R.E.I.N.A.
Towards Pervasive Interface Agents that Transcend the Physical-Digital Worlds

by

Elena Chong Loo Kodama

B.S., Rose-Hulman Institute of Technology (2016)

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning
in partial fulfillment of the requirements for the degree of

Master of Science in Media Arts and Sciences

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2020

Author . . . .
Program in Media Arts and Sciences
August 17, 2020

Certified by . . . .
Joseph A. Paradiso
Alexander W Dreyfoos (1954) Professor
Program in Media Arts and Sciences
Thesis Advisor

Accepted by . . . .
Tod Machover
Academic Head, Program in Media Arts and Sciences

R.E.I.N.A.

Towards Pervasive Interface Agents that Transcend the Physical-Digital Worlds

by

Elena Chong Loo Kodama

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning on August 17, 2020, in partial fulfillment of the requirements for the degree of

Master of Science in Media Arts and Sciences

Abstract

Our generation is spending more time in front of computer screens, in part due to the onset of the COVID-19 pandemic. In front of our screens, we see multiple notes, folders, windows, and applications that somehow replicate a metaphoric desk. The way we navigate this digital system has not changed much in the past four decades. However, in the last two years, the technological landscape is showing sign of a potential shift that could enable novel ways of navigating the physical and digital information spaces. Augmented and virtual reality systems are spinning off from laboratories around the world, promising to offer better ways of storytelling, learning, working, and interacting. Just like how window interfaces became established with the development of computer graphics, and voice assistants with smartphones, augmented reality could generate a new natural multi-modal interface for enhancing our interactions with the physical and digital worlds.

In this thesis, I am introducing the concept of a Pervasive Interface Agent, a cross-platform agent-based interface that acts as a lifelong companion, assisting us autonomously in both physical and digital worlds. Additionally, I am presenting R.E.I.N.A., a phase 1 prototype of a Pervasive Interface Agent together with the Media Lab Tour Guiding System. A set of experiments called the R.E.I.N.A. experiment were designed and conducted on 101 international participants to gain insights on the effect of agent embodiment in a remote tour guiding experience. The results are also discussed in this work.

Thesis Advisor:
Joseph A. Paradiso
Alexander W Dreyfoos (1954) Professor
Program in Media Arts and Sciences

R.E.I.N.A.
Towards Pervasive Interface Agents that Transcend the Physical-Digital Worlds

by

Elena Chong Loo Kodama

This thesis has been reviewed and approved by the following committee members

Joseph A. Paradiso . . . .

Thesis Advisor
Alexander W Dreyfoos (1954) Professor
MIT Media Lab, Responsive Environments Group

Mark Billinghurst . . . .

Thesis Reader
Professor of Human-Computer Interaction
University of South Australia and University of Auckland

Cynthia Breazeal . . . .

Thesis Reader
Professor of Media Arts and Sciences
MIT Media Lab, Personal Robots Group

Acknowledgments

I am truly grateful for all the people who helped me move this thesis forward. First and foremost, I wish to express my sincere gratitude to my advisor Joe Paradiso, for believing, supporting, and advising me for the past two years and for giving me a medium to push forward my vision of the future. To my thesis reader Cynthia Breazeal for making me see more clearly what I wanted to build and for providing great feedback to improve this work. To my thesis reader Mark Bilinghurst for helping me cover the historical view of this work more broadly.

I want to express my thanks to Nan Zhao for taking me under her wing since day one at the MIT Media Lab, and for constantly encouraging me and providing valuable feedback on my work for the past two years. Thanks to Abhi Jain, Patrick Chwalek, Belén Saldías Fuentes, Pedro Colón Hernandez, Pat Pataranutaporn, Vik Parth, Peter Sherman and the entire Responsive Environments Group for allowing me to see things with a diverse perspective, providing me with tools and resources, and for your friendship. To the whole Media Lab Community for giving me a place to meet smart and talented individuals who have one way or another contributed to my work. To my talented collaborators Jiani Zeng and Liz B. De la Torre for teaching me about sophisticated design and for making the class at Harvard School of Design much more fun. To my wonderful and awesome undergrad interns, Malvika Jain, Emily Mattlin, Valeria Yang, Amani Toussaint, and Bowen Wu for making my works shine brighter with your brilliance and talents.

To all my mentors and friends from Panama, Rose-Hulman, and beyond, for their unconditional friendship and support throughout all these years.

To my parents Guillermo Chong and Yuly Loo, and my siblings Javier Chong and Jhoany Chong. To my entire Chong (张), Loo (罗), and Kodama families for being who you are and shaping who I have become. Thank you all for the love and support sent to me across space and time. Finally, I want to express my deepest gratitude to my husband, editor, and best friend Chris Kodama for giving me all your unconditional love and support.

Contents

List of Figures

List of Tables

"The hope is that, in not too many years, human brains and computing machines will be coupled together very tightly, and that the resulting partnership will think as no human brain has ever thought and process data in a way not approached by the information-handling machines we know today."

J. C. R. Licklider

Chapter 1

Introduction

"The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it."

Mark Weiser

1.1 Towards Pervasive Interface Agents

Our physical world is becoming more connected with the virtual one thanks to advances in communication technology and ubiquitous computing. This will result in an expanding information space that is complex and difficult to navigate. Our main gateway to the digital world right now is through computer screens for displaying and navigating the virtual information space. However, this interface deprives us from experiencing and connecting with the physical environment since it is constrained to the dimensionality and affordances of a screen.

In 1960, J. C. R. Licklider published his influential work on Man-Computer Symbiosis [52], in which he stated that,

"Symbiotic partnership [between a man and a computer] will perform intellectual operations much more effectively than man alone can perform them."

Throughout history, researchers have explored agent-based interfaces for augmenting our cognitive abilities and assisting us with tedious computational tasks. In the 1990's, software interface agent systems was explored extensively by Maes's Group at the MIT Media Lab. Some the topics her group worked on were collaborative interface agents [59, 49], learning interface agents [47], agent embodiment [46], and much more [57, 62, 40]. In 1996, Billinghurst et al. proposed a virtual reality (VR) intelligent agent interface with rule-based expert system (procedural), multimodal input, and natural language processing for simulating natural interaction between a human and a virtual agent [10]. A year later, Henry Lieberman and Brenda Laurel introduced the idea of Autonomous interface agents [53, 50], and in 2003, Lieberman and Ted Selker expanded the idea to further describe using agents as the new interface [54]. These papers could be summarized as describing computer programs (interface agents) that incorporate Artificial Intelligence (AI) for assisting and collaborating with us to complete tasks in our work environment.

Ubiquitous computing presents new opportunities for improving these agent-based interfaces by decoupling them from the computer screen and the virtual world, and have them manifest in our physical environment using augmented reality (AR) technology. AR enables us to interact in real-time with digital content overlaid on top of the physical world. In recent years, AR hardware has matured enough to make applications for localizing and mapping of the physical environment more accessible. This makes it feasible to build situational awareness and context awareness into these agent interfaces, so they can keep augmenting us at any time and in any space in our physical and digital environments.

Furthermore, this new interface agent can manifest as a 3D embodied agent in the physical world. Cognitive and interaction researchers have found that embodiment plays an important role in influencing our cognition and behaviors. Specifically, an agent that can traverse in our physical environment can directly stimulate our visual perception [45, 41].

Agent embodiment is nothing new. In fact, researchers has been studying, designing, and evaluating systems for creating embodied mobile social robots and intelligent virtual agents for accompanying, guiding, and augmenting us for decades. Commercial mobile robots such as, Nao[1] [37], Pepper[2] [73], Temi[3], Buddy[4], and Roomba[5] [33] have been helping us with house-cleaning services and task processing in the physical environment. Similarly, commercial social robots, such as Jibo[6] and Kuri[7], have been helping researchers understand how people interact with and use their social companion at home. Extensive reviews on social robots have also been investigated for use in education [65, 7], human-robot interaction [32, 6, 36, 24, 23, 4], assistive care [13, 64, 74, 3], and much more. They have shown to be effective at enhancing our cognition and affection towards them in part due to their physical corporeal presence.

In the virtual world, one iconic example of a virtual assistant is Microsoft Clippy, deployed in 1996 and was intended to help Microsoft Office workers. Prior to its deployment, Microsoft conducted several focus-group testing and found that women disliked the design of Clippy because they found it to be very male-focused, impolite,

[1]https://www.softbankrobotics.com/emea/en/nao
[2]https://www.softbankrobotics.com/emea/en/pepper
[3]https://www.robotemi.com/
[4]https://buddytherobot.com/en/buddy-the-emotional-robot/
[5]https://www.irobot.com/roomba
[6]https://jibo.com/
[7]https://www.heykuri.com/life-with-kuri/

and disturbing. Microsoft had a very male-heavy leadership back then and ignored the results from the focus-group[8]. Eventually, Clippy became known as one of the most annoying user interfaces in the history of virtual assistants. However, it provided valuable lessons that are helping researchers figure out what the next generation of virtual agent might need to be effective assistant [80, 72, 56]. Some of the issues with Clippy were that it was not optimized for repeated use, lacked diversity in its embodiment, and did not possess context-awareness that could have helped it understands users feedback and provide more personalized assistance.

Virtual agents have evolved greatly since then. Nowadays, especially in the gaming community, we have seen the presence of “Non-Player Characters" (NPC), which are embodied characters or avatars controlled by the game (computer), used for assisting or providing information to players in the virtual space [16]. In games like Destiny[9], we even have an embodied AI called a Ghost, who serves as our personal assistant and travels the virtual world with us. In addition to entertainment, developers have created web-based and virtual reality-based digital assistants (3D avatars, embodied conversational agents, chatbot, etc) [18, 2, 19] for tutoring, improving well-being [43], and processing information.

Robots and virtual agents have had a clear world division in terms of assistance across space and time: robots can navigate in the physical, while virtual agents in the virtual world. In other words, although robots can be effective companions and can navigate in the physical world, they cannot easily provide us with assistance everywhere yet due to their constraining physical corporeal presence. Similarly, virtual agents usually stay in their virtual environment, constrained by their application and game environment.

Pokemon Go, one of the most successful augmented reality applications with more than one billion downloads as of early 2019, has attempted to put digital embodied characters in the physical world. The novelty generated a fair amount of press and exposed issues that were not previously considered, for example contributing to some

[8] https://www.theatlantic.com/technology/archive/2015/06/clippy-the-microsoft-office-assistant-is-the-patriarchys-fault/396653/

[9] https://www.bungie.net/

fatal accidents and nuisances [85, 79, 75]. Although the digital characters in Pokemon Go were not designed to be sophisticated social agent or companion, it served as a good starting point of what pervasive interface agents could become. We know now that people are open to the idea of having a digital agent as a companion, thus with deeper understanding of the spatio-temporal (time and space) effect of these virtual assistants on us, we can eventually develop modern interface agents with situational awareness and context awareness that can inhabit and interact with us in our physical and digital worlds. This pervasive interface agent could one day streamline the physical/digital worlds by creating a symbiotic partnership with us, augmenting our cognitive abilities like never before.

Augmented Reality (also known as spatial computing) is poised to become ubiquitous in the next few decades. Similar to AI, it is believed that AR will have a profound effect on our daily life. Currently, more than 88% of all mid-sized businesses are already using AR[10], and the global market size is projected to be around $198 billion US by 2025[11]. As more investment is put into AR technology, we can fast-track and push forward research to help us better understand which context should a robot or virtual agent be used.

One proposed approach to make artificial assistants accessible across space and time is Migratable User Interfaces (UI). One example of this a UI that migrates from a desktop computer to a handheld device [39]. Recently, Tejwani et. al conducted a study on migratable AI in which they migrated a conversational AI agent across multiple embodiments, such as a smart speaker device, a smart display, a mobile robot, etc [89]. In the self-reported questionnaires, they found that users reported "highest trust, competence, likeability, and social presence towards their conversational agent when both identity and information were migrated across embodiments."

The concept of a Pervasive Interface Agent, which will be introduced further in this thesis, aims to push the theory of Migratable UI/AI forward by using augmented reality technology to sustain the identity and embodiment of our personal virtual

[10] https://www.vxchnge.com/blog/augmented-reality-statistics
[11] https://www.statista.com/statistics/897587/world-augmented-reality-market-value/

agents to better assist us across space and time.

As we move towards Pervasive Interface Agents, we need to deepen our understanding of the spatio-temporal effect of digital agents inhabiting our physical environment. Thus, this thesis contributes towards this goal. The contributions of this work are as follows:

1.2 Contributions

In this thesis, I am presenting the vision of a Pervasive Interface Agent. I will introduce R.E.I.N.A., a responsive and embodied indoor navigation agent that assists in tour-guiding. R.E.I.N.A. can manifest itself as a disembodied or embodied agent with non-humanoid features. A virtual demonstration was built to showcase R.E.I.N.A. freely navigating in a 3D-scanned model of the MIT Media Lab building. This demonstration can also work at the MIT Media Lab by positioning the 3D-scanned model to its real world coordinates (situational awareness). R.E.I.N.A. can reliably navigate not just in the X-Y plane, but can also move up in the Z-direction, essentially "climbing" up stairs. R.E.I.N.A. can also understand specific destinations and can move towards them when told (basic context-awareness).

Additionally, a collection of 10 fictional projects were scripted, 3D rendered and animated to form part of an interactive "Gallery" used for a set of user studies designed to provide a better understanding of the effect of different features of the tour guiding agent on memory retention, attention, social presence, and memorability in a static setting and a physical environment. This set of studies is called, The R.E.I.N.A. Experiment, which consists of an online pilot study (n=20), a remote gallery tour user study (n=101), and an in-person user study based on the MIT Media Lab Tour Guiding Experience. Two of the studies have been completed and the results will be discussed in this thesis.

Due to the ongoing COVID-19 Pandemic, the in-person user study will remain as a future work. The MIT Media Lab Tour Guiding Experience user study will provide insights on the spatio-temporal effect of having R.E.I.N.A. accompanying, guiding, and assisting us across space and time.

In summary, this thesis contributes an introductory design framework for a Pervasive Interface Agent called R.E.I.N.A. with situational awareness and context awareness to help study spatio-temporal effect on our cognition while it interacts with us in our physical environment. This work also provides a demonstration of the navigation system and insights into how R.E.I.N.A. could provide a memorable tour guiding experience in the future.

1.3 Chapter Summaries

The remainder of this thesis is arranged as follows: Chapter 2 describes my motivation for this thesis and provides a brief history of user interface and software agent. Chapter 3 reviews previous work done on embodied agents, augmented reality navigation and tour guiding systems, and human factor metrics. Chapter 4 introduces my vision of a Pervasive Interface Agent. Chapter 5 describes the R.E.I.N.A. system and the MIT Media Lab Tour Guiding System and the implementation. Chapter 6 talks about the R.E.I.N.A. Experiment and discusses the results of the experiments. Chapter 7 concludes this work.

Chapter 2

Background

> “The screen is a window through which one sees a virtual world. The challenge is to make that world look real, act real, sound real, feel real.”
>
> Ivan Sutherland

2.1 Motivation

In this brave new world, where massive amounts of information are readily available across space and time, we have reached a tipping point where we feel the need to break out of the constraints and affordances set by the computer screen. Personally, as I write this thesis from home, I find myself using two large computer screens (27" and 24"), and at some point even two laptops at the same time - not to mention my cell phone screen. I fill these screens with documents, websites, notes, and folders, in a desperate attempt to organize my thoughts as I write. Even with hundreds of tabs open in my browser, I sometimes end up staring blankly in front of my screens, as if I had no clue where to even start looking. Wouldn't it be wonderful to have some sort of magical creature acting as my mental secretary, helping me talk through and organize my thoughts and ideas? Maybe the creature would sift through the facts I told it and quickly find patterns, much faster than I could, about my research, and I wouldn't have needed to spend what felt like half a year feeling blocked, grasping for the next step forward. At the very least, I though, it would have been a great companion to have as I struggled through the many phases of my academic journey.

As early as the 1700s, mathematicians and logicians had already pondered around the idea of externalizing human thinking by creating artificial non-human machines. But it was not until the onset of World War II that great mathematicians like Claude Shannon and Alan Turing came forth to make feasible this idea of an "electronic brain". In 1948, Shannon's work in "A Mathematical Theory of Communication" laid out the foundation for the field of information theory and set out the stage for the digital age [86]. In 1950, Alan Turing published "Computing Machinery and Intelligence", in which he proposed "The Imitation Game", a question game that asks if machine can think like a human [91]. This game involves a human interrogator, a human player, and a machine. The role of the interrogator is to discover which was the human being and which was the machine through a series of five-minute keyboard Question-and-Answer. The objective of the machine is to be indistinguishable from a human being. A machine is said to pass the test if it can convince the interrogator more than 30% of the time that it is a human. This became later known as the Turing Test and will be passed six decades later in 2013 by a computer program called Eugene Goostman simulating a 13 year old Ukranian boy.

The term "artificial intelligence" (AI) was coined by John McCarthy in 1956 after delivering a workshop of the same name [77]. Two years later, he will develop Lisp, a high-level programming language that is still used extensively for AI research in the present time. The following year, Arthur Samuel, a computer scientist that developed the first self-learning checkers-playing computer program in 1952, coined the term "Machine Learning" to refer to a computer program that could play a game of chess better than the human who wrote it. After these events in the 50s, AI popularity rose to such a great height that it diffused into other fields such as robotics and became more pronounced in the American pop culture.

One of the most famous sci-fi films that featured an AI is *2001: A Space Odyssey*, released in 1968. It presented an initially dependable AI called HAL 9000 (Heuristically programmed ALgorithmic computer) which became erratic and began to kill its crew. The HAL 9000 interface is shown in Figure 2-1. This influential film showcased the danger and power of a rogue AI, leading to more films on AI takeovers and causing fear amongst the general population towards the technology. In the 21st Century, while the dystopic models persist in our AI storytelling, AI has been portrayed in a more positive light as a friendly companion or savior to humanity (albeit with complexity).

AI companionship is predominantly featured in sci-fi and fantasy themes. For example, we see a fairy-like creature called Navi following the main character Link on his journey to save princess Zelda. Another example is in Destiny, where AIs called Ghosts would follow the main character on an adventure, helping them solve puzzles, inspect the perimeter, translate digital information, and would generally provide assistance in completing missions.

In my experience, sci-fi films can be an extremely influential medium, much more so than game alone. Science fiction films played a very influential role in my upbringing, as it enabled me to escape the world I lived in and immerse myself into another.

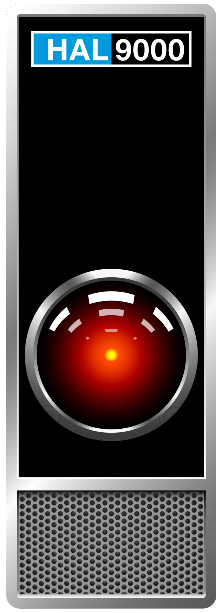


Figure 2-1: The HAL interface as shown in the 1968 sci-fi film *2001: A Space Odyssey*

The film that first truly conveyed the power of an AI companion to me was the original Iron Man movie released in 2008 and the sequels. Throughout the story, it showcased Tony Stark seamlessly interacting with his AI companion called J.A.R.V.I.S. as seen in Figure 2-2. This AI companion constantly assists and enhances Tony at any place and time, with augmented reality visuals that were sometimes floating in mid-air, as shown in Figure 2-3. JARVIS exhibited human-like natural language processing, which enabled him to collaborate in complex tasks with Tony, making this the perfect man-machine symbiosis. The natural and expressive gestures that Tony manifested as he controlled the digital information around him with his bare hands also highlighted what constitutes an intuitive human-computer interface to me.

Augmented reality technology has become truly accessible only in the last two years. There is still a lot of research needed to achieve a human-computer interface with an AI agent in its foreground that can naturally understand us and collaborate with us on complex tasks. I envision this kind of agent-based interface as an externalization of ourselves that will one day pervasively traverse the physical and digital worlds with us. I will introduce this idea of a Pervasive Intelligent Interface Agent further in the next chapter.

Figure 2-2: J.A.R.V.I.S. is an AI companion to Tony Stark in the film Iron Man.

2.2 User Interface

Historically, humans have been able to communicate and control machines through the use of a User Interface (UI). Initially, the purpose of the user interface was simply to make it intuitive for humans to use and control a computer system or machine. The definition of user interfaces started to take shape in 1962 with Douglas Engelbart's "Augmenting human intellect: A conceptual framework" in which he set out to be "focused on making the user more powerful, not simply on making the system easier to use" [29]. In order to pursue this vision, he led the Augmentation Research Center (ARC) in the Stanford Research Institute (SRI) to develop NLS, a computer collaborative system that uses the computer mouse, video monitors, multiple windows, presentation programs, pointers, hypertext, and much more, which he demonstrated in 1968 (see Figure 2-4). This demonstration became known as "The Mother of All Demos", as it featured many innovations in computer hardware and software at the time. This early work would then inspire further advances that led to today's graphical user interface, also known as GUI.

In the 70's, Xerox Palo Alto Research Center (PARC) was established to create "The Office of the Future" [5]. Alan Kay joined PARC and ran the Learning

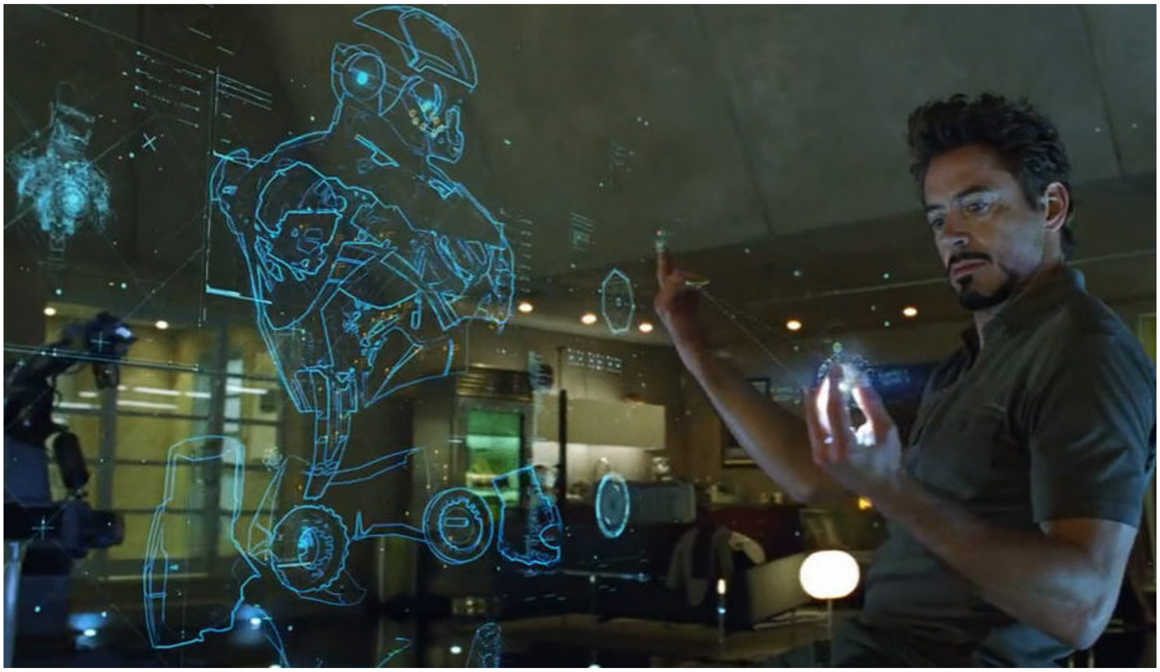


Figure 2-3: Tony Stark interacting with augmented reality visuals displayed in the physical environment with the help of JARVIS.

Research Group. Influenced by the work of Seymour Papert and Jerome Bruner on educational psychology, his team worked on improving the existing technology in user interface design and input devices to make it much like a medium that even children could operate. In 1972, they developed Smalltalk, an object-oriented programming language designed partially for educational use [44]. The Smalltalk language and environments extended the user interface design introduced by Engelbart by adding multiple windows and icons. Subsequently, Xerox PARC released the Alto PC in 1973, which implemented a basic GUI with windows and icons as shown in Figure 2-6, the bitmap WYSIWYG (What You See is What You Get) cut & paste editor, network and mouse (refer to Figure 2-5). In 1975, Xeroc PARC demonstrated for the first time a WIMP interface (WIMP stands for windows, icons, menus, pointers) with "icons, pop-up menus, and overlapping windows that can be controlled easily using a point-and-click technique" [1]. This is essentially the precursor to the GUI that we are so familiar with and still use 45 years later.

The Xerox Alto PC introduced the first computer screen, and ever since then we

Figure 2-4: Douglas Engelbart using an NLS. Source: computerhistory.com

have been attached to the constraints of window system UI design. In 1984, Apple released the Macintosh that popularized the "window-and-mouse-driven" graphical user interface as shown in Figure 2-7. Microsoft will eventually join in and introduce Windows 2.0 in 1987, a GUI-based operating environment that allowed users to resize and overlap windows (Refer to Figure 2-8). The windows interface was attractive because it was a "metaphor" to what a desk would look like. It helped take off cognitive load from our brain by displaying several pieces of information on the screen and allowing us to switch between tasks, which essentially allowed it to hold "external memories that are an extension of one's internal memory" [35].

The Graphical user interface became the most prominent user interface thanks to the ubiquity of the computer screen, however, there are numerous other modalities of interactions, such as sound, touch, and gesture. The goals of a modern user interface, as I understand, are to augment our human intellect using computer software that is likeable, intuitive, and effective; moving from direct manipulation by the user to

Figure 2-5: A Xerox Alto PC on display

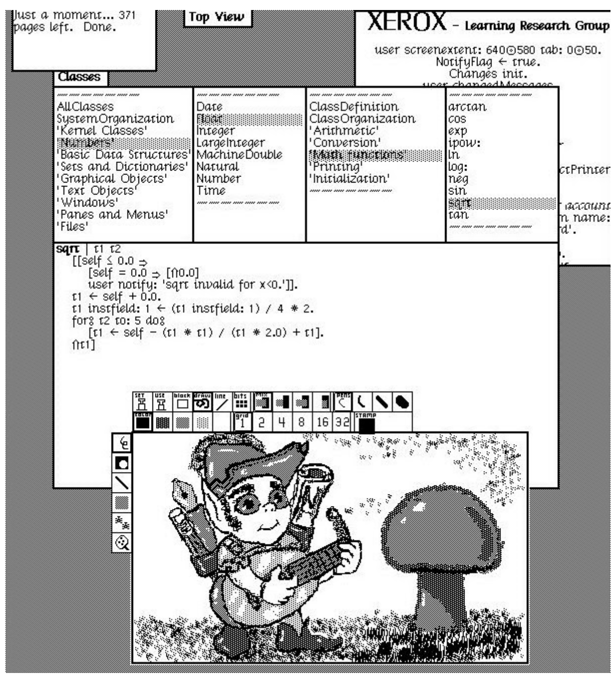


Figure 2-6: GUI on a Xerox Alto PC

Figure 2-7: Macintosh Computer and GUI

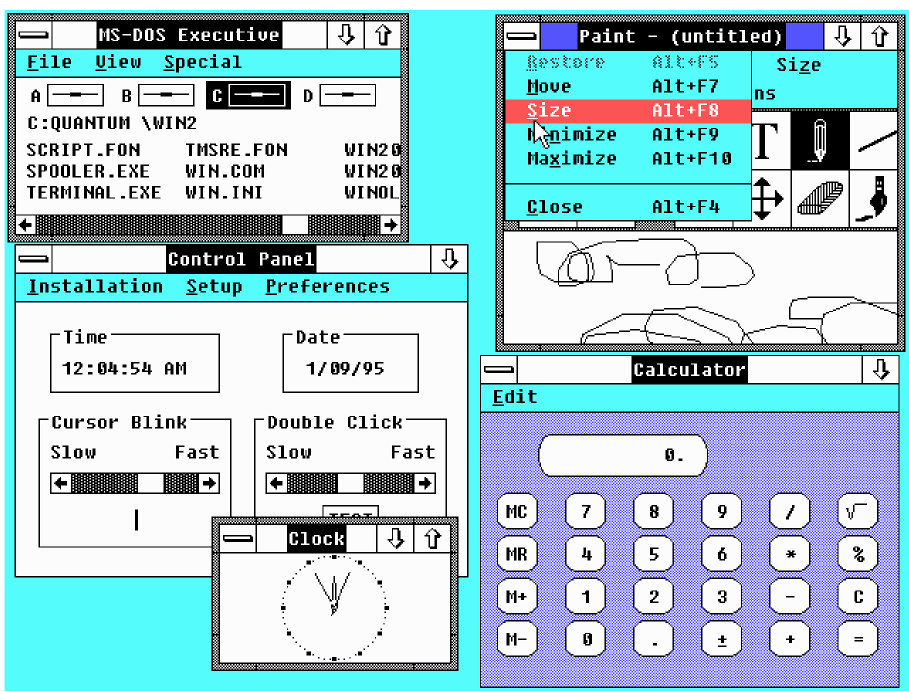


Figure 2-8: Windows 2 GUI

delegating the tasks to an agent-based computer software, so that we can focus on doing more important tasks. In the past few years, augmented reality technologies have become more accessible, and with this we began to accelerate our exploration for a deeper and more natural multi-modal interaction between ourselves and our computer to better navigate the physical and digital information spaces.

