação do voluntário: ____________

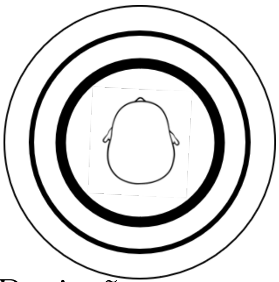
Os dados deste questionário são anônimos e somente serão utilizados para fins acadêmicos e de pesquisa, ficando proibido seu manuseio ou uso sem consentimento do coordenador da pesquisa.

Parte 3 - Questionário sobre consciência situacional (SAGAT).

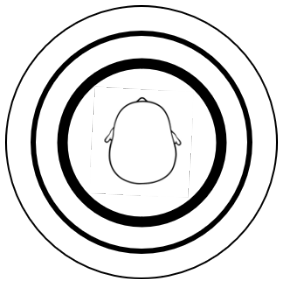
TESTE ( ) BASE ( ) ÁUDIO ( ) CINTO HÁPTICO ( ) BENGALA ( ) MISTO

Nível 2 – Compreensão

7) Sinalize em que direção está a recepcionista:



8) Sinalize em que direção está a saída:



Nível 3 – Projeção

9) O que você espera encontrar no caminho ao retornar à mesa de recepção?

____

____

____

____

10) Qual será a sua trajetória para sair? Descreva seus próximos movimentos.

____

____

____

____

11) Qual a distância você imagina que existe entre você e a mesa da recepção?

____

____

____

____

12) Qual a distância você imagina que existe entre você e a saída?

____

____

____

____

Código de identificação do voluntário: ____________

Os dados deste questionário são anônimos e somente serão utilizados para fins acadêmicos e de pesquisa, ficando proibido seu manuseio ou uso sem consentimento do coordenador da pesquisa.

Parte 3 - Questionário sobre consciência situacional (SAGAT).

TESTE ( ) BASE ( ) ÁUDIO ( ) CINTO HÁPTICO ( ) BENGALA ( ) MISTO

Retorno

Nível 1 – Percepção

1) Existe algum objeto próximo de você?

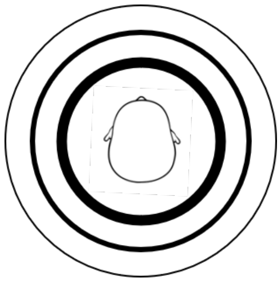
____

____

____

____

2) Sinalize onde o objeto está:



3) Existe alguém perto de você?

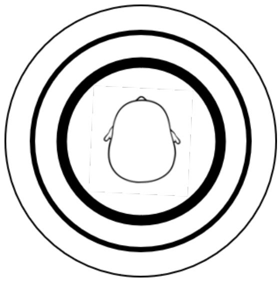
____

____

____

____

4) Sinalize onde a pessoa está:



5) Você percebeu alguma fonte de som característica do lugar onde você se encontra?

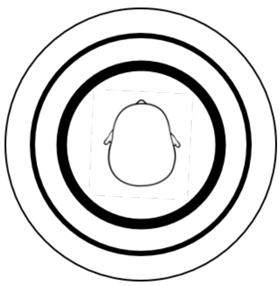
____

____

____

____

6) Sinalize de onde vem o som:



Código de identificação do voluntário: ____________

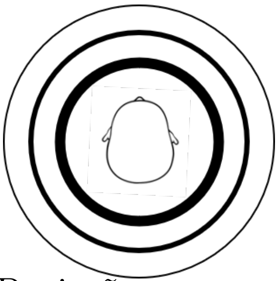
Os dados deste questionário são anônimos e somente serão utilizados para fins acadêmicos e de pesquisa, ficando proibido seu manuseio ou uso sem consentimento do coordenador da pesquisa.

Parte 3 - Questionário sobre consciência situacional (SAGAT).

