Politecnico di Torino

Corso di Laurea in Automotive Engineering
A.a. 2020/2021
Sessione di Laurea Ottobre 2021

Mixed Reality Application for Inspection and Validation in Industrial Environments

Human Performance and Brain-Computer Interface Advantages over Gestures

Relatore
Dr. Andrea Sanna

Candidato
Silvio Da Col

ABSTRACT

Shortening product development cycles demand increasingly efficient methods and tools for the planning of complex production systems. Recently, augmented reality (AR) technologies have been introduced in manufacturing planning functions. Working in augmented environments, users usually select virtual objects with hand gestures that are associated with arm fatigue.

In the first study, a mixed reality (MR) application for inspection and validation of a production line has been developed. The application allows to show the virtual environment in Augmented/Mixed Reality, and it allows to import three-dimensional (3D) CAD with information and structure in a specific position. The application makes the user interacting with the scene through a user-friendly user interface (UI) and changing the position of 3D objects in the space. The application also allows taking a measurement between two points in the space. The Measurement Tool has been validated, and the absolute average error for dimensions that are lower than 100 cm is 3.59%, while it is 1.30% for dimensions that are higher than 100 cm.

In the second study, a steady-state visual evoked potentials (SSVEP) brain-computer interface (BCI) for “hologram” selection in AR is proposed. The usefulness of the BCI was demonstrated with one experiment in dense and dynamic tasks, a NASA TLX test, and a usability test. On the one hand, the BCI is 2.5 seconds slower than the hand gestures in the static tasks, while the time of selection for the two interfaces is comparable in the dynamic environments. On the other hand, the BCI is more precise, with close to 100% accuracy for all tasks. In addition, the BCI resulted in having a lower overall workload (38.52) compared to hand gestures (52.40) and final usability of 77.8 on the System Usability Scale. The results indicate the potential of a BCI in dense and dynamic environments, demonstrating a possible application for AR technologies in industrial settings.

DEDICATION

Dedicated to my family and my closest friends.

ACKNOWLEDGEMENTS

This work is the highest achievement of my university career and life journey. The willingness to come to Canada for the double-degree program started in my first year of the bachelor. I was very motivated and, despite the multiple issues due to the COVID-19 pandemic, I managed to travel to Canada.

First, a big thank goes to the three entities that organized the “Windsor & Oakland” project: Politecnico di Torino, University of Windsor, and Stellantis. Second, I would like to show my gratitude to all the people involved in this. My University of Windsor advisor Dr. Eunsik Kim, my Politecnico di Torino advisor Dr. Andrea Sanna and my Stellantis advisors Keenan O’Brien, Giuseppe Marino, Raffaele Paradiso and Nicoletta Tataranni. I want to thank my advisors’ collaborators Francesco De Pace and Leah Barton. A big thank also goes to the program coordinators Dr. Giovanni Belingardi, Dr. Maria Pia Cavatorta, Dr. Jennifer Johrendt, and Marie Mills.

In Canada, I had an incredible experience, and I knew people that will always be in my heart. Thanks to Mike, who welcomed me into his house and made me live adventures that could not have been possible without him; he introduced me to many of his friends and his lovely family. Thanks to my program mates, Gabby and Fil; with them, I developed a true friendship that will last in the years, and we have always supported each other during the various lockdowns in Ontario. Thanks to Ellie for everything you have done for me; this would not have been the same without your support and initiative. Thanks to Zac, Yihong, and Jiaqi who warmly welcomed me in Canada.

The greatest thank goes to my family: my father Alessandro, my mother Vittoria, and my brother Dami; they have always supported me in my life choices, and coming here was no exception. Thanks to my best friends from Italy, I always felt them close even if we were more than 7000 km far away. Thanks to Monti il Vez, Giambo, Andrew, Ele, Karlos, Mary, Vale, Gendes, Kwdwuko, Ale, Laura, Luca, Deba, Nat, Miky, Simo, Ste, Sugar, Dany, Giuli, Edo, Fede, Lore, Marco, and Salvo. I cannot wait to see you guys again.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS/SYMBOLS

3D	Three-Dimensional.	19, 20, 25, 26, 27, 29, 42, 74, 85
AR	Augmented Reality.	vii, viii, xi, xv, 16, 18, 19, 20, 21, 25, 26, 27, 28, 31, 32, 33, 37, 38, 39, 40, 42, 52, 53, 59, 62, 70, 71, 74, 77, 81
BCI	Brain-Computer Interface.	vii, viii, 18, 20, 21, 22, 28, 29, 30, 31, 33, 34, 35, 36, 37, 38, 39, 53, 55, 59, 63, 64, 65, 67, 70, 71, 72, 73, 74, 75, 77
CAD	Computer-Aided Design.	31, 48, 52, 53, 74
CNS	Central Nervous System	20, 33
EEG	Electroencephalogram.	33, 34, 35, 36, 38, 53
ERP	Event-Related Potential.	36
MR	Mixed Reality.	16, 18, 21, 41, 46, 52, 69, 74
MRTK	Mixed Reality Toolkit.	41, 52, 53
RV	Reality-Virtuality.	16
SSVEP	Steady-State Visually Evoked Potential.	18, 29, 36, 37, 53, 70, 74
SUS	System Usability Scale.	ix, x, 59, 60, 62, 67, 73, 74, 78
UI	User Interface	20, 27, 41, 87, 93, 96, 105
UWP	Universal Windsows Platform.	40, 42
VR	Virtual Reality.	16, 53

CHAPTER 1 - INTRODUCTION

1.1. Background

Companies are experiencing a dynamic and fast-changing customer demand and a consequent reduction of product life cycles, resulting in the need for fast and flexible re-engineering strategies for production lines. Digital-planning tools, such as augmented reality (AR) headsets, help reduce the time workers spend completing tasks by superimposing computer-generated information onto the real environment. AR applications were initially developed on mobile-based AR devices such as smartphones or tablets. The device camera was used as a mobile system augmentation mean to show the real world while augmenting the virtual objects in the device screen. For example, one of the most popular AR mobile applications is a videogame developed by Nintendo: Pokémon GO. It uses a GPS location-based service allowing the user to hunt the creatures that can be augmented through touch gestures (Pokémon GO, 2021). IKEA's catalog app is another example of a mobile-based AR application. Early versions used the paper catalog as a marker in the space so that, once users want to place furniture in their house, the mobile device recognizes the catalog and places the selected object. With the development of computer vision technology, new versions of this application can also detect floors and walls so that furniture can be placed in the desired position without markers (IKEA, 2021). From an automotive industry point of view, also Volvo Cars is moving towards AR. As the chief technology officer Henrik Green said: "Instead of the usual static way of evaluating new products and ideas, we can test concepts on the road immediately, as well as identifying priorities and clearing bottlenecks much earlier in the design and development process." (Lopez, 2019). Hence, this technology results being feasible also in an industrial environment to shorten product and process development.

The most significant limitation of mobile-based AR applications is that the user must continue to hold the device to superimpose the virtual object (Ro et al., 2019). Since the first AR mobile applications were developed, glasses and head-mounted displays (HMD) have been released. Unlike mobile AR, which shows virtual objects with a narrow display screen to be held with hands, HMD-based AR can be applied to a wide variety of fields because the latter can be worn. Microsoft has further improved the concept of AR with the HoloLens, which provides immersive three-dimensional (3D) screens with high-quality realism.

In this device, the AR concept is integrated with other tools, including spatial mapping, spatial sound, gestures, gaze, or voice interactions to provide visual interface advantages (Bury, 2019).

With these additional tools, the AR experience is enhanced and becomes mixed reality (MR). MR is a larger class of technologies that includes AR and virtual reality (VR). AR and VR are related, and they are concepts that can be considered together (Milgram et al., 1995). On the one hand, AR allows the user to be still present in the real world with some virtual information added. On the other, VR is a technology that allows the user to be immersed in the virtual world. However, instead of regarding these two concepts as distinct, it is more appropriate to view them as part of the same continuum at two opposite ends (Milgram et al., 1995). The reference is to the reality-virtuality (RV) continuum, as shown in Figure 1. An example of spatial mapping for the Microsoft HoloLens is then displayed in Figure 2. This functionality is within the MR reference and allows to enhance the augmented experience. The device tracks the surrounding environment and creates a mesh so that the user can interact with it.

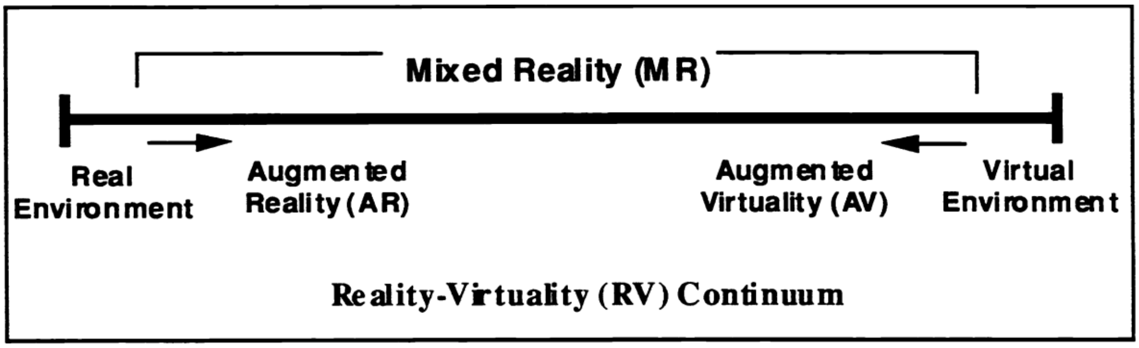


Figure 1: Simplified representation of a RV continuum (Milgram et al., 1995).

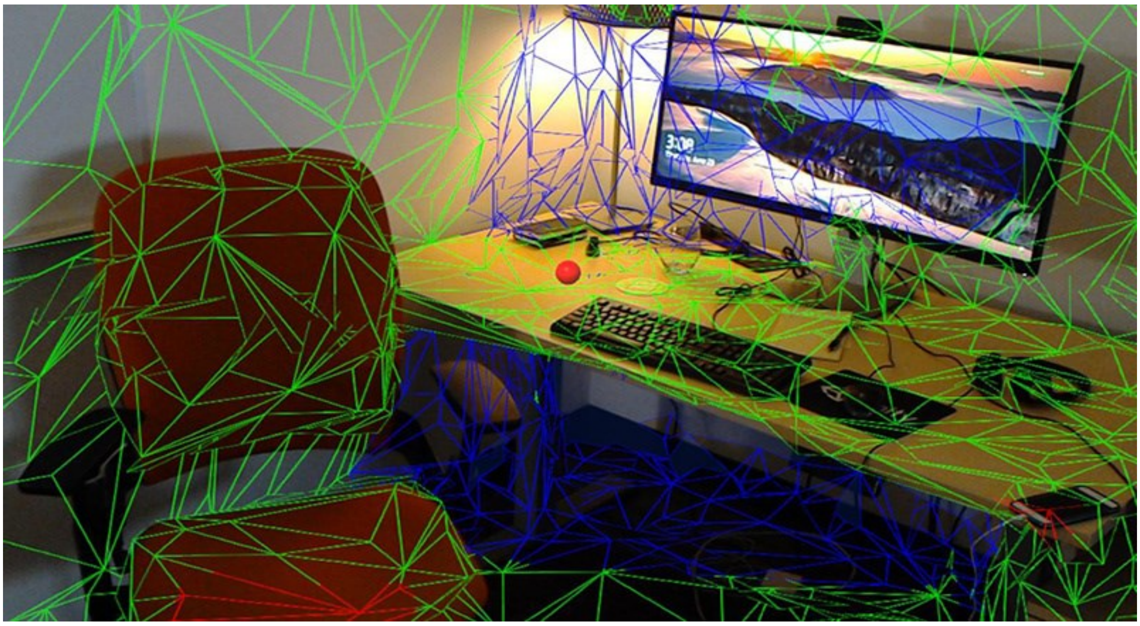


Figure 2: Example of the Spatial Mapping applied to a real-world space (Unity Technologies, 2021)

A natural input method must be found to obtain this experience. This input should be the closest to already present interfaces like smartphones or mouse buttons (Rönkkö et al., 2009). Different possibilities are being explored, such as external controllers, voice recognition, and hand gesture recognition (Zeller, 2019). However, industrial environments are not so suited to external controllers or voice recognition because operators cannot work efficiently with a controller in their hand, and the voice signals are affected by work environment noises as machine tools or human operators (Silaghi et al., 2014).

The most popular input means present on the market are then hand gestures. With this input, the operator can work with free hands, and there is no need for voice inputs. In particular, the Microsoft HoloLens implements the virtual asset selection with hand gestures by Air Tap. A ray is cast from user's head; once the object to be selected is detected, the user keeps the head still while performing the Air Tap gesture, as shown in Figure 3.

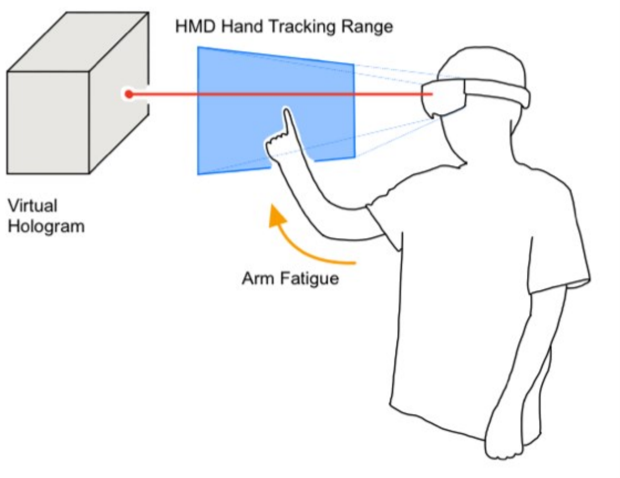


Figure 3: Microsoft HoloLens Air Tap technique.

This input mean is intuitive enough from a usability point of view. The Air Tap technique implemented in the Microsoft HoloLens is similar to the movement of clicking a mouse or tapping a touch screen (Vogel & Balakrishnan, 2005). Nevertheless, this interaction technique comes with some problems. First, hands must be in the field of view of the device, which is usually small. Due to its reduced dimension, users must keep their hands up in the proper position, and prolonged use could be tiring. In addition to this, it is hard and tiring to focus by gazing and keeping the head still while selecting the virtual object (Yilmaz & Kilinc, 2018). In the literature, there are many examples of high perceived workloads while using gestures in AR. There is the need of another alternative to this and so, the research can go in the direction of selecting virtual objects with brain signals by means of a brain-computer interface (BCI).

Firstly, this research aims to introduce MR technologies in the industrial environment using an HMD (i.e., Microsoft HoloLens) for training/support operators in maintenance activities. Secondly, this research focuses on the feasibility of using a steady-state visual evoked potential (SSVEP) BCI for virtual asset selection in an AR environment. The comparison is between a BCI and classic hand gestures. The proposed research is based on accessing if it is possible to increase the usability of an AR interface and thereby increase the speed and accuracy, thus reducing the workload of a selection task.

The following paragraphs will cover more in detail the reason why this research has been employed (Problem Statement), the research idea (Purpose), and the hypotheses behind this idea (Hypotheses). After that, the research will continue with the literature review, the two studies' assets, the discussion about the found results, and the conclusions.

1.2. Problem Statement

Some AR applications have led to a higher than normal workload for users due to their reliance on gestures and the near-constant visual attention required to interact with the holographic interface (Looker, 2015; Ro et al., 2019; Yilmaz & Kilinc, 2018). Even the most advanced HMD AR device, the HoloLens, lacks an interface that allows users' easy interactions (Ro et al., 2019). Although the precise selection and manipulation of virtual assets are among the most important issues in 3D augmented environments (Bellarbi et al., 2017), the gaze-assisted selection interface of the HoloLens is somewhat inaccurate and has poor usability (Chaconas & Höllerer, 2018). As previously anticipated, the asset selection task with Microsoft HoloLens is done by casting a ray from the user's head to the virtual asset object and performing an Air Tap (hand gesture) while keeping the head still. Yilmaz and Kilinc (2018) interviewed participants about their map zoom/pan/rotate methods. Reported comments include: "Air Tap is hard to learn at first […]" and "It is hard and tiring to focus by gaze. Focus point is not stable, and there is a precision problem. […]" (Yilmaz & Kilinc, 2018). Looker (2015) assessed the ergonomics of Air Tap gestures. After reviewing ergonomics literature, they found that the Air Tap movements required by AR interfaces were outside of known anthropometric and biomechanical limits and tolerances (Looker, 2015). Selection techniques that use gestures are more likely than others to cause arm and wrist strain (Argelaguet & Andujar, 2013). Gestures can be used if the user's hand is positioned in the range of vision of the depth camera. As shown in Figure 4, the fundamental limitation of AR is that the field of view of the HMDs is small, and so users must continuously raise their hands up. This causes fatigue, and even if HMDs' field of view is getting bigger every new release, its dimensions are still not enough.

Figure 4: Classic Gestures in AR selection method. The rectangle represents the field of view.

1.3. Purpose

Solutions like external controllers or voice recognition can be implemented to alleviate the arm and wrist strain that results from Air Tap gestures. These solutions would keep the arm in either a lower position or a position that would not even use gestures at all. However, these alternatives are not suitable for industrial environments because operators must work with free hands, and voice signals are affected by work environment noises such as machine tools or human operators (Silaghi et al., 2014).

As a universal interaction task in user interfaces (UI) and a basic interaction technique in AR, 3D virtual object selection has been extensively studied (LaViola et al., 2017). Developers can follow multiple guidelines, but research has clearly shown that there is no best selection technique for all situations. Selection can vary by specific task requirements, work conditions, and user preferences and experience (Cashion et al., 2012). One possible alternative to the classic selection interaction in AR is a BCI.

A BCI is a system that measures the activity of the central nervous system (CNS) and converts it into an artificial output. This output could replace, restore, enhance, supplement, or improve natural CNS output. Thereby this output changes the ongoing interactions between the CNS and its external or internal environment (Wolpaw & Wolpaw, 2012). BCI systems are stable with few head movements, and this is a necessary requirement because AR is usually implemented through HMD (Si-Mohammed et al., 2020).

The purpose of the first study (STUDY 1) is the introduction of MR technologies in the industrial environment. Specifically, use an HMD see-through display for training/support operators in maintenance activities. The main project phases are developing the MR application, testing and optimizing the app in a laboratory environment, and validating it in the industrial environment. The key benefits result in reducing the non-added value activities, knowledge-based assistance, and training of specialists.

The purpose of the second study (STUDY 2) in this research is to compare AR gestures with an active BCI, which is assessed in terms of speed, accuracy, workload, and usability. The proposed research is based on the following question: "Is it possible to use a BCI to increase the usability of an AR interface and thereby increase the speed and accuracy, thus reducing the workload of a selection task in AR?" After deep research, it could be stated that no previous studies have attempted this. For this study, the contribution to the field is expected to be the assessment of an alternative to classic gestures for virtual asset selection in AR. This alternative should fit industrial environments and should work in dense and dynamic virtual objects conditions. A user study has been conducted to determine which is the best performing selection method.

1.4. Hypotheses

Before conducting the study, some hypotheses have been established. The hypotheses are four and are listed in the following.

- H_1: BCI will increase the speed associated with the selection task.

This statement is based on Human Performance Theory. There may be many reasons why one interface is faster than another. A model of human information processing can help identify different design solution features by recording mental and motor processes in users (Dix et al., 2003). The key difference between hand gestures and BCI is in response selection. On the one hand, hand gestures involve both brain and muscles in the selection task. On the other hand, BCI involves only the brain. Therefore, by removing arm movement, the speed of selection should increase.

- H_2: BCI will increase the accuracy associated with the selection task.

This statement is based on the Error Theory of commission and omission errors. The error of commission is a mistake that consists of doing something wrong, thereby committing an action erroneously. As hand gestures and raycasting are always associated with movement difficulties, the accuracy should increase, and commission error will be eliminated by eliminating the movement.

- H_3: BCI will perform worse in the mental demand (i.e., have higher workload), but it will perform better in the physical demand (i.e., lower workload).

Workload is one of the most widely-invoked concepts in human factors research and practice (Wickens et al., 2015). It is constrained by the limited information processing capacity of the brain. High workload can lead to worker stress, errors, and performance decline when a user's sense of effort is maximum (high workload) (Wickens et al., 2015).

- H_4: BCI will increase the usability of the selection task.

Usability is not a one-dimensional property of a UI. Rather, it can be distinguished by five key attributes: learnability, efficiency, memorability, errors, and satisfaction (Karwowski et al., 2003). Based on the previous hypotheses and the key attributes of usability, interface usability is expected to increase with BCI.

CHAPTER 2 - LITERATURE REVIEW

2.1. Industry 4.0

Technological advances have driven dramatic increases in industrial productivity since the beginning of the Industrial Revolution (Rüßmann et al., 2015). In the 19th century, the steam engine power factories started spreading; in the early 20th, the electrification for mass production; in the 70s of the same century, the industrial automation. As displayed in Figure 5, industry manufacturing advanced from water and steam power machines to electrical and digital automated production. This last environment made manufacturing processes more complicated, automatic, and sustainable so that people can operate machines simply, efficiently, and persistently (Qin et al., 2016). The term "Industry 4.0" stands for the 4th industrial revolution, which is defined as a new level of organization and control over the entire value chain of the life cycle of products (Rüßmann et al., 2015).

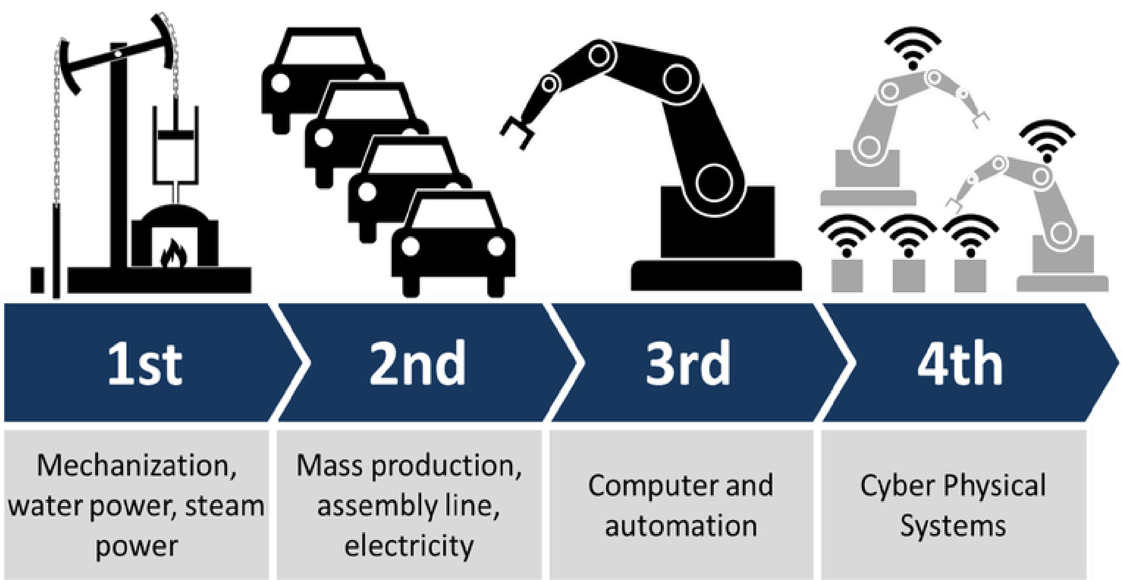


Figure 5: Four Industrial Revolutions (Stilgherrian, 2018).

The central objective of Industry 4.0 is fulfilling individual customer needs, which affect areas like order management, research and development, manufacturing commissioning, delivery up to the utilization, and recycling of products. As displayed in Figure 6, there are nine pillars of technological advancement that give form to Industry 4.0: Big Data and Analytics, Autonomous Robots, Simulation, Horizontal and Vertical System Integration, Internet of Things, Cybersecurity, Cloud, Additive Manufacturing, AR (EITC, 2021). These tools, from isolated and optimized cells, can be used together as a fully integrated, automated, and optimized production flow, leading to greater efficiencies and changing traditional production relationships among suppliers, producers, customers, as well as between human and machine (Rüßmann et al., 2015).