2.3 Software Agent

We cannot dive into interface agents without defining what agency means. In computer science, the term "agent" is used very broadly. Generally, it refers to a computer software that can perform a task on behalf of the user or another program. Many attempts have been made to provide a formal definition of the term "Agent", but the word is used so widely among the general public and academic communities, it is difficult to converge to a single universally agreed definition.

One of the most well-known papers on agent taxonomy is S. Franklin and A. Graesser's *Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents*, [34] published in 1996. In this paper, the authors collected a set of definitions given by prominent researchers in the field at the time, such as Russell and Norvig, Maes, Smith, Cypher and Spohrer, and Hayes-Roth, among others. They used them to help formulate a single definition that would "capture the essence of agency" by distinguishing more clearly the difference between a software agent and an arbitrary program. They called this "essence of being an agent" an autonomous agent, which they formally defined as "a system situated within and a part of an environment that senses that environment and acts on it, over time, in pursuit of its own agenda and so as to effect what it senses in the future." By their definition, an autonomous agent could be a person, a robot, or even a software program. However, in order for a software program to be called an agent, it must be able to sense and act in an environment appropriate for them and must have a sense of time. An example they gave explains that a spell checker embedded in a word processor is not necessarily an agent, but it could be considered one if it automatically correct typos as one

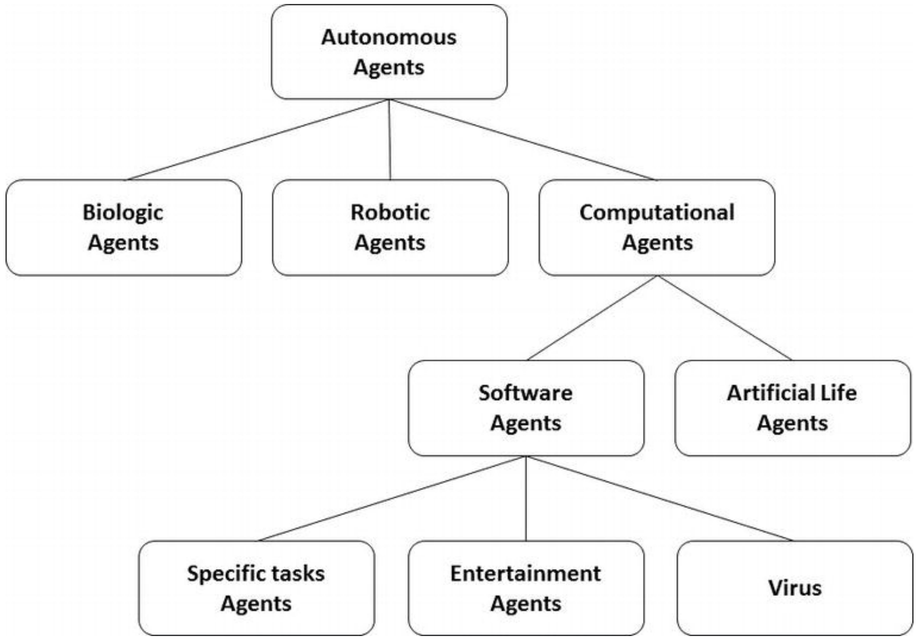


Figure 2-9: Natural Kinds Taxonomy of Agents based on biological model.

types. Additionally, they provided a classification of autonomous agents based on a biological model as seen visually in Figure 2-9. Starting at the kingdom level, they classified autonomous agents as either biological (animate organism), robotic (artifacts), or computational (abstract concepts). At the phylum level, they sub-classified Computational Agents into Software Agents and Artificial Life Agents. Finally, at the class level, they subdivided Software Agents into Task-specific Agents, Entertainment Agents, and Viruses.

Another highly cited work on agents that was also published in 1996 was H. S. Nwana's *Software agents: an overview* [69]. In this work, Nwana surveyed the field of agency and provided an overview of the work done until 1996. According to the author, the term "agent is really an umbrella term for a heterogeneous body of research and development." Since the definition varies among researchers, new synonyms have been created, which also generate more confusion when we want to classify them. Nwana briefly described a few of the synonyms,

> "those that inhabit the physical world, some factory say, are called robots; those that inhabit vast computer networks are sometimes referred to as

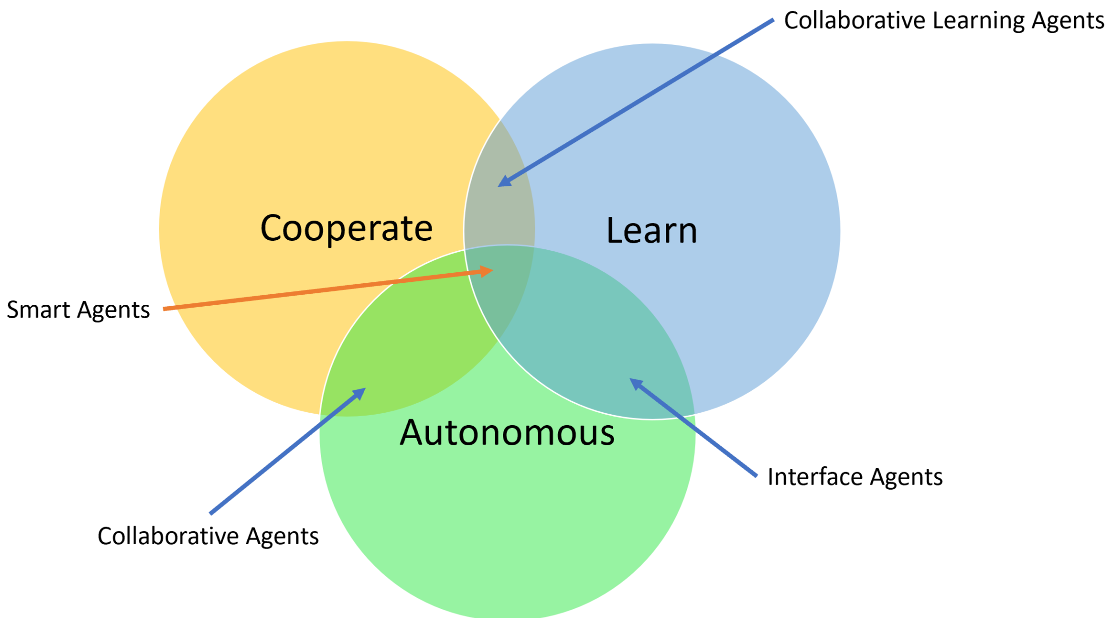


Figure 2-10: Nwana's software agent classified by different dimensions.

> softbots; those that perform specific tasks are sometimes called taskbots; and autonomous agents refer typically to mobile agents or robots which operate in dynamic and uncertain environments. Secondly, agents can play many roles, hence personal assistants or knowbots, which have expert knowledge in some specific domain. Furthermore, due to the multiplicity of roles that agents can play, there is now a plethora of adjectives which precede the word 'agent', as in the following drawn only from King's (1995) paper: search agents, report agents, presentation agents, navigation agents, role-playing agents, management agents, search and retrieval agents, domain-specific agents, development agents, analysis and design agents, testing agents, packaging agents and help agents."

One of the most influential aspect of this work is the attempt to classify agents into several dimensions as seen in the Venn diagram shown in Figure 2-10. The author further described the typology of agents, which was classified into seven types (refer to Figure 2-11) summarized below:

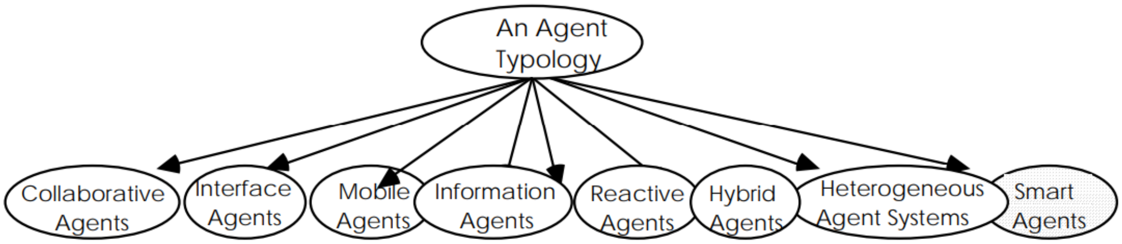


Figure 2-11: Nwana's agent typology.

Collaborative Agents

These agents posses social skills, responsiveness and pro-activeness and have a strong emphasis on autonomy and cooperation (with other agents). They could perform learning but not extensively complex ones. "They can act rationally and autonomously in open and time-constrained multi-agent environments." Refer to Figure 2-12 for an example of a collaborative agent architecture, extracted from Nwana's paper.

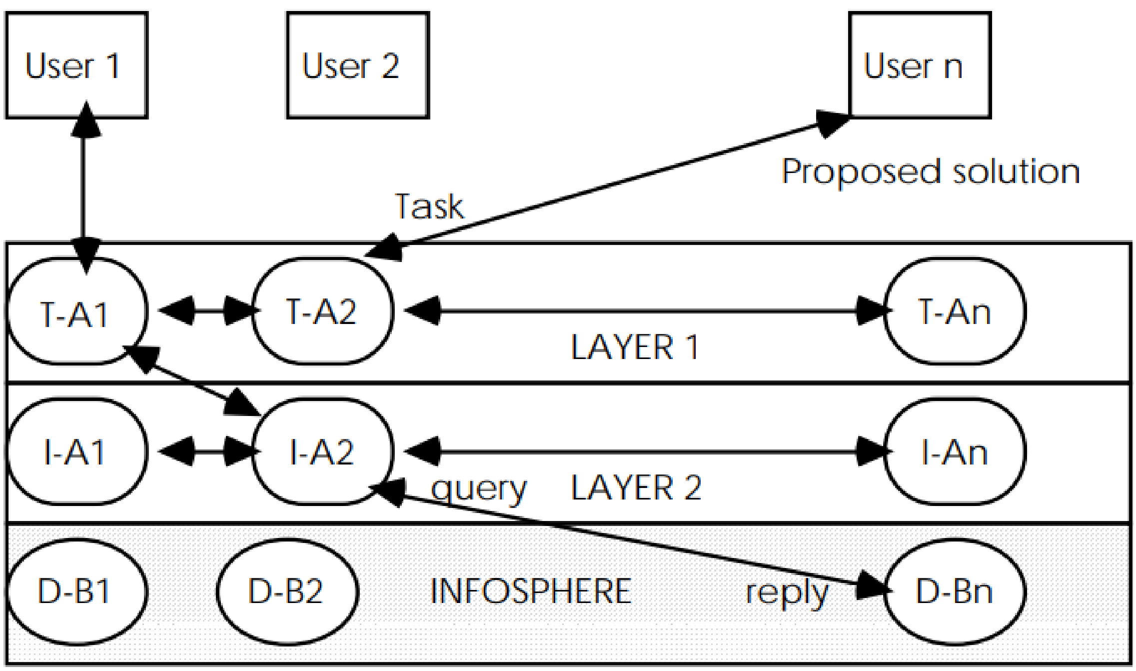


Figure 2-12: An adapted view of the Pleiades Distributed System Architecture used in mobile agents. Source: [69]

Interface Agents

These agents put an emphasis on autonomy and learning. They collaborate with the users instead of with other agents to provide them with personal support and assistance. These agents will watch their users closely in order to learn new "shortcuts" and recommend better approaches for doing the task. They will learn from the user by "observing and imitating, receiving feedback and explicit instructions, and by asking other agents for advice" with the ultimate goal of better assisting the user. Refer to Figure 2-13 for a diagram on how interface agents work, adapted from Pattie Maes' work from 1994, extracted from the Nwana's paper. This can of agent can learn over time and adapt to its user's preferences.

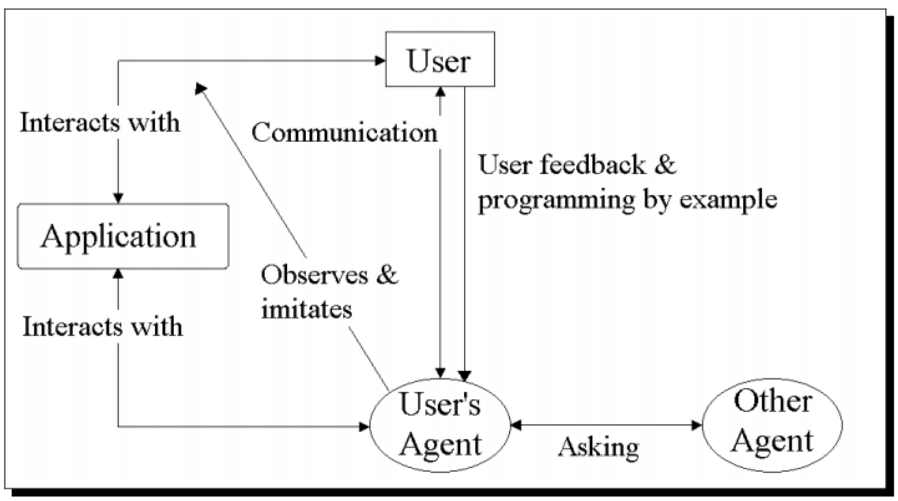


Figure 2-13: An adaptation of Pattie Maes's work on the functionality of interface agents from 1994.

Mobile Agents

"Mobile agents are computational software processes capable of roaming wide area networks (WANs) such as the WWW, interacting with foreign hosts, gathering information on behalf of its owner, and coming 'back home' having performed the duties set by its user. These duties may range from a flight reservation to managing a telecommunications network. However, mobility is neither a necessary nor sufficient

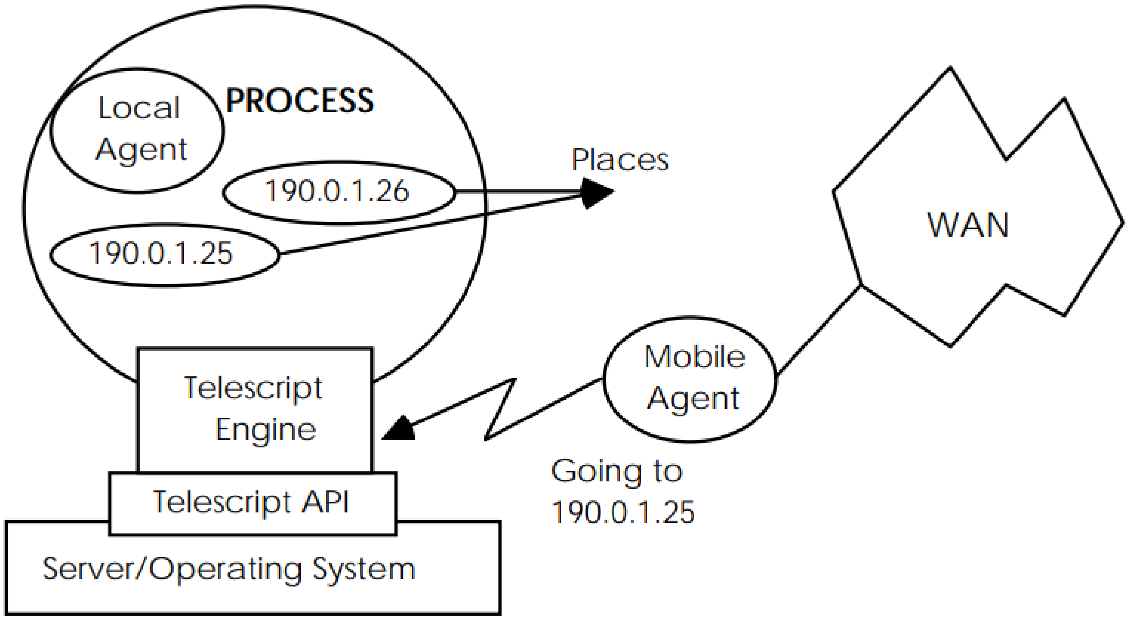


Figure 2-14: An adapted view of how mobile agents Work. Source: [69]

condition for agency. Mobile agents are agents because they are autonomous and they cooperate, albeit differently from collaborative agents. For example, they may cooperate or communicate by one agent making the location of some of its internal objects and methods known to other agents. By doing this, an agent exchanges data or information with other agents without necessarily giving all its information away. This is an important point, not least because the public perception of agents (thanks to the popular computing press) is almost synonymous with mobile agents." Refer to Figure 2-14 for a view of how mobile agents work, extracted from Nwana's paper.

Information/Internet Agents

These agents are deployed on the WWW and their objectives are to manage, manipulate, collect and combine information from various sources. Refer to Figure 2-15 for a view of how information agents work, extracted from Nwana's paper.

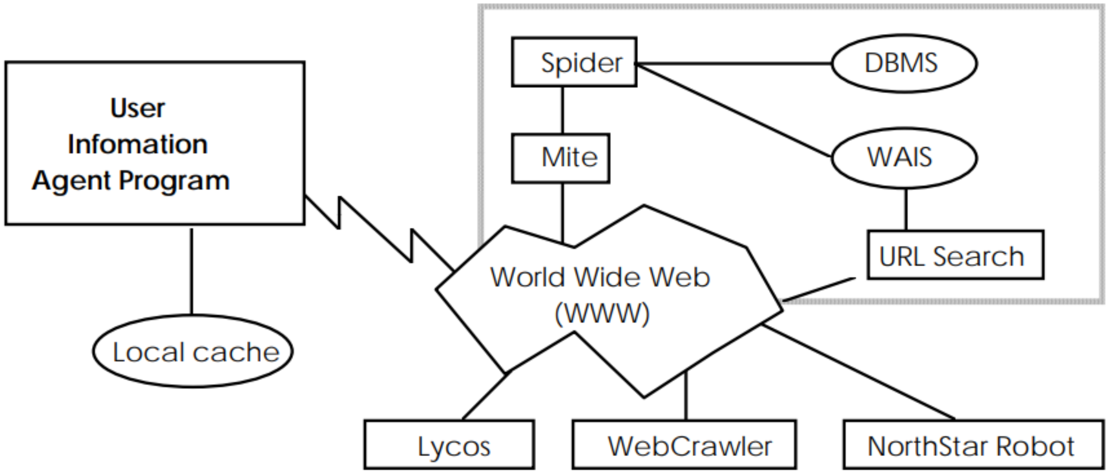


Figure 2-15: An adapted view of How Information Agents Work. Source: [69]

Reactive Agents

"Reactive agents represent a special category of agents which do not possess internal, symbolic models of their environments; instead they act/respond in a stimulus-response manner to the present state of the environment in which they are embedded." These agents have three main characteristics: (1) They have "emergent functionality"; as they interact, they would lead to properties they did not have before. (2) They are composed of several modules that enable them to operate autonomously to accomplish specific tasks. (3) They tend to have a physical representation in the real world that enables them to gather raw data. Refer to Figure 2-16 for a view of one of the reactive agent architecture extracted from Nwana's paper.

Hybrid Agents

These agents are a combination of two or more agent architectures or "philosophies" (e.g. mobile philosophy, interface agent philosophy, collaborative agent philosophy, etc).

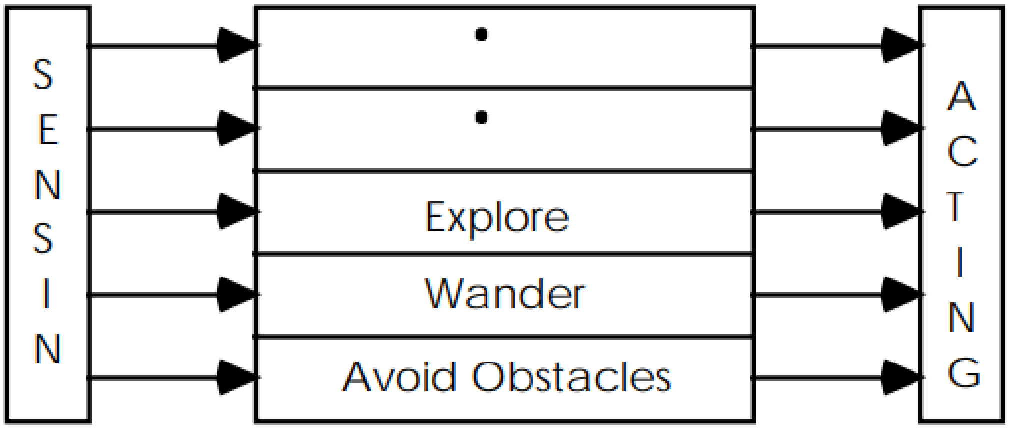


Figure 2-16: Reactive agents tend to depend on reactive agent architecture. One of the mentioned ones is Brook's Subsumption Architecture. Source: [69]

Heterogeneous Agent Systems

If an agent falls into multiple categories, then it is referred to as "heterogeneous agent systems". According to Nwana, these agents, unlike hybrid systems, "refers to an integrated set-up of at least two or more agents which belong to two or more different agent classes." This kind of agent can be composed of one or more hybrid agents.

Smart Agents

Intelligent/smart agents did not exist back in the 1990s. Even in our present time, these "intelligent" agents are still science fiction. Similar to the term "agent", the word "intelligent" is also very subjective and contains different levels of expectations. Thus, these agents are what we want software agents to become, they all aspire to be "smart" or "intelligent". The word itself oversells the technology and create hypes. Nwana provided a statement regarding the word "intelligent agent",

> "There are several things which the serious agent researchers can do. Firstly, they can drop the 'intelligent' in intelligent agents as we have done in the title of this paper: its connotation, and hence expectations,

are much less. Secondly, they could attempt to ensure, where possible, that dilettantes do not publish articles on agents in the popular press, at least not until it has been reviewed by someone whose interest in the area is more than superficial. This is sometimes possible because some experts usually get asked to review such articles before they go to press. Thirdly, they (i.e. experts) should not themselves engage in overselling the domain and, lastly, they must always be critical of the progress in the area in order to provide a more realistic appraisal of the state-of-the-art."

Chapter 3

Related Works

This thesis looks at previous work in three areas of research: 1. Embodied agents, 2. Navigation systems with agents, and 3. Human-factor Metrics for AR.

3.1 Embodied Agents

To date, several studies have investigated the complexity of the effects of agent embodiment on our cognition and behavior. One of those well-studied factors is social presence. Robots naturally exhibit physical embodiment and social presence. Matsumura et al. has shown that the behavior of an embodied tour guide robot can influence visitors to start conversations between themselves [60]. The gender of the agent might also play a role in influencing our behavior, for instance, Vossen et al. showed that women who interacted with an embodied agent (iCat Robot [1]) were persuaded to consume less electricity than those who interacted with a computer [92]. In another study, Siegel et al. noticed that a female robot could influence male users to donate money to the museum [87]. They also noted that "subjects also tended to rate the robot of the opposite sex as more credible, trustworthy, and engaging." Recently, Deng et al. conducted a review on embodiment in socially interactive robots [28] to help classify the extensive embodiment research for social interaction in robotics, assistive robotics, and service robotics.

Interaction researchers believe that physical corporeal presence could engender higher engagement, trust, and perception of social presence from users. Seo et al. showed that people empathized and perceived social presence of a physical robot more than with a virtual one [84]. However, Thellman et. al. argued that the extent of our perception of social presence of an embodied agent might not be affected by the physical corporeal presence (physical representation) [90]. Virtual agents have shown positive effects on social presence and a higher sense of spatial as well. In 2018, Kim et al. conducted a study using augmented reality to analyze the influence of an embodied virtual agent on users [45]. They found that the embodied virtual agent affected users' sense of engagement, social richness, and perception of social

[1] https://www.roboticstoday.com/robots/icat

presence by exhibiting natural human social behaviors.

In 2019, Schmidt et al. conducted two user studies to understand the effect of a virtual agent embodiment and thematic closeness [81]. In this study, they compared four different virtual guides: a theme-related embodied virtual guide (humanoid with astronaut suit), a theme-related disembodied guide (astronout voice only), a genetic embodied virtual guide (humanoid - civilian), and a traditional disembodied audio guide (civilian voice only) for presenting artifacts in virtual and real exhibits. To summarize some of their results, in the virtual exhibit user study, they found that participants reported that "the embodied guides caused more emotional responses such as laughing or smiling". One of the main differences between embodied and disembodied guides was in the scores of presence as social actor, also known as parasocial interaction, which indicate how much the participants thought of "crossing the border between the actual physical environment and the mediated environment in order to interact with the agent in real time." They reported a higher score for those who saw the embodied guides, indicating that they "felt that their presence was noted by the agent and that he was establishing a connection with them." Additionally, they also noticed a positive effect of the thematic closeness (astronaut humanoid vs. generic humanoid) on social presence, and argued that it could be because participants felt that the astronaut guide was inviting them to take part of his personal story. They also found that embodied guides were significantly perceived as having more credibility, indicating that they might be more competent and trustworthy. In term of learning, they found that voice-only guide helped people remember better than the scenarios with an embodied agent, and argued that this could be due to the fact that users tend to pay more attention to details about the agent such as lip movement rather than the actual scene itself. Even thought the data showed otherwise, participants still perceived that they gained higher knowledge with the embodied agent. No significant effects were found between the generic guide and the astronaut in the social questions, as they argued that some participants reported that there was too much noise in the astronaut voice. In the real exhibit user study, they found that embodied guides received a higher score than in the virtual exhibit in the presence-as-social-actor measure, while the voice-only guide was rated higher in the virtual exhibit than the real one. More results can be found in their paper [81].

In a most recent work published this year, Techasarntikul et al. implemented an AR guide system using a Microsoft Hololens and studied five different guiding interfaces with and without an anime embodied agent: Audio, Arrow, Circle, Arrow-Move, and Circle-Move to point specific part of an exhibit and to explain the information. They found that the embodied agent significantly improved the attractiveness, stimulation and novelty of the system [88]. Similar to the findings in Schmidt et al., they reported that the lowest memory recall was from the condition with only the audio interface with the embodied agent (38,75% correct answer), and that the highest score was obtained by participants who were guided with the Circle interface with the embodied agent presence (67.75%) followed by the Circle-Move interface at (60%).

Furthermore, Holz et al. argued that the move towards embodied agents was mainly due to prior works that highlighted the importance of situatedness and embodiment for AI [14, 15, 42]. He stated that, "Brooks' popularisation of the reactive approach served as a catalyst for the creation of a more embodied approach to AI, where an agent must be structurally coupled with its environment if it is to be seen as intelligent [42]." They defined Mixed Reality agents (MiRA) as agents that possess interactive capacity ("sense and act on the virtual and physical environment") and corporeal presence (virtual or physical representation that follows the laws of physics and includes occlusion and collision detection).

In the 90s, Human-computer interaction researchers found that people apply social norms to computer agents if they exhibit anthropomorphism (human characteristics, features, or traits) [66, 67]. One of the early proponent of embodied conversational agent is Justine Cassell, who worked on developing embodiment for conversational interfaces, toolkits for generating behavior expression in animation and much more together with her team [61, 20, 19, 21, 18, 9].

According to prior studies, appearance or embodiment is one of the most salient characteristics that affects our trust towards a computer agent [25, 27, 38]. Human-like agents have shown to inhibit interaction where sensitive questions are being asked.

For example, a comparison between a conversational agent (CA) and a human-like conversational agent showed that agent with no human features was able to get answers to sensitive questions from the users [82]. Visual human-like features remind us of real humans and could trigger social and cultural biases that create a barrier that prevents some human-agent interactions. Users expect virtual human agents to have accurate gestures and behaviors like real humans do since they are perceived as having a mind and a soul [38]. Thus, HCI researchers have begun to study augmented reality non-humanoid agents and have shown that they too can influence our cognition and behavior. For instance, Norouzi et. al. showed that an AR dog was able to affect participants' walking behavior and perception of co-presence, physicality, and animalism [68].

Based on these works, we have a general idea that embodied agents can enhance social presence, engagement, and overall experience of the users even if it is virtual. Interesting findings have seen that embodied agents could take the focus away from participants, leading to lower memory retention when asked about the presented artifacts. In literature, not much work has been done to understand the effect of non-humanoid abstract embodied agent. It will be interesting to see if a more abstract version of an embodied agent could enhance memory retention. This work could contribute towards understanding when (in what context) to appropriately use embodied and disembodied agents.

3.2 Augmented Reality-Based Agent Navigation and Tour Guiding Systems

10% of the travel industry market is made up of tours and activities said to reach $183 billion bookings by 2020. The main purpose of tour-guiding is to provide information about a specific artifact or event in the physical environment to educate people on a subject matter. From the industry standpoint, a successful tour guiding experience could enhance the sustainability and competitiveness of the company, and increase interested partners to invest time and money into it. Naturally, psychologists have found that tackling prior knowledge from individuals can enhance the subjective perception of a memorable tourism experience [22, 26]. Additionally, researchers have also found that by increasing the individual involvement and engagement during the tour experience, they can influence them to revisit the site and recommend it to others [95].

In prior studies, mobile AR-enabled interactions have shown to have an effect in memory retention. For example, Lu et al. showed that by visually augmenting a museum painting with animation, after 24 hours, visitors were still able to remember the features of the painting better compared to the use of physical labels and descriptions [93]. They also reported that 80% of the subjects stated that AR was distracting. However, they argued that the AR “distractions" could be due to novelty effect of AR technology and subject’s prior perception and experience of art. They further found that this distraction might have actually enhanced the subject’s ability to objectify the art piece.