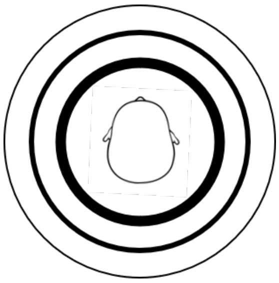
TESTE ( ) BASE ( ) ÁUDIO ( ) CINTO HÁPTICO ( ) BENGALA ( ) MISTO

Nível 2 – Compreensão

7) Sinalize em que direção está a recepcionista:



8) Sinalize em que direção está a saída:



Nível 3 – Projeção

9) O que você espera encontrar no caminho ao retornar à mesa de recepção?

____

____

____

____

10) Qual será a sua trajetória para sair? Descreva seus próximos movimentos.

____

____

____

____

11) Qual a distância você imagina que existe entre você e a mesa da recepção?

____

____

____

____

12) Qual a distância você imagina que existe entre você e a saída?

____

____

____

____

A.4 Guidance method evaluation

Código de identificação do voluntário: ____________

Os dados deste questionário são anônimos e somente serão utilizados para fins acadêmicos e de pesquisa, ficando proibido seu manuseio ou uso sem consentimento do coordenador da pesquisa.

Parte 4 - Questionário sobre o método de navegação

TESTE 1 - ORIENTAÇÃO VIA ÁUDIO

1) Os sons foram de fácil interpretação?

1	2	3	4	5	6	7
Muito fácil			Médio			Muito díficil

2) Os sons causaram algum incômodo durante o uso?

1	2	3	4	5	6	7
Nenhum incômodo			Médio			Muito incômodo

3) Os comandos foram claros?

1	2	3	4	5	6	7
Não foram claros			Médio			Muito claros

4) Como você avalia a quantidade de comandos?

1	2	3	4	5	6	7
Muito pouca			Médio			Muito excessiva

5) Os momentos em que os comandos foram reproduzidos foram inadequados?

1	2	3	4	5	6	7
Não foram adequados			Médio			Muito adequados

6) O som ambiente atrapalhou a reprodução de algum comando

1	2	3	4	5	6	7
Não atrapalhou			Médio			Atrapalhou muito

7) Você sentiu falta de alguma informação durante a orientação via áudio?

____

____

____

Código de identificação do voluntário: ____________

Os dados deste questionário são anônimos e somente serão utilizados para fins acadêmicos e de pesquisa, ficando proibido seu manuseio ou uso sem consentimento do coordenador da pesquisa.

Parte 4 - Questionário sobre o método de navegação

TESTE 2 – CINTO HÁPTICO

8) A informação da vibração do cinto foi precisa?

1	2	3	4	5	6	7
Pouco precisa			Médio			Muito precisa

9) A vibração do cinto causou algum incômodo durante o uso?

1	2	3	4	5	6	7
Nenhum incômodo			Médio			Muito incômodo

10) A vibração do cinto causou alguma confusão durante a navegação?

1	2	3	4	5	6	7
Nenhuma confusão			Médio			Muita confusão

11) A vibração do cinto trouxe alguma segurança na navegação?

1	2	3	4	5	6	7
Nenhuma segurança			Médio			Muita segurança

12) O cinto causou algum incômodo durante o uso?

1	2	3	4	5	6	7
Nenhum incômodo			Médio			Muito incômodo

13) Você sentiu falta de alguma informação durante a orientação usando o cinto?

____

____

____

Código de identificação do voluntário: ____________

Os dados deste questionário são anônimos e somente serão utilizados para fins acadêmicos e de pesquisa, ficando proibido seu manuseio ou uso sem consentimento do coordenador da pesquisa.

Parte 4 - Questionário sobre o método de navegação

TESTE 3 - BENGALA VIRTUAL

14) A informação da vibração do cinto foi precisa?

1	2	3	4	5	6	7
Pouco precisa			Médio			Muito precisa

15) A bengala virtual funcionou semelhantemente à bengala tradicional?

1	2	3	4	5	6	7
Pouco semelhante			Médio			Muito semelhante

16) O uso da bengala foi intuitivo?

1	2	3	4	5	6	7
Não foi intuitivo			Médio			Muito intuitivo

17) A bengala causou algum incômodo durante o uso?

1	2	3	4	5	6	7
Nenhum incômodo			Médio			Muito incômodo

18) A bengala causou alguma confusão durante a navegação?