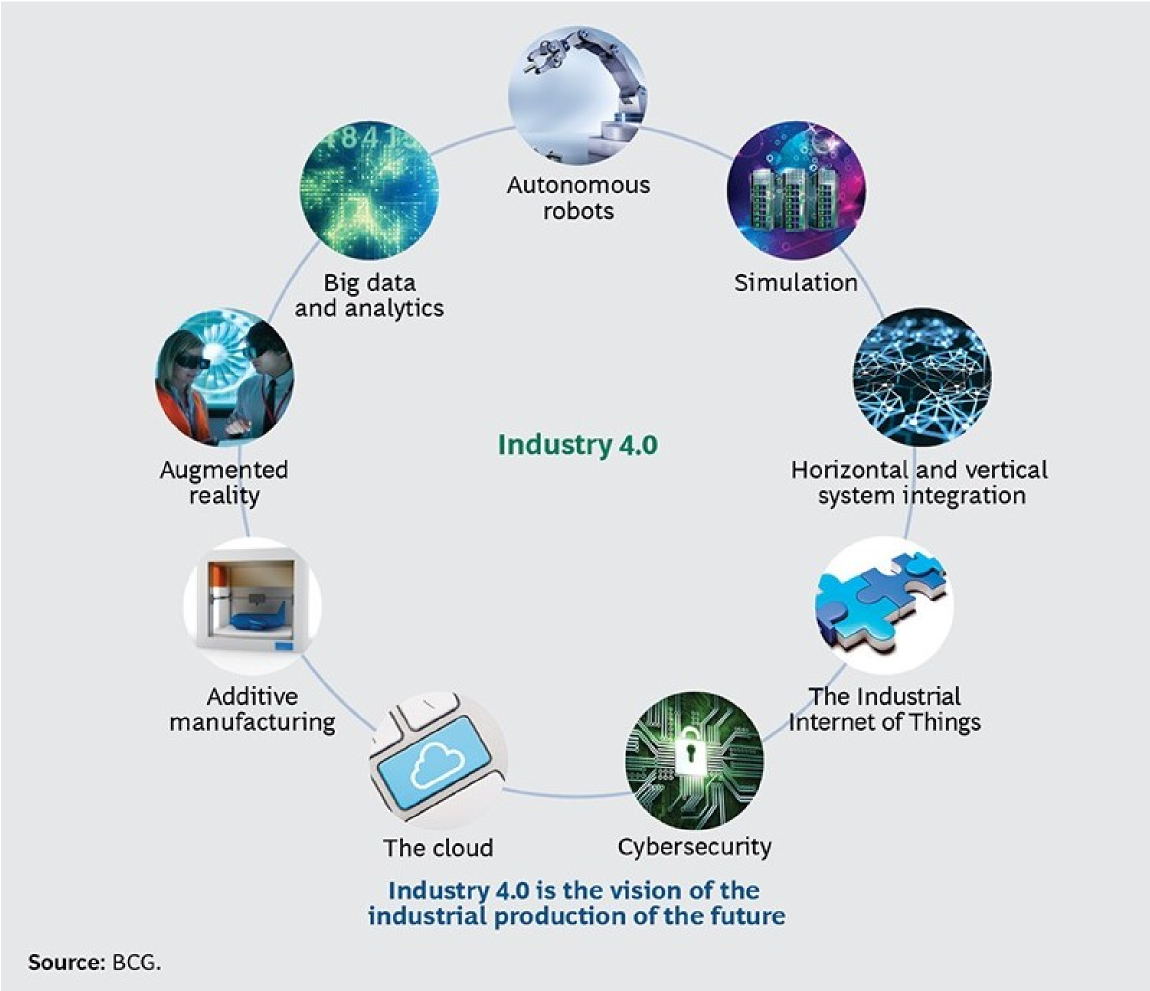


Figure 6: The 9 Pillars of Technological Advancement in Industry 4.0 (Rüßmann et al., 2015).

Big Data and Analytics: this block includes all the collection and comprehensive evaluation of data from many different sources. The data could come from both production equipment and system or the customer management system.

Autonomous Robots: these types of robots can interact with each other, and they can also be collaborative and work side by side with humans.

Simulation: simulations mirror the physical world and are extensively used in plant operations and prototyping. These models allow operators to test and optimize several variations, increasing quality, and reducing development time.

Horizontal and Vertical System Integration: company departments and functions can become more cohesive, creating universal data-integration networks.

Industrial Internet of Things: this process allows devices to communicate and interact with each other and with more centralized controllers.

Cybersecurity: this block aims to protect critical industrial systems and manufacturing lines from cyber attacks. It is essential to have secure and reliable systems.

Cloud: all the information must be available in real-time. Companies continuously deploy machine data and analytics to the cloud, enabling more and more data-driven services.

Additive Manufacturing: one of the typical examples of this is 3D printing. Companies can produce small batches of customized products that are fast to be manufactured and lightweight.

2.2. Augmented Reality

AR is a technology that generates and superimposes virtual information on a user's view of the real world, and it supports a variety of services. This technology is key for enabling industry 4.0 concepts and driving the development of industrial environment concepts. AR allows the employees to link the existing gap between the physical world and more important digital environments. The annual growth rate of the industrial AR market is constantly rising, and it is forecasting to grow more within 2025 (Egger & Masood, 2020). This important growth is likely to be enhanced because AR technology is becoming more reliable, and more applications are being developed. While the general importance of AR is widely recognized, recent research has shown that the implementation for the industry is challenging (Egger & Masood, 2020).

Nevertheless, AR technology has many potential applications (De Pace et al., 2018; Kim et al., 2017). Wang et al., 2019 analyze the use of AR in *computer-aided surgery*. In this paper, a marker-less image registration method is presented. In particular, AR is used to align the virtual scene with reality in guided maxillofacial surgery. The virtual scene is created and coupled with reality to guide surgical operations or provide the surgery outcome without any marker. A 3D scanner is used to acquire the patient's teeth shape model to obtain the augmented environment. This model is then registered with a stereo camera system using an algorithm. By using together all these elements, the surrounding anatomical models and virtual implants can be superimposed on the camera's viewpoint to navigate in the AR environment (Wang et al., 2019). Some experiments were employed to understand the error of the system. The average error was less than 0.50 mm, and clinical feasibility has been shown with a volunteer. This application allows overcoming the misalignment difficulty caused by the patient's movement because no markers are involved, and therefore it is non-invasive and practical (Wang et al., 2019).

Computer-assisted instruction (Deshpande & Kim, 2018; Tang et al., 2003) is another interesting application where AR assists the user in an assembly task. Tang et al. (2003) tested the relative effectiveness of AR instructions in an assembly task. They compared a printed manual with computer-assisted instructions (using a monitor-based display) and an AR system. In comparison to the other methods, results indicate that superimposing 3D instructions on the actual work pieces reduced the error rate of the assembly task by 82%. What is reduced is especially cumulative errors, the ones that are due to previous assembly mistakes. It is worth noting that the researchers also measured the mental workload which decreased in the AR condition, implying that some mental activities are offloaded by the system.

Another application for AR can be *interior design modeling*. Chang et al. (2020) presented a mobile AR application that supports teaching activities in interior design. The aim is to support students in learning interior layout design, interior design symbols, and the effects of different layout decisions. Using this application, users can place 3D models of virtual objects on their mobile devices. The test has been employed by comparing an experimental group with a sample of candidates in a control group. The comparison has been made under the points of view of learning interest, confidence, satisfaction, and motivation of the students. The results have been determined with a t-test and significance has been found. Therefore, it can be concluded that this technology does enhance students' learning ability for interior design.

Laboratory experiments, in general, indicate that AR-supported tasks are more efficient in task completion time and error rates (Sanna et al., 2015). It has been found that the increase in performance through AR depends on the complexity and nature of the task (Henderson & Feiner, 2011). Different tasks give different performances, which are influenced by the experience of the workers (Funk et al., 2017).

AR technology has some limitations coming when talking about input means. The classic AR interactions are obtained with hand gestures that usually lead to a higher than normal workload (Ro et al., 2019). Hand gestures, together with the near-constant visual attention required to interact with the virtual assets, cause fatigue (Looker, 2015; Ro et al., 2019; Yilmaz & Kilinc, 2018). Yilmaz & Kilinc (2018) reported feedbacks from the participants of their experiment. They pointed out how difficult was the use of the Air Tap (hand gesture) and the focus by gaze on the virtual objects for selection purposes. Also, Looker (2015) assessed the ergonomics of Air Tap gestures. After reviewing ergonomics literature, they found that the Air Tap movements were outside the known anthropometric parameters.

2.2.1. Universal Tasks in AR

Selection, positioning, rotation, and scaling are common manipulation tasks in 3D interfaces. *Selection* is the task of getting or identifying a particular asset or subset of assets from the entire set of assets available; *positioning* is the task of changing the 3D position of an object; *rotation* is the task of changing its orientation; *scaling* is the task of changing its size (LaViola et al., 2017). For each of these tasks, there are many possible interaction techniques like different combinations of input devices and UI software. Universal tasks may be based on real-world actions or virtual interactions that enhance capability (LaViola et al., 2017). How these basic tasks are usually implemented is relevant for the quality of the entire 3D UI. The human hand is a considerable tool because it allows to select and manipulate virtual objects with precision and speed. Therefore, the research for new interaction techniques is an important driver for 3D UIs (LaViola et al., 2017). The goal is to develop new techniques or reuse the existing ones, trying to enhance user-manipulation performance and comfort (Knight, 1987).

2.2.1.1. Selection

One of the fundamental tasks in 3D interfaces is virtual asset selection (LaViola et al., 2017). This is also the initial task for the most common user interactions in a Virtual Environment. Manipulation tasks often depend on, and are preceded by, selection tasks. Therefore, poorly designed selection techniques often have a significant negative impact on the overall user performance. For each canonical task, many variables significantly affect user performance and usability (Foley et al., 1984). Each task defines a “task space” that includes multiple variations of the same task. Some variables influence user performance like the distance and direction to target, the target size, the density of objects around the target, the number of targets to be selected, or the target occlusion (Poupyrev et al., 1997). Different selection techniques allow obtaining different results in different environments.

Many studies have attempted to classify and improve the selection of virtual assets in 3D environments (Argelaguet & Andujar, 2013). In the literature, the first research group tackling this topic is Bowman et al. (1999), who proposed a classification schema for 3D selection techniques. In particular, 3D selection techniques for mobile AR HMD can be done by visual raycasting or by hand-controlled 3D cursor (Özacar et al., 2016). In the first case, a ray is cast from the point of origin with a direction. These two values can be found by tracking the position and orientation of a controller (Feiner, 2003), the user’s hand (Bowman et al., 2001), or the user’s head (Tanriverdi & Jacob, 2000). In the second case, the user’s hand is the focus for the selection task. Using quick hand gestures is another approach to confirm the selection. Grossman & Balakrishnan (2006) proposed a thumb trigger gesture where the thumb finger moves in and out toward the index. A similar approach is the one adopted by Vogel & Balakrishnan (2005) that developed the Air Tap technique. This selection method is conceptually close to the previous, but the trigger is the index that closes on the thumb for the selection, as displayed in

Figure 7. This selection technique is adopted in the Microsoft HoloLens, and it will be the hand gesture that will be compared with the BCI in this study.

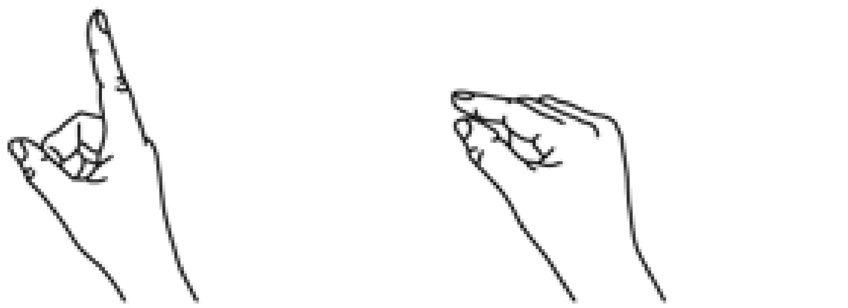


Figure 7: Air Tap gesture implemented in the Microsoft HoloLens (Zeller, 2019).

In addition to these selection techniques, other methods can be implemented. These are the selection by voice recognition (Silaghi et al., 2014), gaze-assisted selection technique (Sidenmark et al., 2020), mobile phone selection technique (Ro et al., 2019), or external controllers selection techniques (Grossman & Balakrishnan, 2006). All of them have some benefits, but they cannot be used in an industrial environment. For example, to alleviate the arm and wrist strain that results from hand gestures, these selection techniques would keep the arm in either a lower position or in a position that would not use gestures at all. However, these alternatives are not suitable for industrial environments because operators must work with free hands, and voice signals are affected by work environment noises such as machine tools or human operators (Silaghi et al., 2014).

In this context, Bowman et al. (1999) study was particularly important because it suggested several taxonomies to classify existing 3D selection techniques. This research study classified different interaction techniques and decomposed them into subtasks. This classification is done according to Figure 8. A selection technique must provide a way to indicate an object (*object indication*), a mechanism to confirm its selection (*confirmation of selection*), and visual, haptic, or/and audio feedback to guide the user during the selection task (*feedback*). The taxonomy provides a broad view of selection techniques, even if it considers a relatively small number of design variables.

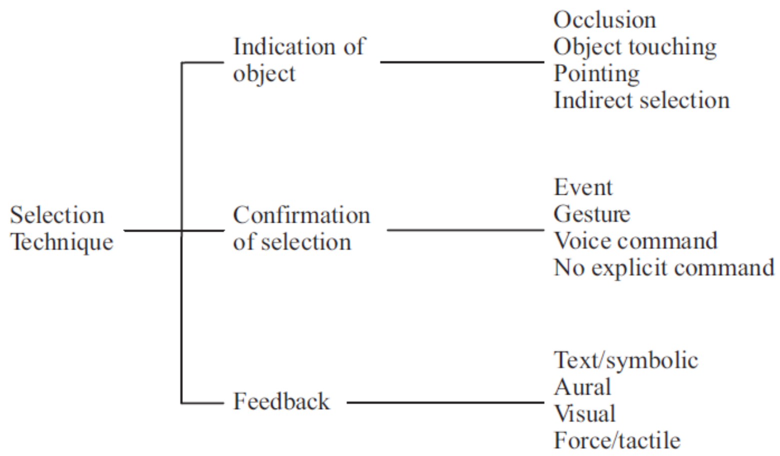


Figure 8: Classification of selection techniques by task decomposition proposed by Bowman et al. (1999).

The employed experiment was a selection and manipulation testbed. It consists of users selecting the correct object within a group of them and manipulate it with different selection techniques. The variables were the distance from the user, the size of the objects, and the density of objects. These factors are important in determining speed, accuracy, ease of use, and comfort for selection techniques (Bowman et al., 1999). The comparison has been made among nine different selection/manipulation techniques composition: one is the Go-Go technique while the other eight compositions are made up with two selection techniques raycasting and occlusion; two attachment techniques move hand and scale hand; and two manipulation techniques, linear mapping, and buttons. Raycasting was the fastest technique staying within the selection field of investigation (Bowman et al., 1999). In this research (STUDY 2), the comparison is between hand gestures (Microsoft HoloLens) and an SSVEP BCI (NextMind). In the first case, the indication is done by casting the ray from the head to the virtual asset, and the confirmation is done by gesture (Air Tap). This selection technique is in line with the experiment of Bowman et al. (1999), in which the raycasting method was the fastest. In the second case, the indication is done by gazing at the virtual asset, while the confirmation is done by no explicit command because the BCI is triggered, and it depends on the users' visual focus on the asset. According to the literature review, most previous studies are based on hand gestures and Air Tap assessment. STUDY 2 will broaden the selection channels in an augmented environment, adding an alternative to the classic hand gestures with a BCI.

2.2.2. Human Performance

Human performance can be explained under three main measures: speed, accuracy, and attentional demand (Wickens et al., 2015). Generally, for a certain task, the higher the speed, the better, the higher the accuracy, the better, and the lower the attentional demand, the better. Thus, it could be thought that the perfect engineering principle allows obtaining the higher speed and accuracy, and the lower attentional demand. But, of course, these measures are always a trade-off, and in some applications, one measure is better to be optimized with respect to the others (Wickens et al., 2015).

The different dimensions of performance, namely speed, accuracy, and attentional demand, assist the understanding of how much the performance is changed in the different design conditions. In our case, the difference is between hand gestures and a BCI. There could be more reasons why an interface is faster than another or why it is less accurate. Mental and motor processes can help to identify different design solution features. What is needed is a model of human information processing, as depicted in Figure 9.

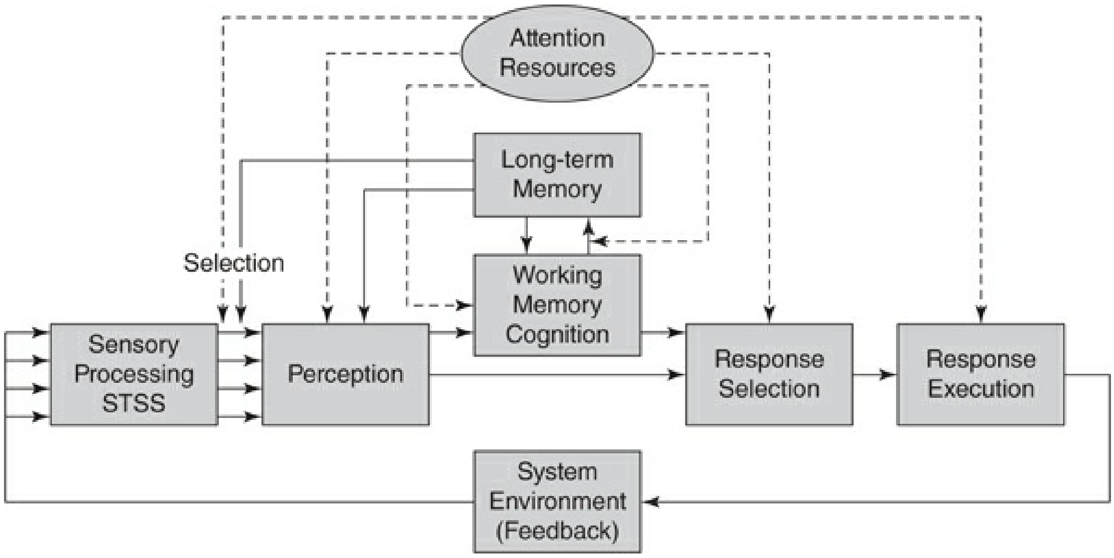


Figure 9: Model of human information processing (Wickens et al., 2015).

This model provides a useful structure to analyze the different mental and motor processes that typically characterize the flow of information. The difference between hand gestures and the BCI is in the “Response Selection” and “Response Execution” blocks. On the one hand, hand gestures involve both brain and muscles in the selection task for both response selection and execution. On the other hand, BCI involves just the brain, so the selection speed should increase as anticipated in the first hypothesis (H_1).

The errors that users make in this circumstance are errors of commission. As stated by the American Psychological Association, an error of commission is a category of human error in which an operator performs an incorrect or additional action, such as pressing a control button twice, leading to inappropriate or duplicate performance of a function (APA, 2021). Hand gestures are always associated with difficulties in movement (Ro et al., 2019). Hence, errors should decrease by eliminating the hand movement, as stated in the second hypothesis (H_2).

2.2.3. Advantages of AR in Industrial Environment

Product life cycles are shrinking due to a market that requires product launches and updates rapidly. Therefore, production lines must become flexible and fast to be re-engineered. The 3rd industrial revolution brought computers in the development stages and, with it, electronic planning tools help decrease the re-engineering time of plants and plant equipment. Different tools are used to plan plants and manufacturing processes. Examples of this are computer-aided design (CAD) systems or simulation systems to simulate material flow and plant layout design systems. In this phase, there is the need for a high-quality representation of the plant layout. Every detail must be considered, obtaining the best reproduction of the real factory, and so succeeding in planning activities like material-flow simulation. Frequently, the layouts used in the CAD and simulation tools do not perfectly fit the real industrial environment, which leads to defects within the planning process and, therefore, to costly replanning activities. Validation tools can tell how close the layout is to the real factory, but this is a time-consuming operation (Doil et al., 2003).

The market-driven rapid product changes oblige manufacturers to find new solutions to shorten the product development and reduce as much as possible time-consuming activities. One solution to have faster processes is AR. The areas of use of AR in the industrial environment range from a higher accuracy in the assembly procedures and flexibility in training with a live guidance system (Egger & Masood, 2020). The use of AR makes it easier for the workforce to obtain real-time data, and the use of all the validation tools to access the factory layout accuracy can be avoided. When working with AR, the real environment is still visible, and it can be used as a direct reference, as displayed in Figure 10.

Currently, numerous companies are seeing AR as a very important tool to provide new services (Zubizarreta et al., 2019) and, in fact, this technology can have multiple applications (Egger & Masood, 2020). For example, AR can be used for assembly operations in intelligent manufacturing, either in training (Sorko & Brunnhofer, 2019) or as a guidance system for operators (Reyes et al., 2020). The industrial planning process can benefit from AR, thanks to the reliable representation of the surrounding industrial space (Doil et al., 2003). Another area where AR can be used is for in-house logistics (Wang et al., 2020) or also in quality assurance (Segovia et al., 2015). As soon as operators depend on or can profit from real-time information, AR can be used to intuitively display this information on site.

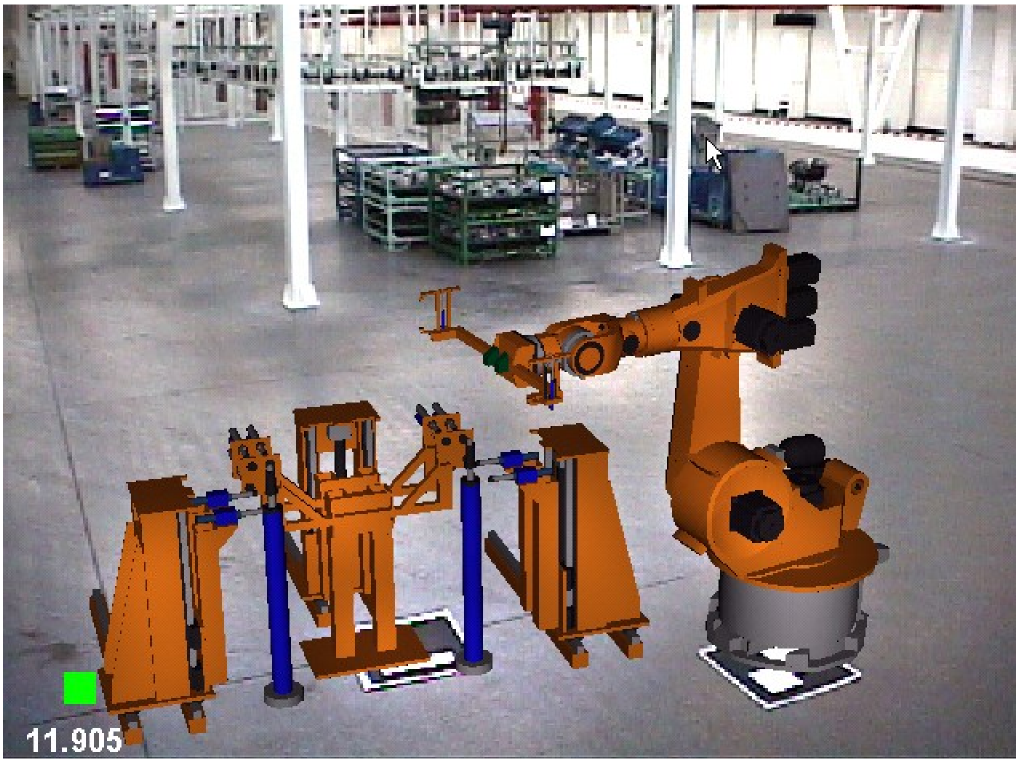


Figure 10: Visualization of virtual robots and machinery in a plant-environment (Doil et al., 2003)

From this perspective, the most relevant study is the one conducted by Doil et al. (2003), who proposed an AR system to improve the industrial planning process that would allow planning tasks to be validated without modeling the surrounding environment of the production site. Their AR model improved factory interfaces by superimposing virtual planning objects on the work environment (Doil et al., 2003). This led to significant advantages in the visualization of the final production line. Planning tasks can thus be validated without modeling the surrounding environment of the production site.