Mobile AR-based navigation applications such as HotStepper[2], overlays a digital humanoid agent in the physical world that shows the way to destination. This agent’s main purpose is to entertain the user by dancing while guiding the user towards the desired destination. The agent does not seem to have situational-awareness and context-awareness of the space it inhabits; the agent will walk on roads with cars and through bushes and other people. In addition, it does not exhibit conversational or sensory modality necessary for creating immersive storytelling experience for enhancing route retention. Similarly, another mobile application called GuideBOT[3] displays an embodied AI chatbot that guides the user towards specific items in the physical environment. It is advertised as having AR indoor and outdoor navigation capabilities as well as AI Chatbot features. The bot asks a few questions on the item that the user wants to look for and then take the user from point A to point B. Although this is a great move towards having agents as our new interfaces, this

[2] http://hotstepper.dance/
[3] https://www.viewar.com/template/guidebot/

agent still lacks natural interaction that a human-human interaction would exhibit. For example, the agent does not have storytelling capabilities and does not engage with the user during the navigation tasks.

Agent-based navigation systems have been explored in the past. Seo et al. built an AR-based, context-aware tour guiding system for overlaying 3D characters in the physical environment for narrating events and presenting scenarios in Gyeongbokgung, South Korea [83]. Miyawaki et al. presented an AR agent using printed markers for assisting users during cooking navigation [63].

In 2014, Campbell et al. conducted an experiment with an AR humanoid agent as a navigation aid and found that the agent was able to provide faster navigation along a short distance compared to other kind of AR visual cues, such as directional arrows or an AR “bubble" highlighting the destination point [17]. Similarly, in 2019, Kuwahara et al. presented a questionnaire evaluation and human-behavior analysis on the perception of a campus navigation application using an AR tour guide [48]. They found that their AR tour guide has the potential to provide easy route navigation.

It will be interesting to reproduce these studies using a fully functional AR tour guiding application with current hardware like the Hololens device and a non-humanoid abstract embodied agent instead to see if the results still hold true. It will also be interesting to compare the spatio-temporal effect by having some participants navigate a one floor environment (X-Y plane) and have another group of participants navigate a multi-floor environment (X-Y-Z dimension).

3.3 Human-Factor Metrics

3.3.1 Social Presence

Social presence has been studied extensively in the virtual reality (VR) and augmented reality (AR) communities. Social presence could be defined broadly as the “sense of being with another [human or artificial intelligence]” [11]. Today, we have chatbots and virtual assistants mediating our digital experience. The theory of social presence looks at providing a clear way to assess performance of new technologies for mediating social interaction.

The social presence of an AR agent could be defined as the notion of “coexistence in the same space" as the user [94]. When interacting with a computer, humans react unconsciously to them as if they were real. Prior work has shown that social agents affect persuasion and attraction [31, 51]. Oh et al. provided a systematic review of all the prior work done on social presence as well [71].

Through the use of questionnaires, we could measure user’s perception of social presence [70]. Previous work has also used objective measures such as sound pressure level (SPL), electrodermal activity (EDA), and posture responses.

3.3.2 Memory Retention

Back in 500 B.C. ancient Greece, people did not have easy access to technology that could help them memorize several pieces of information at the same time. The Method of Loci (“place" in Latin) technique, also known as Memory Palace, was developed around that time to enhance memory retention by using visual cues in a familiar physical environment for creating spatial memory to quickly recall information.

Modern day technology enable new ways to experience the memory palace technique for memory enhancement. Fassbender et al. conducted a study on a Virtual Memory Palace (VMP) using a virtual reality 3D architectural model [30]. In this within-subject study, 15 subjects completed a two part experiment. The first part, they were told to memorize a list of 10 words (“guitar, “anchor", etc), and in the second part, they were provided with visual representations (icon) of similar words that can be placed in the VMP. After a week, the researchers interviewed the subjects to analyze how much they remembered from the two parts. They noted that on average, the subjects remembered 7.5 items in the VMP compared to 5 items in the wordlist, hence presenting the potential of VMP for memory enhancement.

In recent studies, The Method of Loci was used in AR applications to study its potential for memory enhancement. For instance, Rosello et al. presented a preliminary comparative study with 14 subjects using their NeverMind augmented reality interface for memorizing the 10 Super Bowl champions from 1967 to 1976 and using a paper-based technique for remembering the champions from 1977 to 1986 [78]. In the NeverMind condition, users walked a specific route in the university campus, visualizing AR content along the way that helped prime the information. Their results showed consistent memory recall for the NeverMind condition after 2 minutes, 24 hours, and 7 days. The paper-based condition, however, showed a drastic decline in memory retention after 24 hours and 7 days.

Other visual and auditory cues have also shown to enhance memory retention. A number of studies have examined the influence of embodied conversational agents in learning and memory retention. For example, Lusk et al. found that students who worked examples with animations of a fully embodied voice agent outperformed those who experienced the static examples with voice-only condition [55]. In a within-subject study, Bergmann et al. showed that a virtual agent that exhibits gesture during vocabulary training helped learners achieve higher memory scores [8]. However, as mentioned before, Schmidt et al. found a contradicting results, they reported that the participants in the voice-only condition achieved a higher score in learning compared to those with an embodied agent in a virtual exhibit [81]

It will be interesting to see if an abstract embodied agent could achieve a higher memory retention since it is less distracting than a humanoid embodied agent and more engaging than a disembodied agent. It will also be compelling to see how memory retention is compared by adding more visual cues and interactive elements coupled with an embodied agent. Since we can now deploy embodied agents that can navigate in a multi-floor environment, it will be great to analyze the spatio-temporal effect of the method of loci in a previously unknown environment as well.

Chapter 4

Pervasive Interface Agents

4.1 Definition

Nowadays, at least in the United States, we have access to voice assistants (e.g. Alexa, Cortana, Siri, Google) that help us perform tasks such as setting alarms, controlling IoT devices, ordering pizza, and much more. These voice assistants currently live inside our mobile phone or in a device at home. However, they lack a sense of "being present" in the same space as us. Embodiment could enhance social presence and trustworthiness of the agent. As mentioned in the introduction, robots are great companions, but they cannot yet travel and accompany us everywhere due to their physical constraints.

Based on past agent research, I have consolidated the definition set by Maes's extensive work on Interface Agents [58], Nwana's Agent [69], Franklin & Graesser's Taxonomy for Autonomous Agents [34], Lieberman's Autonomous Interface Agents [53], Holz et al. Mixed Reality Agents [42], and have incorporated lessons learned from Microsoft Clippy's downfall in order to form the definition of a Pervasive Interface Agent.

Pervasive Interface Agents aspire to have the following attributes:

- Autonomous: Agents can control their own actions and internal state without direct manipulation from the users.
- Sociable: Agents can simulate emotion and empathy and can communicate and interact with their users and other people and agents.
- Adaptive: Agents can identify the current environment the user is in (virtual or physical), learn and adapt its behavior to the context of that environment, and can respond in a timely manner.
- Pro-active: Agents can take initiative and exhibit goal-oriented behavior. (e.g. if in a virtual game environment, it knows its goal is to help its user completes a mission.)
- Anthropomorphous: Agents can morph into different versions of themselves (disembodied, embodied, etc) and can evolve based on their stages of life.

- Pervasive: Agents can navigate and accompany their users in the physical and digital realms with or without a physical corporeal presence (situated).

The creation of computers led to the Graphical User Interface (GUI); the rise of smartphones prompted the Voice Assistants we use today. There have recently been significant advances in virtual world development and augmented reality devices; as our generations and future generations begin to spend more time with these devices, I believe cross-platform agent-based user interfaces, which I call Pervasive Interface Agents, will play a crucial role in the way we navigate these new virtual and augmented realities.

4.2 Cognitive Model

The cognitive model of a Pervasive Interface Agent can be built based on the widely known Beliefs-Desires-Intentions (BDI) Agent Architecture adapted by Rao and Georgeff (1995) [76] from Bratman's 1987 BDI model of human practical reasoning [12]. The three mental attitudes behind this model are described below:

- Beliefs: Information available about the agent's environment that is believed to be true.
- Desires: Information of the goals that the agent wants to accomplish, which include the properties and cost associated with them.
- Intentions: The chosen action plan containing a subset of desires for the short term that follows the same belief system.

Figure 4-1 shows an adapted version of the BDI model for Pervasive Interface Agents. The agent perceives the physical and digital environments and extracts information from them. These pieces of information go into a belief revision function that determine a new set of beliefs, which then go into the option generation function. Intentions are taken into account in the option generation function for determining the agent's desires and possible course of actions. The current beliefs, desires, and

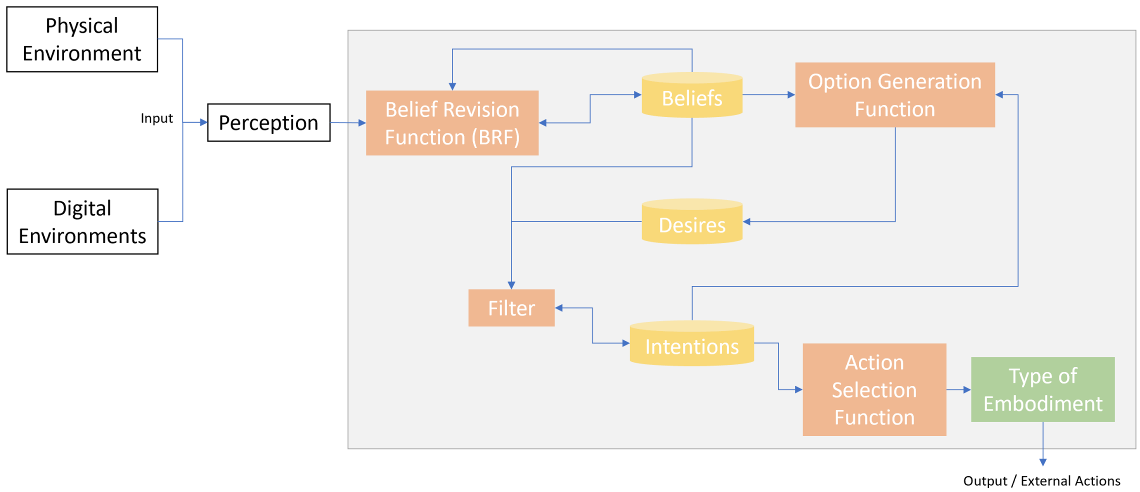


Figure 4-1: Modified BDI model for Pervasive Interface Agents

intentions are then filtered to determine the agent's intentions, which then get fed into the action selection function. Finally, the agent determines an action to perform and chooses the type of embodiment for delivering that action.

The pervasive interface agent will hold several sets of desires, beliefs, and intentions, the first of those sets being the one shaped from its interactions with the user in the physical world. Other virtual worlds or gaming environments will have a different set of desires, beliefs, and intentions that are fed temporarily to the agent when the user inhabits that world. For example, in the real world, if the user is visiting a place he or she has never been, the agent will pull information from the web, synthesize them, and deliver the information to the user as a personal tour guide. When the user visits a virtual world such as a gaming environment, like The Legend of Zelda, the agent will accompany the user there and understand the desires, beliefs, and intentions that the user holds pertaining to that world, such as those related to missions or side quests, as well as those related to "beating the game."

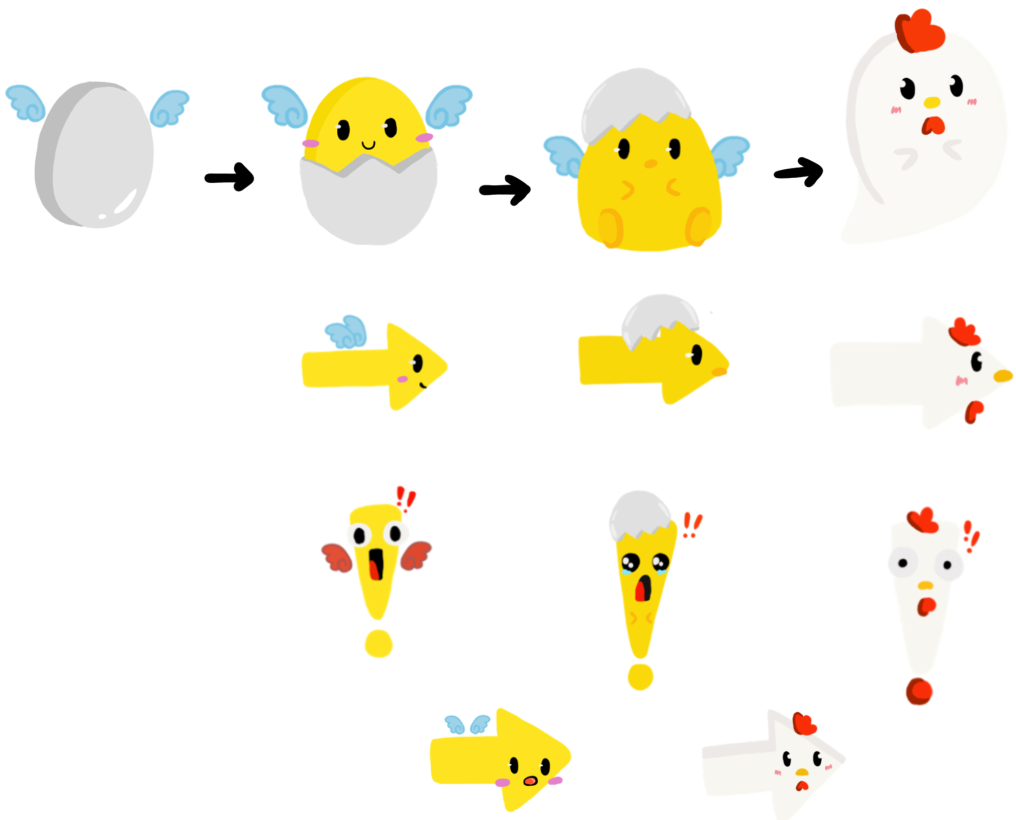


Figure 4-2: An example of a Pervasive Interface Agent with its embodied forms and evolution based on their stage of life.

4.3 Type of Embodiment

These Pervasive Interface Agents will have a fluid form, morphing into different shapes based on context and intention. In addition, the agents, in part as embodied representations of the individual (externalization of self), can evolve its own fundamental form based on the growth of themselves and their user. This is illustrated in Figure 4-2. Much more research must be done on agent embodiment across a diverse context to clearly identify and use them effectively for influencing our cognition and behavior.

Some examples of how a Pervasive Interface Agent would use its different embodiment are illustrated in Figure 4-3 to 4-6

Figure 4-3: Warning embodiment is shown as an exclamation mark to emphasize danger ahead.

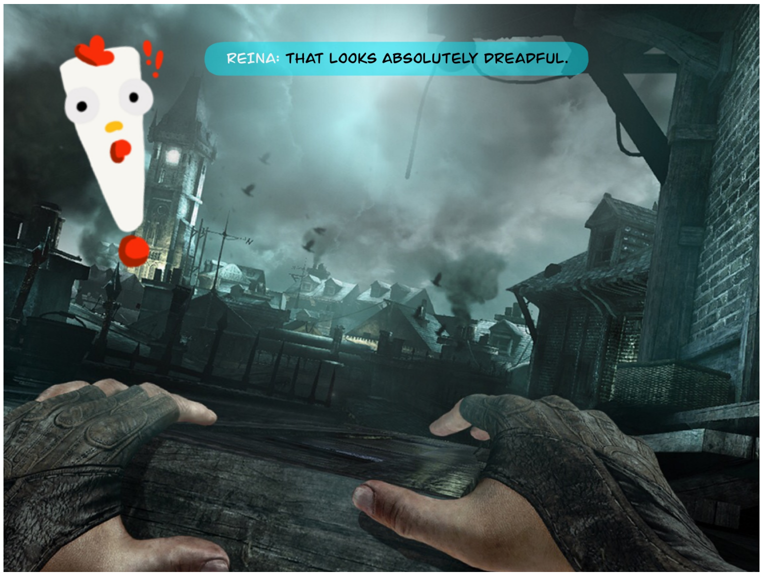


Figure 4-4: An adult version of the same Pervasive Interface Agent warning the user in one of the digital worlds (game environment).

Figure 4-5: The arrow embodiment guides the user in navigating the physical or digital worlds.

Figure 4-6: The fundamental form is used when it wants to make a statement or during general interaction with the user.

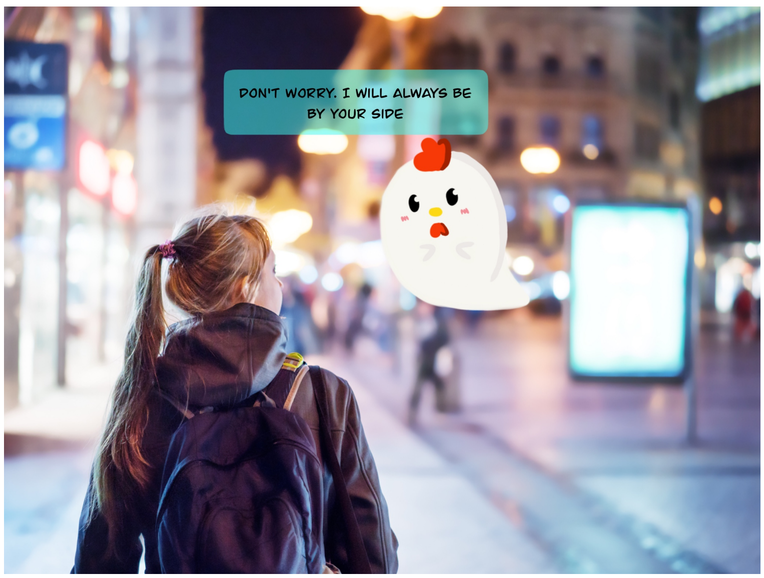


Figure 4-7: The fundamental form changes based on the stage of life of the agent. All users start with an egg version and end with an adult version of the same pervasive interface agent.

Figure 4-8: The disembodied version is used when the user is performing an action that does not require the presence of the pervasive interface agent.

Chapter 5

Media Lab Tour Guiding System

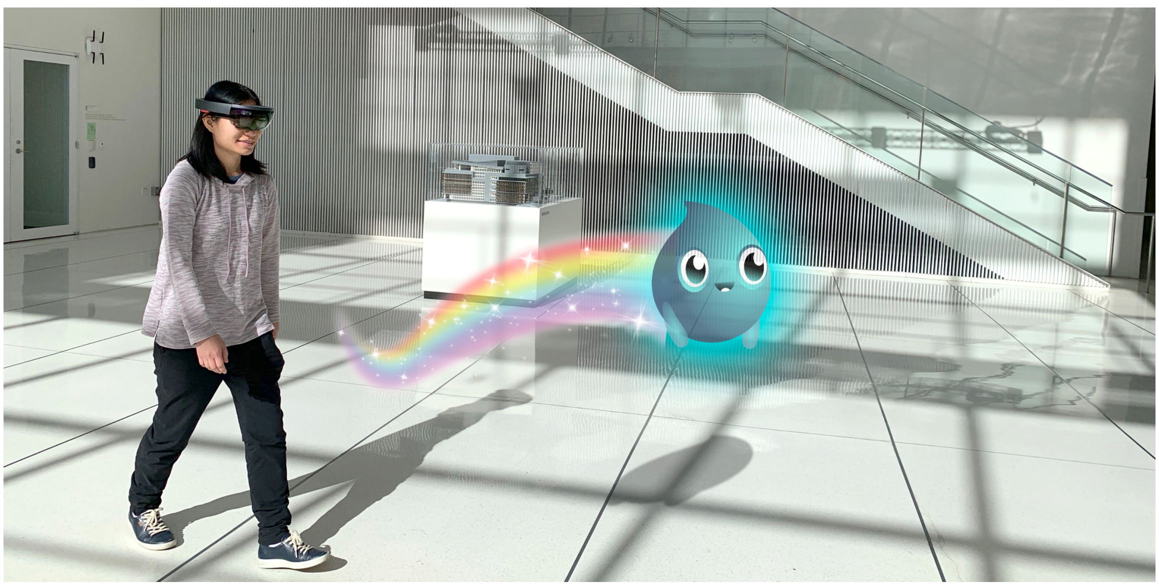


Figure 5-1: Example of what the Media Lab Tour Guiding Experience could look like.

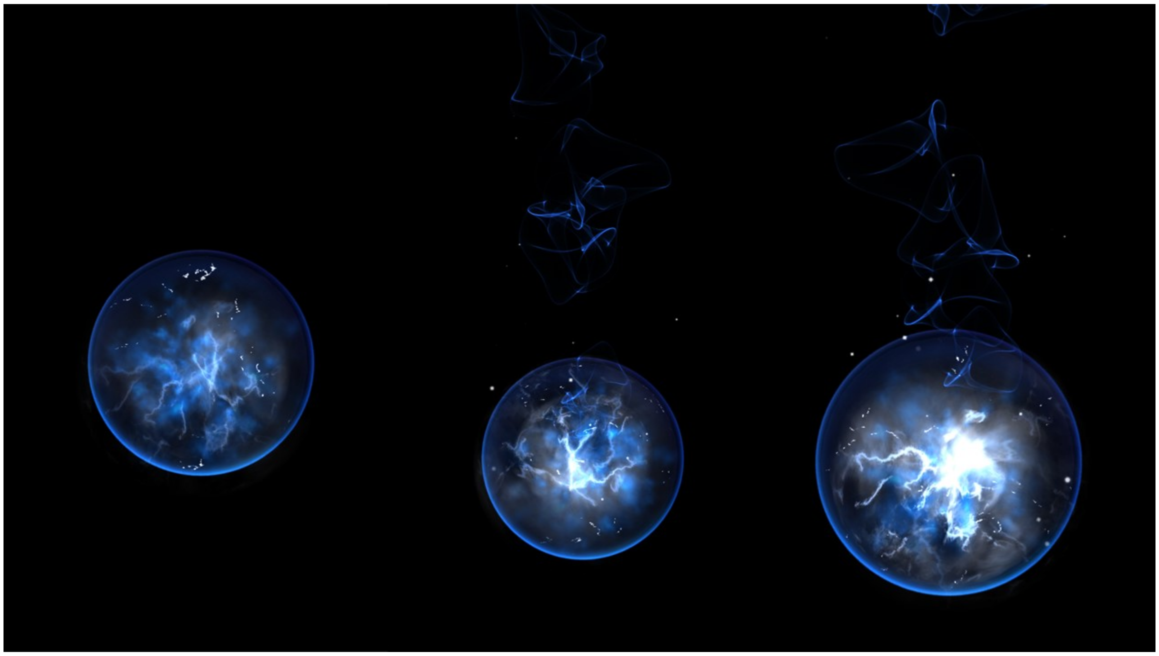


Figure 5-2: R.E.I.N.A. when it is idle (left) and when it is talking (right).

5.1 R.E.I.N.A., Tour Guiding Agent

In this thesis, I am interested in studying the effect of a non-humanoid abstract embodied agent on memory retention, social presence, and attention. I believe this could provide insights into the agent embodiment spectrum. Henceforth, I will refer to an embodied agent as having abstract non-humanoid embodiment throughout the rest of this thesis.

R.E.I.N.A., which stands for Responsive & Embodied Indoor Navigation Agent, was designed to possess a non-humanoid abstract embodiment that resembles that of an orb with particle systems that fluctuate when it talks. This particle system is modulated by the frequency of the audio programmed into it. Figure 5-2 shows R.E.I.N.A. embodiment and the particle system displayed when talking.

I consider R.E.I.N.A. the first phase towards a Pervasive Intelligent Agent. It has been programmed to navigate in the virtual and physical environments and to understand basic voice commands. Further cognitive features can be integrated later. See Figure 5-3 for a diagram of the R.E.I.N.A. system.

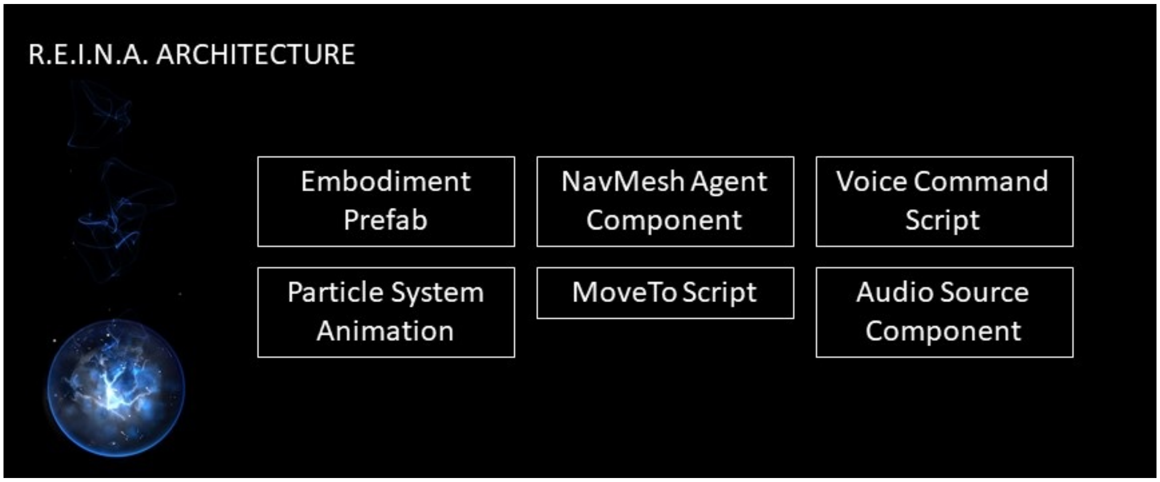


Figure 5-3: R.E.I.N.A. is composed to several components and can be further expanded to have more functionality and cognitive abilities.

5.2 Implementation

In order to build towards the vision of a Pervasive Interface Agent, I created a demonstration in which REINA is shown navigating freely across multi floors in a 3D scan of the MIT Media Lab building. This will be later deployed in the real physical environment for performing user studies. A summary of the architecture is illustrated in Figure 5-4.

5.2.1 3D Scanning the MIT Media Lab

The MIT Media Lab is composed of two buildings, the original Wiesner Building (E15) designed by I. M. Pei and the expanded building (E14) integrated in 2009. The E14 building has six-floors with more than 160,000 square feet of space. In order to 3D scan the E14 building, a Matterport Pro2[1] was borrowed from the MIT.nano Immersive Lab[2]. The Matterport Pro2 contains a set of infrared cameras that capture 3D information of the space in 360 degrees (refer to Figure 5-5). The Matterport software automatically stitches each scan together to form a larger scanned space that can then be viewed on the Matterport website under a professional account. For

[1]https://matterport.com/cameras/pro2-3D-camera
[2]https://mitnano.mit.edu/research-capabilities/immersion-lab

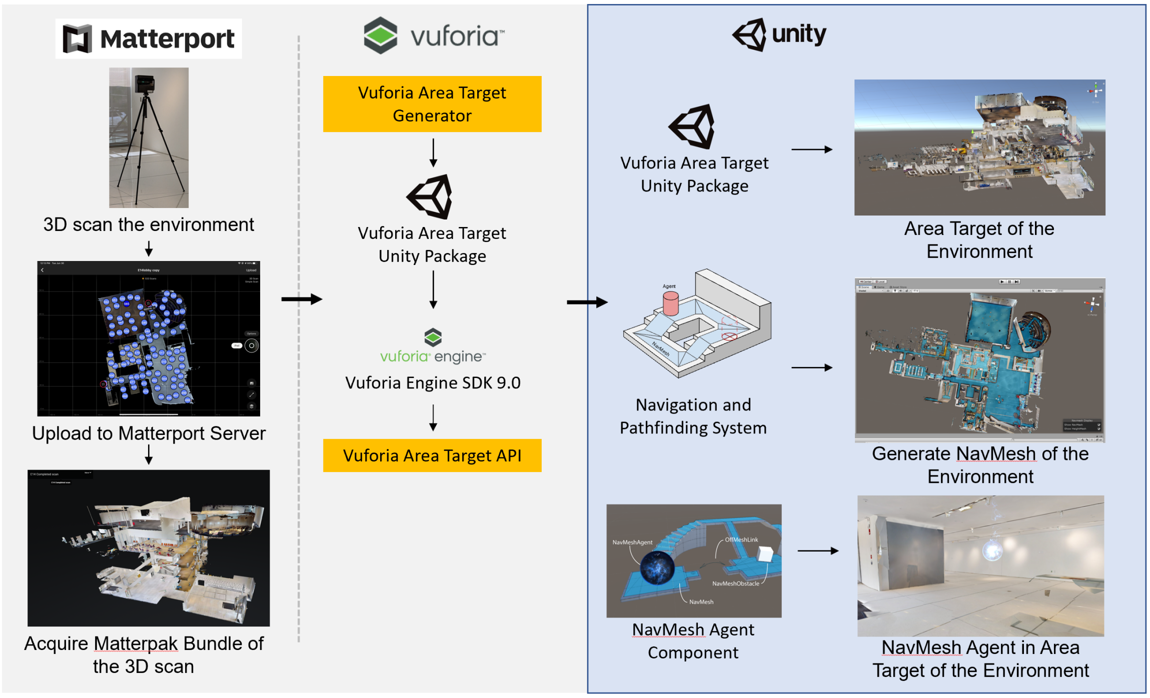


Figure 5-4: Architecture of the Media Lab Tour Guiding System.

scanning the entire building, we start by placing the scanner in the environment, pair it with an ipad pro with the Matterport software installed and hit the scan button on the software (See Figure 5-6). For the software to successfully stitch the scan together, each scan must be a few feet apart; this will ensure that there are enough matching point cloud data to merge the scans together. Refer to Figure 5-7 for a close up of the scans made for a single floor). Thus, for multifloor scanning, it is important to have continuity in the scan, for example, scanning every few steps of a staircase leading to another floor, as seen in Figure 5-8. In total, 565 scans were made to generate a partial 3D scan of the MIT Media Lab (Refer to Figure 5-9). The 3D scan mostly contained just the common areas; most of the research spaces were excluded due to privacy and safety concerns. A dollhouse view of the generated MediaLab scan is shown in Figure 5-10 and can also be interacted with on the Matterport website[3].