1	2	3	4	5	6	7
Nenhuma confusão			Médio			Muita confusão

19) Você sentiu falta de alguma informação durante a orientação?

____

____

____

20) Você prefere mais a bengala virtual ou o cinto? Por quê?

____

____

____

Código de identificação do voluntário: ____________
Os dados deste questionário são anônimos e somente serão utilizados para fins acadêmicos e de pesquisa, ficando proibido seu manuseio ou uso sem consentimento do coordenador da pesquisa.

Parte 4 - Questionário sobre o método de navegação

TESTE 4 - MISTURADO (ÁUDIO E CINTO)

21) Os sons foram de fácil interpretação?

1	2	3	4	5	6	7
Muito fácil			Médio			Muito díficil

22) Os sons causaram algum incômodo durante o uso?

1	2	3	4	5	6	7
Nenhum incômodo			Médio			Muito incômodo

23) Os comandos foram claros?

1	2	3	4	5	6	7
Não foram claros			Médio			Muito claros

24) Como você avalia a quantidade de comandos?

1	2	3	4	5	6	7
Muito pouca			Médio			Muito excessiva

25) Os momentos em que os comandos foram reproduzidos foram inadequados?

1	2	3	4	5	6	7
Não foram adequados			Médio			Muito adequados

26) O som ambiente atrapalhou a reprodução de algum comando

1	2	3	4	5	6	7
Não atrapalhou			Médio			Atrapalhou muito

27) A informação da vibração do cinto foi precisa?

1	2	3	4	5	6	7
Pouco precisa			Médio			Muito precisa

28) A vibração do cinto causou algum incômodo durante o uso?

1	2	3	4	5	6	7
Nenhum incômodo			Médio			Muito incômodo

Código de identificação do voluntário: ____________

Os dados deste questionário são anônimos e somente serão utilizados para fins acadêmicos e de pesquisa, ficando proibido seu manuseio ou uso sem consentimento do coordenador da pesquisa.

Parte 4 - Questionário sobre o método de navegação

29) A vibração do cinto causou alguma confusão durante a navegação?

1	2	3	4	5	6	7
Nenhuma confusão			Médio			Muita confusão

30) A vibração do cinto trouxe alguma segurança na navegação

1	2	3	4	5	6	7
Nenhuma segurança			Médio			Muita segurança

31) O cinto causou algum incômodo durante o uso?

1	2	3	4	5	6	7
Nenhum incômodo			Médio			Muito incômodo

32) A mistura de comandos de áudio com vibração te trouxe mais segurança?

1	2	3	4	5	6	7
Nenhuma segurança			Médio			Muita segurança

33) De forma geral, como você avalia a quantidade de informação recebida?

1	2	3	4	5	6	7
Muito pouca			Médio			Muito excessiva

34) A informação da vibração da bengala foi precisa?

1	2	3	4	5	6	7
Pouco precisa			Médio			Muito precisa

35) A bengala virtual funcionou semelhantemente à bengala tradicional?

1	2	3	4	5	6	7
Pouco semelhante			Médio			Muito semelhante

36) O uso da bengala foi intuitivo?

1	2	3	4	5	6	7
Não foi intuitivo			Médio			Muito intuitivo

Código de identificação do voluntário: ____________

Os dados deste questionário são anônimos e somente serão utilizados para fins acadêmicos e de pesquisa, ficando proibido seu manuseio ou uso sem consentimento do coordenador da pesquisa.

Parte 4 - Questionário sobre o método de navegação

37) A bengala causou algum incômodo durante o uso?

1	2	3	4	5	6	7
Nenhum incômodo			Médio			Muito incômodo

38) A bengala causou alguma confusão durante a navegação?

1	2	3	4	5	6	7
Nenhuma confusão			Médio			Muita confusão

39) Você sentiu falta de alguma informação durante a orientação?

____

____

____

40) Considerando a emissão de comandos via áudio e via cinto háptico, você considera que algum deles é desnecessário?