The Major Automotive Company provided some information for this study about the advantages of using AR in an industrial environment. In defining a new production line or modifying an already existent production line, AR technology brings many advantages in resources optimization. The major benefits are the reduction of time and cost of the line, as well as increased safety. AR-systems allow anticipating as much as possible the criticalities associated with the project realization. It is possible to make verifications of the project plan directly on the production site and verify the equipment's correct positioning and overall dimensions relative to the factory layout. In this way, it is possible to highlight technological issues with respect to the initial project idea: details that were not considered. These details may be space taken for equipment handling, possible interferences with machines that are already present on the line, or dimension enslavements. Modifications can be done in advance with respect to the actual realization, reducing the number of in-person inspections, late modifications that would be time and money-consuming, and increasing the safety on the worksite, limiting the time spent on the line by operators (De Pace et al., 2018).

2.3. Brain-Computer Interface

In 1924, Hans Berger, Professor of Psychiatry at the University of Jena in Germany, discovered that electrical signals of the brain could be recorded from the scalp. Berger published in 1929 an article that established electroencephalography (EEG) as a basic tool for clinical diagnosis and brain research. The possibility of individuals acting through brain signals rather than muscles has fascinated people for many years. Almost a century after Berger's discovery, BCIs have become a reality.

The central nervous system (CNS) function is the starting point to explain what a BCI is. Its purpose is to respond to events in the outside world of the body by producing outputs that serve the organism's needs (Wolpaw & Wolpaw, 2012). All the outputs of the CNS are neuromuscular or hormonal. A BCI allows the CNS to have additional outputs rather than the previous two. A BCI is a system that processes the activity of the CNS and converts it into an artificial output. This output could replace, restore, enhance, supplement, or improve natural CNS output. Thereby this output changes the ongoing interactions between the CNS and its external or internal environment (Wolpaw & Wolpaw, 2012). BCI interfaces collect brain signals, understand them and interact with an external device or machine (Placido, 2021).

As displayed in Figure 11, BCIs can *replace* the natural output of the CNS that could not be used anymore; for example, a person who can no longer speak might use a BCI to type words that are then spoken by a speech synthesizer. This system could also *restore* a lost natural output; for example, when person's limbs are paralyzed, the BCI could stimulate the muscles with implanted electrodes so that the limbs can move. A BCI can *enhance* natural CNS output by warning a driver to stay focused over a prolonged period. This system can then *supplement* users by allowing them to control a robot arm to reach far spots or lift heavyweights. Finally, a BCI can *improve* natural CNS output, for example, in a person with impaired arm movements who uses the system to measure brain signals which stimulate muscles or control an orthotic device to improve arm movements.

BCI research started back in the 1970s when Vidal (1973) made the first systematic attempt to implement an EEG-based BCI to record the evoked electrical activity of the cerebral cortex from the intact skull. The research applications of BCI technology evolved significantly over the years (Saha et al., 2021), including detecting drowsiness for improving human working per-

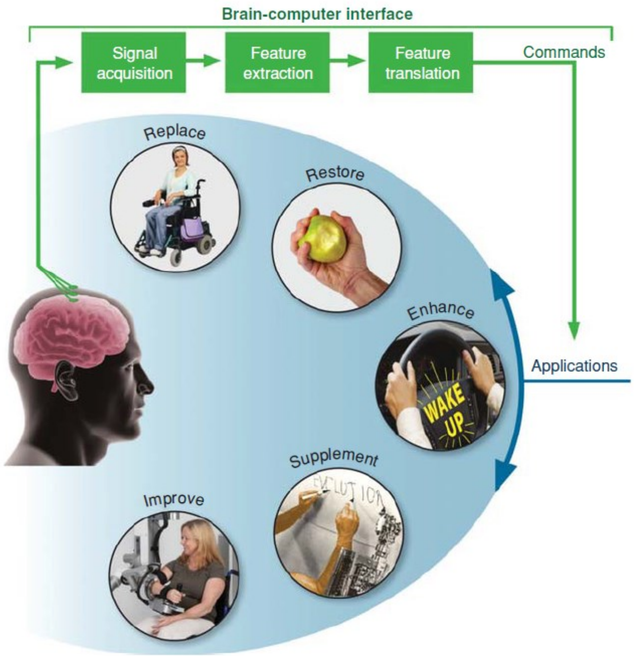


Figure 11: Basic design and operation of a BCI system (Wolpaw & Wolpaw, 2012).

formances (Wei et al., 2018), access the selection of a virtual object (Felton et al., 2012), estimating reaction time (Wu et al., 2017), controlling VR (Vourvopoulos et al., 2019), videogames (Singh et al., 2020), or driving humanoid robots (Spataro et al., 2017). Figure 12 shows the exponential trend of BCI-related publications over the years (Saha et al., 2021).

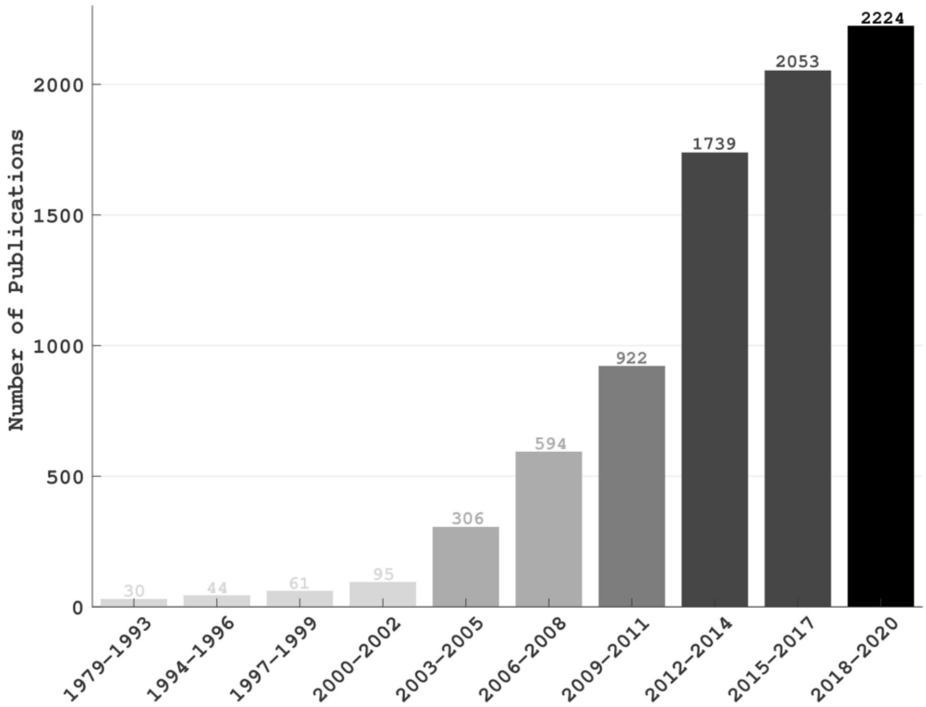


Figure 12: Number of BCI-related publications over the years (Saha et al., 2021).

It is of particular relevance the study conducted by Felton et al. (2012). They proposed a study to access people's perception in performing BCI tasks, which is the mental workload aspect contributing the most, and if there is a perception difference in perceived workload between able and disabled participants. They have also done this by comparing the brain and manual controllers such as mouse or joystick. The employed BCI was an electrode cap that recorded EEG signals, whereas the tool used to assess mental workload was the NASA TLX. The results concluded that able and disable participants have the same perception with similar NASA TLX scores. From the point of view of the comparison between BCI and manual controllers, the major contribution in the NASA TLX scores was the mental demand for the BCI, while it was the physical demand for the joystick. This result is correlated with the third hypothesis (H_3), according to which the BCI will have a higher mental workload but lower physical workload with respect to hand gestures.

2.3.1. Type of BCI

According to Placido (2021), a BCI can generally be divided into three classes: invasive, semi-invasive, and non-invasive interface. The *invasive* interfaces are electrodes positioned into the grey matter, which can measure the activities of the neurons. The quality of the signal is the highest, but there is a risk connected to the surgery. In a *semi-invasive* BCI, the electrodes are implanted inside the skull, but they are kept outside the brain. The *non-invasive* interfaces are instead easy-to-wear devices that do not require surgery and are placed on the scalp. According to Putze et al. (2020), BCIs can also be classified as passive and active. *Passive* BCIs focus on two aspects which are often tackled for adaptive technology: attention and workload. For example, when drivers become distracted or inattentive, the BCI measuring their neural activity can warn about focusing on the task again. *Active* BCIs are instead interfaces used for active control and used to trigger an action like moving a robotic arm or controlling a virtual environment.

For what concern the non-invasive BCIs, there are several types of these techniques as magnetoencephalography (MEG), positron emission tomography (PET), functional magnetic resonance imaging (fMRI), near-infrared spectroscopy (fNIRS), and EEG. The last one is the most commonly used because of cost and hardware portability. EEG provides the recording of the brain's electrical activity from the surface of the scalp, and it is possible to use it in many applications. The first example is BCIs which use *P300 event-related potentials* (ERP). ERPs are manifestations of neural activity that are triggered by specific events. The P300 is a positive deflection that occurs in the EEG recorded in the scalp after a stimulus. This often occurs at a latency of about 300 msec. Current P300-based BCIs allow users to select items displayed on a screen and the selection method is similar to a standard computer keyboard. Another example is

BCIs that use *sensorimotor rhythms,* which are the evidence that the execution or imagination of limb movement changes the potential recorded in the sensorimotor cortex. These signals are extracted and translated, and the resulting output is then used for moving robotic arms in one, two, or three dimensions. The last example is BCIs that use *SSVEP* that are distinctive patterns of positive and negative voltage deflections. The most prominent deflections include N70 and P100, which occur about 70 and 100 msec after the visual stimulus, respectively. With this kind of interface, the user typically selects by gazing at the stimulus that represents the desired BCI output. The BCI evaluates the signal's frequency spectrum, and the latter usually matches the rate of stimulus on which the user is fixating. In this way, users can select objects or buttons which give them a specific visual stimulus. The BCI device used in the second study (STUDY 2) will be of this last kind, and it will rely on machine learning (Protalinski, 2020).

2.3.2. Coupling of AR and BCI

BCIs have enabled individuals to control devices such as robotic arms, drones, or wheelchairs. With the advent of AR systems, these technologies can be coupled to offer immersive scenarios through induced illusions of an artificially perceived reality. The two systems together can provide additional communication channels by increasing the possibilities of interaction of the human with AR. According to Putze et al. (2020), regarding passive BCI, research concentrates on two aspects that are always touched for adaptive technologies: attention and workload of the user. On the other side, active BCI research is less about the potential improvement of the AR interface. However, it focuses on the enhancement of the BCI paradigm in different applications like rehabilitation. The second study (STUDY 2) of this thesis deepens the possible benefits of using an active BCI to select objects in an AR environment. Many other potential applications of this coupling have yet to be explored.

Research on the combination between AR and BCIs has focused on different aspects. Angrisani et al. (2020) and Coogan and He (2018) analyzed *Internet of Things* applications of the coupling between AR and BCI. For example, the study conducted by Angrisani et al. (2020) reported a measurement instrument for a wearable BCI. This AR-BCI system is employable in industrial inspection, where users can communicate with smart transducers without using hands or voice. In this paper, the researchers assessed the minimum stimulation and acquisition time for the SSVEP in a single-channel BCI. Even if the paper proposes a BCI to replace the classical input interface of AR platforms, the focus is more on the interaction with the real plant equipped with smart sensors. Furthermore, they do not analyze and compare the performances of hand gestures and BCI. The experimental results demonstrate that the minimum stimulation and acquisition time for the SSVEP signals in a single-channel BCI can be as low as 2.0 s with accuracies up to

90.0% and 100.0% for some subjects. This paper suggests that SSVEP BCIs can have a decent selection speed even if they can lead to some delays.

The coupling between AR and BCI could also be used to interact with superimposed virtual assets to trigger an action with brain signals, as presented by Putze et al. (2019). This paper proposed HoloSSVEP, a smart home control system that uses the Microsoft HoloLens camera to position controllable elements within the environment, marked by visual identifiers automatically detected as displayed in Figure 13. The BCI records users' EEG signals while another system tracks their eye gaze. The comparison between the performances of hand gestures and BCI is not present. This hybrid interface with AR, BCI, and eye-tracking, ended up having high accuracy in the selection stage. The accuracy of the BCI is on average 76.1%, the one of eye tracking is 82.1%, while the accuracy of the two systems combined is 89.3%.

Figure 13: Window Blinds control via AR using the SSVEP-BCI paradigm in a room (Putze et al., 2019).

Coupling a BCI to AR can also improve user attention while performing a task, as suggested by Vortmann and Putze (2020). They follow the idea that HMD has a distraction problem caused by an unavoidable display of control elements even when focused on internal thoughts. In this paper, they tried to reduce the distraction by including information about the current attentional state. They divided the attentional state into internal and external and adapted an application by making it respond just in case the attention state would have been classified as “external.” They concluded that systems would benefit from this attention division because the application responds just if the user is “active.”

Most previous studies focus on accessing the positive influence that a BCI can have on the user while working in an AR environment. Nevertheless, after deep research, it could be stated that no previous studies have attempted to compare hand gestures and BCI selection in an augmented or virtual environment. STUDY 2 will contribute to assessing the feasibility of using a BCI in dense and dynamic conditions, focusing on increasing the speed and accuracy, thus reducing the workload of a selection task in AR. Together with the increasing popularity of AR in the industry, longer tasks in augmented environments will have to be completed. Therefore, this study is needed to access an alternative to classical hand gestures to reduce the effort while operating.

CHAPTER 3 - STUDY 1

3.1. Technology

3.1.1. Microsoft HoloLens

The Microsoft HoloLens (Microsoft Corporation, United States) is the first HMD running the Windows Mixed Reality platform under Windows 10, which supports Universal Windows Platform (UWP) apps. It is worn like a helmet, and with it, it is possible to project onto the real environment holograms which enhance the experience. The Microsoft HoloLens has been selected for this study because it represents the highest standard for the AR technology. An example of this device is displayed in Figure 14.

Figure 14: Microsoft HoloLens (1st Generation) (Zeller, 2019).

The Microsoft HoloLens is more than just an AR HMD. Its technology is called "Mixed Reality", which utilizes multiple sensors and holographic processing optimized with its environment. The virtual assets can display information, represent the real world with a mesh, or even simulate the virtual world. This device is based on an Intel 32-bit architecture with TPM 2.0 support. It is equipped with 64 GB of flash memory and 2 GB of RAM; the network connectivity features Wi-Fi 802.11ac and Bluetooth 4.1 LE. The mounted sensors are 1 inertial measurement unit, 4 environment understanding cameras, 1 depth camera, and a 2MP photo/HD video camera. The display has a holographic resolution producing 2.3M total light points and a holographic density greater than 2.5k radiants. The device capabilities are Gaze Tracking, Gesture Input, and Voice Support (Zeller, 2019).

3.1.2. Unity

Unity is a game engine (Unity Technologies, United States) that allows cross-platform integration. It utilizes what is called game objects, components, and scenes. A game object can be of any form like shape, model, and UI element, or it can be empty to contain just a component that has a specific function. In addition, components can be added to game objects to provide functionality, such as C# scripts, position/rotation values, or standard Unity pre-build components. A scene, in the end, is a collection of game objects with which the user can interact. An example of the Unity environment is displayed in Figure 15.

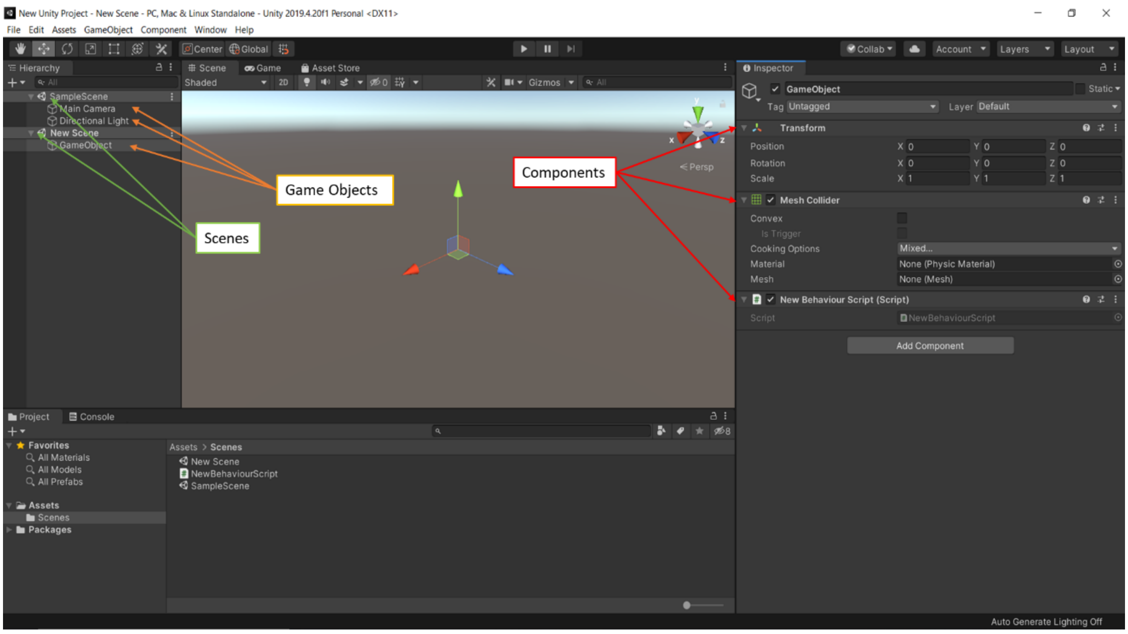


Figure 15: Unity game engine interface.

Unity can be integrated with assets that can be downloaded from the Asset Store or the official websites of the provider. These assets are called Software Development Kit (SDK), and they facilitate developers' development speed and creation time. For the application development, Unity Version 2019.4.20f1 has been employed.

3.1.2.1. Mixed Reality Toolkit (MRTK)

From the official Microsoft website (Semple, 2021): MRTK-Unity is a Microsoft-driven project that provides a set of components and features used to accelerate cross-platform MR app development in Unity. The kit provides a cross-platform input system and building blocks for spatial interactions and UI. Having at hand many standard game objects, this enables rapid prototyping via in-editor simulation that allows to see changes and swap out core components quickly.

3.1.2.2. Vuforia

Vuforia Engine is a popular AR development platform (Vuforia, 2021), and it allows the creation of AR applications. This engine, coupled with Unity, allows developers to easily add advanced computer vision to UWP apps and create AR experiences that realistically interact with objects and the environment. This engine uses computer vision to acquire and track planar markers, called image targets, and 3D objects in real-time. This feature allows developers to position and orient virtual 3D holograms in relation to real-world objects using the camera of the used device. In this application, Vuforia is used to place the robot's holograms in the scene in a specific position using the image targets, as shown in Figure 16. From that initial point, the virtual robot is then tracked by the HoloLens coordinate system, and it is possible to select and manipulate it.

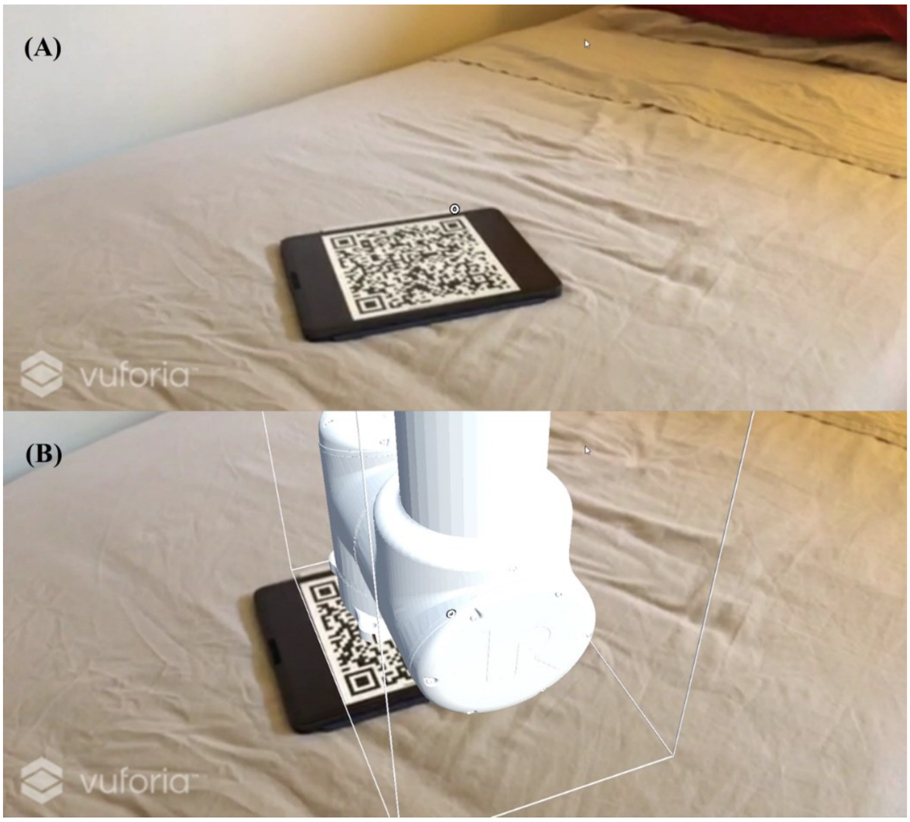


Figure 16: Marker recognition with Vuforia engine. In particular, the robot that appears is a full-scale collaborative robot UR10. (A) The marker just before it is recognized by the Microsoft HoloLens, (B) The robot appears when the HMD recognizes the marker.

3.2. Application Layout

The main task of the AR system is the reception of the information about the marker and the derivation of different 3D positions of virtual objects from this basis. This task must be performed very fast and without any manual assistance. The Vuforia engine and its capabilities have been employed to obtain this, as displayed in Figure 16.

The application is structured such that all the controls derive from a hand-menu. This menu can be used to enable or disable the Vuforia capabilities and robot manipulation. Through it, the robot can also be removed, and the measurement tool can be activated. The application layout is schematized in Figure 17.

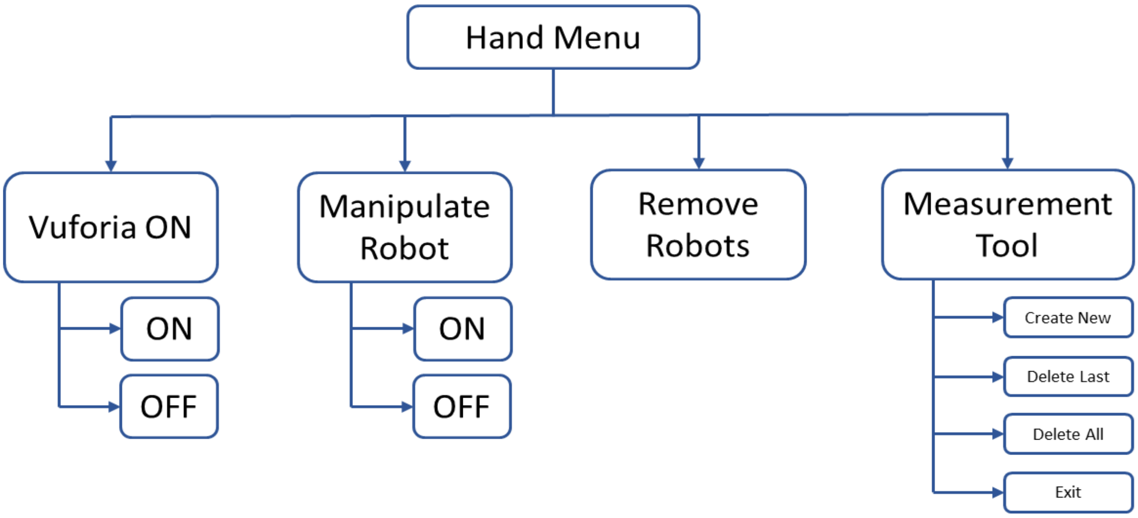


Figure 17: Study 1 Application Layout.

The main menu is activated by hand, and the movement is similar to extracting the phone from a pocket and use it. If the Microsoft HoloLens detects your flat hand, this hand-menu made up of buttons to control the scene will appear. A Unity simulation in Figure 18 can make understand this movement better.