[3]https://my.matterport.com/show/?m=ZLNysZSdD8Tbrand=0

Figure 5-5: A Matterport Pro2 camera used for 3D scanning an environment.

Figure 5-6: For scanning the environment, we first setup the matterport in a tripod and place it in the space we want to scan.

Figure 5-7: A close up of the scans made for a single floor.

Figure 5-8: We must ensure continuity between floors, so we have to scan every few steps in a staircase.

Figure 5-9: 565 scans were made to generate a complete 3D scan of the MIT Media Lab.

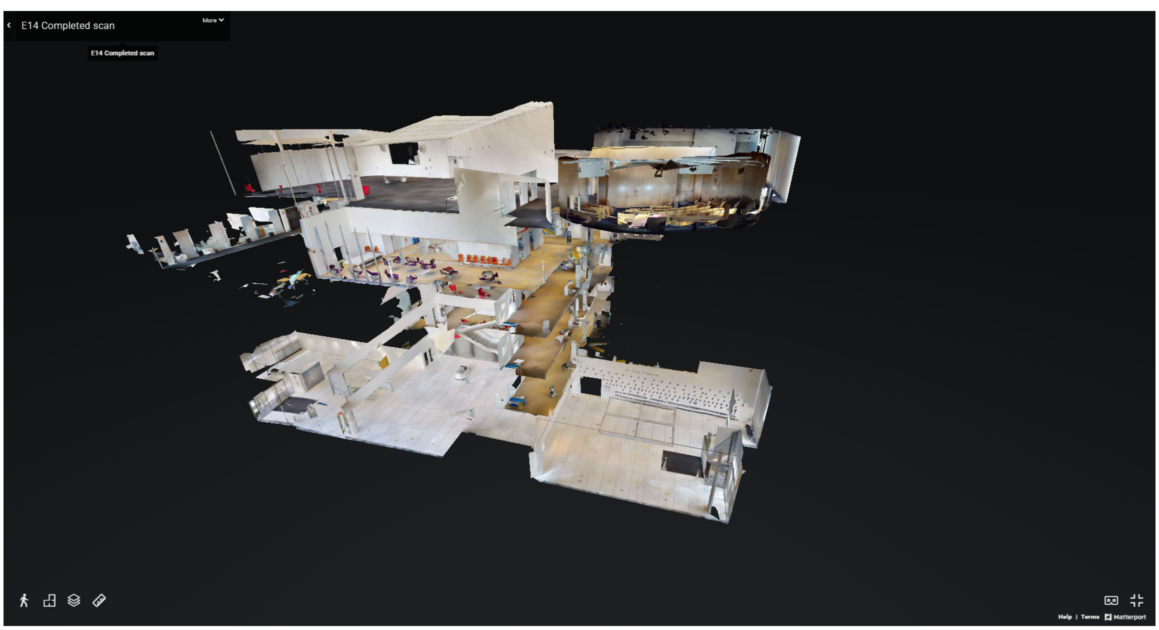


Figure 5-10: A dollhouse view of the scanned MIT Media Lab building.

5.2.2 Generating Vuforia Area Target

A Vuforia Area Target is an environment tracking feature that enables developers to track and augment 3D-scanned spaces. This Area Target can then be mapped into the real world coordinate using augmented reality technology provided by their Vuforia Engine 9, which will then display AR content into the real world. In our case, it could display REINA accurately in the physical environment. REINA will also have a prior knowledge of where everything is located at, and thus, be able to navigate freely, as well as give a tour and guide users around.

After scanning the Media Lab, the scans were uploaded to the Matterport server. In order to create a Vuforia Area Target, the Matterpak Bundle must be purchased from their website, which can be done once the scanned space is processed (See Figure 5-11. Next step is to request Vuforia to enable the Area Target API through their customer service email vuforiafeedback@ptc.com. Once it is ready, we request the Token by accessing the Matterport website, log in, navigate to settings/manage account/request TOKEN.

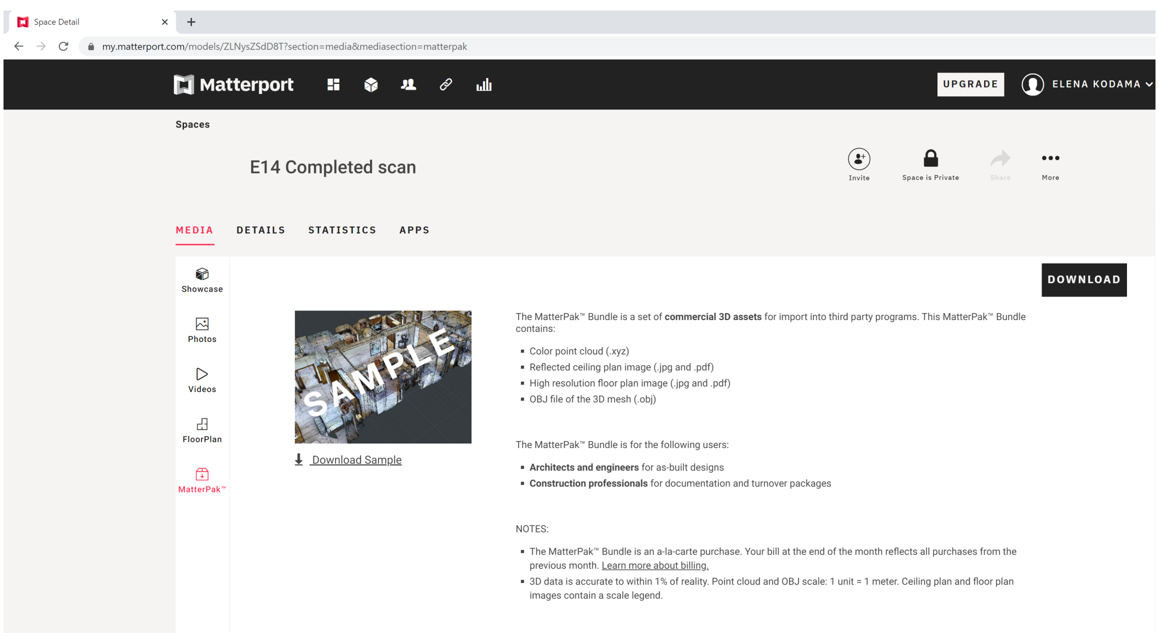


Figure 5-11: The Matterpak bundle can only be purchased off a Professional license account. MIT.nano has one.

Figure 5-12: The area target generator provided by Vuforia.

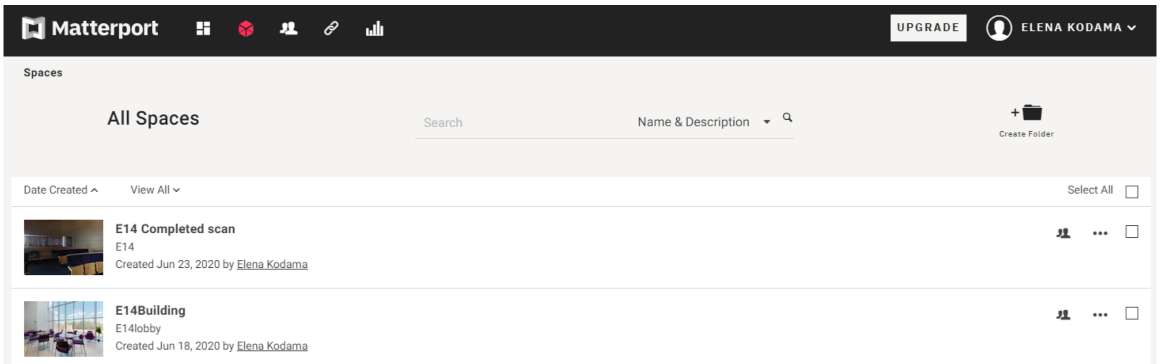


Figure 5-13: The All Spaces section found on Matterport website after logging into an account.

The Area Target Generator[4] must be installed before we proceed. After installing the application, it will request the Matterport Token details which was generated previously. Vuforia gives its users 10 free area target generations. Figure 5-12 shows the Area Target Generator screenshot.

To generate an area target, press the "New Area Target" button, then under Scanned Space, we look for the scanned space ID, which can be found by going to the Matterport "All Spaces" section on your Matterport Professional License account as shown in Figure 5-14. Then select the scanned space you want to generate the area target on, copy the entire URL and paste it in the field (See Figure 5-13). It will then automatically extract the 11-digit SpaceID. Once the Area target is generated, you will see four files (.DAT, .XML, .UnityPackage, and .glb) as shown in Figure 5-15.

[4]https://developer.vuforia.com/downloads/tool

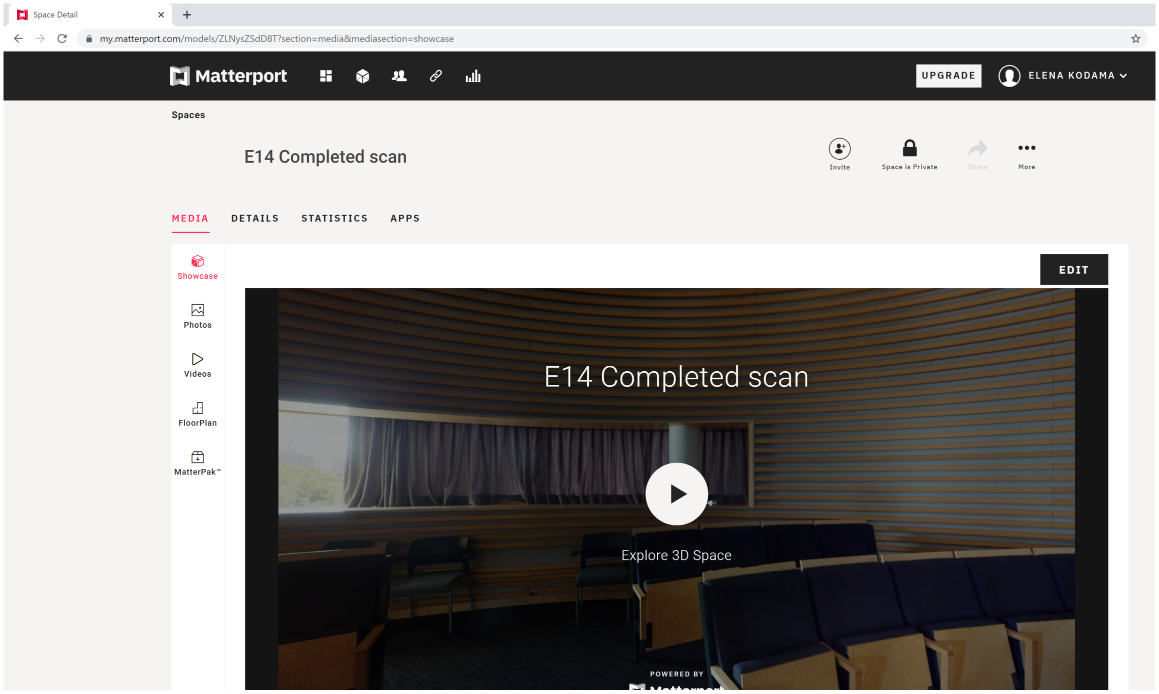
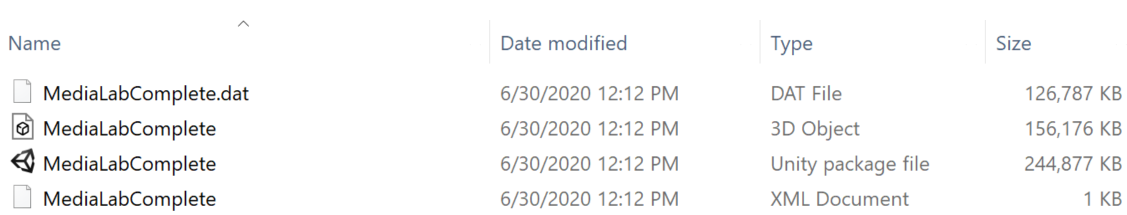


Figure 5-14: The SpaceID can be found in the URL of the selected space.



Name	Date modified	Type	Size
MediaLabComplete.dat	6/30/2020 12:12 PM	DAT File	126,787 KB
MediaLabComplete	6/30/2020 12:12 PM	3D Object	156,176 KB
MediaLabComplete	6/30/2020 12:12 PM	Unity package file	244,877 KB
MediaLabComplete	6/30/2020 12:12 PM	XML Document	1 KB

Figure 5-15: The area target generator generates four files that will be used on Unity.

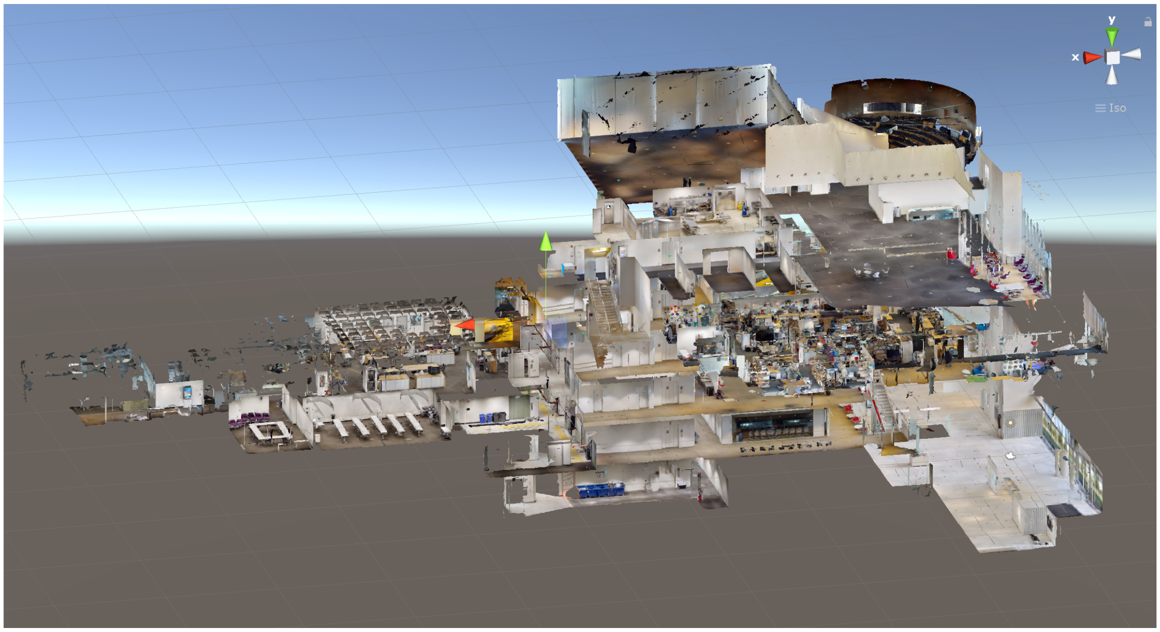


Figure 5-16: A view of the Scanned MIT Media Lab model on unity.

5.2.3 Navigation

The area target unity file can be imported into Unity3D for generating a navigation mesh for an embodied agent to walk on. Figure 5-16 shows the 3D-scanned model of the MIT Media Lab in a Unity3D project. A complete tutorial on how to create a Unity3D project for an embodied agent to walk on the 3D-scanned model can be found in Appendix A. An AR view of the 3D-scanned model in simulation mode is shown in Figure 5-17. REINA is then imported into the same Unity3D project and placed into the 3D scan as seen in Figure 5-18

5.2.4 Navigating in the Real World

Due to the COVID-19 pandemic, I have not been able to test REINA while walking freely in the physical Media Lab building. Thus, this will remain as a future work and should eventually be used through the planned in-person user studies.

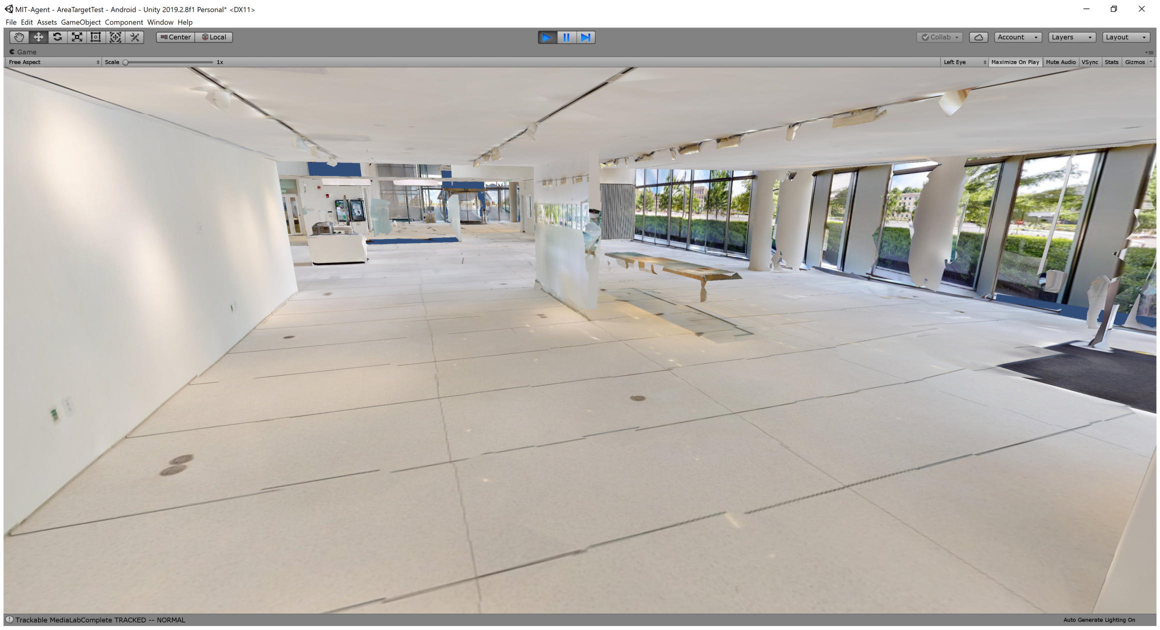


Figure 5-17: An AR perspective view of the Scanned MIT Media Lab model on unity simulation mode.

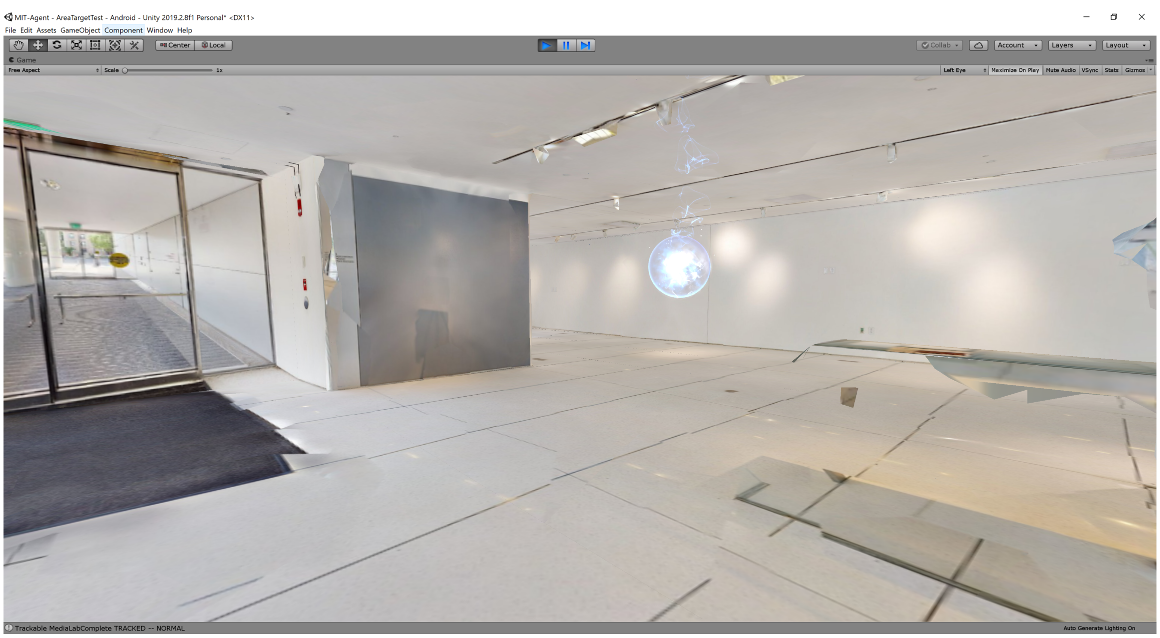


Figure 5-18: REINA in the scanned model of the MIT Media Lab in simulation mode.

Chapter 6

The R.E.I.N.A. Experiment

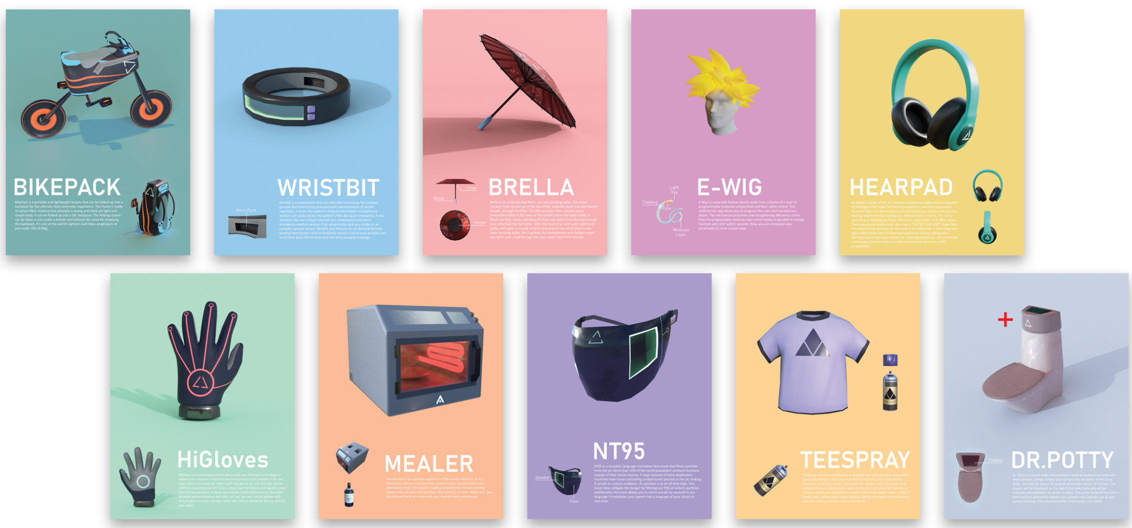


Figure 6-1: 10 fictional projects were designed from scratch to ensure that the information was new for all participants. The preparation included a poster design, 3D rendering and animation, audio description and text description for each project.

The R.E.I.N.A. Experiment consists of a pilot study, a remote user study and an in-person user study based on the Media Lab Tour Guide Experience. However, due to the ongoing COVID-19 Pandemic, the in-person user study has been postponed and remains as a future work. All these user studies follow the approved MIT IRB PROTOCOL #2002000103.

Each of the user studies were designed to guide the next experimental design. The pilot study was conducted to help refine the remote user study, and the remote user study was designed to improve the in-person user study experiment. Three initial research questions will explore the interplay between agent embodiment, memory retention, social presence, attention, and cognitive dissonance in the remote user study. Based on the results and discussions, three additional research questions will be examined in the in-person user study to provide new knowledge about the spatio-temporal effect of having a tour guiding agent in the physical environment, considering effects on agent embodiment, memory retention, social presence, attention and cognitive dissonance.

For the R.E.I.N.A. Experiment, we brainstormed and designed 10 different fictitious projects to ensure that the material used for the experiments were new for all participants. For each project, we devised a text description, an audio script, a 3D rendering, and a 3D animation. A poster was designed for each project containing an image of the 3D rendering with smaller side views, the name in bold, and the text description as seen in Figure 6-1. The English audio script was recorded using an iPhone with the voice of one of the undergraduate research assistants who worked in this project.

6.1 Pilot Study

The purpose of the pilot study was: to identify which of the 10 projects were better refined, to get feedback for improving the projects, and to evaluate the difficulty level of the questions that were going to be used to test memory retention in the remote and in-person user studies.

20 participants (n=20, 35% female, 85% in age bracket 18-29 years old) were recruited through social platforms such as Facebook group and Instagram. Each participant was given a survey containing 5 videos, each with a single poster accompanying the audio description. They were asked to preview each video in full screen and review 7 questions per video about the projects. They rated each question according to their perceived difficulty levels (Easy, Moderate, Difficult). They were told to rate "Easy" if they can recall the answer exactly, "Moderate" if they think they know the answer, and "Difficult" if they have no clue what the answer could be. They were also asked to provide open-ended feedback on each project design (e.g. feasibility of the project, clarity of the audio description, etc).

Based on the results and feedback, we improved the text and audio descriptions, sped up the audio speech to a more normal speed, and selected 5 projects that had a balance of questions in various difficulty levels to be used in the remote user study.

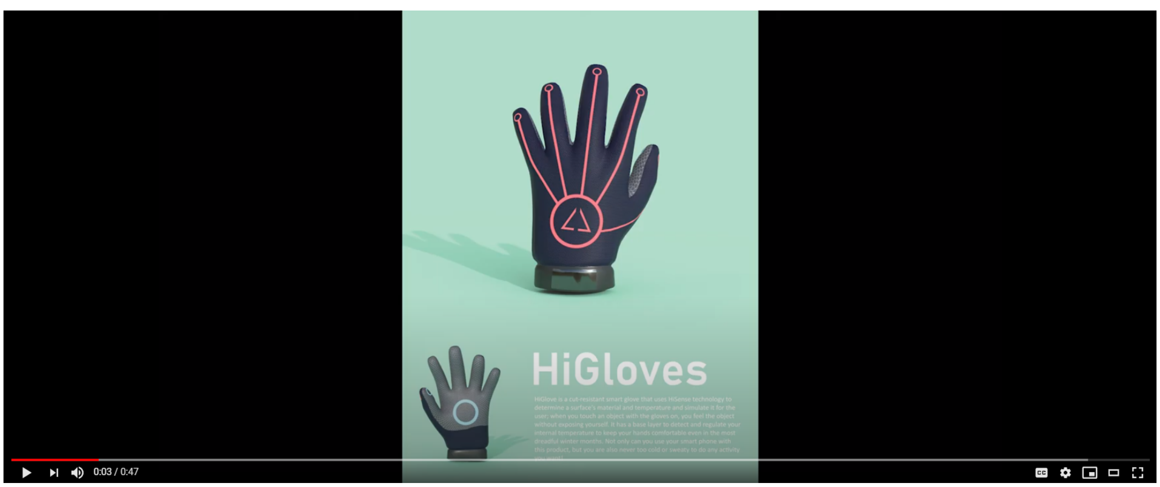


Figure 6-2: Example of a video with the poster and audio description presented during the pilot study.

6.2 Memorable Tour Guiding Experience User Study (Remote)

6.2.1 Research Questions & Hypotheses

In Chapter 3, we have seen previous work related to agent embodiment and memory retention and memorability. However, none of the prior work has compared a disembodied with an abstract non-humanoid embodied agent in a remote gallery tour with several posters and explored their effects on memory retention (learning), social presence, attention and cognitive dissonance on a large pool of international participants.

Main Research Questions:

- **R1:** Is there a significant difference in people's ability to retain information naturally based on what kind of tour guiding agent they observe? (e.g. disembodied vs. abstract-embodied agent).

- **R2:** If an abstract-embodied tour guiding agent poses an open-ended question, will it affect information retain-ability?

- **R3:** Which group will have a higher cognitive dissonance, similar to the Stroop effect, if the audio and the text descriptions do not match?

Hypotheses:

- **H1:** In prior work, we saw that despite people's preference for embodied agent, the disembodied agent was easier to pay attention to, and thus enhanced memory retention of participants [81]. Since an abstract-embodied agent has less distracting elements than an embodied humanoid agent, I hypothesize that an abstract-embodied agent might be able to enhance people's ability to retain information naturally similar to those of a disembodied version of itself.

- **H2:** I hypothesize that by posing a question, people will be able to remember information better because it gives people an action to do after they receive the information. It will remind them to process the information before they move to the next project.

- **H3:** I hypothesize that Group A will have the highest cognitive dissonance because they do not have an embodied agent to pay attention to. Thus, they will eventually shift their focus away to the poster and text, making it difficult to process the audio information given by the agent since the audio and text do not match. Meanwhile, since Groups B and C have an embodied agent, if participants pay attention to the agent, there will not be much cognitive dissonance.

6.2.2 Experiment Design

The purpose of this remote user study was to test our hypotheses H1-3 and attempt to answer research questions R1-3. The results were used to help improve the experiment design for the in-person user study.

All participants were randomly assigned into three different groups A, B, and C. They were sent three surveys to assess their ability to retain information in three phases: right after the tour guiding experiment, 24 hours after, and 72 hours after. Once all participants completed the experiment, an additional survey was sent out to assess the usability, likability, and the perception of social presence of the tour guiding agent.

The experience was built and deployed into iOS and Android applications using a Vuforia image target. The target was used to summon the tour guiding agent when the corresponding image was scanned. A customized application was sent to each participant based on the group they were randomly assigned into. An example of what participants saw during the experience is shown in Figure 6-3. Each group experienced the following conditions:

- **Group A**: Application that summons a disembodied tour guiding agent that only provides an audio description of the project without any physical manifestation (Audio only).

- **Group B**: Application that summons a sphere-like tour guiding agent that radiates light particles when talking, as well as shows lighting modulation inside the sphere based on audio frequency. The audio descriptions are exactly the same as the one in Group A.