____

____

____

FOLHA DE REGISTRO DO DOCUMENTO			
1. CLASSIFICAÇÃO/TIPO DM	2. DATA	3. DOCUMENTO Nº DCTA/ITA//	4. Nº DE PÁGINAS 126
5. TÍTULO E SUBTÍTULO: Virtual reality for the human-centred design of assistive devices			
6. AUTOR(ES): **Ivan de Souza Rehder**			
7. INSTITUIÇÃO(ÕES)/ÓRGÃO(S) INTERNO(S)/DIVISÃO(ÕES): Instituto Tecnológico de Aeronáutica – ITA			
8. PALAVRAS-CHAVE SUGERIDAS PELO AUTOR: Realidade Virtual; Fatores Humanos; Simulação			
9. PALAVRAS-CHAVE RESULTANTES DE INDEXAÇÃO: ; ;			
10. APRESENTAÇÃO: ( ) **Nacional** (**X**) **Internacional** ITA, São José dos Campos. Curso de Mestrado. Programa de Pós-Graduação em Engenharia Aeronáutica e Mecânica. Área de Materiais, Manufatura e Automação. Orientadora: Prof^a^. Dr. Emilia Villani. Coorientador: . Dr. Edmar Thomaz da Silva. Defesa em 04/08/2022. Publicada em .			
11. RESUMO: A sociedade alcançou tecnologia para criar veículos autônomos e conectar diferentes aparelhos e máquinas umas às outras a fim de trocar informações e otimizar a eficiência de produção. Com essa tecnologia, logo será possível obter melhores métodos para orientar usuários cegos e deficientes visuais (CDV) nas suas atividades diárias. Os produtos que estão disponíveis no mercado hoje em dia possuem um número de limitações e não agradam os usuários CDV. Acredita-se que uma das razões desse problema é a ausência do envolvimento de indivíduos CDV no desenvolvimento desses produtos. A falta de uma solução eficiente para a navegação desse público tornou-se mais grave com a pandemia da SARS-CoV 2, quando pessoas eram instruídas a praticar isolamento social e evitar contato em superfícies que possam estar contaminadas. O objetivo desse trabalho é propor um método para avaliação de opções de design para produtos assistivos para CDV baseados em Realidade Virtual (RV). A ideia é usar o RV como um campo de teste, onde o usuário pode experimentar diferentes soluções em diferentes cenários. Com isso, ele se torna integrante do design e da avaliação, resultando em um produto melhor e com uma interface mais simples. O método proposto inclui, além da montagem do ambiente virtual, o uso de sensores fisiológicos e testes subjetivos que aferem a carga mental e a consciência situacional nas diferentes situações e produtos que estão em desenvolvimento. Para ilustrar o método proposto, é estudado a navegação de indivíduos CDV em um hospital que usa protocolos COVID-19. Esse estudo de caso foi escolhido devido a ocorrente pandemia e a situação crítica que ela causa à população CDV. O cenário virtual foi feito usando Unity3D, uma plataforma de desenvolvimento de aplicações para realidade virtual largamente utilizada. O aparelho RV é o Tobbi Eye Tracking VR. São óculos que foram desenvolvidos usando o HTC VIVE. Esses óculos são utilizados para definir a posição e orientação do usuário no ambiente virtual do Unity. Para inferir a carga mental, foram utilizados os sensores fisiológicos da TEA Capitv T-Sens. Eles são o eletrocardiograma (ECG), usado para coletar a frequência cardíaca e a variância cardíaca, e o GSR (Galvanic skin reaction, reação galvânica da pele), para captar a condutância da pele. Além desses sensores, os voluntários também responderam os testes NASA-TLX, também para verificar a carga mental, e uma versão adaptada do SAGAT, para determinar a consciência situacional. Entre os benefícios esperados pelo método é a flexibilidade e a agilidade para se criar diferentes cenários e também a possibilidade de testar eles no mesmo espaço físico. Isso pode acelerar o design de novas soluções e melhorar a qualidade dos produtos. Outro resultado esperado da pesquisa é a identificação de características chaves dos produtos que causam o aumento ou diminuição da carga mental ou da consciência situacional nos usuários CDV.			
12. GRAU DE SIGILO: (**X**) **OSTENSIVO** ( ) **RESERVADO** ( ) **SECRETO**			