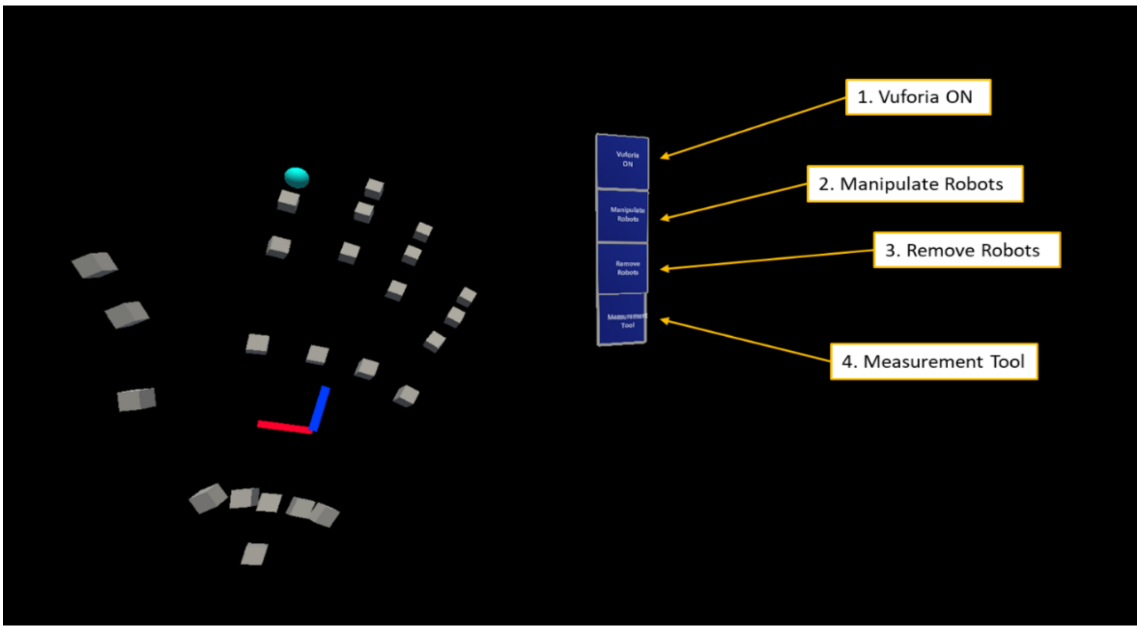


Figure 18: Application Main Menu with 4 buttons: 1. Vuforia ON, 2. Manipulate Robots, 3. Remove Robots, 4. Measurement Tool.

The "Vuforia ON" button allows to switch on and off the Vuforia capability. This operation is done to save some processing time when the application is running. Vuforia uses the HoloLens cameras to track the markers in the surrounding environment. By disabling Vuforia, it is possible to reduce the CPU overload.

The “Manipulate Robots” button allows to activate a bounding box around the robots in the scene and give them the possibility to be manipulated. A button of this kind is necessary to avoid unwanted hologram manipulation in the scene.

The “Remove Robots” button is used to cancel the robot from the scene. This button could be needed when users make a mistake or want to replace the robot with respect to the marker. Users can select the correspondent “Remove Robots” button and Air Tap on the robot to be removed.

Last is the “Measurement Tool” button, which allows users to enter and use the correspondent tool. This function will be explained in detail in the following paragraph (Measurement Tool).

Due to the COVID-19 pandemic, it was impossible to access the facilities and to have a Microsoft HoloLens available. In the following, some figures help to describe the application flow. Figure 19 displays an operator in front of a box conveyor. It has been decided that a collaborative robot should be installed in the line to load the boxes on the conveyor, reducing the operator's effort. First, there is the need to understand where to place the robot in order to obtain an efficient solution. A printed QR Code has been put on the floor, and this is the marker that the application should search for. No robots are displayed because the HMD still has to detect the marker.

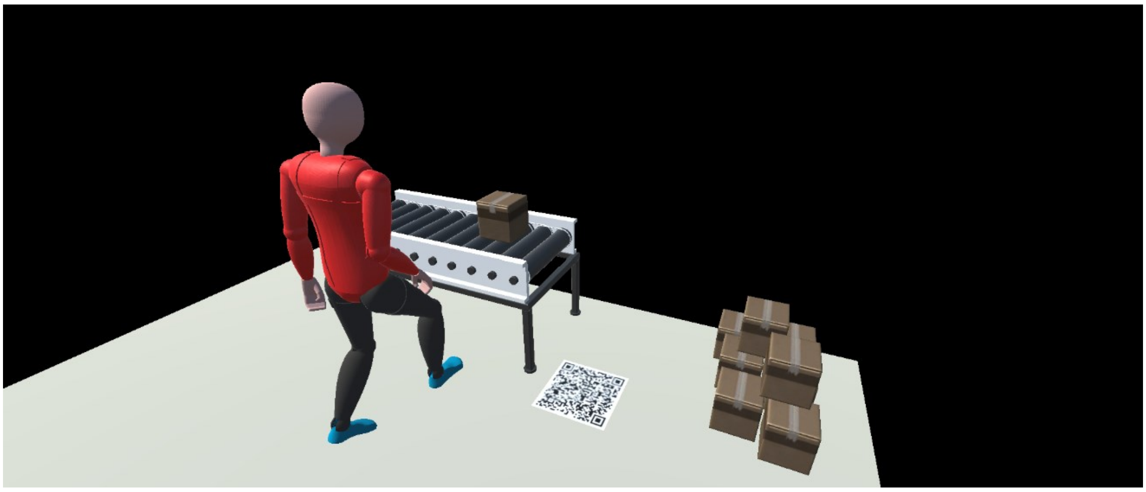


Figure 19: Example of scene for collaborative robot virtual visualization (the robot is not visible).

The virtual robot appears in the scene once the HMD detects the marker with its cameras, as shown in Figure 20. The operator can now use the application functionalities to move or rotate the virtual robot.

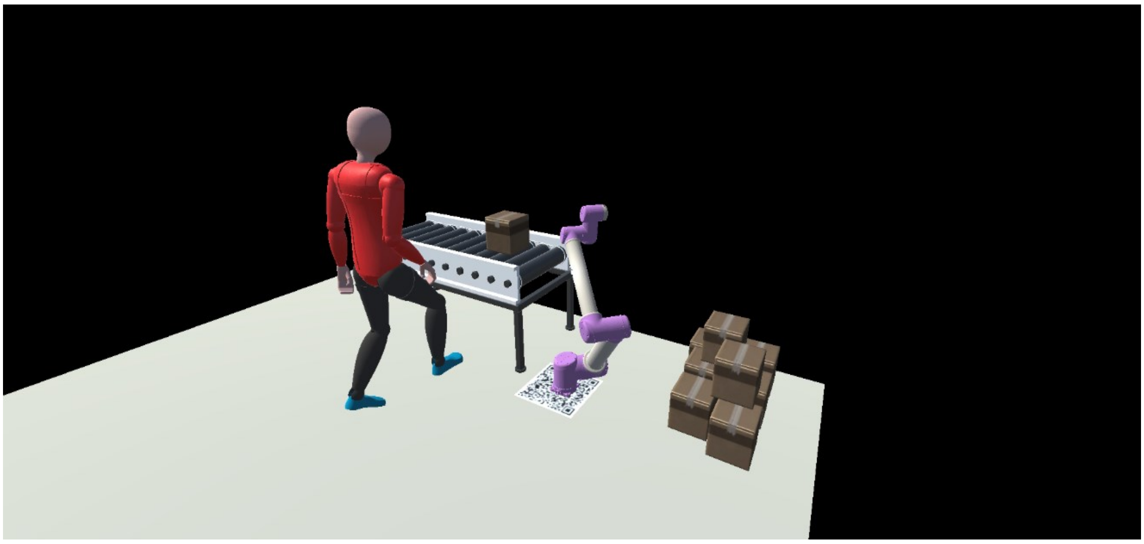


Figure 20: Example of scene for collaborative robot virtual visualization (the robot is visible).

When the robot first appears in the scene, it is not possible to manipulate it. However, the hand-menu button enables the manipulation, and a “wireframe box” appears around the virtual robot, as displayed in Figure 21.

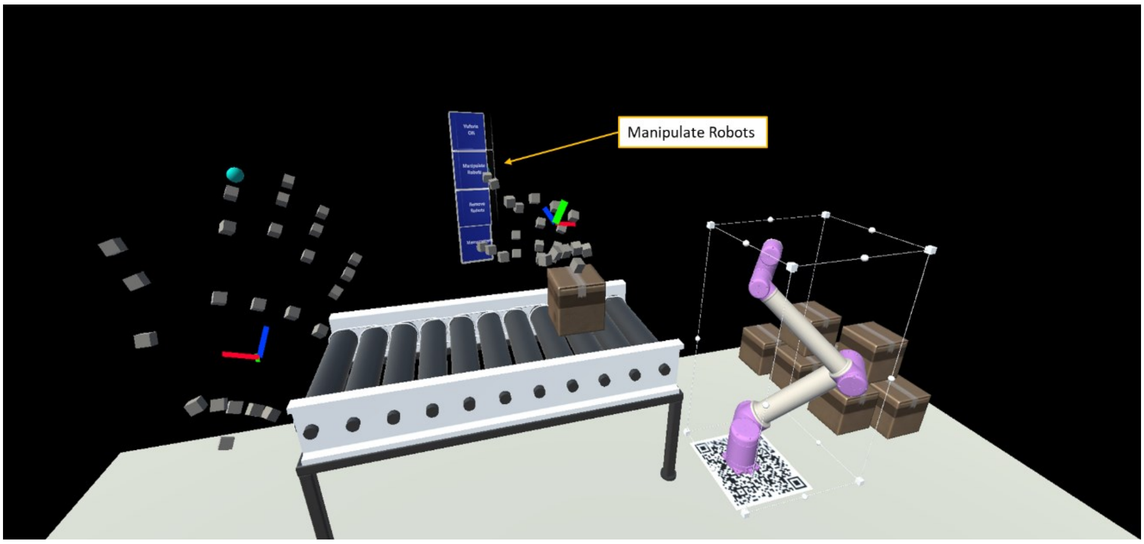


Figure 21: First person scene. Hand-menu to enable the virtual robot manipulation.

Grabbing the virtual robot in the center of one of the “wireframe box” faces makes it possible to move it around. This operation allows understanding which is the best robot position to obtain an efficient and feasible solution. As displayed in Figure 22, the robot is grabbed and moved behind the boxes. The latter is not the best solution, but it allows to appreciate the occlusion of the virtual robot behind the real boxes, thanks to the Microsoft HoloLens MR capabilities.

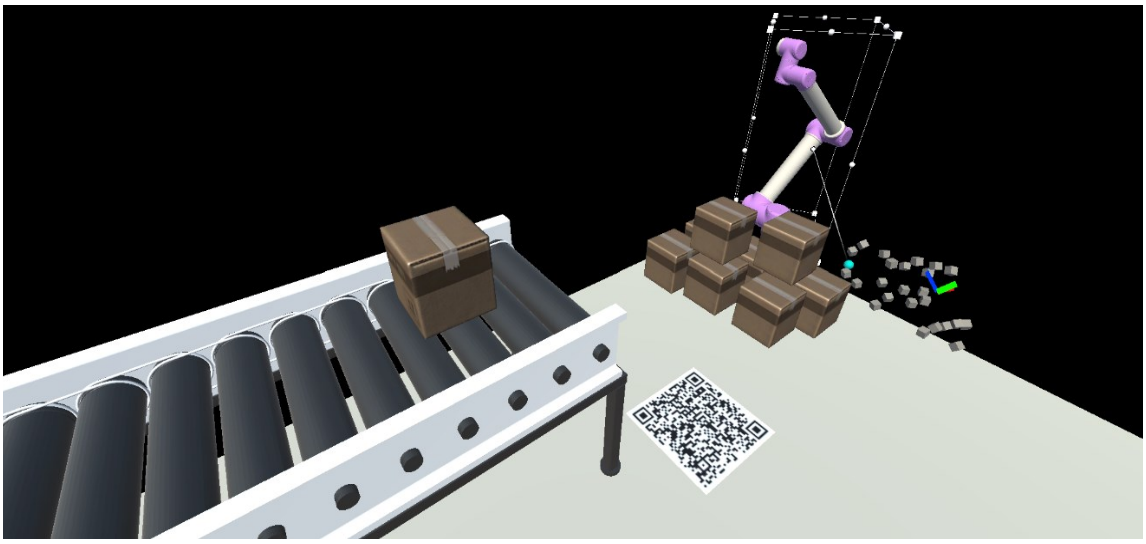


Figure 22: First person scene. Virtual robot positioning.

Grabbing one of the little balls on the “wireframe box” edges makes it possible to rotate the virtual robot, as shown in Figure 23.

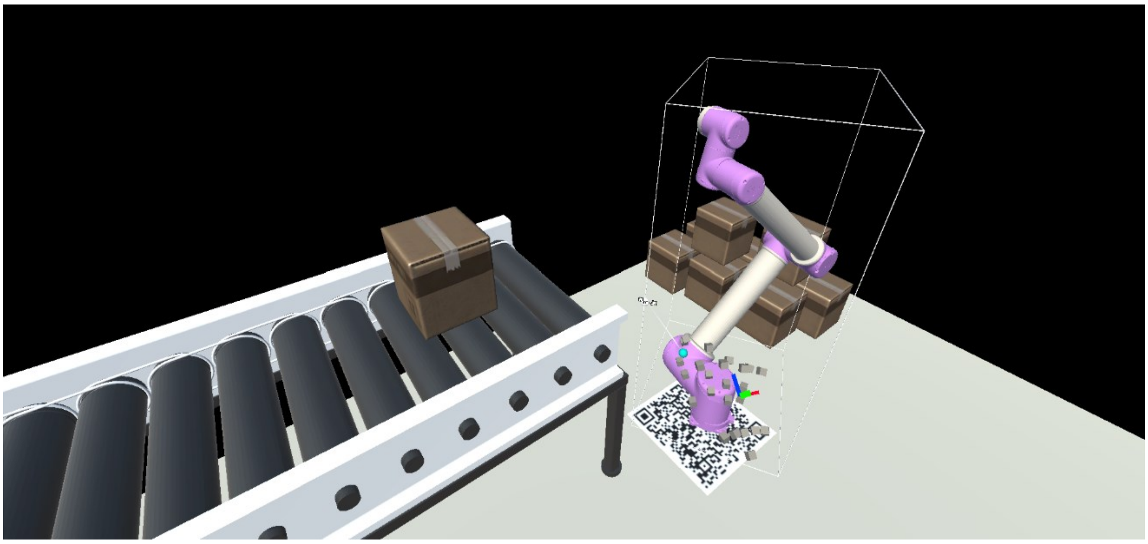


Figure 23: First person scene. Virtual robot rotation.

Once the position and orientation of the virtual robot are the right ones, the operator can click the manipulation toggle button again on the hand-menu and disable the manipulation. As displayed in Figure 24, the virtual asset is now fixed in the scene, and it is possible to use other application capabilities.

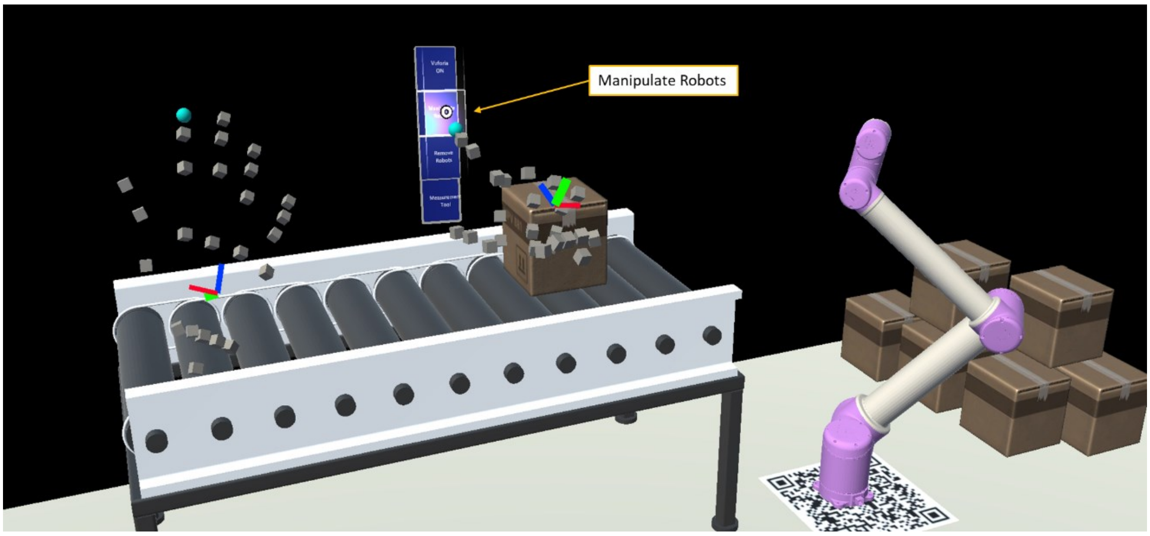


Figure 24: First person scene. Hand-menu to disable the virtual robot manipulation.

Suppose the analysis terminates or the operator wants to place the virtual robot on the marker at its initial position. In that case, the latter can be removed with the correspondent hand-menu button, as displayed in Figure 25.

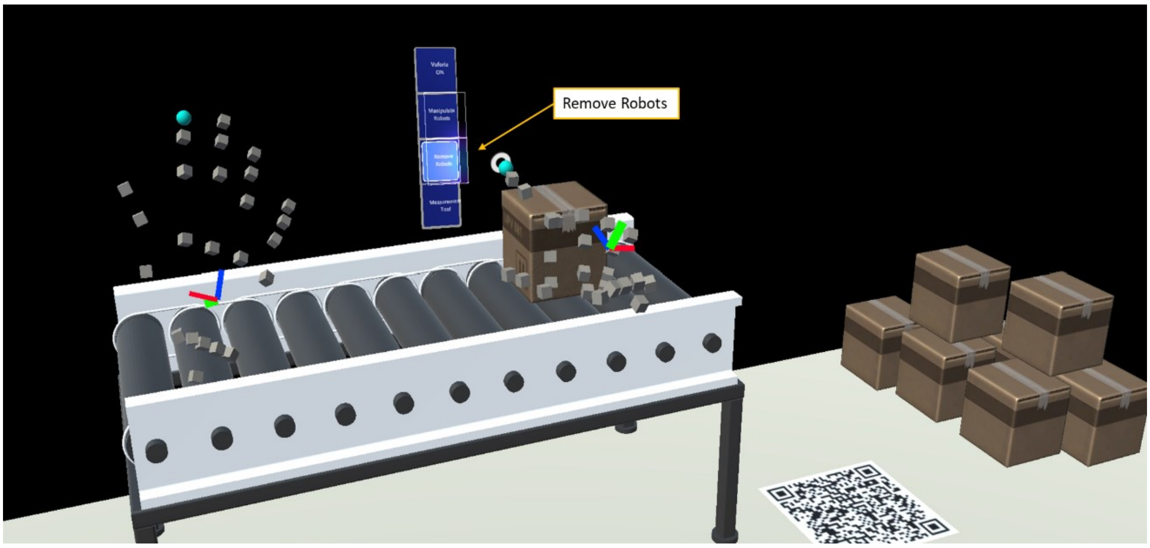


Figure 25: First person scene. Hand menu to remove the virtual robot from the scene.

3.2.1. Provided Equipment

The Major Automotive Company provided some CAD files of robots and equipment to be used for application testing. In this report, just the 2 robots that are public and available online will be presented: the UR10 and the Comau NJ-220-2.7. A clamping station used to keep attached metal parts in car body assembly was provided as well, but it will not be included in this report for confidential reasons. The first robot is the UR10 (Universal Robots, Denmark) that is displayed in Figure 26. This robot is a 6-axis articulated arm that can perform different tasks as pick-and-place, machine tending, palletizing, and packaging (Universal Robots, 2021). Being a collaborative robot, the UR10 can work in close collaboration with human workers and so there is no need to restrict it in a cage. The robot can be programmed by simply grab it, move it around, and tap on a touchscreen to record the desired positions and actions (Guizzo, 2015). The UR10 has a reach of 1.3 m, a supported payload up to 12.5 kg, a footprint of 19 cm, and a weight of 33.5 kg.

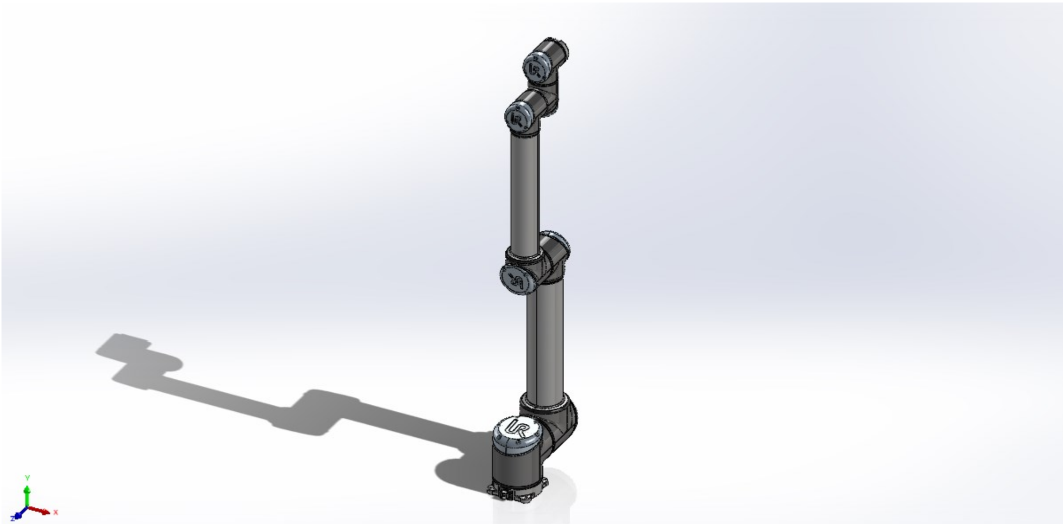


Figure 26: UR10 Collaborative Robot.

The second robot is the Comau NJ-220-2.7 (Comau, Italy) that is displayed in Figure 27. This robot is slightly different from the previous one because it cannot work in close contact with humans. It is a 6-axis robot arm that offers 220 kg of wrist payload and 2.7 m reach. Its functions are assembly, processing machining, spot welding, handling/packaging, and much more (Comau, 2021).

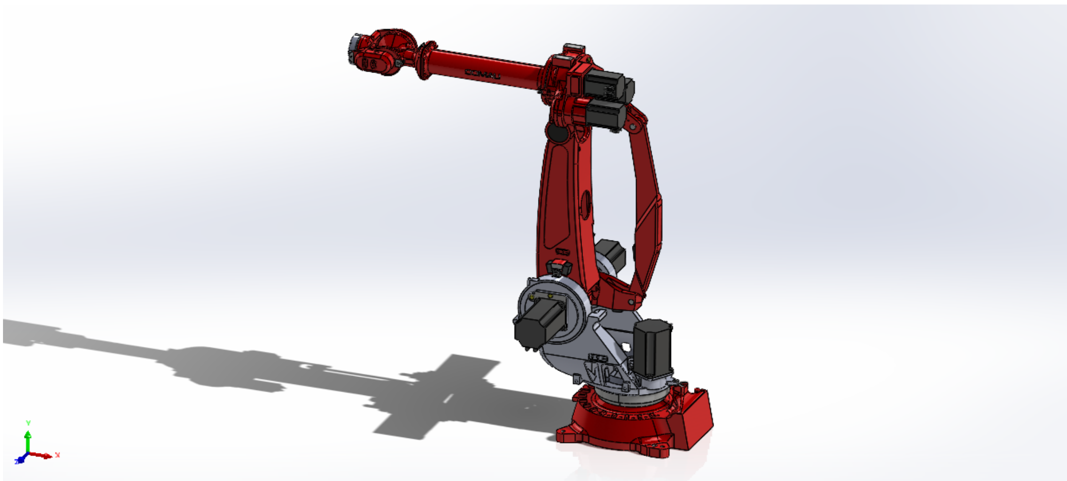


Figure 27: Comau NJ-220-2.7 Robot.

3.2.2. Measurement Tool

The measurement tool is one of the core features of the application. It allows measuring the distance between 2 points on the X, Y, and Z directions, and it also calculates the absolute value. The implementation in Unity is a standard prefab file made up of two little red spheres and a panel containing the information about the measure. The game objects that can be detected by the measurement tool are the ones that integrate the “Mesh Collider” component.

The measurement tool is integrated with the HoloLens capabilities and gestures. First of all, an hand-menu allows selecting what to do, as displayed in Figure 28.