- **Group C**: Experiences the same conditions as Group B, but at the end of the audio script, the tour guiding agent poses an open-ended question that the participants need to think about but do not need to say it out loud.

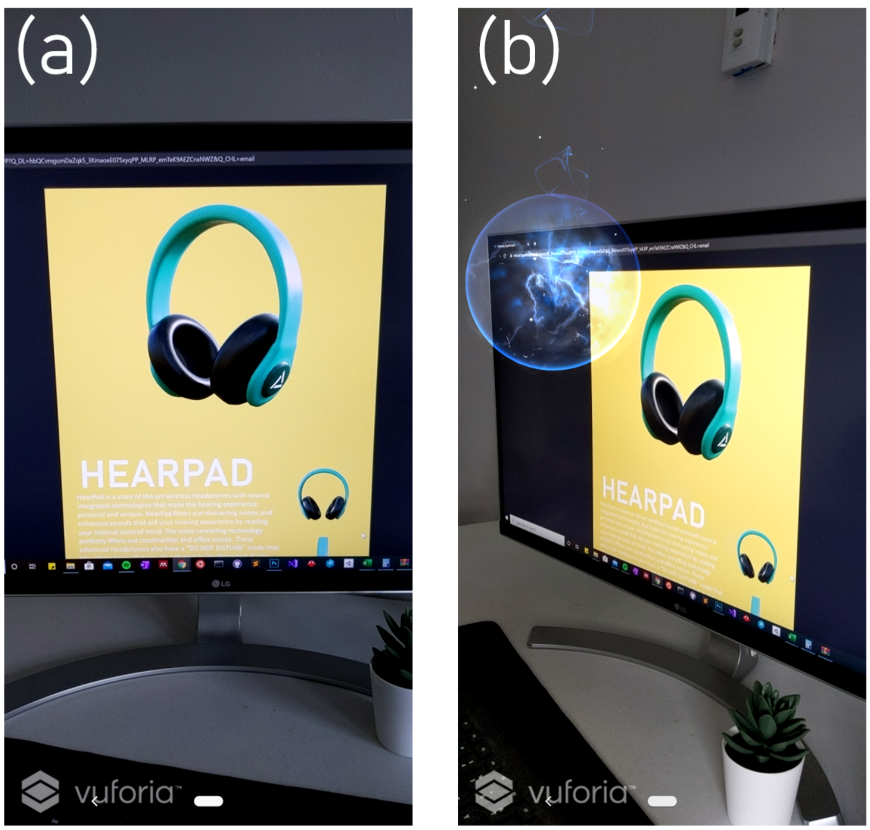


Figure 6-3: (a) Group A's mobile app only triggers the audio description after scanning the poster image. (b) Group B and Group C applications trigger the non-humanoid embodiment of the tour guiding agent (R.E.I.N.A.), manifested as a dynamic sphere-like orb

6.2.3 Procedure

Potential participants first filled out a sign-up form distributed via social platforms such as Facebook, Instagram, LinkedIn, and WhatsApp. They were then randomly selected into one of the three groups A, B, or C. An introduction email was sent to each participant instructing them to read the consent form and to install the provided iOS or Android application on their mobile device. They were also given a 30-minute time slot for a Zoom call with the author to set up and explain the procedure as seen in Figure 6-4. During the Zoom call, they were told to act as if they were in a Gallery Tour and were specifically told not to take notes or force themselves to memorize information that they otherwise would not. Additionally, the application was tested with each participant to ensure proper use. They were told about the Gallery Tour and that they will see five different projects total. They were asked to only summon the tour guide once per project, and that they could read the poster text description as many times as natural for them. However, the author also said that on average, they should not spend more than 3 minutes looking at the poster to avoid forced memorization. Finally, they were given their Subject number and first survey, containing the experience and the questionnaire. The second and third surveys were scheduled to be sent automatically 24 hours and 72 hours after the first survey, using Qualtrics. Participants were also instructed to fill out the other surveys as close as possible to the time of day that they received the first survey, to ensure a more consistent 24hr and 72hr difference across all participants. Once all questions were answered, the participants got off the Zoom call and started the experience.

6.2.4 Evaluation

Since this is a remote user study, the data will be collected through self-reporting questionnaires delivered as four different surveys. Sample of the surveys can be found in Appendix D. All surveys were done without supervision from the researcher. They were already explained the procedure during the setup call, and all the surveys contain detailed instructions on what they should do.

Figure 6-4: All participants downloaded and installed the application on their phone and tested it during the Zoom call. They were also given their subject number and a summary of the procedure before they started the Gallery Tour experiment.

- **Survey 1 and Experiment (S1):** The first survey asks for some background information about the participants and contains a short tutorial for testing the application. It will then display one project at a time. When the participant is ready, they can press the "next" button to see the next project. After exploring the five projects, they are then asked to delete the application. Right after, they see an image of one of the projects, but with no text on it, followed by two multiple choice questions or one fill-the-blank question and one multiple choice question. Usually, the fill-the-blank asked for the name of the project. In total, they see 10 questions, two per project. The aggregated results will be the Total Correct Answers used to compare memory retention.

- **Survey 2 (S2):** This survey is automatically scheduled and sent 24 hours after the experiment for each participant. It contains another 10 questions, two per each project.

- **Survey 3 (S3):** This survey is automatically scheduled and sent 72 hours after the experiment for each participant. It contains another 10 questions, two per each project. It also asks the participants to provide a description of the tour guiding agent if they saw it, their general thoughts about the experiment, the projects, the tour guiding agent, and if they had any technical difficulties.

- **Survey 4 (S4):** This survey was sent after all participants (n=101) had completed the previous three surveys. It dives deeper into what the participants were doing while running the experiment. It asks participants for additional background information that the first survey did not cover, their perceived attention level, a three-item social presence questionnaire adapted from Obaid et al. [70], a question on cognitive dissonance, and more space for general comments.

6.2.5 Participants

114 participants signed up for the study and were randomly assigned to Group A, B, or C. However, 13 of them dropped out due to scheduling issues, technical issues with their mobile device, or did not respond to their emails regarding the setup call. Thus, 101 participants (n=101, 55 female and 46 male) were recruited. The average age was 25 (M=25.3, SD=5.8) with the youngest participant at 18 years old and oldest at 59 years old. 52.5% of the participants are USA-based, 31.7% are Panama-based. Additional background information was collected from the participants for analysis. 53.5% said they rarely use or talk to a voice assistant (e.g. Cortana, Siri, Alexa), while 12.9% said they use one several times a day, and 17.8% indicated they have never used a voice assistant before. 88.2% of the participants spoke fluent English while 11.8% expressed that they were not regular English speakers but were able to understand it. 69% stated they have used an augmented reality (AR) or virtual reality (VR) device or application before, while 31% have never been exposed to AR/VR before this study. Refer to Table 6.1 for a more detailed description about the background of the participants and Table 6.2 and Table 6.3 for detail on the groups. All participants received a $15 Amazon Gift Card as compensation.

Table 6.1: General information about the participants

Age	M=25.3, SD=5.8
Gender	54.5% female (n=55), 45.5% male (n=46)
Location (Currently Based)	52.5% USA (n=55), 31.7% Panama (n=32), 5.9% India (n=6), 4% Canada (n=4), the remaining 5.9% are from Chile, Colombia, Nepal, Japan, Korea, and UK.
Education (Possess or currently pursuing)	50.5% Bachelor's, 28.7% Master's, 10.9% Doctorate, 6.9% High School Diploma, and 3% Associate's
English Fluency	88.2% fluent in English, 11.8% not a regular English speaker.
AR/VR Usage	69% Yes, 31% No
Average amount of meetings or classes on a weekly basis (0, 1, 2, 3, 4 5+)	M=3.7, SD=1.6 48.5% have five or more meetings or classes per week, 12.9% have four, 14.9% have three, 11.9% have two, 4% have one, and 7.9% have none.
Perceived memory retention 24hr after a class, meeting or tour (1= a little, 2= moderate, 3= a lot, 4= a great deal)	M=2.1, SD=0.66 3% a great deal, 16.8% a lot, 65.3% a moderate amount, 14.9% a little
Frequency of usage of voice assistant (e.g. Cortana, Siri, Alexa) (Never used before = 1, Rarely =2, Once a week =3, Several times a week =4, Once a day =5, Several times a day =6)	M=2.7, SD= 1.6 17.8% never used before, 53.5% rarely, 3% once a week, 5% Several times a week, 7.9% once a day, 12.9% several times a day
Type of learner (Strongly visual =5, mostly visual =4, neutral =3, mostly auditory =2, strongly auditory =1)	M=3.9, SD=0.67 66% Mostly Visual, 16% Neutral (50/50), 14% Strongly Visual, and 4% Mostly Auditory.

Table 6.2: Group basic information

	Group A	Group B	Group C
Number of participants	29	37	35
Age	M=26.1, SD=4.9	M=23.8, SD=3.5	M=26.2, SD=7.9
Gender	48.3% female (n=14), 51.7% male (n=15)	51.4% female (n=19), 48.6% male (n=18)	62.9% female (n=22), 37.1% male (n=13)
Location (Currently Based)	41.4% USA, 44.8% Panama, 10.3% India, 3.4% Others	56.8% USA, 27% Panama, 8.1% India, 8.1% Others	57.1% USA, 25.7% Panama, 8.1% India, 8.6% Others
Education (Possess or currently pursuing)	13.8% High School Diploma, 3.4% Associate's, 31% Bachelor's, 41.4% Master's, 10.3% Doctorate	5.4% High School Diploma, 2.7% Associate's, 62.2% Bachelor's, 18.9% Master's, 10.8% Doctorate	2.9% High School Diploma, 2.9% Associate's, 54.3% Bachelor's, 28.6% Master's, 11.4% Doctorate
English Fluency	80% fluent, 20% not a regular speaker.	95.2% fluent, 4.8% not a regular speaker.	86.7% fluent, 13.3% not a regular speaker.
AR/VR Usage	72.4% Yes, 27.6% No	69.4% Yes, 30.6% No	65.7% Yes, 34.3% No

Table 6.3: More detail about the groups

	Group A	Group B	Group C
Average amount of meetings or classes on a weekly basis (0, 1, 2, 3, 4 5+)	M=3.5, SD=1.6 41.4% five or more meetings or classes per week, 13.8% four, 17.2% three, 13.8% two, 6.9% one, and 6.9% none.	M=3.9, SD=1.5 56.8% five or more meetings or classes per week, 8.1% four, 13.5% three, 13.5% two, 2.7% one, and 5.4% none.	M=3.6, SD=1.7 45.7% five or more meetings or classes per week, 17.1% four, 14.3% three, 8.6% two, 2.9% one, and 11.4% none.
Perceived memory retention 24hr after a class, meeting or tour 1= a little, 2= moderate, 3= a lot, 4= a great deal	M=2.0, SD=0.7 3.4% great deal, 13.8% a lot, 65.5% moderate, 17.2% little	M=2.1, SD=0.5 0% great deal, 18.9% a lot, 70.3% moderate, 10.8% little	M=2.1, SD=0.8 5.7% great deal, 17.1% a lot, 60% moderate, 17.1% little
Frequency of usage of voice assistant (e.g. Cortana, Siri, Alexa) 1= Never, 2= Rarely, 3= Once a week, 4= Several times a week, 5= Once a day, 6= Several times a day	M=2.9, SD=1.7 10.2% never, 55.2% rarely, 6.9% once a week, 3.4% once a day, 6.9% Several times a week, 17.2% several times a day	M=2.6, SD=1.7 24.3% never, 51.4% rarely, 0% once a week, 8.1% once a day, 2.7% Several times a week, 13.5% several times a day	M=2.7, SD=1.6 17.1% never, 54.3% rarely, 2.9% once a week, 11.4% once a day, 5.7% Several times a week, 8.6% several times a day
Type of learner 5= Strongly visual, 4= Mostly visual, 3= Neutral, 2= Mostly auditory, 1= Strongly auditory	M=4.1, SD=0.6 20.7% Strongly Visual, 65.5% Mostly Visual, 13.8% Neutral (50/50), 0% Mostly Auditory	M=3.8, SD=0.8 16.7% Strongly Visual, 52.8% Mostly Visual, 22.2% Neutral (50/50), 8.3% Mostly Auditory	M=3.9, SD=0.5 5.7% Strongly Visual, 80% Mostly Visual, 11.4% Neutral (50/50), 2.9% Mostly Auditory

6.3 Remote User Study Analysis & Results

For data analysis, all participant responses were inspected to identify issues with their data. Responses from 13 participants were dropped due to image tracking issues with the application during the experiment, incomplete survey responses, or submission of survey response more than 24hr past the intended deadline. In total, data from 88 participants were used in the analysis. The results from the four surveys were collected into a .csv file and processed using Python data analysis libraries, such as Pandas[1], statsmodels[2], etc. Matplotlib[3] was used for plotting. The jupyter notebook code can be found in Appendix C.

First, we tested the data for normal distribution using the Shapiro Wilk test. The total number of correct answers, found by summing the three surveys (S1+S2+S3), was used as the input data. Refer to Figure 6-5 for the histogram. We got the following, $W = 0.977, p - value = 0.126$. Since the P-Value of the Shapiro Wilk Test is greater than 0.05, we can assume a normal distribution.

Bartlett's test was used to check for homogeneity of variances. Similar to the Shapiro Wilk Test, the total number of correct answers (S1+S2+S3) per group was used as the input data. This yielded the following results: $W = 0.701, p - value = 0.704$. Since the p-value is non-significant, we can assume that the data for the three groups have equal variances.

6.3.1 Project Questionnaire

The difficulty of each question (Q) in each project was measured by using the mean (easy =1, hard =0) and standard deviation for each survey (S). Refer to Table 6.4, Table 6.5, and Table 6.6. All three surveys contained two questions about each of the five projects (HiGloves, Brella, Dr.Potty, HearPad, and BikePack). A Boxplot of the correct answers per project per group is shown in Figure 6-6.

[1] https://pandas.pydata.org/
[2] https://www.statsmodels.org/stable/index.html
[3] https://matplotlib.org/

Table 6.4: The Mean and Standard Deviation of the Questions in Each Project in Survey 1

S1 Questions	A	B	C
HiGloves S1Q1	M=0.962, SD=0.196	M=0.967, SD=0.183	M=0.938, SD=0.246
HiGloves S1Q2	M=0.385, SD=0.496	M=0.533, SD=0.507	M=0.469, SD=0.507
BikePack S1Q1	M=0.769, SD=0.430	M=0.767, SD=0.430	M=0.906, SD=0.296
BikePack S1Q2	M=0.731, SD=0.452	M=0.667, SD=0.479	M=0.781, SD=0.420
Dr.Potty S1Q1	M=0.846, SD=0.368	M=0.733, SD=0.450	M=0.844, SD=0.369
Dr.Potty S1Q2	M=0.654, SD=0.485	M=0.700, SD=0.466	M=0.656, SD=0.483
HearPad S1Q1	M=0.231, SD=0.430	M=0.267, SD=0.450	M=0.281, SD=0.457
HearPad S1Q2	M=1.000, SD=0.000	M=0.967, SD=0.183	M=0.969, SD=0.177
Brella S1Q1	M=1.000, SD=0.000	M=1.000, SD=0.000	M=1.000, SD=0.000
Brella S1Q2	M=0.808, SD=0.402	M=0.833, SD=0.379	M=0.812, SD=0.397
S1 TOTAL (MAX=10)	M=7.385, SD=1.388	M=7.433, SD=1.278	M=7.656, SD=1.473

Table 6.5: The Mean and Standard Deviation of the Questions in Each Project in Survey 2

S2 Questions	A	B	C
HiGloves S2Q1	M=0.962, SD=0.196	M=0.800, SD=0.407	M=0.906, SD=0.296
HiGloves S2Q2	M=0.846, SD=0.368	M=0.867, SD=0.346	M=0.781, SD=0.420
BikePack S2Q1	M=0.577, SD=0.504	M=0.567, SD=0.504	M=0.375, SD=0.492
BikePack S2Q2	M=0.962, SD=0.196	M=0.900, SD=0.305	M=0.969, SD=0.177
Dr.Potty S2Q1	M=0.885, SD=0.326	M=0.733, SD=0.450	M=0.812, SD=0.397
Dr.Potty S2Q2	M=0.231, SD=0.430	M=0.267, SD=0.450	M=0.406, SD=0.499
HearPad S2Q1	M=0.731, SD=0.452	M=0.733, SD=0.450	M=0.812, SD=0.397
HearPad S2Q2	M=0.885, SD=0.326	M=0.767, SD=0.430	M=0.812, SD=0.397
Brella S2Q1	M=0.577, SD=0.504	M=0.567, SD=0.504	M=0.688, SD=0.471
Brella S2Q1.1	M=0.615, SD=0.496	M=0.633, SD=0.490	M=0.688, SD=0.471
S2 TOTAL (MAX=10)	M=7.269, SD=1.251	M=6.833, SD=1.533	M=7.250, SD=1.814

Table 6.6: The Mean and Standard Deviation of the Questions in Each Project in Survey 3

S3 Questions	A	B	C
HiGloves S3Q1	M=0.231, SD=0.430	M=0.167, SD=0.379	M=0.219, SD=0.420
HiGloves S3Q2	M=0.577, SD=0.504	M=0.600, SD=0.498	M=0.656, SD=0.483
BikePack S3Q1	M=0.385, SD=0.496	M=0.367, SD=0.490	M=0.281, SD=0.457
BikePack S3Q2	M=0.692, SD=0.471	M=0.700, SD=0.466	M=0.562, SD=0.504
Dr.Potty S3Q1	M=0.808, SD=0.402	M=0.867, SD=0.346	M=0.875, SD=0.336
Dr.Potty S3Q2	M=0.808, SD=0.402	M=0.733, SD=0.450	M=0.906, SD=0.296
HearPad-S3Q1	M=0.885, SD=0.326	M=0.667, SD=0.479	M=0.781, SD=0.420
HearPad S3Q2	M=0.731, SD=0.452	M=0.800, SD=0.407	M=0.719, SD=0.457
Brella S3Q1	M=0.692, SD=0.471	M=0.700, SD=0.466	M=0.656, SD=0.483
Brella S3Q2	M=0.462, SD=0.508	M=0.500, SD=0.509	M=0.438, SD=0.504
S3 TOTAL (MAX=10)	M=6.269, SD=1.511	M=6.100, SD=1.242	M=6.094, SD=1.614

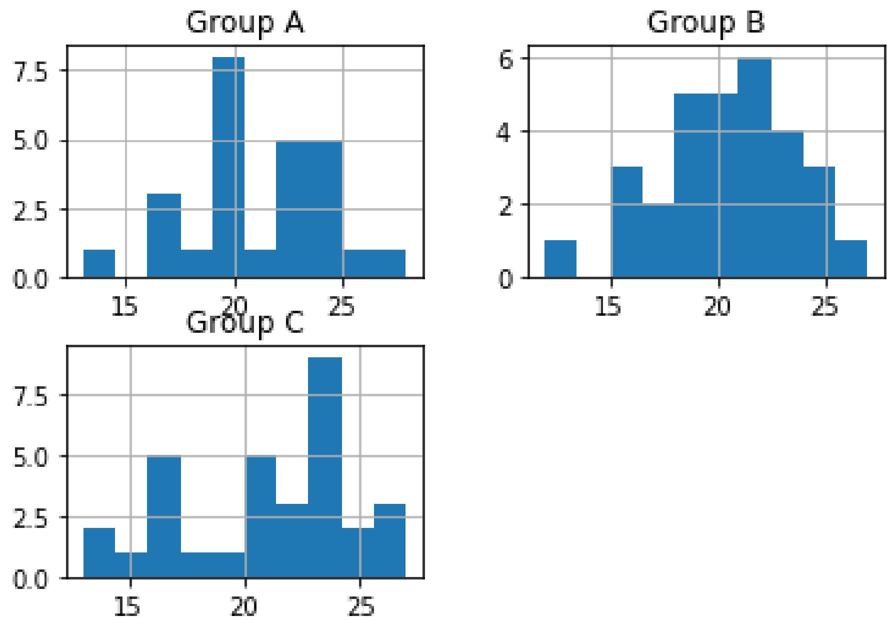


Figure 6-5: Histograms of the sum of the total correct answers (S1+S2+S3) by group

6.3.2 Memory Retention

In order to answer research question R1 and R2, a one-way analysis of variances (ANOVA) was computed for different conditions comparing the three groups (between subject analysis) to identify significant differences in their mean values. Refer to Table 6.9 for a one-way ANOVA analysis comparing the total correct answers in survey S1, survey S2, survey S3, and the sum of all three surveys (S1+S2+S3). Refer to Table 6.10 for a one-way analysis comparing the total correct answers in survey S1 (TOTAL CORRECT S1) for different filtered conditions, Table 6.11 for survey S2, Table 6.12 for survey S3, and Table 6.13 for the sum of all three surveys (S1+S2+S3). The mean and standard deviation for the total correct answers in survey S1, S2, and S3 is shown in Table 6.8. A Boxplot of the total correct answers per survey per group is shown in Figure 6-8, and a boxplot of the total correct answers (also referred as Total Memory Retention) for different filtered conditions is shown in Figure 6-9.

In the survey S4, the one sent after the entire experiment, participants were asked to rate how much they think they remembered from the tour guiding agent compared to the text in a 5-point Likert Scale. The question had the labels: "Remember mostly

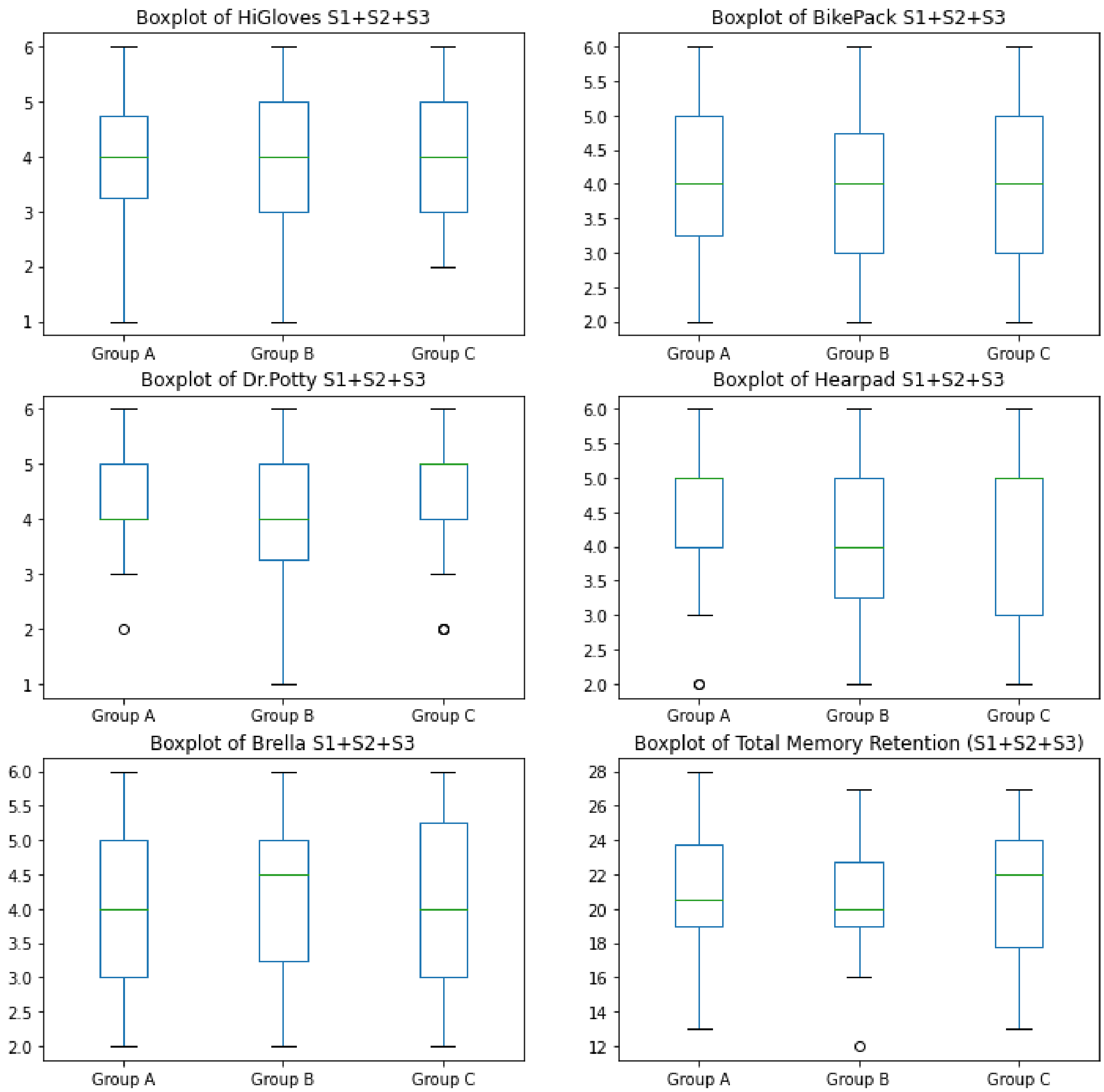


Figure 6-6: Boxplot of Memory Retention based on Projects

Table 6.7: The Mean and Standard Deviation of all the surveys summed together per project

S1+S2+S3 Questions	Population	A	B	C
HiGloves S1+S2+S3	M=3.955, SD=1.164	M=3.962, SD=1.113	M=3.933, SD=1.285	M=3.969, SD=1.121
BikePack S1+S2+S3	M=3.977, SD=1.154	M=4.115, SD=1.143	M=3.967, SD=1.098	M=3.875, SD=1.238
Dr.Potty S1+S2+S3	M=4.261, SD=1.218	M=4.231, SD=0.908	M=4.033, SD=1.450	M=4.500, SD=1.191
HearPad S1+S2+S3	M=4.341, SD= 1.183	M=4.462, SD=1.174	M=4.200, SD=1.095	M=4.375, SD=1.289
Brella S1+S2+S3	M=4.227, SD=1.230	M=4.154, SD=1.347	M=4.233, SD=1.073	M=4.281, SD=1.301
Total Memory Retention (S1+S2+S3) (MAX=30)	M=20.761364, SD=3.380	M=20.923, SD=3.273	M=20.367, SD=3.200	M=21.000, SD=3.690

from the Agent" (5pt), "Remember a bit more from Agent" (4pt), "Remember about the same from Both" (3pt), "Remember a bit more from Text" (2pt), and "Remember mostly from Text" (1pt). Participants in Group A, Group B, and Group C had a mean scores of M=2.73, SD=1.19, M=2.23, SD=1.07, and M=2.31, SD=0.931 respectively.

6.3.3 Usability

Prior to the start of the experiment, participants were asked how often they used their voice assistant (e.g. Cortana, Alexa, Siri) in the survey using a 6-point Likert scale with the labels: "Never used before" (1pt), "Rarely" (2pt), "Once a Week" (3pt), "Several times a week" (4pt), "Once a Day" (5pt), and "Several times a day" (6pt). The mean scores and standard deviations for Group A, B, and C are M=3.000, SD=1.744, M=2.400, SD=1.589, and M=2.656, SD=1.558 respectively. The mean and standard deviation for the entire population (100 responses, all groups) is M=2.670 and SD=1.624.

Additionally, participants were asked, "If you were going to a tour guiding ex-

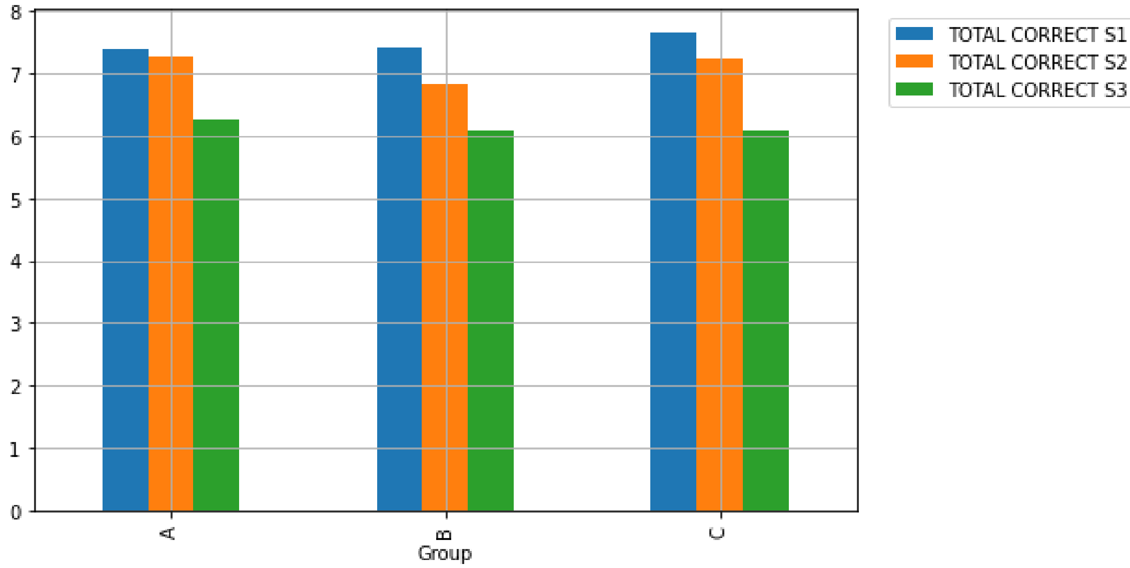


Figure 6-7: Bar plot showing correct answers of each group right after experiment (S1), 24hr after (S2), and 72hr after (S3)

Table 6.8: The mean and standard deviation of group A, B, and C in survey s1, s2, and s3

Survey	Group A	Group B	Group C
Total Correct Answers in Survey S1 (Right after Gallery Tour Experience)	M=7.385, SD=1.388	M=7.433, SD=1.278	M=7.656, SD=1.473
Total Correct Answers in Survey 2 (24hr after Gallery Tour Experience)	M=7.269, SD=1.251	M=6.833, SD=1.533	M=7.250, SD=1.814
Total Correct Answers in Survey 3 (72hr after Gallery Tour Experience)	M=6.269, SD=1.511	M=6.100, SD=1.242	M=6.094, SD=1.614
Total Memory Retention (S1+S2+S3)	M=20.923, SD=3.273	M=20.367, SD=3.200	M=21.000, SD=3.690

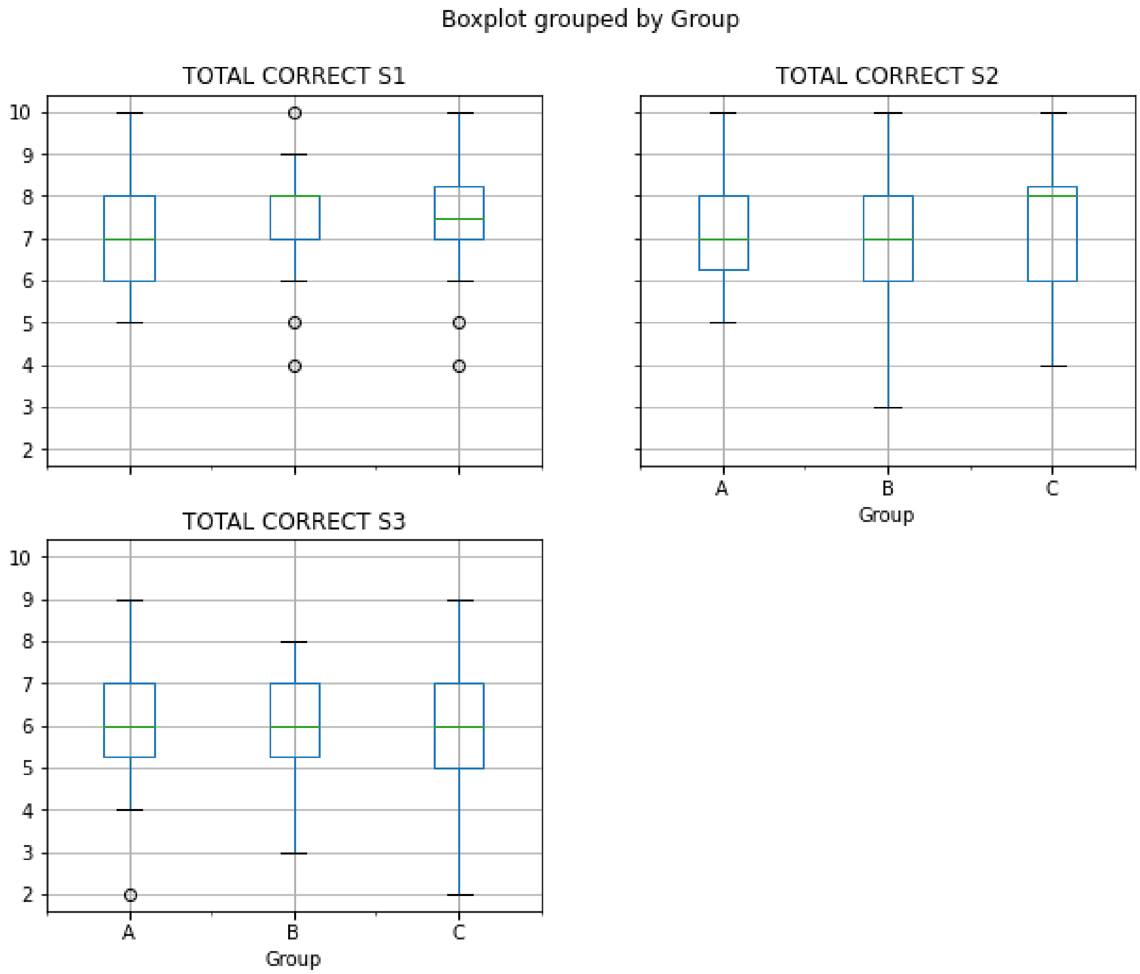


Figure 6-8: Boxplot of total correct answers of each survey compared between group

Table 6.9: One-way ANOVA analysis for comparing memory retention between group A, B, and C

Condition	F	PR(>F) (p-value)
Total Correct Answers in Survey 1 (Right after Gallery Tour Experience)	0.329	0.720
Total Correct Answers in Survey 2 (24hr after Gallery Tour Experience)	0.727	0.486
Total Correct Answers in Survey 3 (72hr after Gallery Tour Experience)	0.127	0.881
Total Memory Retention (S1+S2+S3)	0.309	0.735

Table 6.10: One-way ANOVA analysis comparing total correct answer from survey 1 between group A, B, and C on filtered data

Analyzed Data	Filtered Condition	Group Count	F	PR(>F) (p-value)
TOTAL CORRECT S1	Gender=Female	A=12, B=17, C=21	1.332	0.274
TOTAL CORRECT S1	Gender=Male	A=14, B=13, C=11	0.036	0.965
TOTAL CORRECT S1	Country=USA	A=11, B=17, C=17	0.331	0.720
TOTAL CORRECT S1	Country=PA	A=11, B=10, C=9	1.162	0.328
TOTAL CORRECT S1	English Fluency=Yes	A=22, B=24, C=28	0.034	0.966
TOTAL CORRECT S1	Learner Type=4	A=18, B=15, C=25	0.139	0.871

Table 6.11: One-way ANOVA analysis comparing total correct answer from survey 2 between group A, B, and C on filtered data

Analyzed Data	Filtered Condition	Group Count	F	PR(>F) (p-value)
TOTAL CORRECT S2	Gender=Female	A=12, B=17, C=21	0.949	0.394
TOTAL CORRECT S2	Gender=Male	A=14, B=13, C=11	0.069	0.934
TOTAL CORRECT S2	Country=USA	A=11, B=17, C=17	0.124	0.884
TOTAL CORRECT S2	Country=PA	A=11, B=10, C=9	1.903	0.169
TOTAL CORRECT S2	English Fluency=Yes	A=22, B=24, C=28	0.477	0.623
TOTAL CORRECT S2	Learner Type=4	A=18, B=15, C=25	1.107	0.338

Table 6.12: One-way ANOVA analysis comparing total correct answer from survey 3 between group A, B, and C on filtered data

Analyzed Data	Filtered Condition	Group Count	F	PR(>F) (p-value)
TOTAL CORRECT S3	Gender=Female	A=12, B=17, C=21	0.934	0.400
TOTAL CORRECT S3	Gender=Male	A=14, B=13, C=11	0.243	0.786
TOTAL CORRECT S3	Country=USA	A=11, B=17, C=17	0.267	0.767
TOTAL CORRECT S3	Country=PA	A=11, B=10, C=9	0.636	0.537
TOTAL CORRECT S3	English Fluency=Yes	A=22, B=24, C=28	0.052	0.950
TOTAL CORRECT S3	Learner Type=4	A=18, B=15, C=25	0.728	0.487

Table 6.13: One-way ANOVA analysis comparing total correct answers from the sum of all three surveys between group A, B, and C on filtered data

Analyzed Data	Filtered Condition	Group Count	F	PR(>F) (p-value)
Total Memory Retention	Gender=Female	A=12, B=17, C=21	0.833	0.441
Total Memory Retention	Gender=Male	A=14, B=13, C=11	0.018	0.983
Total Memory Retention	Country=USA	A=11, B=17, C=17	0.380	0.686
Total Memory Retention	Country=PA	A=11, B=10, C=9	1.365	0.272
Total Memory Retention	English Fluency=Yes	A=22, B=24, C=28	0.150	0.861
Total Memory Retention	Learner Type=4	A=18, B=15, C=25	0.541	0.585

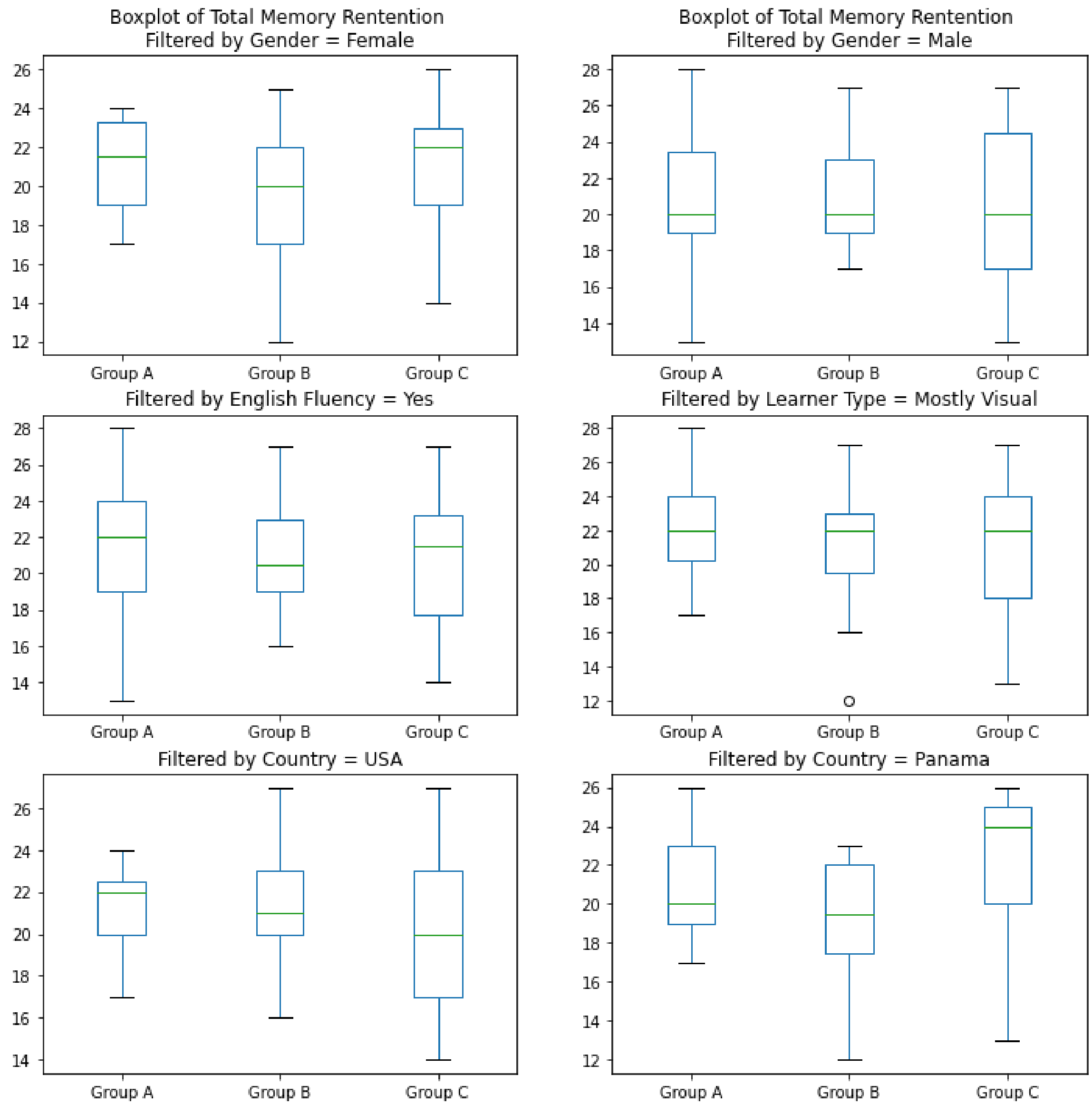


Figure 6-9: Boxplot of Total Memory Retention (sum of total correct answers from all three surveys) filtered.

Table 6.14: The mean and standard deviation for Group A, B, and C for the social presence questions.

Social Presence Question	Population (100 responses)	Group A (29 responses)	Group B (36 responses)	Group C (35 responses)
I would have liked the experience to continue	M=3.966, SD=0.915	M=3.769, SD=1.032	M=3.900, SD=0.923	M=4.188, SD=0.780
I felt REINA provided a memorable experience	M=3.773, SD=0.919	M=3.538, SD=0.989	M=3.667, SD=0.884	M=4.062, SD=0.840
I felt as though I was in the same space as REINA	M=3.330, SD=1.111	M=3.308, SD=1.050	M=3.133, SD=1.332	M=3.531, SD=0.915

perience in the future, would you want to have a similar agent give you a tour?", to which 81% stated yes and 19% no. The population mean and standard deviation is M=0.807, SD=0.397 with a mean of 1 being yes and a mean of 0 being no. Group-wise, the mean and standard deviation are M=0.769, SD=0.430 for Group A, M=0.767, SD=0.430 for Group B, and M=0.875, SD=0.336 for Group C.

Social Presence

There were three questions in survey 4 regarding social presence: (1) I would have liked the experience to continue, (2) I felt REINA provided a memorable experience, and (3) I felt as though I was in the same space as REINA. A 5-point Liker Scale rating was used with the labels "Strongly Agree" (5pt), "Somewhat Agree" (4pt), "Neutral" (3pt), "Somewhat Disagree" (2pt), and "Strongly Disagree" (1pt). The mean and standard deviation of the population (100 responses) is (1) M=3.966, SD=0.915, (2) M=3.773, SD=0.919, (3) M=3.330, SD=1.111. Refer to Table 6.14 for more detail.

6.3.4 Agent Visualization

Participants were asked, "If you can have an augmented reality AI companion, would you want it to manifest visually or would you rather have a voice assistant like cortana, alexa, siri?". A 5-point Likert scale rating was implemented with the labels: "Always Visual (Embodied agent all the time)" (5pt), "Mostly Visual (shows up only when I need it to in physical form)" (4pt), Neutral (I am okay with it being visual or just voice only)" (3pt), "Mostly Voice (rarely should it shows up as physical agent)" (2pt), and "Always Voice (I am good with Cortana, Alexa, Siri. Do not need it to be manifested physically)" (1pt). The population mean is 3.67 with a standard deviation of 0.867. The mean score and standard deviation for Group A, B, and C are M=3.577, SD=0.703, M=3.600, SD=0.932, and M=3.812, SD=0.931 respectively.

Participants in group B and C were asked, "Did you see the guiding agent? If yes, please describe what you saw and what it resembles." Refer to Table 6.15 for highlight of some of the descriptions and Appendix F for full responses. Many described the tour guiding agent as a "bubble", "sphere", or an "orb", some said it reminded them of "Cortana" or "Siri", others also said it resembled "Navi" from Zelda and "Jarvis" from Iron Man and some other game references.

6.3.5 Attention & Cognitive Dissonance

Participants were asked the following questions about attention: (1) "I was able to pay attention to the agent when it was talking" and (2) "If you struggled with paying attention to the agent, can you describe why and what you were doing instead?". For question (1), a 5-point Likert Scale rating was used with the labels: "Strongly agree (Easy to pay attention to)" (5pt), "Agree" (4pt), "Neutral (3pt), "Disagree" (2pt), and "Strongly disagree (couldnt pay attention at all)" (1pt). The mean and standard deviation of the population is M=3.682, SD=0.838. The mean and standard deviation for Group A, B, and C are M=3.962, SD=0.720, M=3.567, SD=0.774, and M=3.562, SD=0.948 respectively. Refer to Table 6.16 for the comments on why participants could not pay attention, only those who responded "Disagree" are highlighted. None

Table 6.15: Description of the tour guiding agent by some of the participants in Group B and C.

Group	"Did you see the guiding agent? If yes, please describe what you saw and what it resembles."
B	It looked like a sphere. It was shining mix of blue-white colors
B	Yes, it looked like an orb, kinda looked like a jellyfish. Particles floated out of it consistently when it spoke.
B	Yes, it is a glowing blue orb that releases electric "stem". It just looks like some kind of sci-fi ball.
B	yes, reminds me of a water droplet or Navi from Legend of Zelda
B	R.E.I.N.A. Circular shaped, with a galactic feel in it. Remind me of Siri. When it is talking, the middle will start lighting up.
B	I saw R.E.I.N.A. i saw her like a blue floating orbe that reminds me a little bit of Jarvis from "Iron Man".
B	Yes it look like a fairy of the legend of zelda
C	It looks like a floating 3D Cortana. Or from Maplestory, the "Bonus XP" that can be obtained.
C	Yes, it was a little blue orb that would float and vibrate as it spoke. It flowed and had blue squiggly lines coming off it. It resembled a guiding light, kinda like one of those Disney movies (Frozen I think), where the lights guide the character.
C	Yes. It was a translucent sphere that looked like water flowing a bit and it lit up as it spoke.
C	To be really honest, I don't remember. I don't remember seeing anything particular except for the screen, but I can't remember for sure.
C	Yes, it was a floating orb that talked. I don't think it resembled anything specific to me beyond that.
C	Electric glowing vaporous ball
C	Yes, I did. It's on the top left conner of the posters. It's like a transparent glass ball with blue electric spark inside.
C	Yes, it resembled an oracle. It had particles flowing out of it and it floated on the top left corner of each product. It was purple.
C	yes, i saw a sphere blue with lightning bolts , floating in my pc , like a cortana from halo
C	yes, like navi(from Zelda) sparkling.

of the participants responded "Strongly disagree". Refer to the Appendix E for full responses.

They were also asked about cognitive dissonance: "Did you suffer from cognitive dissonance while the tour guiding agent was explaining? For example, you were reading the text while the tour guide was talking, but since the text and the audio do not match, you were not able to retain the audio description. (Similar to Stroop Effect)". 62% of the participants responded "Yes, I suffered cognitive dissonance." and 38% said, "No, I was able to pay attention to the audio description, so did not suffer cognitive dissonance." The mean and standard deviation was calculated by converting the label "Yes" to 1pt and "No" to 0pt. Hence, the scores for the population is M=0.614, SD=0.490, and for Group A, B, and C are M=0.500, SD=0.510, M=0.600, SD=0.498, and M=0.719, SD=0.457 respectively.

6.3.6 Additional Participant Comments

Participants were given a few chances to provide additional feedback in survey 3 and 4. Refer to Appendix G for the full responses.

6.4 Discussion

In H1, I hypothesized that an abstract-embodied agent like REINA could enhance participant's memory retention more than a disembodied agent, although Group C (abstract-embodied agent + question) showed a higher memory retention, it was not significant. All three groups showed very similar memory retention scores regardless of embodiments types. In fact, the average correct answer in Group B was slower than Group A.

In H2, I hypothesized that by posing a question, people will remember information better, and according to the results, Group C (abstract-embodied agent + question) did show a slightly better memory retention (non-significant) compared to Group B (abstract-embodied agent without question) and Group A (disembodied). In H3, I hypothesized that Group A will have the highest cognitive dissonance since they

Table 6.16: Participants who responded "disagree" to paying attention commented on why and what they were doing instead of paying attention to the agent. No participant responded "Strongly disagree".

Group	Comment on Attention
B	"Well part of it was probably because I was hung over, I was focusing on how the orb moved along with the words rather the words themselves."
B	"Didn't know if I should be reading or listening and got distracted."
C	"I think I was taken aback by the agent itself and it took me a little while to not marvel at how cool it was. in the mean time i was missing all of the content the agent was sharing."
C	I was easily able to pay attention to the agent audio, but had no reason to pay attention visually. I was looking at the projects and listening to the agent.
C	"Looking at the product and thinking what can it be used for or what are some flaws."
C	"I struggled to pay attention because I was also trying to read the text at the same time."
C	"I was looking at the shape of the agent and the little particle that were leaving its body. I was also trying to read at the same time to match up the information I was receiving."
C	"I was trying to listen at first and then frustrated by the speed and intonation of the agent I began to read the poster and attempt to take in visual information whilst the agent was talking. Perhaps not wise to try to do both simultaneously! I was also aware of background noise in my house from children but this would be normal for me in a real life gallery experience. The information from the agent was very detailed and also different to the information on the poster. Had the visual and auditory guides been designed to complement each other I think I would have absorbed more information."

do not have an embodied agent to pay attention to, however, the results showed that Groups C and B showed higher cognitive dissonance than Group A. Overall, H3 disagrees with the results, H2 somewhat agreed with the results, and H1 seems inconclusive because of contradicting results between the groups. None of the results were significant, but showed some interesting trends that are discussed below.

6.4.1 Memory Retention

The mean and standard deviation from the total correct answers from survey S1, S2, and S3 showed that right after the experiment, participants in group C were most correct with a mean score of 7.656, SD=1.473, followed by Group B (M=7.433, SD=1.278), and then Group A (M=7.385. SD=1.388) out of the possible 10 points (two questions per project, 5 projects). However, we can clearly see that the means are very close. The results from the ANOVA analysis confirmed that there are no significant differences in memory retention scores between the three groups in survey S1 (F=0.329, p-value=0.720). Survey S2, which was sent to participants 24 hours after the experiment, also contained two questions per project. The questions were different than the ones shown in survey S1 to avoid the training effect. The mean and standard deviation for S2 are M=7.269, SD=1.251 for Group A, M=6.833, SD=1.533 for Group B, and M=7.250, SD=1.814 for Group C. Participants in Group A performed slightly better than Group C and Group B. However, the ANOVA also showed no significant differences in the mean scores (F=0.727, p-value= 0.486). Survey S3, which was sent 72 hours after the experiment, also contained two different questions for each project. The mean and standard deviation are M=6.269, SD=1.511 for Group A, M=6.100, SD=1.242 for Group B, and M=6.094, SD=1.614 for Group C. Group A did slightly better than Group B and C again. The results for the sum of the total correct answers from all three surveys (maximum correct score of 30 points - S1+S2+s3) showed a mean of M=20.923, SD=3.273 for Group A, M=20.367, SD=3.200 for Group B, and M=21.000, SD=3.690 for Group C. Overall, Group C seems to have a slightly higher score than Group A and B, but the ANOVA also showed no significant differences (F=0.309, p-value=0.735). One interesting observation from the filtered data analysis was that female (n=50) had a slightly higher memory retention score than male (n=38) in all of the three surveys and groups. Also, there were more females in group C (n=21) than group A (n=12) and group B (n=17). Females made up 46.2% of Group A, 56.7% of Group B, and 65.6% of Group C. Several factors might have contributed to Group C having a higher memory retention besides the gender composition. Group C saw REINA, the embodied tour guiding agent and were posed an open question by REINA at the end of the audio description for each project. By posing a question at the end of the audio, it might have given participants more time to process the information, which explain the slightly better memory retention. This result contradicts those of Techasarntikul et al. where they reported that the audio-only with an embodied anime agent scored the lowest in memory recall in their 10 conditions user study [88]. However, since there is no statistical significant, I cannot conclude that an abstract-embodied agent can help us retain information better.

Nevertheless, when asked to compare how much participants think they remembered from the Text compared to the Tour Guiding Agent through a 5-points Likert Scale rating (1-Text, 5-Tour Guide Agent), all groups showed that they remembered better from the Text (Group A: M=2.73, SD=1.19; Group B: M=2.23, SD=1.07; Group C: M=2.31, SD=0.931). In fact, participants in Group A (disembodied tour guiding agent with not physical embodiment) showed a slightly higher mean than Group B and C, which could suggest that voice-only agent might be easier to pay attention to, and hence, help participants retain information better than an embodied agent. This result is consistent with prior work where they found that the disembodied agent yielded a higher learning score from their participants [81].

There are many additional variables that could have affected the memory retention score. This remote experiment exposed many limitations of the study, a lot of variables could not be controlled, such as the environment, noise, and timing. One of the participants reported the following:

> "I was trying to listen at first and then frustrated by the speed and intonation of the agent I began to read the poster and attempt to take in visual information whilst the agent was talking. Perhaps not wise to try to do

> both simultaneously! I was also aware of background noise in my house from children but this would be normal for me in a real life gallery experience. The information from the agent was very detailed and also different to the information on the poster. Had the visual and auditory guides been designed to complement each other I think I would have absorbed more information."

Another one stated, "well part of it was probably because I was hung over, I was focusing on how the orb moved along with the words rather the words themselves," highlighting that the agent itself was distracting to some, potentially due to the novelty effect of AR technology. This is consistent with the result found by Lu et al., where a high percentage of the participants thought that the AR was distracting [93], but they attributed it to the novelty effect of AR technology and argued that that distraction might have actually enhanced subject's ability to objectify the art.

Another issue was that the embodied agent did not exhibit interactivity and social skills. Many reported that they got used to the agent talking and complained that the embodied agent was not doing much besides changing the particle system visualization that was modulated by the audio. Since the agent is not interactive or social besides posing a question for those in Group C, they stopped paying attention to it after the first few projects and shifted their attention elsewhere. One participant said, "Needing to pay attention to a virtual agent is harder than a physical human being. I suppose having someone physical that is conscious of my attention and focus makes me pay attention more."

Based on the results on memory retention, R1 can only be partially answered. There was no significant difference in people's ability to retain information naturally based on the types of embodiment they observed. It should be noted that this result is only necessarily true in a "static experience" like the one used in this experiment, where people are not moving in the environment and are only staring at their screen, looking at posters, and using an application to trigger the agents. Since participants did not fully interact with a disembodied or embodied agent, and did not have a tour guiding experience in an actual physical environment, it remains inconclusive on which method, embodied or disembodied, is better at helping people retain information. It does seem that having a more embodied, interactive agent could have enhanced the experience even more for some people. A few participants from Group A stated the following: "Could have been better if the avatar showed up talking or interacting with the product." "I definitely think a visual presence would have been helpful for me personally. Just the voice felt like a listening tour in a museum." Overall, H1 did not agree with the results.

6.4.2 Attention

Furthermore, participants were asked if they were able to pay attention to their tour guiding agent in a 5-point Likert Scale rating question (1-No attention paid, 5-Paid attention). The results showed that Group A (M=3.963, SD=0.720) participants were able to pay more attention to their disembodied agent than those in Group B (M=3.567, SD=0.774) and C (M=3.562, SD=0.948) who saw an embodied agent. This seems to agree with the result from the above question about memory retention comparing text and tour guiding agent. Overall, group A (voice only) reported a slightly higher attention towards their tour guiding agent.

6.4.3 Cognitive Dissonance

As mentioned by the participants, many attempted to read the poster while the agent was talking. The questions regarding cognitive dissonance showed that 62% of the participants suffered it, meaning they were not able to focus on the information from the audio while reading the text, since the text and audio from the agent did not match. Since participants noticed the mismatch, some may have compensated by reading the text several times. Although the audio gave a bit more information about the projects when compared to the text, since many questions were multiple choice and could be guessed correctly, a higher retention score could still be achieved solely based on the text description, because participants could read as many times as they wanted but could only listen to the audio once. This would help explain why the participants attributed their higher memory retention and attention to the text over the agent. The mean and standard deviation across the groups are M=0.500, SD=0.510 for Group A, M=0.600,SD=0.498 for Group B, and M=0.719, SD=0.457 for Group C with a mean of 1 being true for cognitive dissonance and a mean of 0 being false for cognitive dissonance. To answer R3, surprisingly, Group A showed the lowest amount of participants suffering from cognitive dissonance while Group B and C showed a higher level.

According to the responses on why they could not pay attention to their agent, cognitive dissonance seems to have played a major role. A few the participants stated that because the embodied agent was not doing much, they shifted their attention towards the text instead, leading to higher cognitive dissonance. Some participants in Group B and C had to make a conscious decision to ignore the embodied agent to read the text, which might have made them more aware of their cognitive dissonance later on in the experiment. One participant commented that, "there was quite a bit of content to remember overall, and my eyes kept wanting to look at text." Another participant said, "Having to hold my arm up to look at the dot [tour guiding agent] was distracting when I could have essentially the same experience just listening and reading the poster." Yet another participant stated, "The agent wouldn't always read from the text. While I listened to the agent, I also tried reading the text, which caused me to not give either the agent or text 100% of my attention."

Meanwhile, participants in Group A did not have an embodied agent and could only listen to the audio. This might have given these participants more time to adapt to new strategy for retaining information better. They might have realized that the text and the audio do not match right after the first project, which is much earlier than those in Group B and C since these participants tend to focus on the embodied agent movement. Additionally, participants in Group A could have just triggered the audio, put down their phone and focused their attention in the voice. This is illustrated clearly by one of the comments from a participant, "I think I got the version where no one showed up, which made me have to look away from the posters so that I didn't read while the voice talked. It's hard to listen to someone/something when there is similar text in screen, UNLESS there's a person to look at. I was aware of this so I was able to pay attention to the audio but it was not frictionless as I had to actively not look at the text."

6.4.4 Usability and Agent Embodiment

Despite some participants thought that the embodied agent was distracting, they still expressed preference for an embodied tour guiding agent. When asked if they would like to see a similar tour guiding agent in future tour, participants in Group C expressed the highest level of approval with a mean and standard deviation of M=0.875, SD=0.336, followed by Group A (M=0.769, SD=0.430), and finally Group B (M=0.767, SD=0.430). Additionally, when participants were asked to rate their preference between having a disembodied AI companion similar to Alexa, Cortana, or Siri (1pt) vs. having an embodied agent that they could see (5pt), most leaned towards having an embodied AI companion that will show up only when they need it to and will remain disembodied otherwise (M=3.670, SD=0.867). Group-wise, Group C showed the highest preference towards having an agent that "shows up only when I need it to in physical form" (M=3.812, SD=0.931), followed by Group B (M=3.600, SD=0.932), and Group A (M=3.577, SD=0.703).