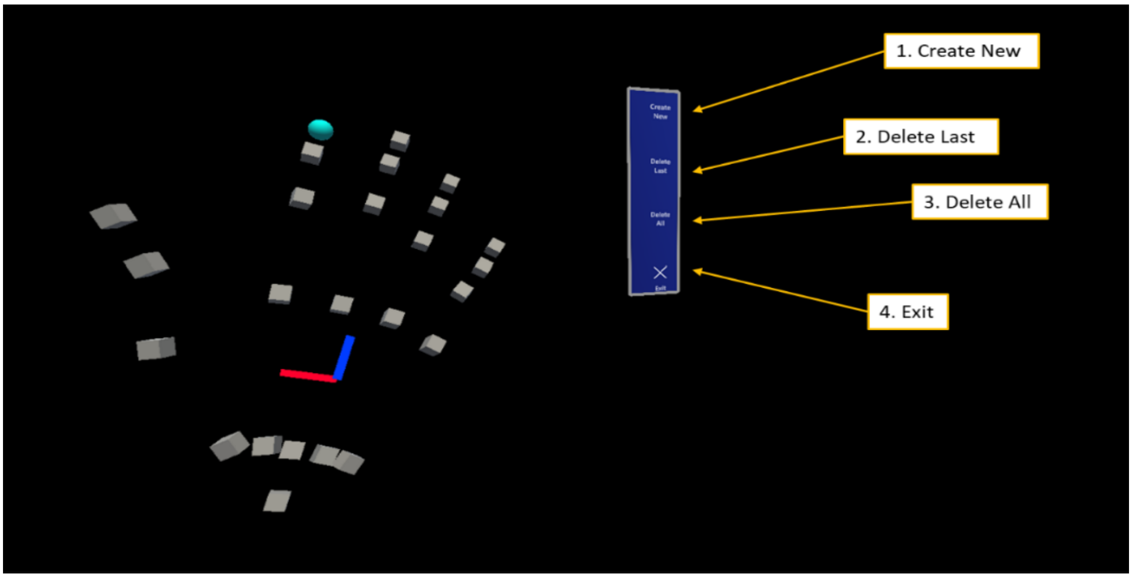


Figure 28: Application Measurement Tool Menu: 1. Create New, 2. Delete Last, 3. Delete All, 4. Exit.

Users can decide, for example, to create a new measure. After clicking this button, users might select two consecutive points in the space and obtain the measure, as displayed in Figure 29.

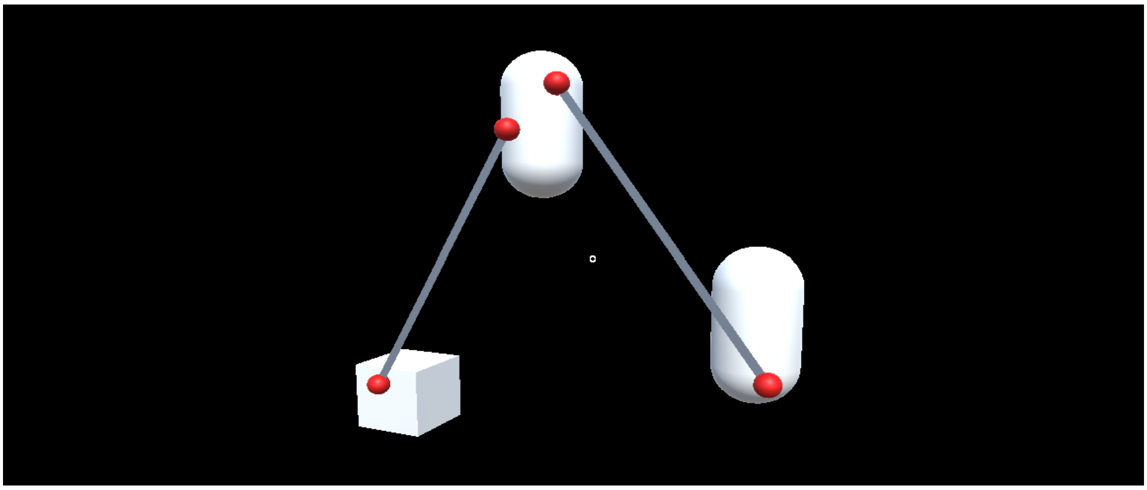


Figure 29: Example of two measures on random shape objects.

In this figure, two measures are displayed: one from the capsule to the cube and one from the first capsule to the second capsule. After creating the first measure, users can also decide to delete the last-done measure or delete all the measures again. The measure information panel is not displayed at first, and the user must point the head Raycast towards one of the two red balls of the measure, as shown in Figure 30. Then, the panel appears for 5 seconds in front of the users, and it turns with respect to their orientation and position.

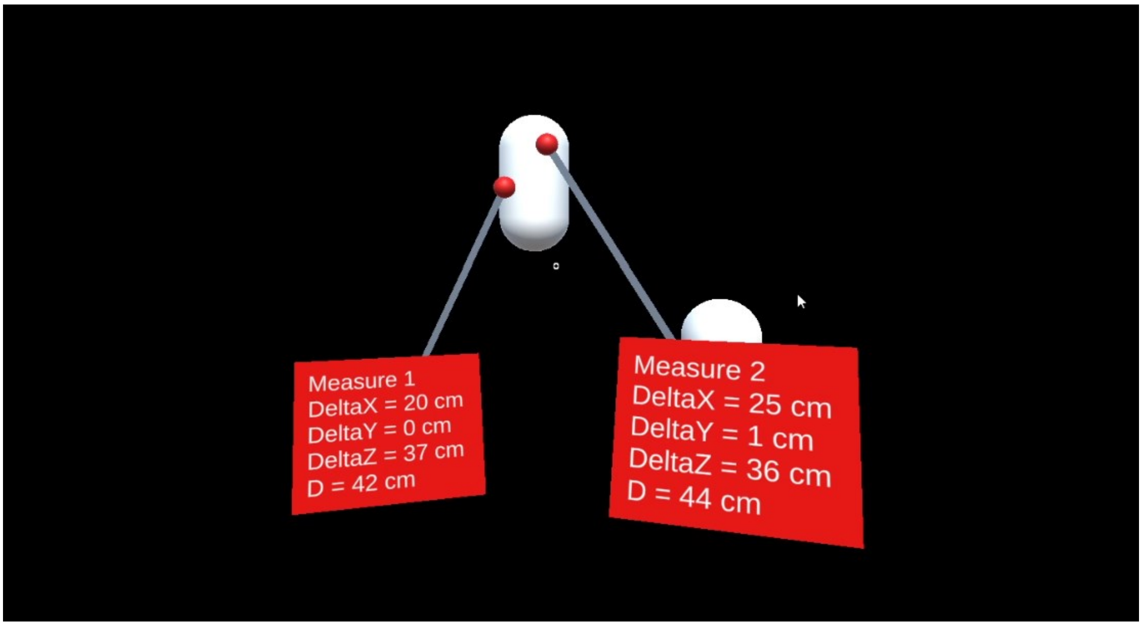


Figure 30: Measurement information along the 3 reference axes and the absolute value.

3.2.3. Integration Issues

In the application development, some issues have been faced. Vuforia is an AR platform that can be coupled with Unity and can run on a wide range of devices, not just the Microsoft HoloLens. For this reason, several integration issues between the MRTK and the Vuforia Engine have been found. For example, certain scripts of the Vuforia Engine override some HoloLens capabilities, and developers do not have the freedom to write scripts easily. Vuforia is a useful tool because it provides AR behavior to the HMD, but, at the same time, it reduces development flexibility. One example of this is the Microsoft HoloLens Spatial Awareness. This device can track the surrounding environment by its cameras and sensors to obtain the MR experience. By adding Vuforia to the application, the Spatial Awareness capability cannot be used. So, the holograms do not appear to be occluded, or the Measurement Tool cannot be used to measure from the external walls to the actual holograms. This issue has been attempted to be solved by doing online research, contacting the Vuforia support multiple times, writing some questions on the

Vuforia Developer Portal (Vuforia, 2021), or writing on the Stack Overflow website (Stack Overflow, 2021), but with no success.

The application should have worked with different virtual robots CAD files. However, importing these files in Unity was a complex procedure because the Major Automotive Company provided .STEP and .jt files. These file extensions cannot be directly imported in Unity which accepts only .fbx files. Moreover, the .fbx files cannot be generated by the CAD software. Therefore, the procedure has been to open the .STEP and .jt files with the CAD software, convert them into .stl files, open the .stl files with a computer-animation software and convert them into .fbx files. In the whole process, the virtual robot should be scaled and oriented to obtain the proper dimension and position in the left-handed reference system of the Unity environment.

Once the application was built on the Microsoft HoloLens, some low-performance issues have arisen. For example, when the three CAD models were all present in the scene simultaneously, the application started lagging. This issue is caused by many game objects, details, and Vuforia capabilities that the device must process. The issue can be partially solved by disabling the Vuforia tracking with the menu displayed in Figure 18.

CHAPTER 4 - STUDY 2

4.1. Methodology

4.1.1. Participants

Twenty participants volunteered to participate in the study for a compensation of 20 CAD each. Participants were undergraduate and graduate students at the University of Windsor, with some external exceptions. Participants were, on average, 25 years old (6 years standard deviation), 175 (8) centimeters tall, and 74 (18) kilograms. Five (25%) participants were female, and 15 (75%) were male. Forty percent of them had previous experience with AR or VR, but none had ever tried a Microsoft HoloLens. Fifteen percent of them had previous experience with BCIs, but none had ever tried the NextMind device. None of the participants had upper limbs discomfort. The protocol for the experiment was reviewed and approved by the University of Windsor Office of Research Ethics (REB Number: 38917). Due to the COVID-19 pandemic, this study was also evaluated and approved by the Research Safety Committee (RSC 2021-304-001-P3).

4.1.2. Hardware and Software – Brief Recall

The different tasks were implemented using the Microsoft HoloLens (Microsoft Corporation, United States) as the AR HMD and NextMind (NextMind SAS, France) as BCI. The software employed was the Unity game engine, coupled with Microsoft Visual Studio. Unity is the most popular cross-platform game engine. It is used primarily to create video games for numerous platforms, including mobile devices, Internet browsers, PCs, and VR and AR devices. Therefore, the application for this research was developed with the two dedicated SDKs for Microsoft HoloLens and NextMind. These packages were, respectively, the MRTK and the NextMind Unity SDK.

4.1.2.1. NextMind

NextMind is a brain-sensing wearable that goes in contact with the scalp on the back of the head. It is in close contact with the brain's visual cortex, where the images we see are projected. What the user focuses on is identified by the sensor in real-time. This sensor is an EEG that measures brain activity, translates it into digital commands, and triggers a correspondent action in the augmented environment (Protalinski, 2020). This device has been selected for the study because it is the only available SSVEP BCI commercially market available. NextMind is an SSVEP BCI, and this means that it can detect stable oscillations in voltage evoked by repetitive stimulations such as a strobe light presented on a monitor. This device has a dimension of 135x66x55mm and a weight of 60 grams, implementing nine high-quality electrodes. The minimal software requirements are Bluetooth LE support (4.0), Intel i5-4590/ CPU, and 8 GB

RAM. It is tested and approved for the Microsoft HoloLens 1st Generation for form factor and software.

4.1.3. Application Layout

The NextMind application can be found on a GitHub repository at the following link: [silvio-dacol/NextMind (github.com)](). It is organized into the main menu screen, tasks screens, a tutorial screen, and a calibration screen for the NextMind device, as shown in Figure 31.

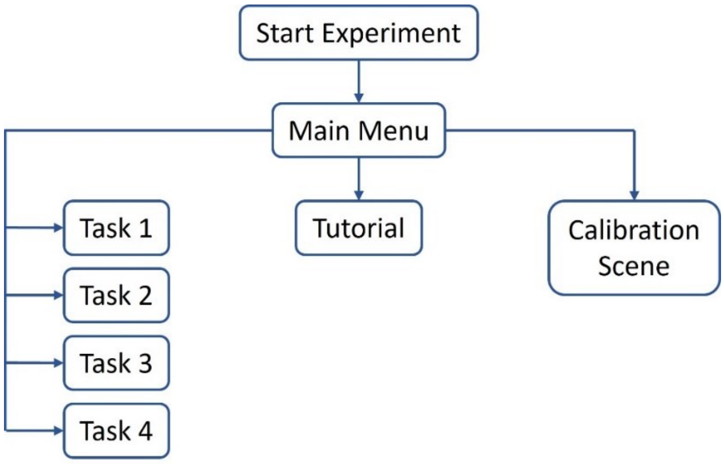


Figure 31: Study 2 application layout.

The main menu screen, displayed in Figure 32, allows the user to easily navigate through the other screens.

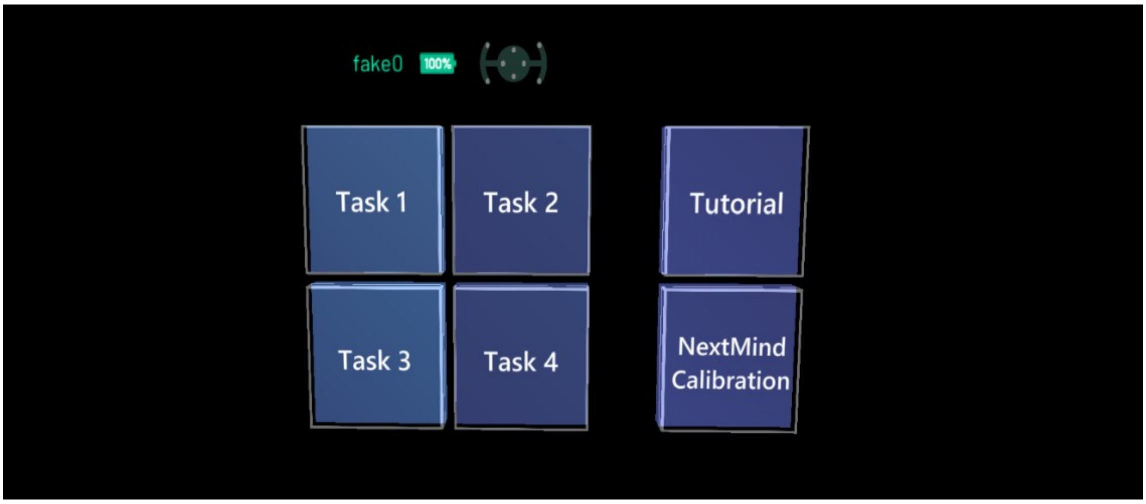


Figure 32: Example of main menu screen.

In the tutorial screen, displayed in Figure 33, the user may practice moving their head as well as their hands and focusing on objects for brain selection. For each selection technique, participants were given an indefinite amount of time to remain in a tutorial session and test the selection techniques until they felt comfortable and became accustomed to the movements needed to complete the upcoming tasks.

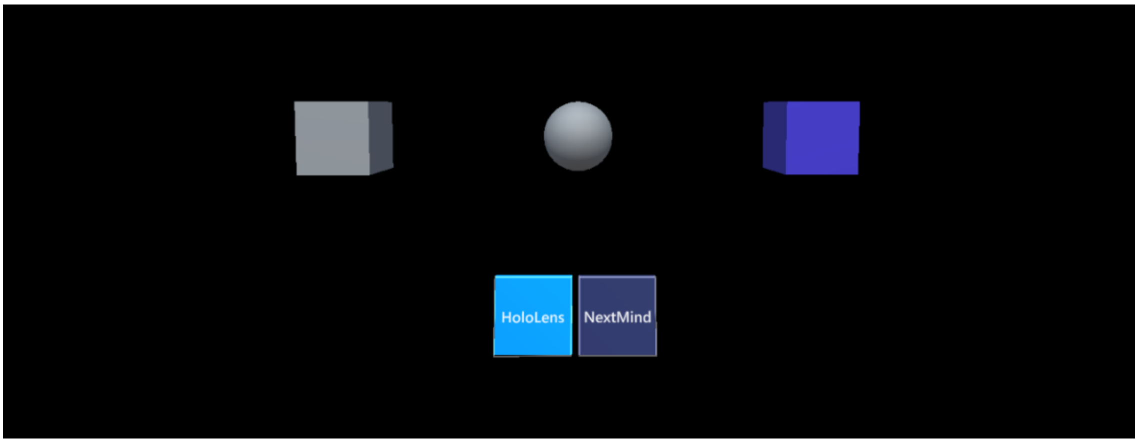


Figure 33: Example of tutorial screen.

In the BCI calibration screen, shown in Figure 34, the user can calibrate the NextMind device based on the signals coming from their brain. To accomplish this, the user concentrates on a gray circle where a blinking texture is displayed. A score from 1 to 5 is displayed at the end of the calibration, indicating its accuracy.

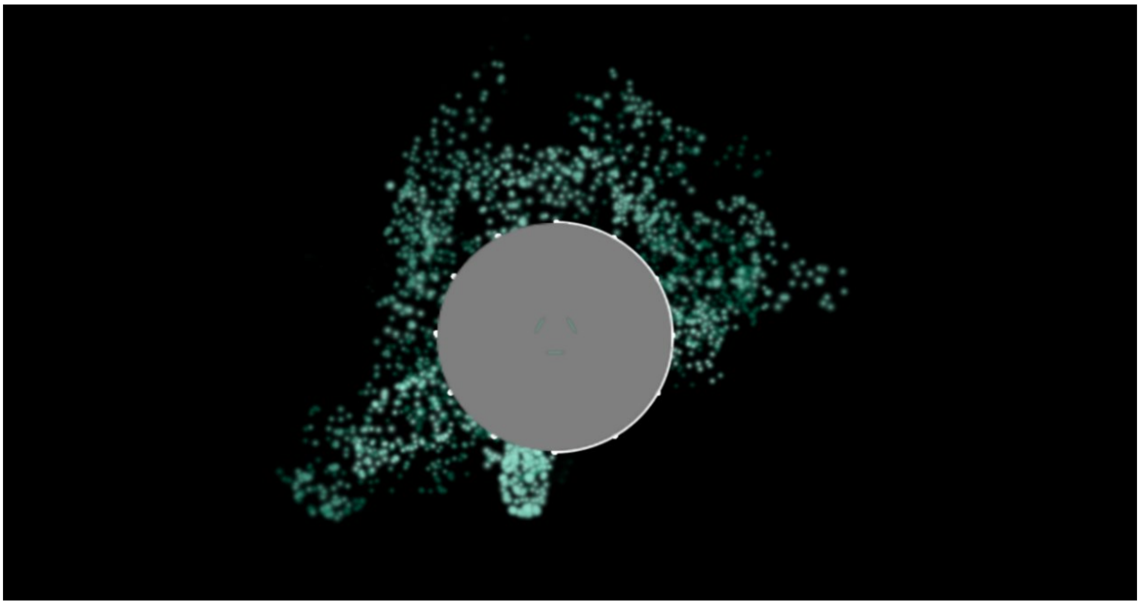


Figure 34: Example of calibration screen.

4.1.4. Task Description

Participants were asked to test the two selection techniques in four different scenarios. Both the order of selection techniques and the scenarios were randomized. The experiment was designed such that every observed difference would be attributable to the independent variable. The provided interface was identical in every way for each participant except for the selection method. In each task, there were a series of virtual objects of different shapes: cubes, cylinders, spheres, and capsules. These objects were scaled at 0.6 with respect to the standard dimension of the Unity environment and were placed at a distance between 5 and 9 meters from the user. The design choices for these experiments come from my literature review. Dense and dynamic environments can be found in Cashion et al. (2012) and Schröder-Kroll et al. (2008), while basic design choices, as the dimension of the virtual objects or their distance from the user, have been found in LaViola et al. (2017).

After the completion of the practice, participants were able to start the actual tasks. For each task, the target object on the screen was uniquely colored violet twenty times. Upon selecting an object, that object would start blinking, and audio feedback would indicate when a selection was made. The tasks are displayed in Figure 35, Figure 36, Figure 37, and Figure 38 and described below:

Task 1, static objects, non-crowded environment: the area in front of the user features 28 floating objects. These are static and do not overlap. Therefore, every object is easily visible to the participants.

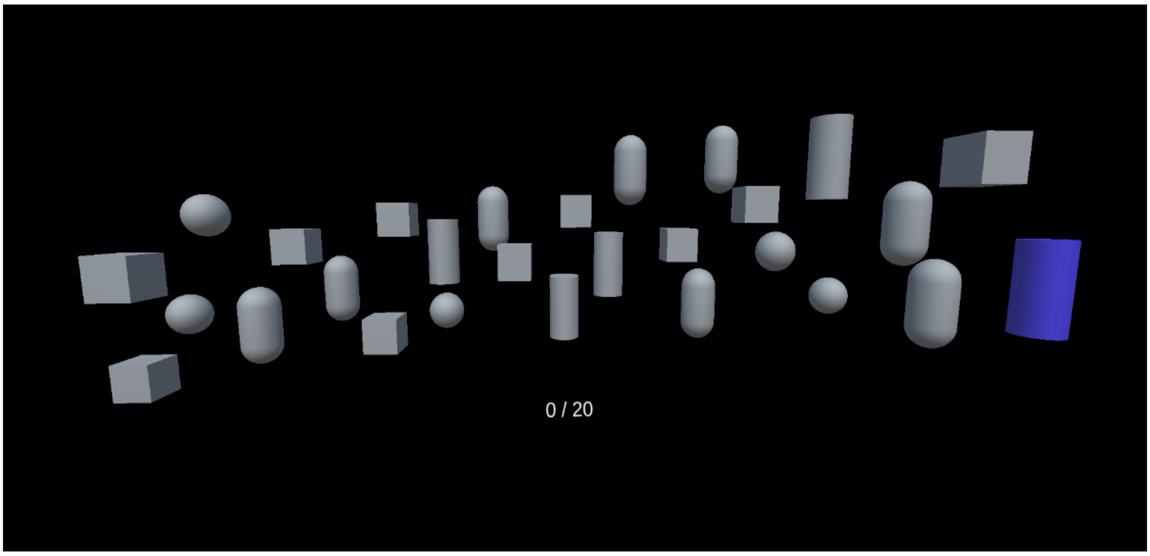


Figure 35: Application task 1 screen.

Task 2, static objects, crowded environment: the area in front of the user features 97 floating objects. These are static, and overlaps exist, so that participants may experience greater difficulties in selecting objects that are partially hidden from others.

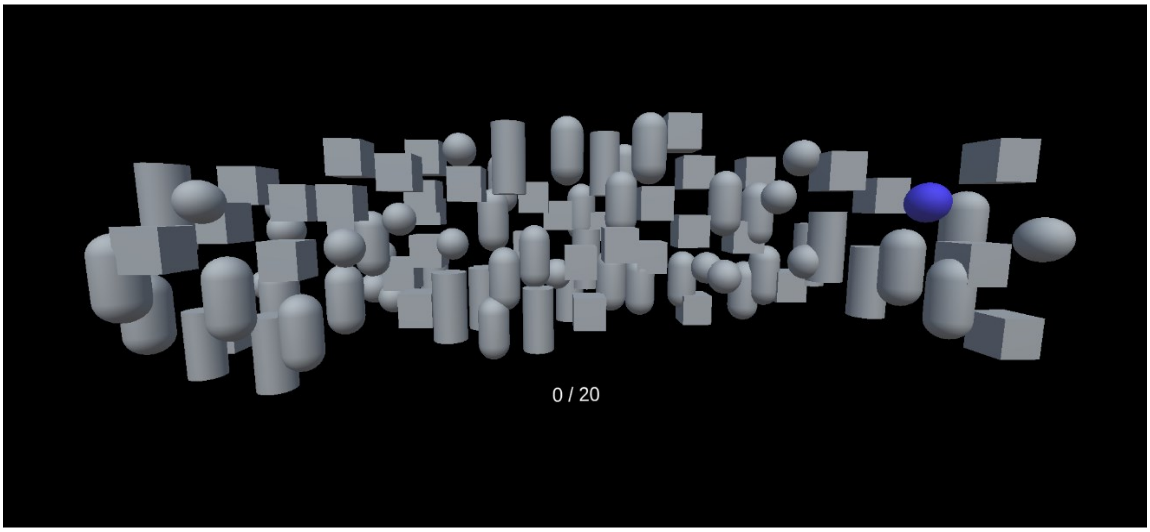


Figure 36: Application task 2 screen.

Task 3, moving objects, slow speed: the area in front of the user features 28 floating objects. These move randomly with periodically changing directions. The speed is relatively low (0.5 m/s), but object movement is unpredictable, and the user is verbally encouraged to focus carefully.