6.4.5 Social Presence

Posing a question might have helped participants in Group C feel that the agent was more socially present. According to the responses to the Social Presence, 5-Points Likert Scale rating items shown in Table 6.14, Group C showed the highest mean among the three groups for all three items. The mean and standard deviation for item 1, "I would have liked the experience to continue," are M=3.966, SD=0.915 for the population, M=3.769, SD=1.032 for Group A, M=3.900, SD=0.923 for group B, and M=4.188, SD=0.780 for Group C. Item 2, "I felt REINA provided a memorable experience", resulted in M=3.773, SD=0.919 for the population, M=3.538, SD=0.989 for Group A, M=3.667, SD=0.884 for Group B, and M=4.062, SD=0.840 for Group

C. Finally for item 3, "I felt as though I was in the same space as REINA", the mean and standard deviation for the population is M=3.330, SD=1.111, M=3.308, SD=1.050 for Group A, M=3.133, SD=1.332 for Group B, and M=3.531, SD=0.915 for Group C. Although the result was not significant, it did seem to suggest that having a more social agent, even if it is just an agent that poses a question, could enhance the social presence of the agent. Thus, the answer to R2 is yes, it does make a subtle difference in people's ability to retain information naturally if the agent poses an open-ended question, but just not significant enough.

6.4.6 What make a memorable tour guiding experience?

Although memory retention is important for tour guiding and companionship, it is more essential for the agent to create memorable experiences in people's lives. Memorability is difficult to define and measure, but the results in this remote user study suggests that overall the embodied agent provided a more memorable experience than voice alone based on the social presence questionnaire and the comments.

In future work, it will be interesting to test a more social and interactive agent to see if it enhances the experience even more. This remote user study has provided some useful information on how to improve the future in-person user study further. A lot was gleaned from the participant's comments. For instance, one of the participants commented that:

> "REINA's voice and ways of addressing us (greetings) and explaining each project made her sound somewhat robotic and monotonous. For me, a more memorable experience would be obtained if the information provided by her would have been more casually introduced, rather than sound as someone/something who reads literally from a source; it sounded with lack of emotion. Maybe some inflections in her voice (and if possibly, she interacting with some features of the project) could make me feel more driven and eager towards her. However, it was good that the information she conveyed was different to the one provided in the texts, since it was

> an interesting info, it was memorable. The memorable trait came from the quality of the info rather than from REINA's voice/ways itself."

Additionally, many expressed that they would have liked to explore on their own pace and be able to repeat the audio freely.

Thus, in order to truly test R1, it will be more meaningful to test fully interactive and sociable disembodied and embodied agents. This means that the research question should be updated to the following: R4: Can a fully interactive and sociable agent affect people's ability to retain information naturally based on the kind of embodiment they possess or would a more interactive project visualization be enough to enhance memory retention? and R4.1: Will a disembodied, interactive, and sociable agent make a more memorable tour guiding experience than an embodied, interactive, and sociable agent in the physical environment?

Chapter 7

Future Work: Media Lab Tour Guiding Experience User Study (In-Person)

The effect of having an intelligent interface agent live and navigate in our physical environment has been explored before, but it has previously been done in a small room under a very controlled environment. Recent advances in augmented reality authoring tools have enabled reliable indoor navigation for intelligent interface agents, providing them with information of physical spaces and making them more dynamic and situation-aware. Therefore, with these new technologies, we can explore the spatio-temporal effect on agent embodiment, memory retention, social presence, attention and cognitive dissonance. Overall, this aims to provide a deeper understanding of what makes a memorable tour guiding experience.

The remote user study provided us with enough lessons to make a better tour guiding experience for an in-person user study. The embodied agent can be improved to be an embodied conversational agent (ECA), hence it will be more interactive and sociable, and will be able to answer some basic questions when prompted by the user. 3D interactive elements can be added to the projects in some conditions.

7.1 Research Questions and Hypotheses

The following research questions remain of interest to me:

Main Research Questions:

- **R4:** Can a fully interactive and sociable agent affect people's ability to retain information naturally based on the kind of embodiment they possess, or would a more interactive project visualization be enough to enhance memory retention?
- **R4.1:** Will a disembodied, interactive, and sociable agent makes a more memorable tour guiding experience than an embodied, interactive, and sociable agent in the physical environment?
- **R5:** Can the spatio-temporal effect enhance memory retention?
- **R5.1:** If so, which elements (disembodied agent, embodied agent, interactive 3D elements) enhance the spatio-temporal effect more for memory retention?

- **R5.2:** How will a multi-floor tour guiding experience affect memory retention, and social presence?

- **R6:** Will an embodied agent be perceived as more socially present if it presents 3D interactive elements and situational-awareness (e.g. space and time awareness)?

Hypotheses:

- **H4:** Based on the results from the remote user study, I hypothesize that having AR 3D interactive elements will enhance the memory retention since participants get to actively interact and play with the presented information, and having an embodied agent will enhance the memorability of the tour guiding experience. Thus, an embodied agent presenting AR 3D interactiveable elements should lead to a higher memory retention compared to a disembodied agent or 3D elements alone.

- **H4.1:** I hypothesize that an embodied agent will be able to deliver a more memorable tour guiding experience in the physical environment than the disembodied version because it is easier to follow an embodied agent around than to follow a disembodied agent's instruction on where to go.

- **H5:** As explained in Chapter 2, the method of loci has been used by people as a way to retain mass amounts of information by mapping them to a physical space in their mind. Thus, I hypothesize that there should be a spatio-temporal effect on memory retention that could yield higher scores than those seen in the remote user study. Essentially, this creates an augmented reality Memory Palace in the physical environment, even when it is their first time visiting the environment. Since information is presented in different space and time, users have different visual cues to trigger memory retrieval.

- **H5.1:** In this experiment, I hypothesize that condition D (described below) might yield the highest effect, while condition 0 will stay fairly low compared to the other conditions. Since we naturally see things such as posters in different space and time, the spatio-temporal effect applies naturally. However, this effect might be unconscious and weak compared to having an embodied agent guiding you in the physical environment and showing you additional visual and interactive media aid for priming information in an unknown environment.

- **H5.2:** I hypothesize that a multi-floor tour guiding experience will enhance memory retention and social presence of the agent because it is the most natural and human experience. The objective measures will be able to tell more about how people feel about having a tour guiding agent that can accompany you anywhere in the building.

- **H6:** I hypothesize that an embodied agent will be perceived as socially present if it can present 3D interactive elements and situational-awareness. Condition D with the social ECA and interactive elements will be perceived as more socially present than the ECA alone (Condition C). The reason for this is that by interacting with the 3D elements presented by the ECA, the users might perceive themselves as being more engaged with the agent.

7.2 Tour Guiding Agent and Interactivity

The tour guiding agent will be upgraded to exhibit basic natural language processing. Some of the prompts that it will be able to process are:

- How are you doing?
- What is my user number?
- Can you take me to the first experiment? / Can we start the experiment?
- Could you repeat the project? / Could you explain the project again?
- Take me to the next project.
- Where are we?

- How many projects are left to see?
- Can we end this experiment?
- Can I drop out of this experiment now?
- How do I play with this interactive element?
- Could you repeat the instruction?
- How long do I have to explore? / Time left

This tour guiding agent will essentially become a social embodied conversational agent (Social ECA) able to talk with the participants when asked.

7.3 Experiment design

The main idea of this experiment is to analyze the spatio-temporal effect on memory retention, and thus identify elements that can create a more memorable tour experience. A user study with between-subject design for comparing spatio-temporal presentation of information in an unfamiliar physical environment will be conducted. The subjects will be randomly assigned into five groups (0, A, B, C, and D).

- **Group 0**: will walk in the physical environment but only one hallway, explore and listen to an audio explaining each of the five project shown in the hallway on their own pace. This is analogous to the audio guides that are commonplace in modern museums. (Control: Physical posters + audio)
- **Group A**: will walk in the multi-floor physical environment, find the 5 posters, explore and listen to an audio explaining the project on their own pace. This is analogous to the audio guides that are commonplace in modern museums. (Control: Physical posters + audio)
- **Group B**: will walk in the multi-floor physical environment, find the 5 posters, press virtual icons and buttons to start the audio and to play with embedded

Augmented reality 3D interactive elements for each poster. (Physical poster + audio + 3D interactive elements)

- **Group C**: will engage with an embodied conversational agent who will present the information with no visual enhancement other than the physical poster present in the multi-floor environment. (Physical poster + Social ECA)

- **Group D**: will engage with an embodied conversational agent who will present the information with 3D interactive elements in the multi-floor environment. (Physical posters + Social ECA + 3D interactive elements)

In all the five conditions, the experiment should take less than one hour to complete. Also, in all five conditions, we will have the same pre-recorded audio and physical posters placement in the environment. All participants will be able to freely tour the Media Lab and interact with the agent and posters as they will with the only condition being that they should finish the tour in 30 min.

7.4 Evaluation

Subjective and objective measures will be implemented in this user study to obtain more information about what the participants perceived and sensed. For the surveys, the same or a modification of the same questionnaires seen in the remote user study will be used here.

7.4.1 Social Presence Questionnaire

This questionnaire is a modification from [70]. The questionnaire will implement a Direct Rating Method using a five-point Likert scale.

1. I felt as though I was in the same space as REINA.

2. I would have liked the experience to continue.

3. I felt REINA responded to my questions.

4. I felt REINA provided a memorable experience

7.4.2 Memory Retention Questionnaire

The participants will respond to a multiple choice quiz at the end of the experiment on the five projects they saw to assess memory retention. The same quiz will be sent out 24 hours after, 72 hours after, and 7 days after. Some of the questions are a modification from the ones in the remote user study and are as follows:

- What is the first project you saw with REINA.
- What is this project about?
- What is the second project you saw with REINA.
- What are all the features you remember from this specific project?

7.4.3 Objective measures

Objective measures will be used to help better understand the participants experience through out the experiment. Participants will wear devices to monitor and record their biosignals, such as electrodermal activity (EDA), gaze and heart-rate variability (HRV) for analyzing arousal to estimate engagement, attention, and objective social presence.

7.5 Plans and Developments

As of now, natural language processing features need to be added into REINA to make it more sociable. Another item that needs to be completed is to test REINA while walking in the physical Media Lab building. In the fall or the spring, when quarantine and MIT restrictions are lifted, we can start running a small in-person pilot study to refine the experience.

Chapter 8

Conclusion

To conclude, in this thesis, I have introduced the concept of a Pervasive Interface Agent, a type of cross-platform agent-based interface that acts as our life companion helping us delegate and perform tasks in both physical and digital worlds. Additionally, I have presented R.E.I.N.A., a phase 1 prototype of a Pervasive Interface Agent and have showed the implementation of the Media Lab Tour Guiding system developed to deploy and showcase R.E.I.N.A. navigation in the physical world.

Finally, two users studies were conducted remotely and the results were discussed. The lessons learned from these user studies were then used as guidelines for improving the MIT Media Lab Tour Guiding System, the experiment design of the in-person user study, and phase 2 of R.E.I.N.A.

As augmented reality and virtual reality become a ubiquity in this fast changing time, I hope this work can contribute towards having cross-platform agent-based interfaces, which I have called Pervasive Interface Agents, as our lifelong companion that can transcend the physical-digital worlds with us.

Appendix A

Tutorial: Building a Unity NavMesh Project

1. Create a new Unity project on Unity2019.3.
2. Install Vuforia Engine from Package Manager, then
3. add the upgraded Vuforia Engine 9.2.7 from Vuforia download site.
4. Delete the Main Camera and add ARCamera from GameObject/Vuforia Engine/ARCamera.
5. Add Area Target Component to the scene by going to GameObject/VuforiaEngine/Area Target.
6. Add Vuforia Developer License.
7. Import your area target unity package.
8. Place the PreviewModel of the area target into the scene, you can find it at Assets/Editor/Vuforia/<YourAreaTargetName>.
9. Add a Mesh Renderer component to your Preview Model and enable Navigation Static.
10. Now go to Window/AI/Navigation/Bake.

11. Setup the agent size and bake.

The step by step is also shown in the Figures below.

MoveTo Script

```
    using System.Collections;
using System.Collections.Generic;
using UnityEngine;
using UnityEngine.AI;

public class MoveTo : MonoBehaviour
{
    public Transform goal;
    // Start is called before the first frame update
    void Start()
    {
        NavMeshAgent agent = GetComponent<NavMeshAgent>();
        agent.destination = goal.position;
    }
}
```

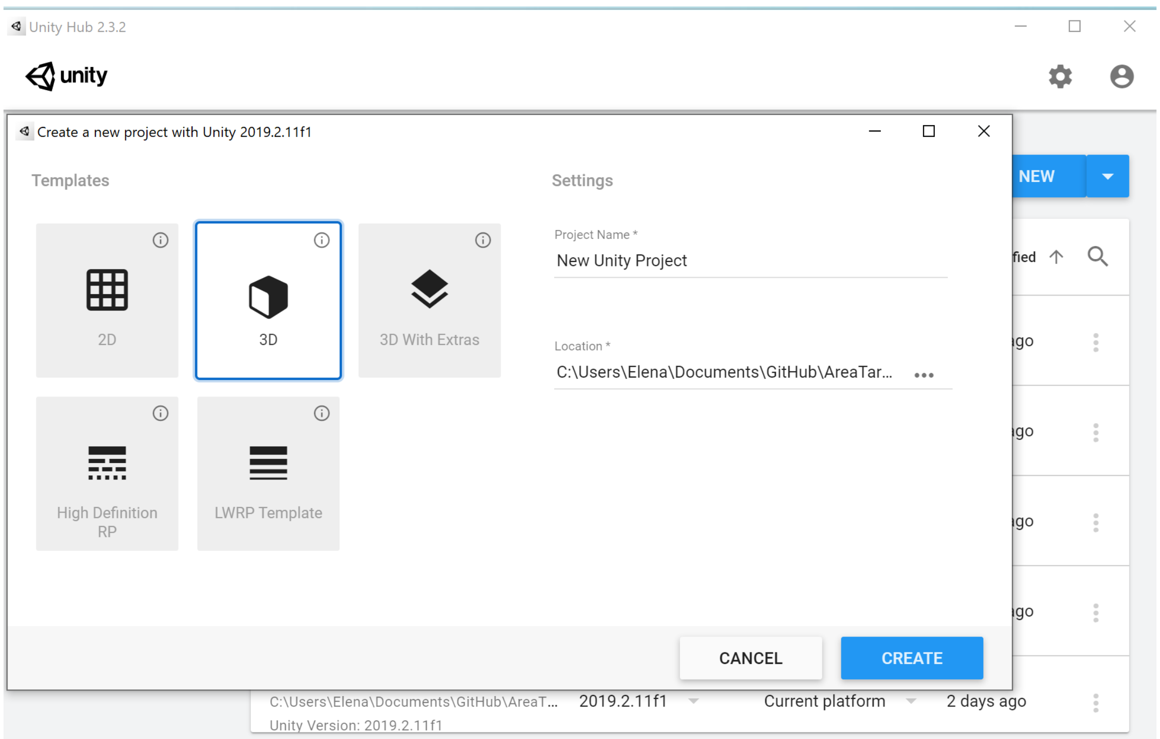


Figure A-1: Create a new unity3D project.

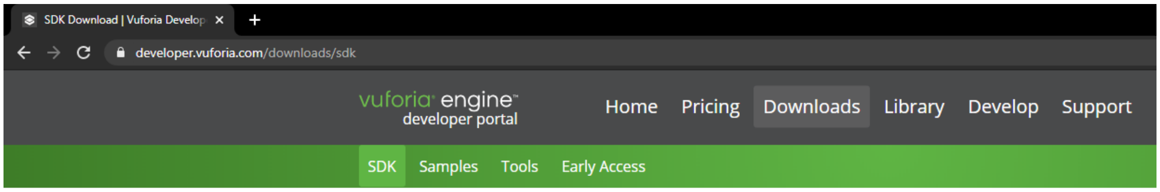


Figure A-2: Download Vuforia Engive 9.3 SDK or the latest one.

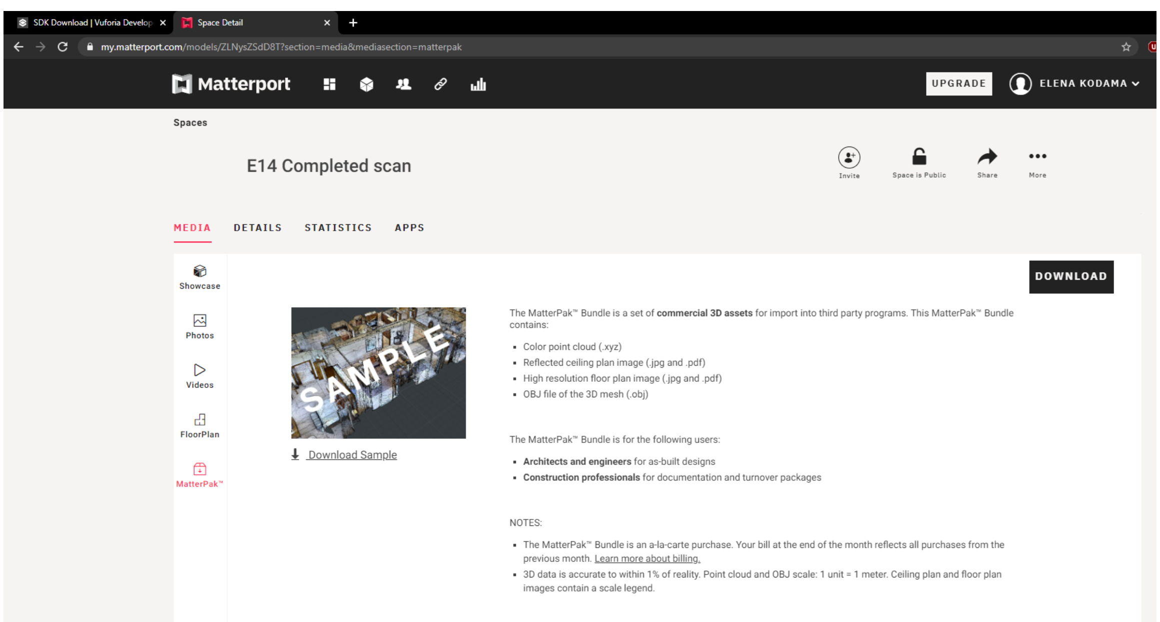


Figure A-3: Download the Matterpark Bundle from Matterport Professional license account.

Figure A-4: Follow instruction in Chapter 5 on how to generate Area Target.

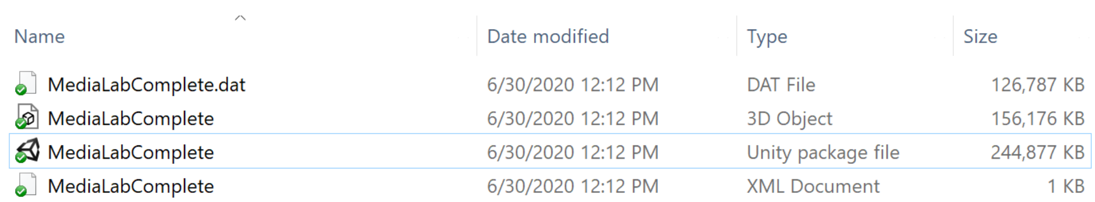


Figure A-5: The Vuforia Area Target Generator should give four files, one of them is a unity package.

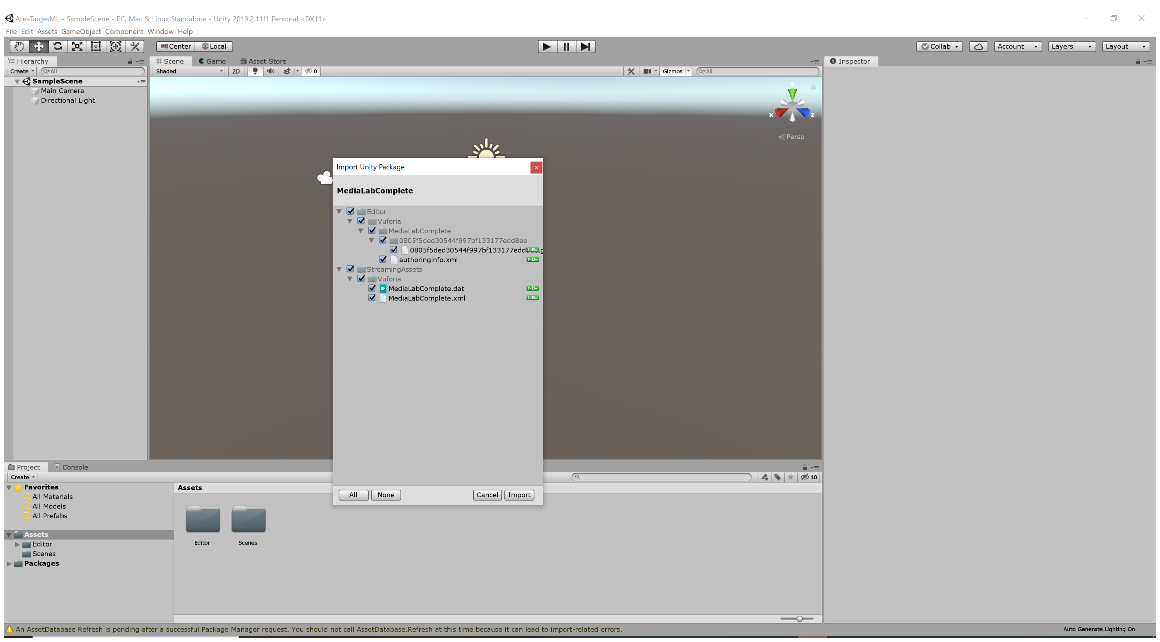


Figure A-6: Import the Area Target Unity package into the Unity project.

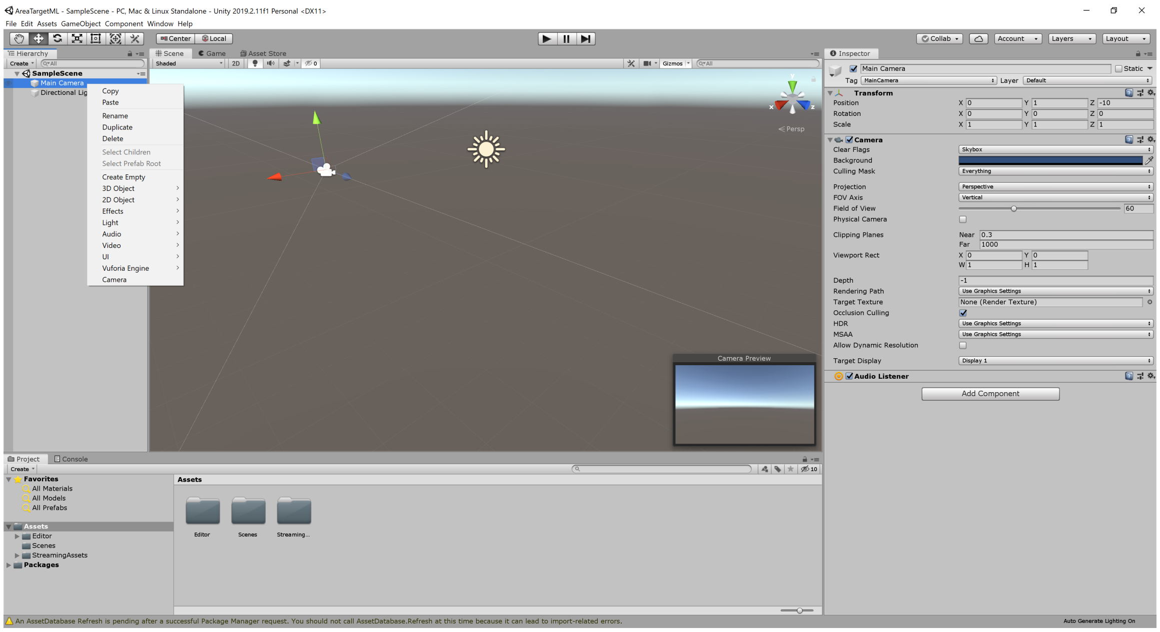


Figure A-7: Delete the Camera on the sample scene.

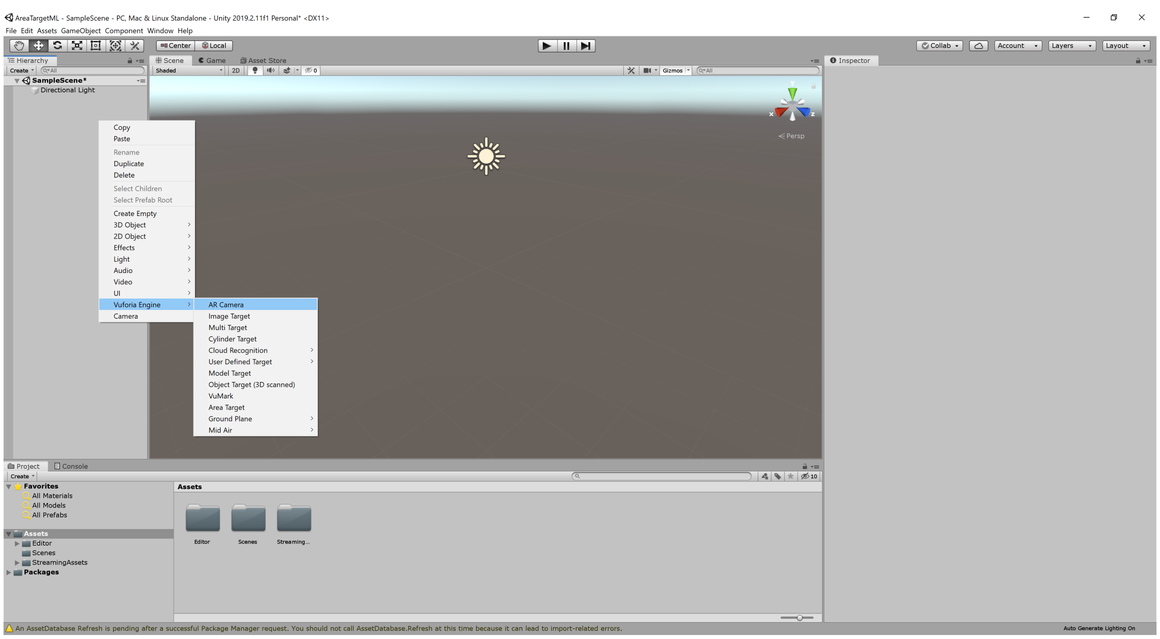


Figure A-8: Import the Vuforia Engine 9.3 SDK if it is not already there. It could be available in the unity package manager too. Add the Vuforia AR camera onto the sample scene.

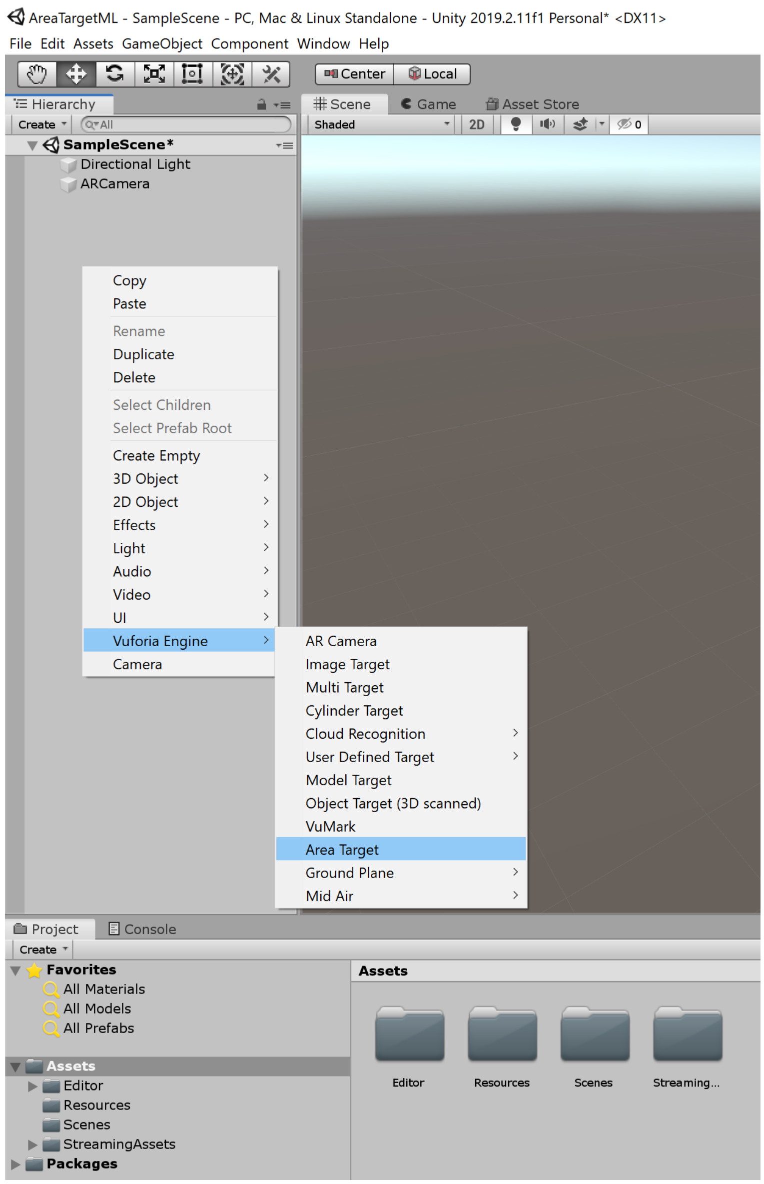


Figure A-9: Add the Area Target object from the Vuforia Engine into the SampleScene.