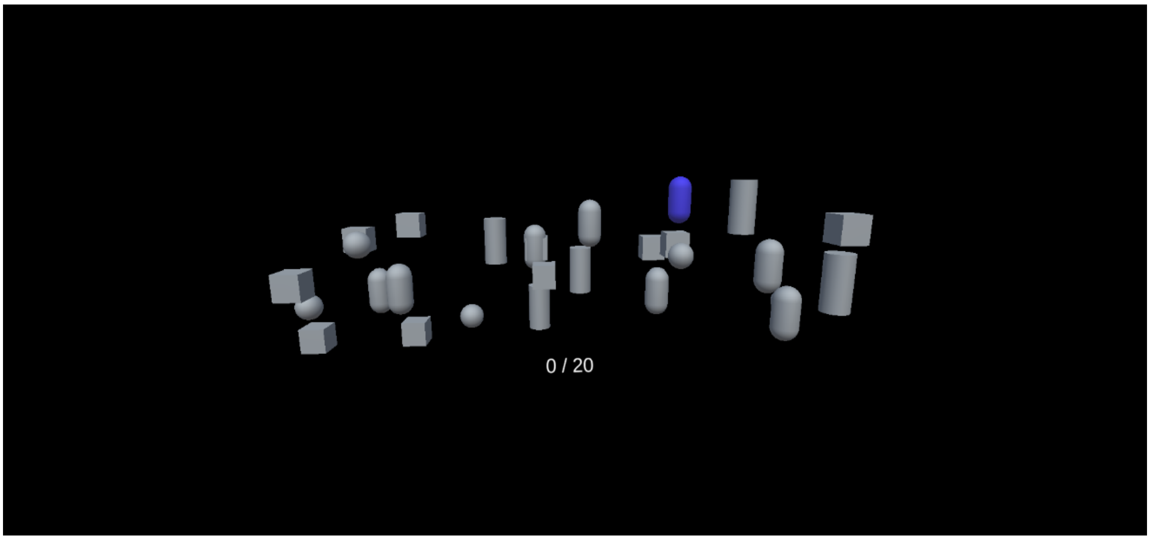


Figure 37: Application task 3 screen.

Task 4, moving objects, high speed: the area in front of the user features 28 floating objects. These move randomly with periodically changing directions. In this case, the speed is high (1 m/s), and the user is again encouraged to focus carefully.

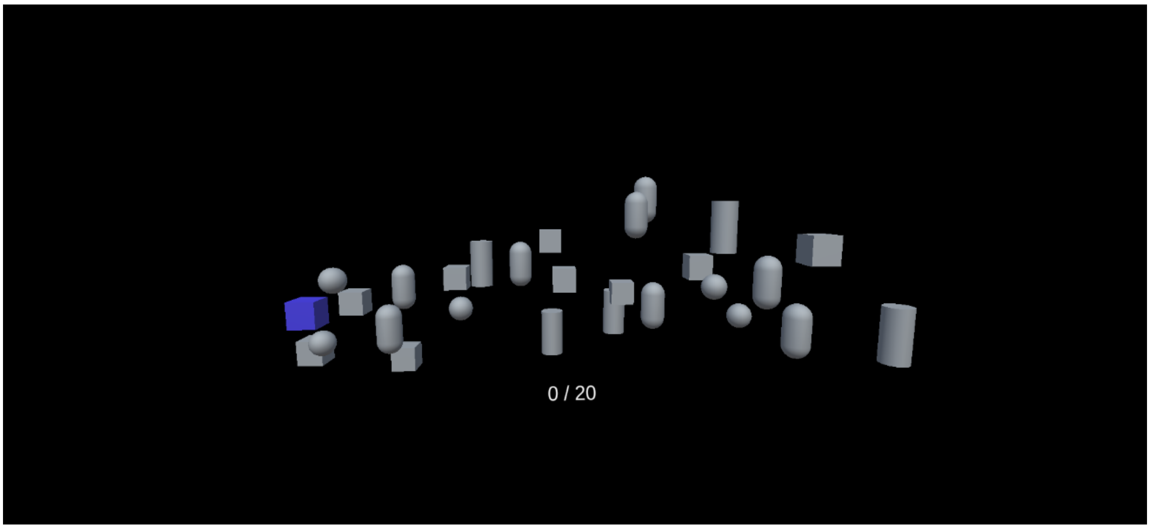


Figure 38: Application task 4 screen.

4.1.5. Experimental Design

Participants were welcomed upon arrival and seated at a desk. They were then asked to sign a consent form and complete a biometrics questionnaire. After this first phase, they were introduced to the experiment and the purpose of the study. Once they were ready, they were asked to wear the Microsoft HoloLens and the NextMind device. Participants were then introduced to the AR interface so they could adjust comfortably to the Microsoft HoloLens. When participants were in the application's main menu, they were told about the selection method to be used first, hand gestures or BCI, which was based on the randomization process. When the NextMind device was used for BCI, participants were invited to enter the calibration screen and calibrate the device. They were then asked to enter the tutorial screen and gain familiarity with the selection method. Once participants felt comfortable with the selection technique, they were asked to enter the tasks screen and start the experiment. All trials were conducted based on random order. Once all 4 tasks were completed, participants were given the NASA TLX and the System Usability Scale (SUS) questionnaire to fill out, focusing on the method they had used. Upon completing the questionnaires, each participant was given a break and, once ready, replicated the same process for the second selection method.

Figure 39 displays the position of the Microsoft HoloLens on the head of the user. The HMD is fixed with the internal support, and it can be consequently adjusted on the nose of each user. In Figure 40, it is shown the NextMind device position on the head of the user. The electrodes go in contact with the scalp on the back of the user’s head, where the visual cortex is.

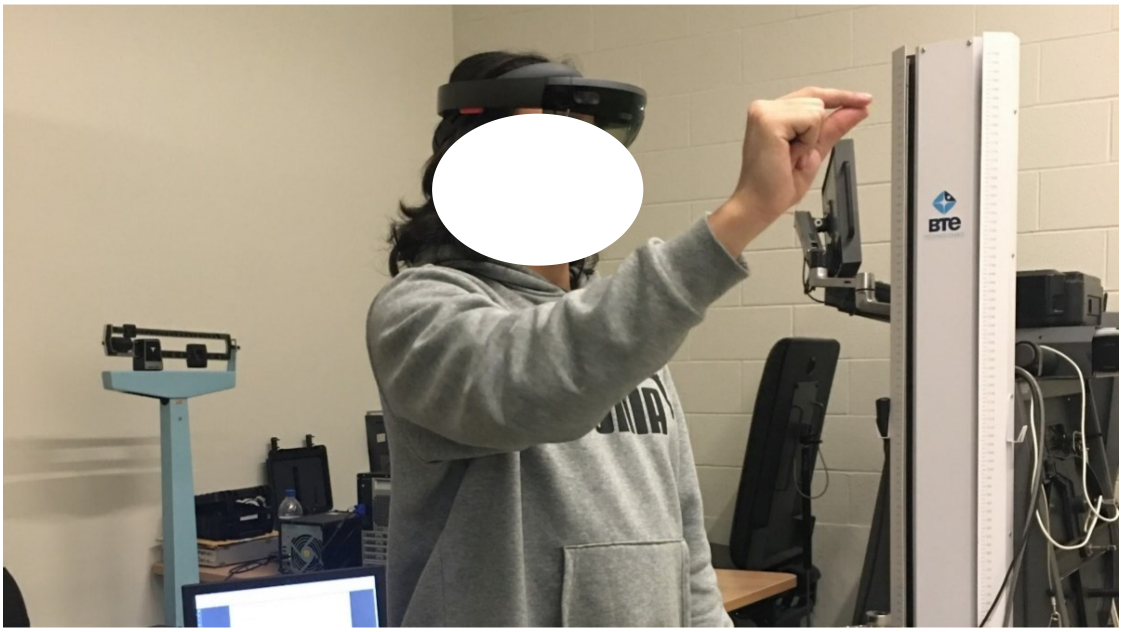


Figure 39: Apparatus to carry out the final experiment: Microsoft HoloLens.

Figure 40: Apparatus to carry out the final experiment: NextMind.

4.1.6. Measurement

A 2x4 within-subject factorial design was used in which the independent variables were the selection techniques of hand gestures and brain signals. The dependent variables were selection time and accuracy, mental workload as measured by the NASA TLX, and usability as measured with the SUS. Selection time was measured by the time it took to move from the previous target onto selecting the subsequent target. With regards to the accuracy, target objects that were not selected on the first attempt were considered errors, as was the selection of different objects other than the target object. This feature was implemented to prevent users from continuously air-tapping haphazardly (Vogel & Balakrishnan, 2005). The NASA TLX asks the user to provide a separate and subjective rating based on the following subscales: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration. For each subscale, the participant provides a numerical rating and a weight; the overall workload is obtained by multiplying the two values (Hart & Staveland, 1988). The System Usability Scale (SUS) is used with an absolute grade scale (Bangor et al., 2009), allowing immediate comparison of two systems. The SUS contains generic usability items and allows the assessment of a large range of interactive systems (Putze et al., 2019).

4.1.6.1. Experimental Design Methods

There are two main experimental design methods: between-subjects and within-subjects. In a *between-subjects* design, each participant is assigned to a different condition. The advantage of a between-subjects design is that any learning effect resulting from the user performing in one condition and the other is controlled. In fact, each user performs just under one condition. In this case, a greater number of participants is required, and a variation between the groups can negate any results (Dix et al., 2003). Due to the pandemic, it was difficult to find many participants, and a within-subject design was used. With this method, each user performs under each different condition. We randomized the order of tasks and devices to remove the learning effect. Thus, a lower chance of variation between participants can be obtained (Dix et al., 2003).

4.1.6.2. NASA TLX

The NASA TLX has been widely used in many research projects involving human-computer interaction (Felton et al., 2012). The Human Performance Group developed this tool at the NASA Ames Research Center to measure the workload (Hart & Staveland, 1988). It asks the user to provide a separate and subjective rating based on six subscales: the first three relate to demands on the participant, namely Mental Demand, Physical Demand, Temporal Demand; the second three relate to how the participant deal with the task, namely Performance, Effort, and Frustration (Choi et al., 2017). Since it is unlikely that individuals keep in mind specific cases of load, absolute conclusions or comparisons across different type of tasks are generally not meaningful (Hart & Staveland, 1988). This is the reason why, between more modi operandi to be compared, one must be the reference. Two methods can be compared if their final target is the same (Hart & Staveland, 1988). For each subscale, the participant marks a line divided into 20 equal intervals, converted to a rating on a 0 to 100 scale. The line is anchored on each side with bipolar descriptors as low and high. This procedure allows to give a numerical rating for each subscale and to obtain a raw workload score. The overall workload score is then calculated on a weighted average on the subscales. Fifteen comparisons between the subscales are made to obtain the weight. The participant should select the most relevant subscale between the presented two. An overall workload score is evaluated with the product of each subscale rating with its correspondent weight (Hart & Staveland, 1988). The weighted rating is then obtained, summing the overall scores of all the subscales, and dividing everything by 15 as explained in Equation 4.1.

Equation 4.1:

$$\begin{matrix} Weighted \\ Rating \end{matrix} = \frac{\begin{matrix} Mental \\ Demand \end{matrix} + \begin{matrix} Physical \\ Demand \end{matrix} + \begin{matrix} Temporal \\ Demand \end{matrix} + Performance + Effort + Frustration}{15}$$

Thanks to its subscales, the NASA TLX evaluation is useful for obtaining more detailed data than a unidimensional scale. The mental workload is a complex concept that is better analyzed by splitting its elements.

4.1.6.3. System Usability Scale

The SUS allows measuring reliably the usability of computer systems on which the users are working (Affairs, 2013). John Brooke invented it in 1986, and it gives the possibility of evaluating products and services, including software, hardware, mobile devices, websites, and applications (Affairs, 2013). It consists of 10 questions at which the respondent can decide between 5 options: from strongly agree (5) to strongly disagree (1). Thus, the final score can be from 0 to 100, and it is calculated in Equation 4.2.

Equation 4.2:

$$SUS = \{[(Q_1 + Q_3 + Q_5 + Q_7 + Q_9) - 5] + [25 - (Q_2 + Q_4 + Q_6 + Q_8 + Q_{10})]\} * 2.5$$

Where Q_n is the score associated with each question (from 1 to 5).

In this research, the SUS has been used to evaluate the selection methods for the virtual objects in AR. Therefore, when filling out the SUS, users are asked to focus on the actual selection method rather than the application or the devices themselves.

As usability is defined by the context in any given instance, it follows that, in general, how you measure usability will also be defined by that context (Brooke, 2013). ISO 9241-11 explains SUS, dividing it into three components: effectiveness, efficiency, and satisfaction (Brooke, 2013). All these refer to users and respectively: how successfully can users achieve their objectives? How much effort and resources are used to reach those objectives? How satisfied are users with the experience?

The benefits of using the SUS after an experiment are that the least skilled participant can well administer the tool. This tool gives reliable results with a small sample, and it can easily differentiate between usable and unusable systems.

4.1.7. Data Analysis

Statistical data for time and accuracy of selection, NASA TLX, and SUS were analyzed using Minitab software version 19 (Minitab, LLC, United States). The Wilcoxon signed-rank test was performed to compare the time, accuracy, NASA TLX, and SUS score between hand gestures and BCI. The confidence level was set at $\alpha = 0.05$ to control for the chance of Type I errors.

4.1.7.1. Wilcoxon Signed-Rank Test

The Wilcoxon Signed-Rank Test is a non-parametric statistics hypothesis test. It is used to compare two related samples and as an alternative to the paired Student's t-test when the distribution of the difference cannot be assumed to be normal. As the t-test, the Wilcoxon test is used as a hypothesis testing tool, allowing testing an assumption applicable to a population. A limitation of this tool is that when the difference between the groups is zero, the observations are discarded (Derrick & White, 2017).

4.2. Results

4.2.1. Time of Selection

Time of selection descriptive-statistics is summarized in Table 1. The first hypothesis (H_1) was not confirmed because BCI was slower than hand gestures in every situation. The static tasks, Task 1 and Task 2 revealed the biggest differences. In particular, Task 1 showed a mean value of 6.25 seconds for BCI and 3.54 seconds for hand gestures, while Task 2 showed a mean of 6.97 seconds for BCI and 4.64 seconds for hand gestures. In the dynamic tasks, Task 3 and Task

Table 1: Selection Time (in seconds) for each task.

	Selection Time [sec]							
	Hand Gesture				BCI			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
Task 1	3.54	1.41	1	12	6.25	1.88	1	28
Task 2	4.64	3.24	1	30	6.97	2.59	2	28
Task 3	7.30	6.07	1	30	7.70	3.26	3	30
Task 4	8.16	6.26	2	30	8.31	4.00	3	30

4, BCI was still slower than hand gestures, but the difference between them was smaller than in the static tasks. For Task 3, the mean was 7.70 seconds for BCI and 7.30 seconds for hand gestures, while for Task 4, the mean was 8.31 seconds and 8.16 seconds for BCI and hand gestures, respectively. As described in Figure 41, the Wilcoxon test revealed a significant effect for Task 1 ($p < 0.001$, $r = -0.790$), Task 2 ($p < 0.001$, $r = -0.636$), Task 3 ($p < 0.001$, $r = -0.210$) and Task 4 ($p = 0.016$, $r = -0.121$). With these results, significant differences in selection time were shown for all tasks.

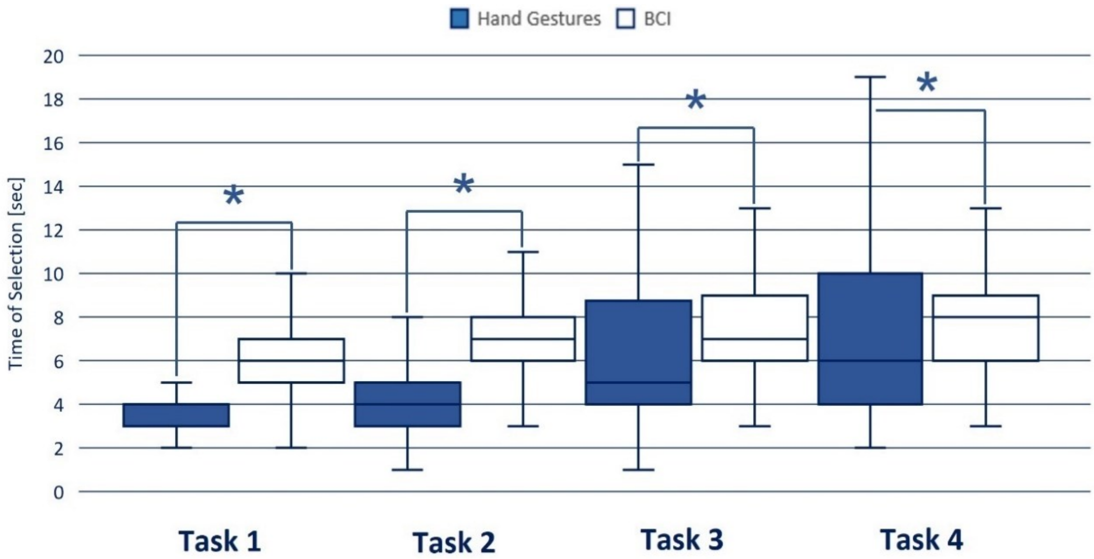


Figure 41: Selection time in each task for both interfaces (: P < 0.05).*

4.2.2. Accuracy of Selection

Table 2 presents the descriptive statistics of the percentage of right selections for each task. The BCI has higher performances than hand gestures for all the tasks. The dynamic tasks are the ones in which the difference between BCI and hand gesture is the highest. In particular, the BCI has a mean of 97% of right selections in Task 3 and 96% of right selections in Task 4 against a mean of 73% and 64% for hand gestures, respectively. The results confirm the second hypothesis (H_2), for which the accuracy of selection is higher in the BCI case.

Table 2: Percentage of right selections for each task.

	Selection Accuracy [%]							
	Hand Gesture				BCI			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
Task 1	93%	8%	80%	100%	97%	8%	90%	100%
Task 2	83%	10%	65%	95%	86%	7%	70%	95%
Task 3	73%	15%	40%	95%	97%	4%	90%	100%
Task 4	64%	18%	30%	95%	96%	4%	90%	100%

As displayed in Figure 42, significant differences have been found for Task 1 ($p = 0.016$, $r = -0.560$), Task 3 ($p < 0.001$, $r = -0.856$) and Task 4 ($p < 0.001$, $r = -0.856$), while there was no significant difference for Task 2 ($p = 0.570$).

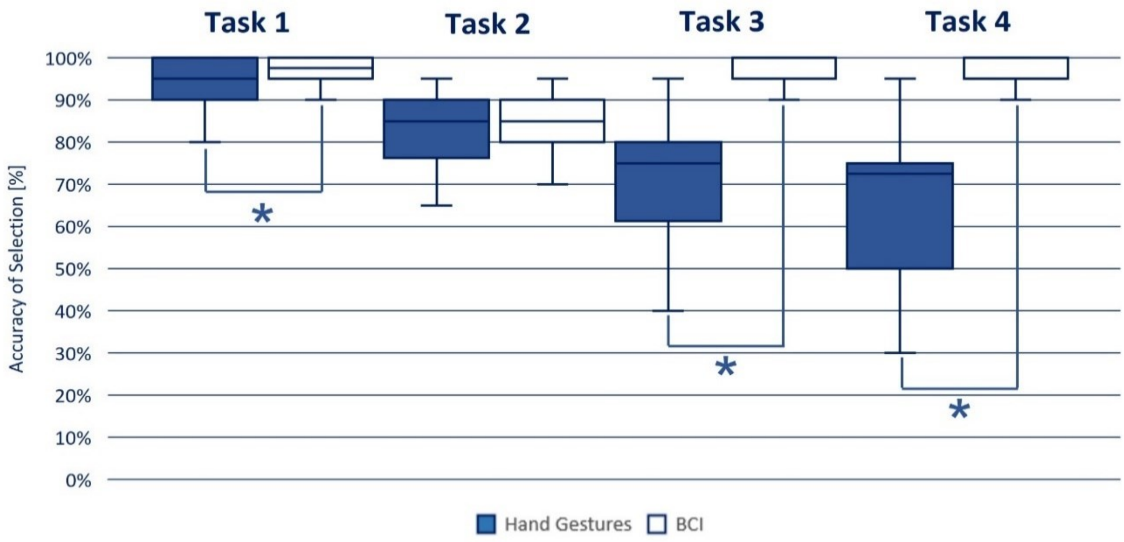


Figure 42: Selection accuracy in each task for both interfaces (: P < 0.05).*

4.2.3. NASA TLX

For the NASA TLX, absolute values are usually not considered a viable way of describing the workload, and a comparison, between two different devices, for example, is normally preferred. The descriptive statistics summary of the NASA TLX is shown in Table 3.

Table 3: NASA TLX questionnaire survey results.

	NASA TLX [ ]							
	Hand Gesture				BCI			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
Mental Demand	60	99	0	340	190	134	0	475
Physical Demand	252	156	0	500	28	51	0	150
Temporal Demand	107	88	0	260	133	112	0	425
Performance	77	75	0	270	37	48	0	225
Effort	180	120	35	400	128	95	20	300
Frustration	110	123	0	340	63	111	0	450
Weighted Rating	52.40	13.64	29.00	75.00	38.52	17.56	8.67	74.67

The results showed that the overall workload was significantly lower ($p = 0.007$, $r = -0.610$) for BCI (mean = 38.52) than for hand gestures (mean = 52.40). While BCI performed worse than hand gestures with respect to Mental Demand (BCI mean = 190, hand gestures mean = 60) and Temporal Demand (BCI mean = 133, hand gestures mean = 107), it nevertheless fared better with respect to Physical Demand (BCI mean = 28, hand gestures mean = 252), Performance (BCI mean = 37, hand gestures mean = 77), Effort (BCI mean = 128, hand gestures mean = 180) and Frustration (BCI mean = 63, hand gestures mean = 110). These results confirm the third hypothesis (H_3) in which we predicted a better performance of BCI in Physical Demand and worse performance in Mental Demand.

As shown in Figure 43, significant differences have been found for Mental Demand ($p = 0.001$, $r = -0.720$), Physical Demand ($p < 0.001$, $r = -0.856$), Performance ($p = 0.040$, $r = -0.464$) and Weighted Rating ($p = 0.007$, $r = -0.610$). However, there was no significant difference for Temporal Demand ($p = 0.360$), Effort ($p = 0.117$) or Frustration ($p = 0.098$).

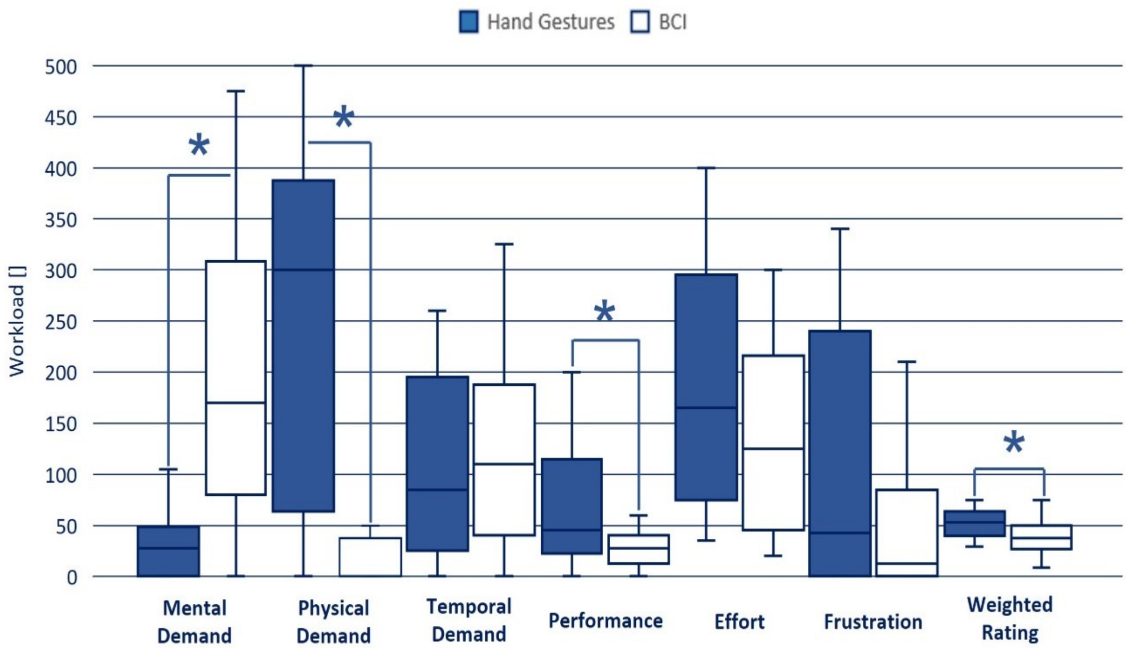


Figure 43: NASA TLX Scores for each subscale and for both interfaces (: P < 0.05).*

4.2.4. SUS Score

The outcomes of the SUS score are summarized in Table 4 and Figure 44. This test included both positive and negative questions. Therefore, when calculating the total score, the low score of the negative questions were reverse scored and appear as high scores in the total. The maximum possible final score was 100 points. Table 4 lists the overall SUS scores for both selection techniques. The mean SUS score for the BCI was 77.8, and for hand gestures was 66.5.

Table 4: SUS Score results.

	Hand Gesture				BCI			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
SUS [ ]	66.5	18.9	32.5	90.0	77.8	14.3	45.0	95.0

As displayed in Figure 44, BCI performed better than hand gestures, but the result of SUS scores did not indicate significance ($p = 0.076$).

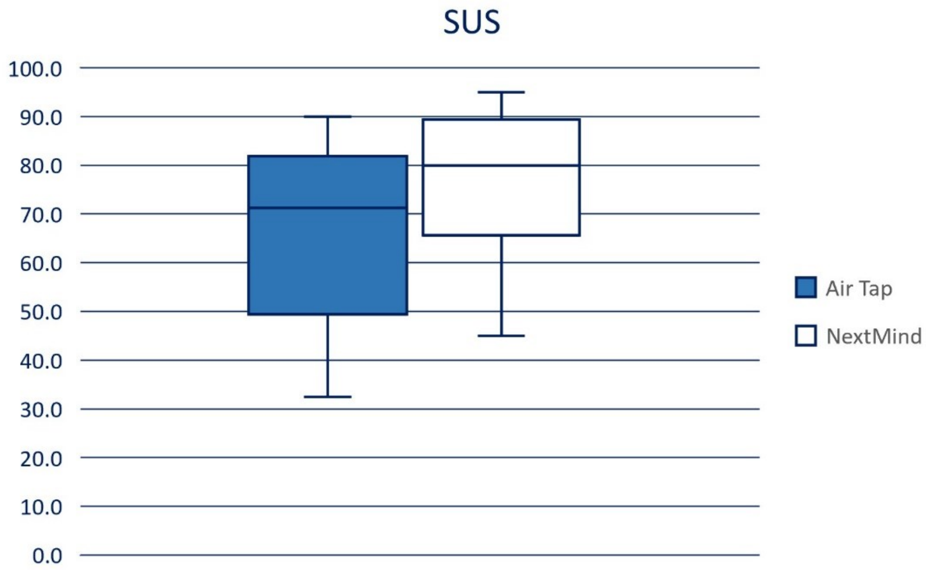


Figure 44: SUS Scores for both interfaces.

CHAPTER 5 - DISCUSSION

In the first study (STUDY 1), a manufacturing planning system has been presented. This first part aims to develop an application to introduce MR technologies in the industrial environment. The application allows the visualization of full-scale holograms to allow operators to see a new workstation while walking through the plant. The operator could inspect the new workstation virtually and quickly, identify discrepancies between the designed workstation and the environment. Possible collisions spotted in advance could save money and resources addressing the problem in advance (De Pace et al., 2018).

Furthermore, the application allows importing in the augmented environment the wanted robot using physical markers. Once in the scene, the robot can be selected and manipulated so that it is possible to change its position and orientation. The application also allows measuring the distance between two points in the space so that it is possible to better understand where the robots can be most efficiently placed.

In the development phase, the Measurement Tool has been validated using a coarse mesh of the environment. All the measured objects had a regular shape and could be well approximated by a coarse mesh. As it is possible to see from Table 5, the Measurement tool works well for big dimensions (> 100 cm), while it has a bigger error for smaller dimensions (< 100 cm). The overall absolute average error is 2.45%. The absolute average error for dimensions lower than 100 cm is 3.59%, while it is 1.30% for dimensions higher than 100 cm.

Table 5: Measurement Tool validation.

Coarse Mesh		Real (cm)	Measured (cm)	Difference (cm)	Error (%)
Bed	Length	185.42	184	1.42	0.77%
	Width	93.98	93	0.98	1.04%
Floor/Ceiling	Height	240.03	240	0.03	0.01%
Desk	Length	180.34	176	4.34	2.41%
	Width	73.66	73	0.66	0.90%
Bedside Table	Height	58.42	55	3.42	5.85%
	Length	60.96	57	3.96	6.50%
Wardrobe	Height	116.84	115	1.84	1.57%
Door	Height	195.58	199	-3.42	-1.75%
	Width	76.2	79	-2.8	-3.67%

In the second study (STUDY 2), an exploratory study was conducted to compare two input means for AR, the classic hand gesture and the novel BCI, with the goals of increasing the speed and accuracy of object selection and the usability of the system while reducing workload. The comparison between these interfaces has shown that BCI is more accurate than hand gestures and has a lower workload and greater usability. On the other hand, BCI has a lower selection speed. This study demonstrates the advantage of using BCI as a communication device in AR, and findings corroborate those of other studies (Putze et al., 2020, 2019; Si-Mohammed et al., 2017). Combining AR and BCI can, therefore, be used in scenarios favoring hands-free interaction and should be further developed. The coupling of these two technologies also favors the technological process of combining them together in the same device. In fact, a weakness of these technologies is that, being currently integrated on difference devices, there must be a compatibility between them.

BCI did, however, have a slower selection speed than hand gestures, with an average time difference of about 2.5 seconds for static tasks, Task 1 and Task 2. This result refutes the first hypothesis (H_1). An additional investigation of this difference was done by analyzing the selection of a single object on the screen using BCI to figure out the delay between the user's decision to select an object and the BCI device's selection of that object based on the signal from the user's brain. As displayed in Table 6 below, the delay was around 2.5 seconds. Being an SSVEP BCI, NextMind is triggered by signals from stimulation of the eyes. BCI requires time to process these signals. If this delay were eliminated, BCI object selection time could perfectly align with that of hand gestures. The NextMind device is based on machine learning and is continuously being optimized by the device development team. The more a machine learning algorithm is optimized, the faster the device response will be. The resulting reduction in delayed response time depends on the amount of data collected (Protalinski, 2020). The difference between BCI and hand gesture object selection times was not present in dynamic environments, Task 3 and Task 4, likely due to BCI's delay in selection and user difficulty selecting moving objects with their hands.

Table 6: BCI delay results on a single object selection.

BCI Delay [sec]				
N for test	Mean	S.D.	Min	Max
50	2.52	0.58	1.25	3.84

Concerning the accuracy, BCI exceeds hand gestures in environments with low object density, like Task 1, Task 3, and Task 4, as predicted in H_2. However, in Task 2, which is a static and high-density environment, BCI surpassed hand gesture accuracy by a smaller percentage (hand gestures: 83%, BCI: 86%) than in the low object density tasks. A possible justification for this finding may be that, as eye signals trigger BCI, when two virtual objects are very close to each other, the selection of an unwanted object can be triggered instead of the target object. For example, as displayed in Figure 45, the violet target object is in the background and is covered by other objects. Therefore, this proximity of triggering signals has led to lower accuracy. Considering that this research focuses on the feasibility of the application of BCI in industrial environments, this task presented an extreme scenario with virtual objects in close proximity that would be unlikely to occur in a real-life task. Nevertheless, this solution is feasible in a dense environment because the accuracy of the BCI is still higher than the one involving hand gestures.

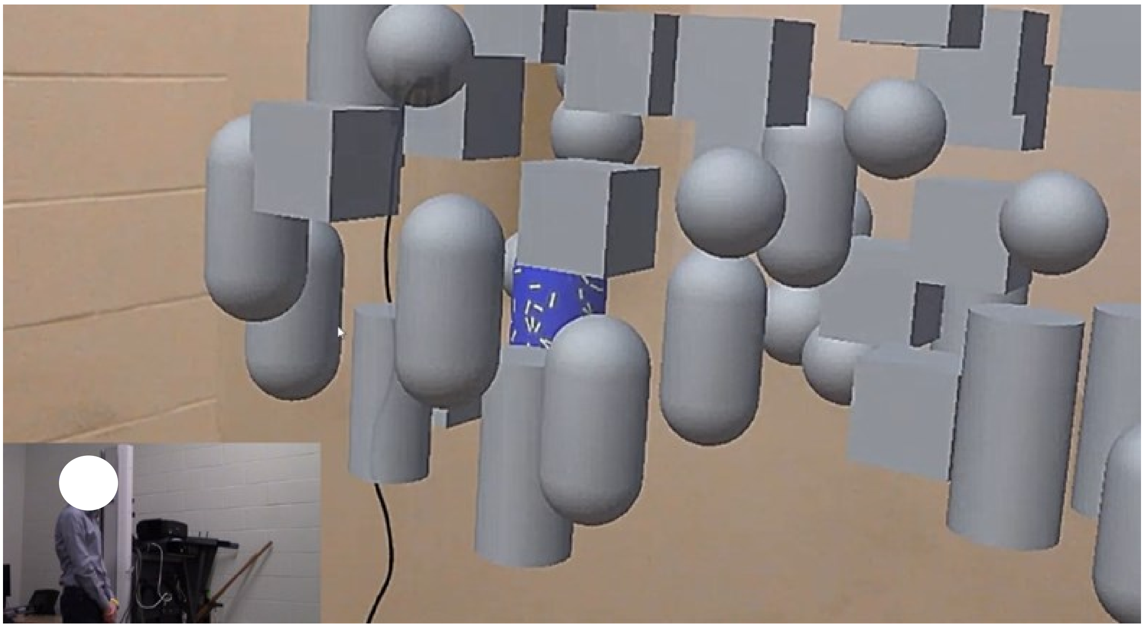


Figure 45: Example of target object in the background covered by other objects.

The NASA TLX gives many insights about the relative difference between BCI and hand gestures. Mental demand was higher for BCI than for hand gestures, as predicted in the third hypothesis (H_3). The mean was 190 for BCI and 60 for hand gestures. However, since the overall score of BCI NASA TLX is lower than the hand gestures' one, the level of mental demand when using BCI should be accepted (Grier, 2015).

Physical demand is, by contrast, lower for BCI when compared to hand gestures. The mean was 28 for BCI and 252 for hand gestures. With BCI, the selection method does not include any upper limb movements in the selection process. This method can prevent the user from keeping their hand up, causing arm and wrist strain (Argelaguet & Andujar, 2013). BCI, together with AR, offers the possibility for immersive scenarios (Putze et al., 2020) controlled just with brain signals.

Temporal demand showed no significant difference, likely because the application used for the experiment had no time pressure for the user while performing the tasks. However, when comparing mean values between BCI and hand gestures, a higher mean value is obtained for the former (BCI: 133, hand gestures: 107). This could be due to participants' perception of a longer selection time, which is especially relevant in static environments. As already mentioned, a delay of 2.5 seconds (Table 6) was recorded for a single object selection. From this, the conclusion might be that the device requires some time to process the brain signals. A better machine learning algorithm should improve the BCI response rate in future updates.

With regard to performance, BCI exceeded hand gestures with a result of 37 against 77. This outcome is consistent with BCI's greater accuracy than hand gestures for all tasks. The participant likely felt more satisfaction while performing the task with BCI rather than with hand gestures because there were fewer errors in the first case. Within dynamic environments, in particular, BCI scored much better than hand gestures with respect to the accuracy of selection.

The effort was not significant, but the BCI score (128) was still lower than that of hand gestures (180). The effort is defined as how much the user had to work mentally and physically to achieve the results. As the BCI is better from a physical standpoint and the hand gestures are better from a mental standpoint, the effort is insignificant. This outcome confirms (again) that even if BCI has a higher mental demand, this does not impact user perception.

The frustration level is also not significant, but BCI once again performed better than hand gestures (BCI: 63, hand gestures: 110). The users were likely more familiar with hand gestures than BCI. The Air Tap technique implemented in the Microsoft HoloLens is similar to the movement of clicking a mouse or tapping a touch screen (Vogel & Balakrishnan, 2005). This non-significance led to the conclusion that participants felt to be familiar with both selection methods.

Finally, the Weighted Rating is significant and lower for BCI than that of hand gestures, with values of 38.52 and 52.40, respectively. Metanalysis of the cumulative frequency distributions of NASA TLX Global Workload Scores (Grier, 2015) has identified the following typical values for a Computer Activity: minimum of 7.46, first percentile of 20.99, mean of 54.00, third percentile of 60.00, and a maximum of 78.00. BCI led to a weighted rating of 38.52, which is just between the 1st percentile value and the mean. This value is not harmful because it lies between 25% and 30% of the Deciles and Quartiles of the Global NASA-TLX Analysis Table of Grier (2015). For the Microsoft HoloLens hand gestures, the weighted rating is 52.40, a result that is very close to what Ro et al. (2019) found in their study (56.50). The main issue with the existing HoloLens interface is that it causes user fatigue. Since the proposed solution was aimed to solve the problem, the NASA TLX confirmed the better performance of the BCI over hand gestures.

The SUS usability test for both interfaces is displayed in Figure 44. Usability showed a higher mean value for BCI than hand gestures, but the SUS Score did not show significance. The p-value was, however, very close to significance ($p = 0.076$). A slightly larger number of participants may lead to a significant difference between BCI and hand gesture usability. Furthermore, in relation to the average SUS score of 68 established in a previous study, BCI scored higher than average while hand gestures scored lower. The major problem with existing hand gestures is user fatigue. Since the proposal was a BCI to solve this problem, the obtained usability test result has been encouraging.

This study has several limitations. First, the total number of study participants was limited due to multiple factors, including the study being performed during the COVID-19 pandemic. In addition, the recruitment of participants during this phase was difficult due to the lockdown in the province of Ontario. Even though the number of participants in this study is relatively low, 20 participants are still enough to achieve significant results. The second limitation is that gender was not considered in the selection of participants. Researchers in psychological and social sciences widely acknowledge that males and females differ in spatial ability (Halpern & Collaer, 2005; Kimura, 1999). However, since only five females participated in this study, the gender difference cannot be analyzed. Furthermore, due to COVID-19, participants were found through University's channels only because it was not allowed to recruit participants from outside of campus. For this reason, the average age was 25 and an age effect could not be considered. The lifting of pandemic restrictions will eliminate this limitation, as greater numbers of participants will be recruited more easily, thereby allowing for equal numbers of male and female participants to be selected for the study. The third limitation is related to the limited field of view of the HMD. The screen that users see inside the device is reduced in such a way their head must turn more to see all the virtual assets in the scene. This is the fundamental weakness of AR devices and a better solution is found every hardware update. The last limitation comes from the simplicity of the selection task. Selection is a universal interaction task (LaViola et al., 2017), as position and orientation are. In this study, the selection was the only interaction analyzed, which is accomplished in less time than position and orientation. Limiting the study to object selection could have led users to spend less time with their hands up than they otherwise would have when interacting with the virtual objects, thereby requiring less effort than positioning and orienting. This problem was solved by making participants select immediately successive virtual objects to approximate the position and orient interaction techniques.

Despite its limitations, this study has shown that the NASA TLX scores can improve the design of BCI applications because it provides data on both overall and individual contributors to workload (Felton et al., 2012). In addition, these scores can compare different contributions to the overall workload for both hand gestures and BCI.

CHAPTER 6 - CONCLUSIONS

This study focuses on the development of an MR application for training/support operators in maintenance activities. This application can show the virtual environment in AR, and the user can interact with the scene through a user-friendly UI. With an appropriate menu, the user can import 3D CAD in a specific position using markers. The robot can be manipulated once in the scene, and the application also allows to measure between two points. This last tool is used to understand the best position for the robot in the augmented scene. The Measurement Tool has been validated, and it has been found that its error is around 1.30% for dimensions higher than 100 cm. This result can be considered acceptable because, in industrial environments, measurements are usually larger than 100 cm.

Furthermore, the feasibility of an alternative to the classic interaction techniques in AR has been analyzed. This alternative is an SSVEP BCI, which was implemented using a device to detect stable oscillations in voltage evoked by repetitive stimulations. Hand gestures are the conventional technique to interact with the virtual assets in an AR screen. Hand gestures, however, usually lead to a high perceived workload in users, who require effort to keep the hand in the correct position to be recognized by the HMD. Therefore, an experiment was performed to demonstrate the high performance and usability of the BCI. This experiment was based on different virtual object densities as low and high and object speeds as static, low, and high. First, the time of single selection for the virtual objects in the screen was measured, and the BCI results were found to be slower than those of hand gestures for all tasks. The difference in speed was larger for static environments and very small for dynamic environments. Secondly, the accuracy of selection was measured and, BCI performed better than hand gestures, especially in dynamic tasks. This result confirmed that with hand gestures, the selection is more difficult for small and moving objects. Thirdly, a NASA TLX questionnaire was provided to the participants and, an overall workload of 38.52 for the BCI and 52.40 for the hand gestures were obtained. Thanks to its subscales, the NASA TLX effectively described the workload associated with the two interfaces. Lastly, a SUS questionnaire was provided to each participant to assess and compare the usability of both systems. BCI performed better with respect to usability than hand gestures, with a score of 77.8 versus 66.5, respectively. The result, while not significant, is still relevant. Future studies shall improve the BCI interface and implement a way to position and orient the virtual assets in the space, not just select them. Further improvement can be employing the experiment with more participants and more female individuals to consider the gender difference.

Nevertheless, the BCI shows good results in object selection accuracy, workload, and usability. There is still some delay for BCI in the selection procedure, and this could be solved in the future with a larger amount of data collected and with an upgrade of the machine-learning algorithm. Therefore, it has been demonstrated that this system can be used in dense and dynamic environments with good performances. The expectation is that there will be future analyses to pursue this research further.

REFERENCES/BIBLIOGRAPHY

Affairs, A. S. for P. (2013, September 6). *System Usability Scale (SUS)*. Department of Health and Human Services. https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html

Angrisani, L., Arpaia, P., Esposito, A., & Moccaldi, N. (2020). A Wearable Brain–Computer Interface Instrument for Augmented Reality-Based Inspection in Industry 4.0. *IEEE Transactions on Instrumentation and Measurement*, *69*(4), 1530–1539. https://doi.org/10.1109/TIM.2019.2914712

APA. (2021). *Error of commission – APA Dictionary of Psychology*. https://dictionary.apa.org/error-of-commission

Argelaguet, F., & Andujar, C. (2013). A survey of 3D object selection techniques for virtual environments. *Computers & Graphics*, *37*(3), 121–136. https://doi.org/10.1016/j.cag.2012.12.003

Bangor, A., Kortum, P., & Miller, J. (2009). Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *J. Usability Stud.*, *4*, 114–123.

Bellarbi, A., Zenati, N., Otmane, S., Belghit, H., Benbelkacem, S., Messaci, A., & Hamidia, M. (2017). A 3d interaction technique for selection and manipulation distant objects in augmented reality. *2017 5th International Conference on Electrical Engineering-Boumerdes (ICEE-B)*, 1–5.

Bowman, D., Johnson, D., & Hodges, L. (1999). Testbed evaluation of virtual environment interaction techniques. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, 26–33.

Bowman, D., Wingrave, C., Campbell, J., & Ly, V. (2001). *Using pinch gloves (tm) for both natural and abstract interaction techniques in virtual environments*.

Brooke, J. (2013). SUS: A retrospective. *Journal of Usability Studies*, *8*(2), 29–40.

Bury, G. (2019). *Mixed Reality Core Concepts—Mixed Reality*. https://docs.microsoft.com/en-us/windows/mixed-reality/design/core-concepts-landingpage

Cashion, J., Wingrave, C., & Jr, LaViola. (2012). Dense and Dynamic 3D Selection for Game-Based Virtual Environments. *IEEE Transactions on Visualization and Computer Graphics*, *18*, 634–642. https://doi.org/10.1109/TVCG.2012.40

Chaconas, N., & Höllerer, T. (2018). An Evaluation of Bimanual Gestures on the Microsoft HoloLens. *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 1–8. https://doi.org/10.1109/VR.2018.8446320

Choi, I., Rhiu, I., Lee, Y., Yun, M. H., & Nam, C. S. (2017). A systematic review of hybrid brain-computer interfaces: Taxonomy and usability perspectives. *PLOS ONE*, *12*(4), e0176674. https://doi.org/10.1371/journal.pone.0176674

Comau. (2021). *NJ-220-2.7: Characteristics and technical specifics | Comau*. https://www.comau.com/en/our-competences/robotics/robot-team/nj-220-27

Coogan, C. G., & He, B. (2018). Brain-Computer Interface Control in a Virtual Reality Environment and Applications for the Internet of Things. *IEEE Access*, *6*, 10840–10849. https://doi.org/10.1109/ACCESS.2018.2809453

De Pace, F., Manuri, F., & Sanna, A. (2018). Augmented reality in industry 4.0. *American Journal of Computer Science and Information Technology*, *6*(1), 17.

Derrick, B., & White, P. (2017). Comparing two samples from an individual Likert question. *International Journal of Mathematics and Statistics*, *18*(3).

Deshpande, A., & Kim, I. (2018). The effects of augmented reality on improving spatial problem solving for object assembly. *Advanced Engineering Informatics*, *38*, 760–775.

Dix, A., Finlay, J. E., Abowd, G. D., & Beale, R. (2003). *Human-Computer Interaction (3rd Edition)*. Prentice-Hall, Inc.

Doil, F., Schreiber, W., Alt, T., & Patron, C. (2003). Augmented Reality for Manufacturing Planning. *Proceedings of the Workshop on Virtual Environments 2003*, 71–76. https://doi.org/10.1145/769953.769962

Egger, J., & Masood, T. (2020). Augmented reality in support of intelligent manufacturing–a systematic literature review. *Computers & Industrial Engineering*, *140*, 106195.

EITC. (2021). *Smart Manufacturing and Industry 4.0—EITC*. http://www.eitc.org/research-opportunities/smart-manufacturing-and-industry-4.0

Feiner, A. O. S. (2003). The flexible pointer: An interaction technique for selection in augmented and virtual reality. *Proc. UIST*, *3*, 81–82.

Felton, E. A., Williams, J. C., Vanderheiden, G. C., & Radwin, R. G. (2012). Mental workload during brain-computer interface training. *Ergonomics*, *55*(5), 526–537.

Foley, J. D., Wallace, V. L., & Chan, P. (1984). The human factors of computer graphics interaction techniques. *IEEE Computer Graphics and Applications*, *4*(11), 13–48.

Funk, M., Bächler, A., Bächler, L., Kosch, T., Heidenreich, T., & Schmidt, A. (2017). Working with augmented reality? A long-term analysis of in-situ instructions at the assembly workplace. *Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments*, 222–229.

Grier, R. A. (2015). How High is High? A Meta-Analysis of NASA-TLX Global Workload Scores. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *59*(1), 1727–1731. https://doi.org/10.1177/1541931215591373

Grossman, T., & Balakrishnan, R. (2006). The design and evaluation of selection techniques for 3D volumetric displays. *Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology*, 3–12.

Guizzo, E. (2015). *Universal Robots UR3 Arm Is Small and Nimble, Helps to Build Copies of Itself—IEEE Spectrum*. IEEE Spectrum: Technology, Engineering, and Science News. https://spectrum.ieee.org/automaton/robotics/industrial-robots/universal-robots-ur3-robotic-arm

Halpern, D. F., & Collaer, M. L. (2005). *Sex Differences in Visuospatial Abilities: More Than Meets the Eye.* Cambridge University Press.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload* (Vol. 52, pp. 139–183). North-Holland. https://doi.org/10.1016/S0166-4115(08)62386-9

Henderson, S. J., & Feiner, S. K. (2011). Augmented reality in the psychomotor phase of a procedural task. *2011 10th IEEE International Symposium on Mixed and Augmented Reality*, 191–200.

IKEA. (2021). *Ikea App Page*. https://www.ikea.com/au/en/customer-service/mobile-apps/say-hej-to-ikea-place-pub1f8af050

Karwowski, W., Rizzo, F., & Rodrick, D. (2003). Ergonomics. In H. Bidgoli (Ed.), *Encyclopedia of Information Systems* (pp. 185–201). Elsevier. https://doi.org/10.1016/B0-12-227240-4/00061-7

Kim, S. K., Kang, S.-J., Choi, Y.-J., Choi, M.-H., & Hong, M. (2017). Augmented-reality survey: From concept to application. *KSII Transactions on Internet and Information Systems (TIIS)*, *11*(2), 982–1004.

Kimura, D. (1999). *Sex and cognition*. MIT press.

Knight, J. L. (1987). Manual control and tracking. *Handbook of Human Factors*, 182–218.

LaViola, J. J., Kruijff, E., McMahan, R. P., Bowman, D., & Poupyrev, I. P. (2017). *3D User Interfaces: Theory and Practice*. Pearson Education. https://books.google.ca/books?id=fxWSDgAAQBAJ

Looker, J. (2015). *Reaching for Holograms: Assessing the Ergonomics of the Microsoft™ Hololens™ 3D Gesture Known as the 'Air Tap'*.

Lopez, P. (2019, June 13). *Using Augmented & Virtual Reality for the Best Volvo Cars Possible*. Volvo Cars New Brunswick. https://www.volvocarsnb.com/using-augmented-virtual-reality-for-the-best-volvo-cars-possible/

Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1995). Augmented reality: A class of displays on the reality-virtuality continuum. *Telemanipulator and Telepresence Technologies*, *2351*, 282–292.

Özacar, K., Hincapié-Ramos, J. D., Takashima, K., & Kitamura, Y. (2016). 3D Selection Techniques for Mobile Augmented Reality Head-Mounted Displays. *Interacting with Computers*, *29*. https://doi.org/10.1093/iwc/iww035

Placido, F. (2021). *Intro to Brain Computer Interface*. NeurotechEDU. http://learn.neurotechedu.com/introtobci/

Pokémon GO. (2021). *Pokémon GO*. Pokémon GO. https://pokemongolive.com/

Poupyrev, I., Weghorst, S., Billinghurst, M., & Ichikawa, T. (1997). A framework and testbed for studying manipulation techniques for immersive VR. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, 21–28.

Protalinski, E. (2020, December). *NextMind is building a real-time brain computer interface, unveils Dev Kit for $399*. https://venturebeat.com/2020/01/05/nextmind-is-building-a-real-time-brain-computer-interface-unveils-dev-kit-for-399/

Putze, F., Vourvopoulos, A., Lécuyer, A., Krusienski, D., Bermúdez i Badia, S., Mullen, T., & Herff, C. (2020). Editorial: Brain-Computer Interfaces and Augmented/Virtual Reality. *Frontiers in Human Neuroscience*, *14*, 144. https://doi.org/10.3389/fnhum.2020.00144

Putze, F., Weiß, D., Vortmann, L.-M., & Schultz, T. (2019). Augmented Reality Interface for Smart Home Control using SSVEP-BCI and Eye Gaze. *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 2812–2817. https://doi.org/10.1109/SMC.2019.8914390

Qin, J., Liu, Y., & Grosvenor, R. (2016). A Categorical Framework of Manufacturing for Industry 4.0 and Beyond. *Procedia CIRP*, *52*, 173–178. https://doi.org/10.1016/j.procir.2016.08.005

Reyes, A. C. C., Del Gallego, N. P. A., & Deja, J. A. P. (2020). Mixed reality guidance system for motherboard assembly using tangible augmented reality. *Proceedings of the 2020 4th International Conference on Virtual and Augmented Reality Simulations*, 1–6.

Ro, H., Byun, J.-H., Park, Y. J., Lee, N. K., & Han, T.-D. (2019). AR Pointer: Advanced Ray-Casting Interface Using Laser Pointer Metaphor for Object Manipulation in 3D Augmented Reality Environment. *Applied Sciences*, *9*(15), 3078.

Rönkkö, K., Winter, J., & Hellman, M. (2009). Inside information 2 – usability and user research: Eight years of research and method development cooperation. *Undefined*. https://www.semanticscholar.org/paper/INSIDE-INFORMATION-2-USABILITY-AND-USER-RESEARCH%3A-R%C3%B6nkk%C3%B6-Winter/0eab217d2d406cf7fadc25d4aeefe89fc3421c6e

Rüßmann, M., Lorenz, M., Gerbert, P., Waldner, M., Justus, J., Engel, P., & Harnisch, M. (2015). Industry 4.0: The future of productivity and growth in manufacturing industries. *Boston Consulting Group*, *9*(1), 54–89.

Saha, S., Mamun, K. A., Ahmed, K., Mostafa, R., Naik, G. R., Darvishi, S., Khandoker, A. H., & Baumert, M. (2021). Progress in Brain Computer Interface: Challenges and Opportunities. *Frontiers in Systems Neuroscience*, *0*. https://doi.org/10.3389/fnsys.2021.578875

Sanna, A., Manuri, F., Lamberti, F., Paravati, G., & Pezzolla, P. (2015). Using handheld devices to sup port augmented reality-based maintenance and assembly tasks. *2015 IEEE International Conference on Consumer Electronics (ICCE)*, 178–179.

Schröder-Kroll, R., Blom, K., & Beckhaus, S. (2008). *Interaction techniques for dynamic virtual environments*. 57–68.

Segovia, D., Mendoza, M., Mendoza, E., & González, E. (2015). Augmented reality as a tool for production and quality monitoring. *Procedia Computer Science*, *75*, 291–300.

Semple, K. (2021). *MRTK-Unity Developer Documentation—Mixed Reality Toolkit*. https://docs.microsoft.com/en-us/windows/mixed-reality/mrtk-unity/

Sidenmark, L., Clarke, C., Zhang, X., Phu, J., & Gellersen, H. (2020). *Outline Pursuits: Gaze-assisted Selection of Occluded Objects in Virtual Reality*. https://doi.org/10.1145/3313831.3376438

Silaghi, H., Rohde, U., Spoială, V., Silaghi, A., Gergely, E., & Nagy, Z. (2014). Voice command of an industrial robot in a noisy environment. *2014 International Symposium on Fundamentals of Electrical Engineering (ISFEE)*, 1–5. https://doi.org/10.1109/ISFEE.2014.7050596

Si-Mohammed, H., Argelaguet, F., Casiez, G., Roussel, N., & Lécuyer, A. (2017). *Brain-Computer Interfaces and Augmented Reality: A State of the Art*. https://doi.org/10.3217/978-3-85125-533-1-82

Si-Mohammed, H., Petit, J., Jeunet, C., Argelaguet, F., Spindler, F., Évain, A., Roussel, N., Casiez, G., & Lecuyer, A. (2020). Towards BCI-Based Interfaces for Augmented Reality: Feasibility, Design and Evaluation. *IEEE Transactions on Visualization and Computer Graphics*, *26*(3), 1608–1621. https://doi.org/10.1109/TVCG.2018.2873737

Singh, A. K., Wang, Y.-K., King, J.-T., & Lin, C.-T. (2020). Extended interaction with a bci video game changes resting-state brain activity. *IEEE Transactions on Cognitive and Developmental Systems*, *12*(4), 809–823.

Sorko, S. R., & Brunnhofer, M. (2019). Potentials of augmented reality in training. *Procedia Manufacturing*, *31*, 85–90.

Spataro, R., Chella, A., Allison, B., Giardina, M., Sorbello, R., Tramonte, S., Guger, C., & La Bella, V. (2017). Reaching and grasping a glass of water by locked-in ALS patients through a BCI-controlled humanoid robot. *Frontiers in Human Neuroscience*, *11*, 68.

Stack Overflow. (2021). *Stack Overflow—Where Developers Learn, Share, & Build Careers*. Stack Overflow. https://stackoverflow.com/

Stilgherrian. (2018). *How brand new science will manage the fourth industrial revolution*. ZDNet. https://www.zdnet.com/article/how-brand-new-science-will-manage-the-fourth-industrial-revolution/

Tang, A., Owen, C., Biocca, F., & Mou, W. (2003). Comparative Effectiveness of Augmented Reality in Object Assembly. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 73–80. https://doi.org/10.1145/642611.642626

Tanriverdi, V., & Jacob, R. J. (2000). Interacting with eye movements in virtual environments. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 265–272.

Unity Technologies. (2021). *HoloLens Spatial Mapping—Unity Manual*. https://docs.unity3d.com/es/2018.4/Manual/SpatialMapping.html

Universal Robots. (2021). *UR10e Collaborative industrial robot—Payload up to 12.5 kg*. https://www.universal-robots.com/products/ur10-robot/

Vidal, J. J. (1973). Toward direct brain-computer communication. *Annual Review of Biophysics and Bioengineering*, *2*(1), 157–180.

Vogel, D., & Balakrishnan, R. (2005). Distant freehand pointing and clicking on very large, high resolution displays. *UIST: Proceedings of the Annual ACM Symposium on User Interface Softaware and Technology*, 33–42. https://doi.org/10.1145/1095034.1095041

Vortmann, L.-M., & Putze, F. (2020). Attention-Aware Brain Computer Interface to Avoid Distractions in Augmented Reality. *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–8. https://doi.org/10.1145/3334480.3382889

Vourvopoulos, A., Pardo, O. M., Lefebvre, S., Neureither, M., Saldana, D., Jahng, E., & Liew, S.-L. (2019). Effects of a brain-computer interface with virtual reality (VR) neurofeedback: A pilot study in chronic stroke patients. *Frontiers in Human Neuroscience*, *13*, 210.

Vuforia. (2021). *Vuforia Developer Portal*. https://developer.vuforia.com/

Wang, J., Shen, Y., & Yang, S. (2019). A practical marker-less image registration method for augmented reality oral and maxillofacial surgery. *International Journal of Computer Assisted Radiology and Surgery*, *14*. https://doi.org/10.1007/s11548-019-01921-5

Wang, W., Wang, F., Song, W., & Su, S. (2020). Application of augmented reality (AR) technologies in inhouse logistics. *E3S Web of Conferences*, *145*, 02018.

Wei, C.-S., Wang, Y.-T., Lin, C.-T., & Jung, T.-P. (2018). Toward drowsiness detection using non-hair-bearing EEG-based brain-computer interfaces. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *26*(2), 400–406.

Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2015). *Engineering psychology and human performance*. Psychology Press.

Wolpaw, J., & Wolpaw, E. (2012a). *Brain-Computer Interfaces: Principles and Practice* (p. 424). https://doi.org/10.1093/acprof:oso/9780195388855.001.0001

Wolpaw, J., & Wolpaw, E. W. (2012b). *Brain-computer interfaces: Principles and practice*. OUP USA.

Wu, D., Lance, B. J., Lawhern, V. J., Gordon, S., Jung, T.-P., & Lin, C.-T. (2017). EEG-based user reaction time estimation using Riemannian geometry features. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *25*(11), 2157–2168.

Yilmaz, M., & Kilinc, I. (2018). An Exploratory Study to Assess Digital Map Zoom/Pan/Rotate Methods with HoloLens. *Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, *22*(2), 458–463.

Zeller, M. (2019). *HoloLens (1st gen) Hardware*. https://docs.microsoft.com/en-us/hololens/hololens1-hardware

Zubizarreta, J., Iker, A., & Aiert, A. (2019). A framework for augmented reality guidance in industry. *The International Journal of Advanced Manufacturing Technology*, *102*(9–12), 4095–4108.

APPENDICES

Appendix A – Consent Form

Title of Study: “HUMAN PERFORMANCE AND MENTAL WORKLOAD IN AUGMENTED REALITY: BRAIN COMPUTER INTERFACE ADVANTAGES OVER GESTURES”

You are asked to participate in a research study conducted by **Mr. Silvio Da Col (Student), Dr. Eunsik Kim (Research Coordinator)**, from the **Mechanical, Automotive & Materials Engineering** at the University of Windsor. The results will be contributed to a graduate thesis project.

If you have any questions or concerns about the research, please feel to contact:

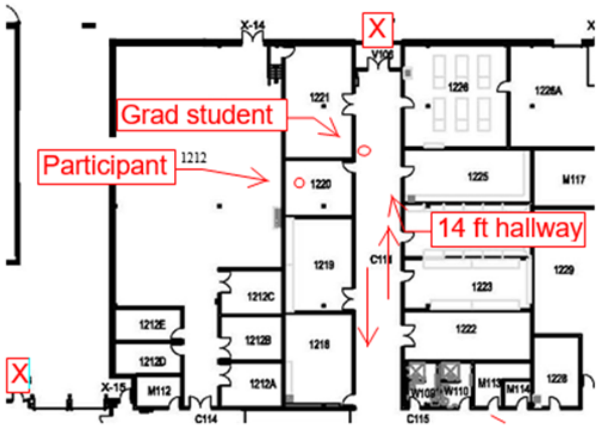
- Mr. Silvio Da Col: dacol@uwindsor.ca
- Dr. Eunsik Kim: eunsik.kim@uwindsor.ca

PURPOSE OF THE STUDY

The research is based on building a Brain Computer Interface (BCI) for hologram selection in Augmented Reality (AR). A BCI is an interface which uses the brain signals to trigger some actions in a virtual environment. For example, if you are focusing on a button in the virtual environment, the interface will acquire your brain signals and will press that button without the need of moving other body parts. To acquire those signals, the NextMind device is used; this is a new mass market device that reads the image projected onto the visual cortex of the user and triggers an action.
Let us go back to the button example. Instead of clicking that button with your brain signals, you can use the actual hand-gestures already implemented in the Microsoft HoloLens. Nevertheless, these hand-gestures are typically not precise and lead to high perceived workloads and low accuracy. The purpose of the research is demonstrating that the new way of selection by brain signals can perform better than the hand-gestures currently working on HoloLens under speed, accuracy, and workload point of view.

PROCEDURES

The building and room number for the study are: Centre for Engineering Innovation (CEI), Room 1220. The following diagram shows where the lab is situated in CEI, entry and exit to the buildings:



All lab members conform with health and safety protocols (https://www.uwindsor.ca/returntocampus/) and they have access to personal protective equipment (PPE). Masks must be worn in all common areas (including hallways, washrooms, etc.). Masks must also be worn in the lab when other personnel are present. The “COVID-19 Mandatory Non-medical Mask Policy” can be found at this link:
https://lawlibrary.uwindsor.ca/Presto/content/Detail.aspx?ctID=OTdhY2QzODgtNjhlYi00ZWY0LTg2OTUtNmU5NjEzY2JkMWYx&rID=MjQw&qrs=RmFsc2U=&q=KFVuaXZlcnNpdHlfb2ZfV2luZHNvcl9DZW50cmFsX1BvbGljaWVzLkFsbFRleHQ6KENPVklEKSk=&ph=VHJ1ZQ==&bckToL=VHJ1ZQ==&rrtc=VHJ1ZQ==

If you volunteer to participate in this study, this is the general followed procedure:

1. Upon arrival, each participant will be welcomed and will receive a brief introduction of the experimental procedure.
2. Participant will be asked about his/her experience in Brain Computer Interfaces and Augmented Reality.
3. Participant will be asked to wear the HoloLens device.
4. The session will be recorded to check if the acquired time and accuracy data are correct. Participant faces will not be included.
5. Participant will be asked to complete a training session on the HoloLens.
6. There will be a break if needed.
7. Whenever the participant will be comfortable, he/she will be asked to complete the assigned tasks on the HoloLens.

8. Participant will be asked to complete the NASA TLX and the SUS about the HoloLens device (NASA TLX: https://en.wikipedia.org/wiki/NASA-TLX and SUS: https://en.wikipedia.org/wiki/System_usability_scale)
9. Participant will be informed of the obtained results and there will be a break.
10. Whenever the participant will be comfortable, he/she will be asked to wear the HoloLens device and the NextMind device.
11. Participant will be asked to complete a training session on the NextMind device.
12. There will be a break if needed.
13. Whenever the participant will be comfortable, he/she will be asked to complete the assigned tasks on the NextMind device.
14. Participant will be asked to complete the NASA TLX and the SUS about the NextMind device.
15. Participant will be informed of the obtained results and there will be a break.
16. The experiment is now completed.

There will be only one session each participant. It is expected that the session will last 2 hours.

POTENTIAL RISKS AND DISCOMFORTS

The risk is always low because the environment is controlled, and the used devices (HoloLens and NextMind) are safe. Physical risks could happen for improper movements during data collection or dizziness because of the devices use. Investigators will show which are the proper movements to be followed. If participant feel discomfort or pain, at any point they can immediately stop the experiment. They will be helped by me (Silvio Da Col) and by the Supervisor (Dr. Eunsik Kim) to solve their problem. The research team will be responsible for what happens to participants.
Session will be recorded and so psychological/emotional risks could happen. Participants could feel uncomfortable or pressed but we will reassure participants to performs as far as physically and mentally possible. The purpose of the research is to access the real performance of participants.
Social risks can be again associated to the recordings because participants could think that this could increase the chances of being recognised. The face will not be visible in the video and moreover, videos will be accessible just to Mr. Silvio Da Col and Dr. Eunsik Kim.

After each session, all the devices, equipment and door handles will be properly sanitised and cleaned before the other participants use/touch them. Each participant will be asked to sanitise his/her hands properly before entering into the lab and participants are required to wear the masks and keep proper social distance of 2 meters.
Possible discomforts could derive from the coupling between the experiments and the respect of all COVID-19 rules. To reduce this risk, participants will be guided in respecting all the procedures.

POTENTIAL BENEFITS TO PARTICIPANTS AND/OR TO SOCIETY

Participants who are not familiar with Augmented Reality or with Brain Computer Interfaces can gain some skills in these environments. These devices will be more and more popular in the future and so participants will have a general overview of them.
This research is innovative from the point of view of the holograms selection in Augmented Reality environments. More and more Companies now are going in the direction of designing and validating production lines in Augmented Reality and this research will try to find an interface to reduce user workload and increase virtual objects selection accuracy and speed.

COMPENSATION FOR PARTICIPATION

Participants will be paid $20 for the session. Source of funding is from Supervisor Dr. Eunsik Kim. There will not be any financial support for transportation to come to campus.

CONFIDENTIALITY

Any information that is obtained in connection with this study and that can be identified with you will remain confidential and will be disclosed only with your permission.
We will store the raw data (recordings and measured data) by number only. We will store them in our personal University of Windsor OneDrive protected cloud. Only the faculty Supervisor (Dr. Eunsik Kim) and me (Silvio Da Col) can access the data. Data will be stored on an Excel spreadsheet and be used for the duration of the study (until August 2021). Names will be removed to ensure confidentiality within participants.

The activity will be videotaped and participants have the right to review the tapes associated to their experiment session. If participants want to delete the recording associated to their experiment session, they have the right to do this. All the recordings will be accessible by Mr. Silvio Da Col and Dr. Eunsik Kim until the end of the study (August 31st, 2021).

PARTICIPATION AND WITHDRAWAL

Participants, who will want to withdraw their data from the study, can stop the session at any point without any penalty. Participants, who want to withdraw their already collected data, need to contact the Supervisor (Dr. Eunsik Kim) and notify that they would like to withdraw the collected data. This procedure shall be done within 2 weeks after completing the experiment. Participants will not

be allowed to withdraw their data after two weeks participating in the study because this could have an impact on the completion of the thesis. In fact, data must be post-processed.
If the participant wants to stop before the first trial (corresponding to the 50% of the experiment session), they can do it but there will not be any compensation. Participants who want to withdraw after the first trial will receive a portion of compensation based on how long they have participated in the study. Data will not be collected for the participants that has withdrawn from study.

FEEDBACK OF THE RESULTS OF THIS STUDY TO THE PARTICIPANTS

Participants will be informed of the overall results of the study via publications from the study or research result summaries on the University of Windsor website.

Web address: https://scholar.uwindsor.ca/research-result-summaries/
Date when results are available: August 31st, 2021

SUBSEQUENT USE OF DATA

These data may be used in subsequent studies, in publications and in presentations.

RIGHTS OF RESEARCH PARTICIPANTS

If you have questions regarding your rights as a research participant, contact: Office of Research Ethics, University of Windsor, Windsor, Ontario, N9B 3P4; Telephone: 519-253-3000, ext. 3948; e-mail: ethics@uwindsor.ca

SIGNATURE OF RESEARCH PARTICIPANT/LEGAL REPRESENTATIVE

I understand the information provided for the study "HUMAN PERFORMANCE AND MENTAL WORKLOAD IN AUGMENTED REALITY: BRAIN COMPUTER INTERFACE ADVANTAGES OVER GESTURES" as described herein. My questions have been answered to my satisfaction, and I agree to participate in this study. I have been given a copy of this form.

____
Name of Participant

______________________________________ ____________________
Signature of Participant Date

SIGNATURE OF INVESTIGATOR

These are the terms under which I will conduct research.

______________________________________ ____________________
Signature of Investigator Date

Appendix B – Biometrics Data Form

Name:

Date of Birth:

Gender:

Stature (approximately):

Weight (approximately):

Please answer these questions:

Question 1: Are you familiar with Virtual Reality or Augmented Reality?

Yes **No**

Question 2: If you have answered **Yes** in Question 1, please answer this question: have you ever used a Microsoft HoloLens device?

Yes **No**

Question 3: Are you familiar with Brain-Computer Interfaces?

Yes **No**

Question 4: If you have answered **Yes** in Question 3, please answer this question: have you ever used the NextMind device?

Yes **No**

Question 5: Do you have any discomfort in using your upper limbs?

Yes **No**

Question 6: Which is your dominant hand?

Right **Left**

Question 7: Do you have any vision problem in normal conditions?

Yes **No**

Appendix C – REB Approval

Today's Date: March 02, 2021
Principal Investigator: Mr. Silvio Da Col
REB Number: 38917
Research Project Title: REB# 21-015: "A Brain Computer Interface to Reduce the Workload Involved in AR Holograms Selection"
Clearance Date: March 2, 2021
Project End Date: August 31, 2021

This is to inform you that the University of Windsor Research Ethics Board (REB), which is organized and operated according to the Tri-Council Policy Statement and the University of Windsor Guidelines for Research Involving Human Participants, has granted approval to your research project. This approval is valid for one year after the clearance date noted above.

An annual Progress Report must be submitted for renewal of the project. The REB may ask for monitoring information at some time during the project's approval period. A Final Report must be submitted at the end of the project to close the file.

During the course of the research, no deviations from, or changes to, the protocol or consent form may be initiated without prior written approval from the REB. Approval for modifications to an ongoing study can be requested using a Request to Revise Form.

Investigators must also report promptly to the REB:
a) changes increasing the risk to the participant(s) and/or affecting the conduct of the study;
b) all adverse and unexpected events that occur to participants;
c) new information that may affect the risks to the participants or the conduct of the study.

Forms for submissions, notifications, or changes are available on the REB website: www.uwindsor.ca/reb. If your data are going to be used for another project, it is necessary to submit a secondary use of data application to the REB.

Sincerely,

Appendix D – RSC Approval

University of Windsor

Phase 3 COVID-19 Research Resumption Plan

Section vi: Signatures

Research Resumption Plan submitted by:

Eunsik Kim

Faculty Name	Faculty Signature

Approved by:

MAJID AHMADI

Dean or Associate Dean Name	Dean or Associate Dean Signature

Appendix E – System Usability Scale

I think that I would like to use this system frequently.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	2	3	4	5

I found the system unnecessarily complex.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	2	3	4	5

I thought the system was easy to use.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	2	3	4	5

I think that I would need the support of a technical person to use this system.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	2	3	4	5

I found the various functions in this system were well integrated.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	2	3	4	5

I thought there was too much inconsistency in this system.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	2	3	4	5

I would imagine that most people would learn to use this system very quickly.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	2	3	4	5

I found the system very awkward to use.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	2	3	4	5

I felt very confident using this system.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	2	3	4	5

I needed to learn a lot of things before I could get going with this system.

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	2	3	4	5

Appendix F – Post-Experiment Questionnaire

How would you rate the Air Tap technique in speed?

1	2	3	4	5

How would you rate the NextMind technique in speed?

1	2	3	4	5

How would you rate the Air Tap technique in accuracy?

1	2	3	4	5

How would you rate the NextMind technique in accuracy?

1	2	3	4	5

How adequate do you feel the 3D interface was?

1	2	3	4	5

How adequate do you feel the time allotted for practice was?

1	2	3	4	5

How comfortable were you using the HoloLens device?

1	2	3	4	5

When determining how much you like using a selection technique, how much influence does ease-of-use have on your decision?

1	2	3	4	5

When determining how much you like using a selection technique, how much influence does speed have on your decision?

1	2	3	4	5

Which was your preferred selection technique between Air Tap and NextMind?

Air Tap	NextMind

VITA AUCTORIS

NAME:	Silvio Da Col
PLACE OF BIRTH:	Pieve di Cadore, BL, Italy
YEAR OF BIRTH:	1998
EDUCATION:	Liceo Scientifico E. Fermi, Pieve di Cadore, BL, Italy, 2011-2016
	Politecnico di Torino, B.Sc., Turin, TO, Italy, 2016-2019
	Politecnico di Torino, M.Sc., Turin, TO, Italy, 2019-2021
	University of Windsor, M.Sc., Windsor, ON, 2020-2021