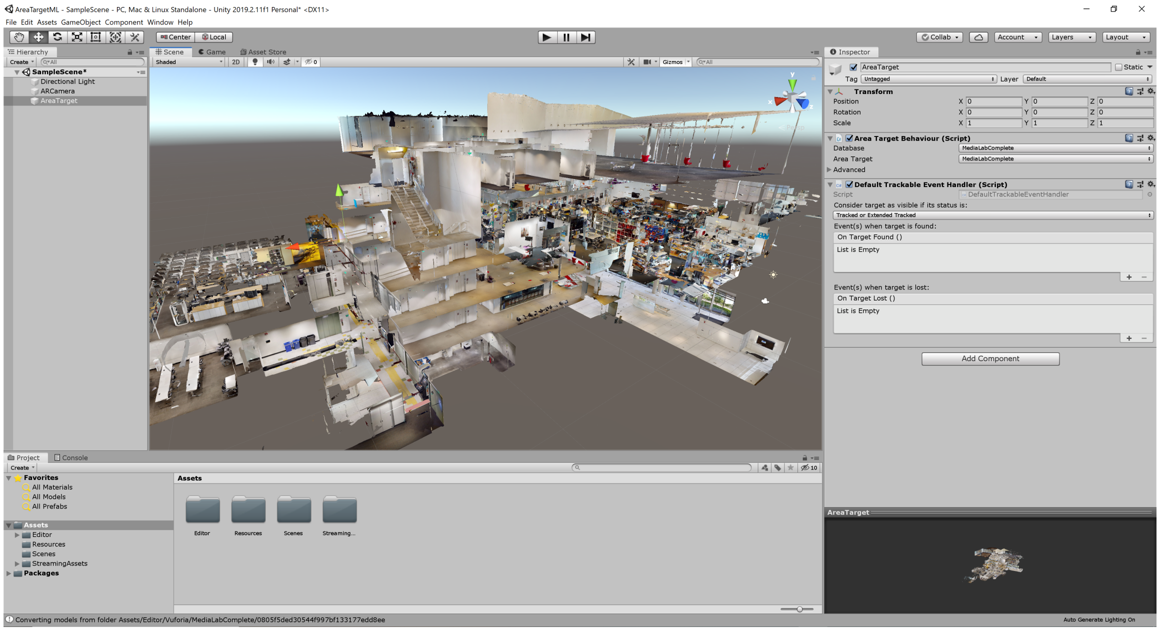


Figure A-10: Find the Area Target of the 3D scan located under the Vuforia folder.

Figure A-11: Create a Sphere or import an embodied character into the scene.

Figure A-12: Add the Unity NavMesh Agent component into the Sphere or the embodied character.

Figure A-13: Play with the agent settings to match the desired dimensions.

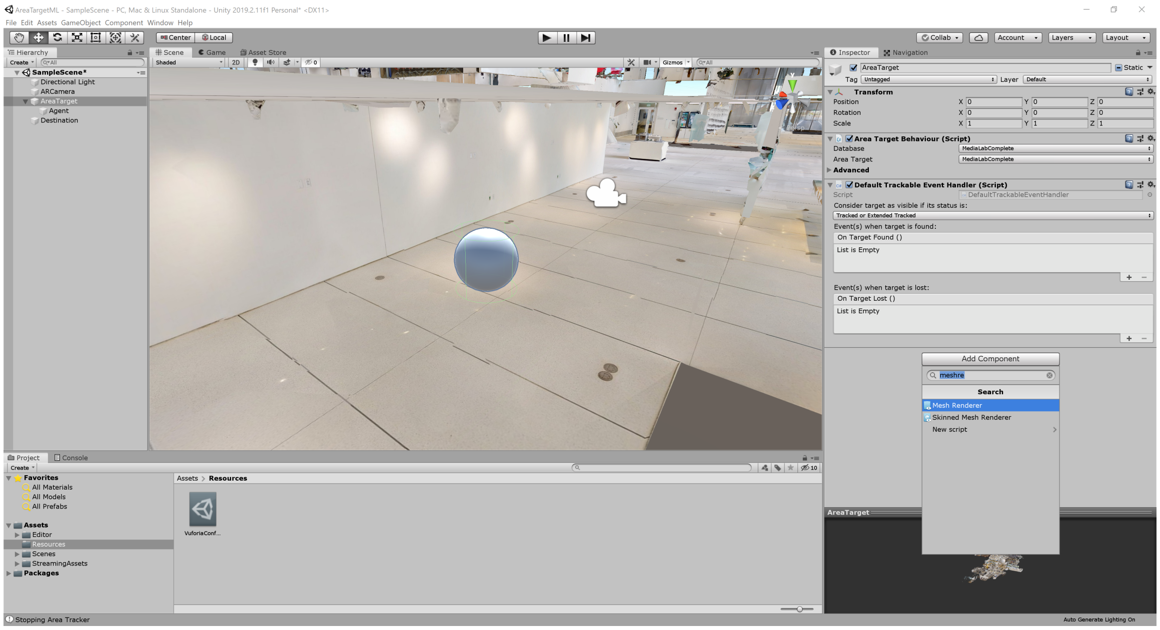


Figure A-14: Add a Mesh Renderer to the Area Target object.

Figure A-15: Enable Navigation Static under the navigation tab next to the Inspector Tab.

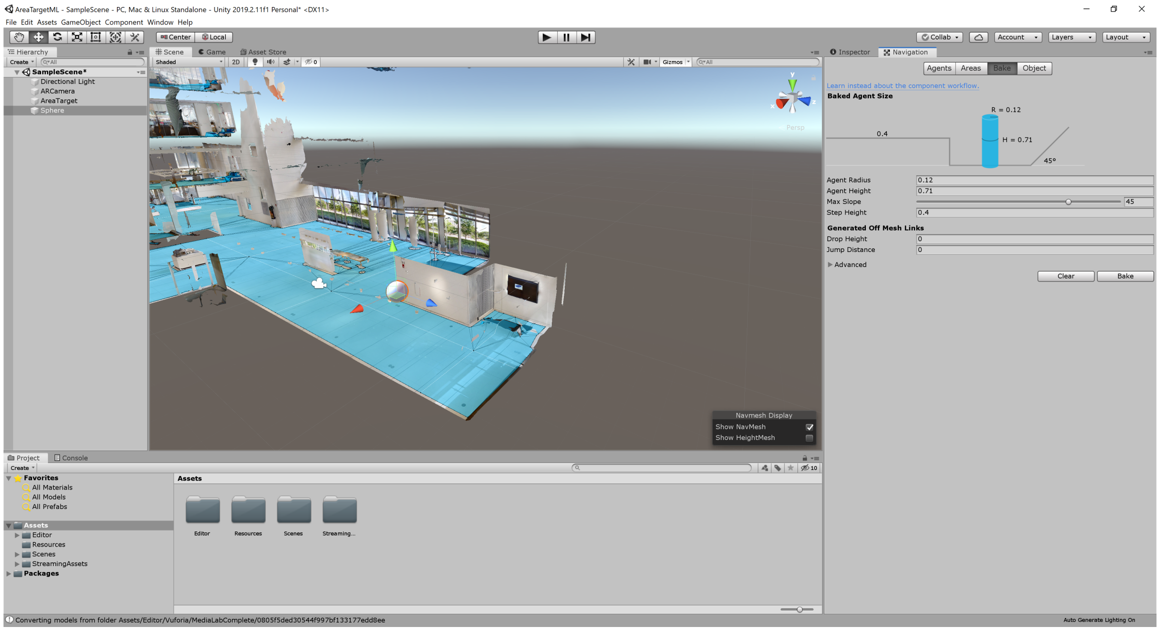


Figure A-16: Bake to obtain navigation mesh of the area target. Play around with the agent settings to generate more appropriate NavMesh surfaces.

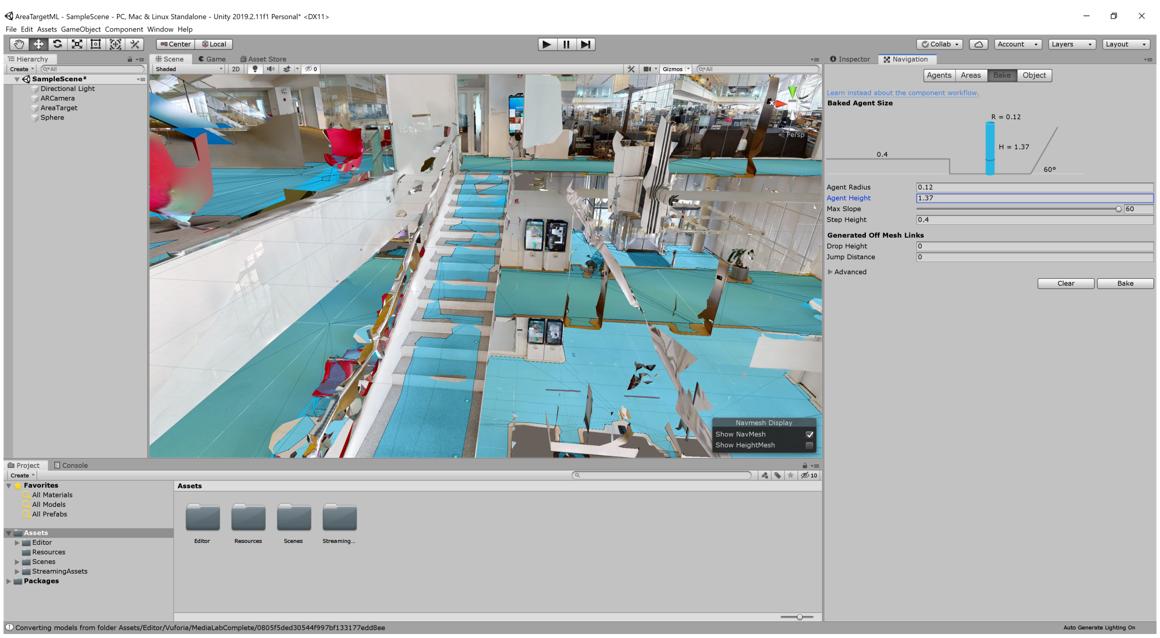


Figure A-17: Adjust the agent height setting to get more walkable area.

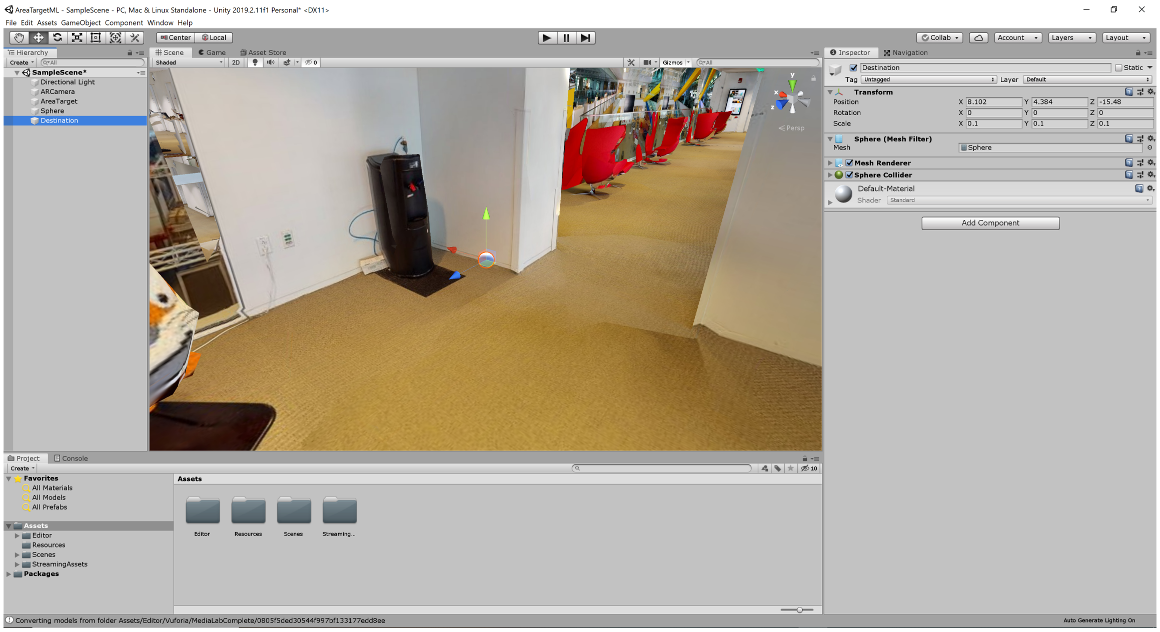


Figure A-18: Create a destination sphere by placing a sphere in the desired destination. This sphere can be made invisible, only need to know the position of it, so the NavMesh agent can navigate towards it.

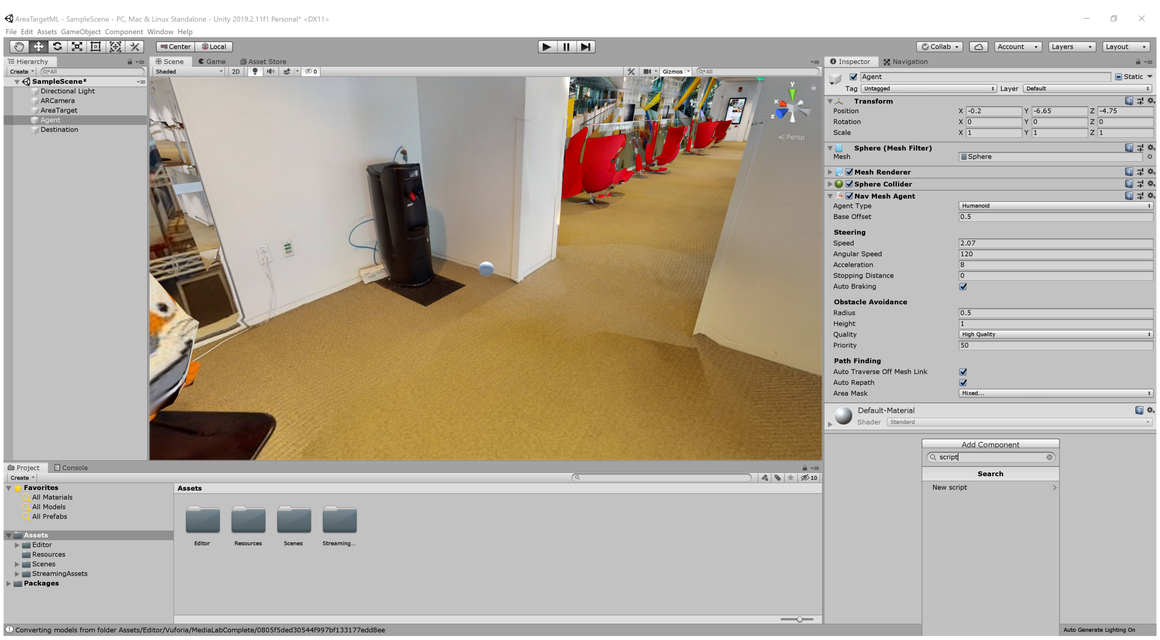


Figure A-19: Create a new script and add the MoveTo code.

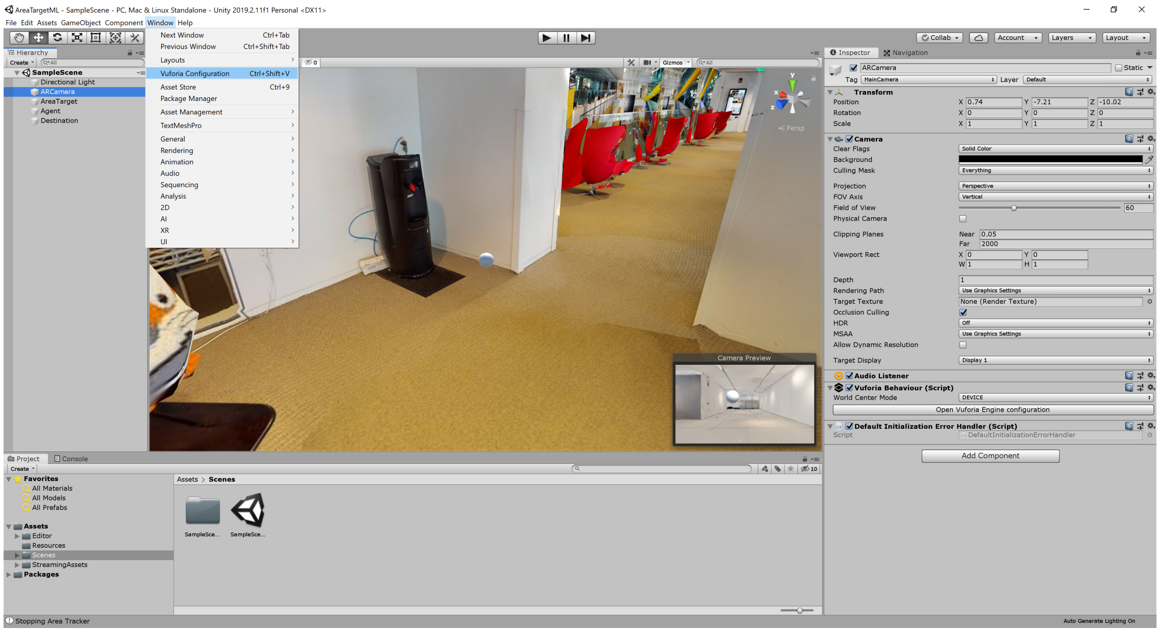


Figure A-20: Open Vuforia Configuration.

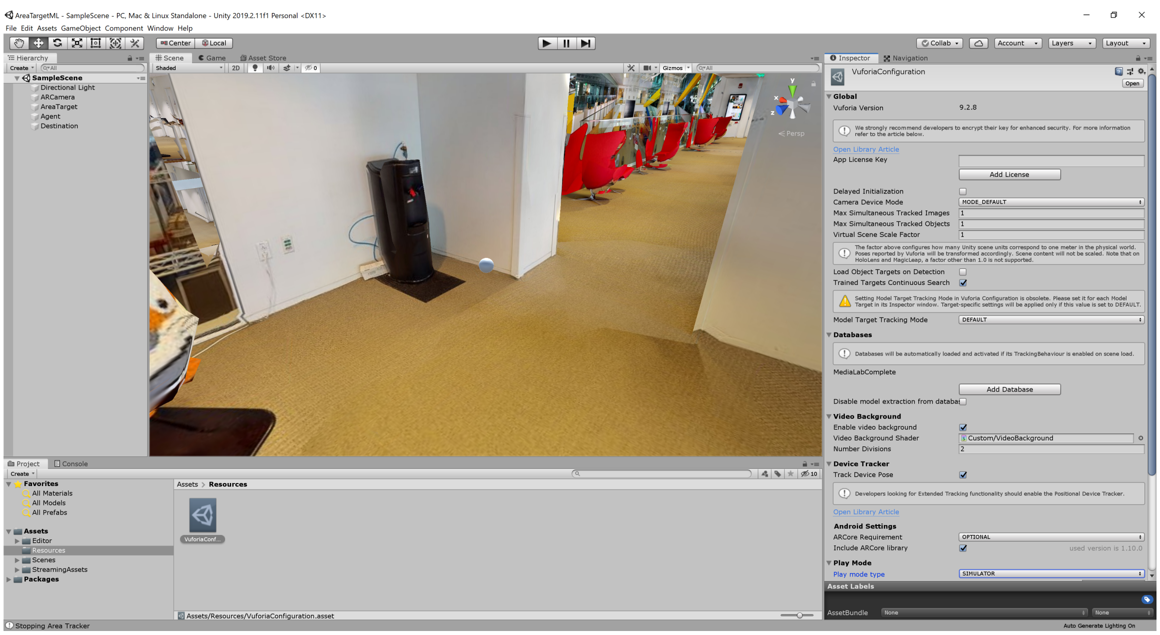


Figure A-21: Add Vuforia License and set play mode to simulator.

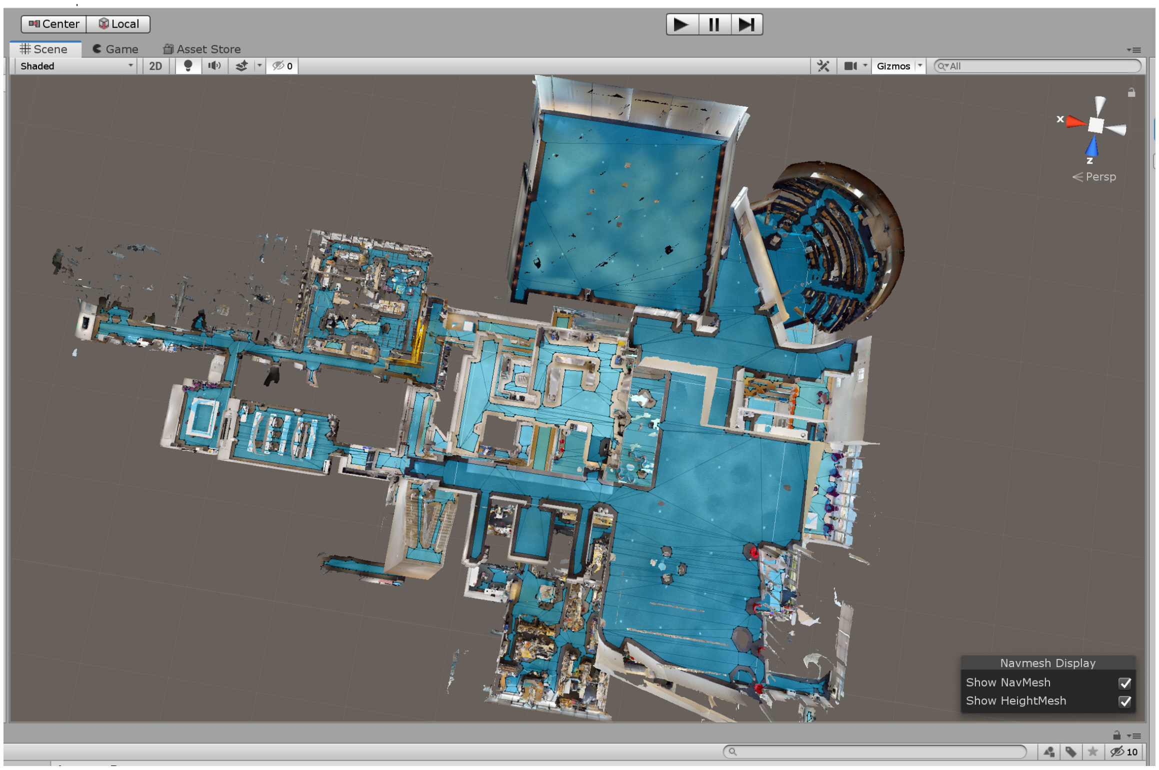


Figure A-22: All the blue area on the Area Target are the surface that the NavMesh agent can walk on.

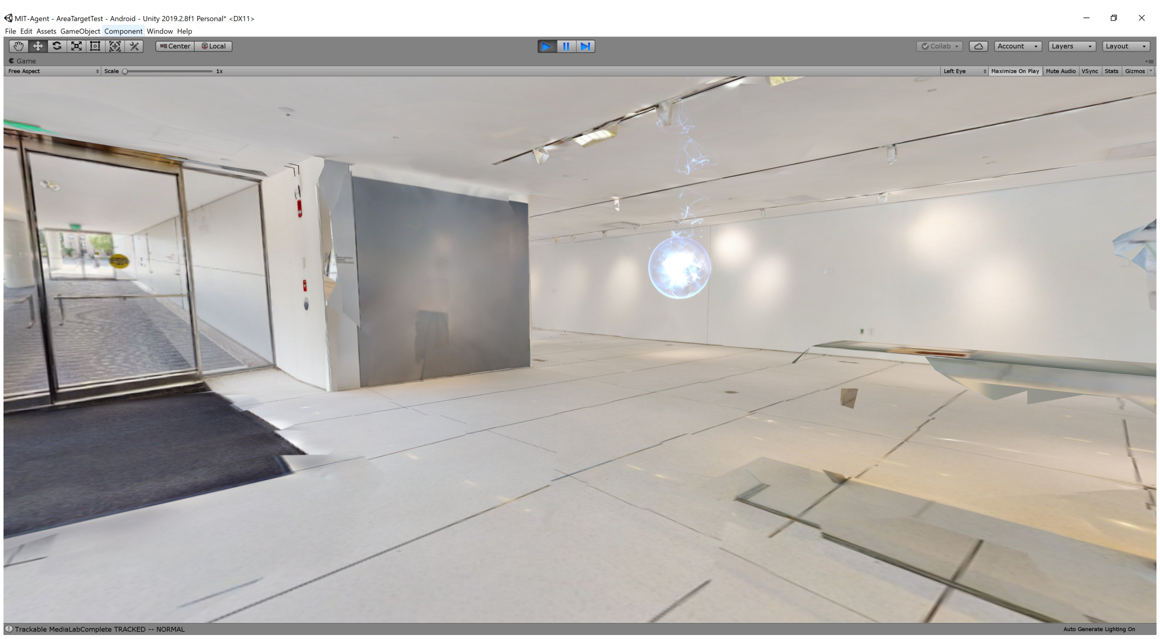


Figure A-23: Example of R.E.I.N.A. in the Media Lab Area Target on simulation mode.

Appendix B

Gallery Tour Posters

Posters for fictional study projects to be encountered during the tour.

Figure B-1: One of the poster shown during the remote gallery tour.

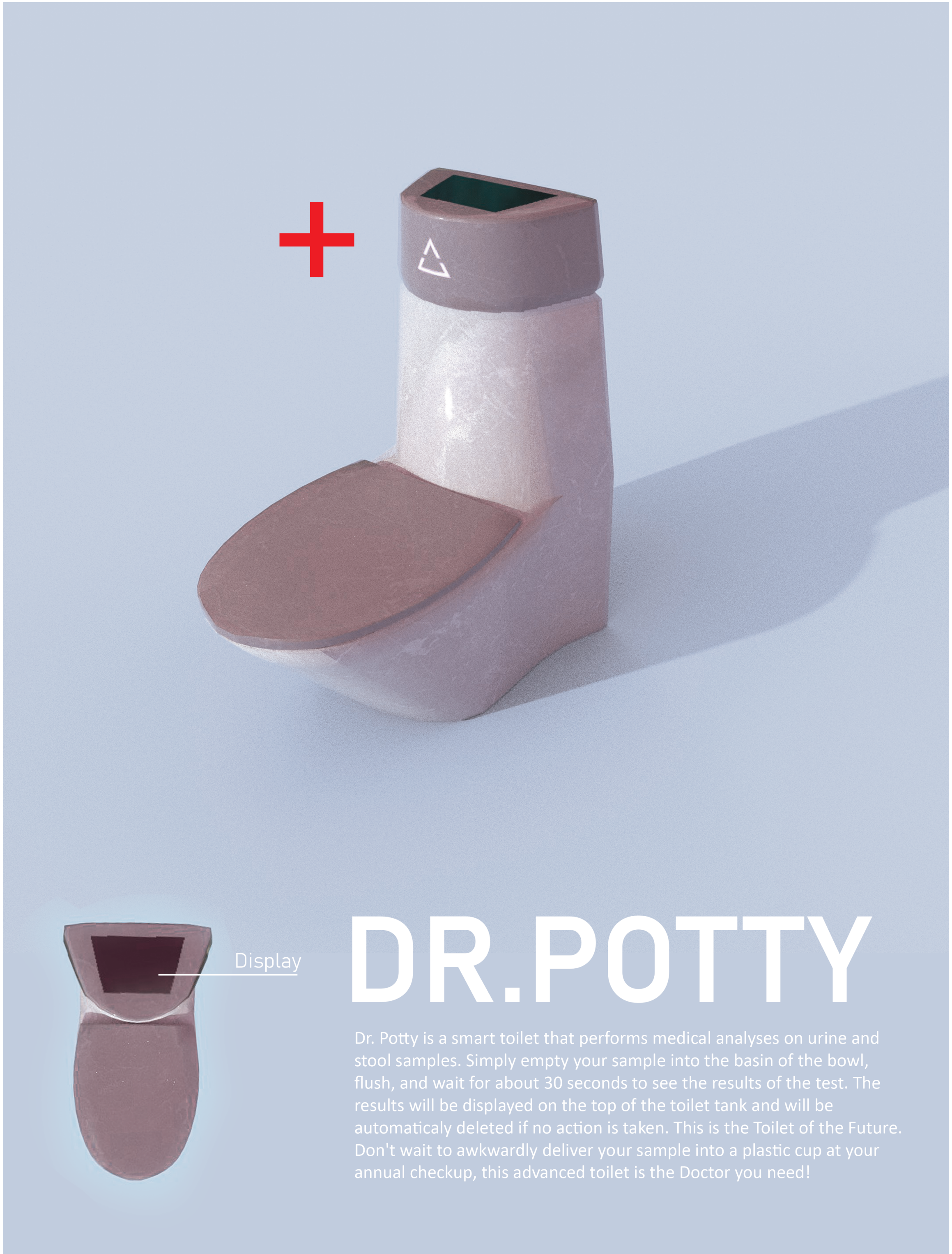


Figure B-2: One of the poster shown during the remote gallery tour.

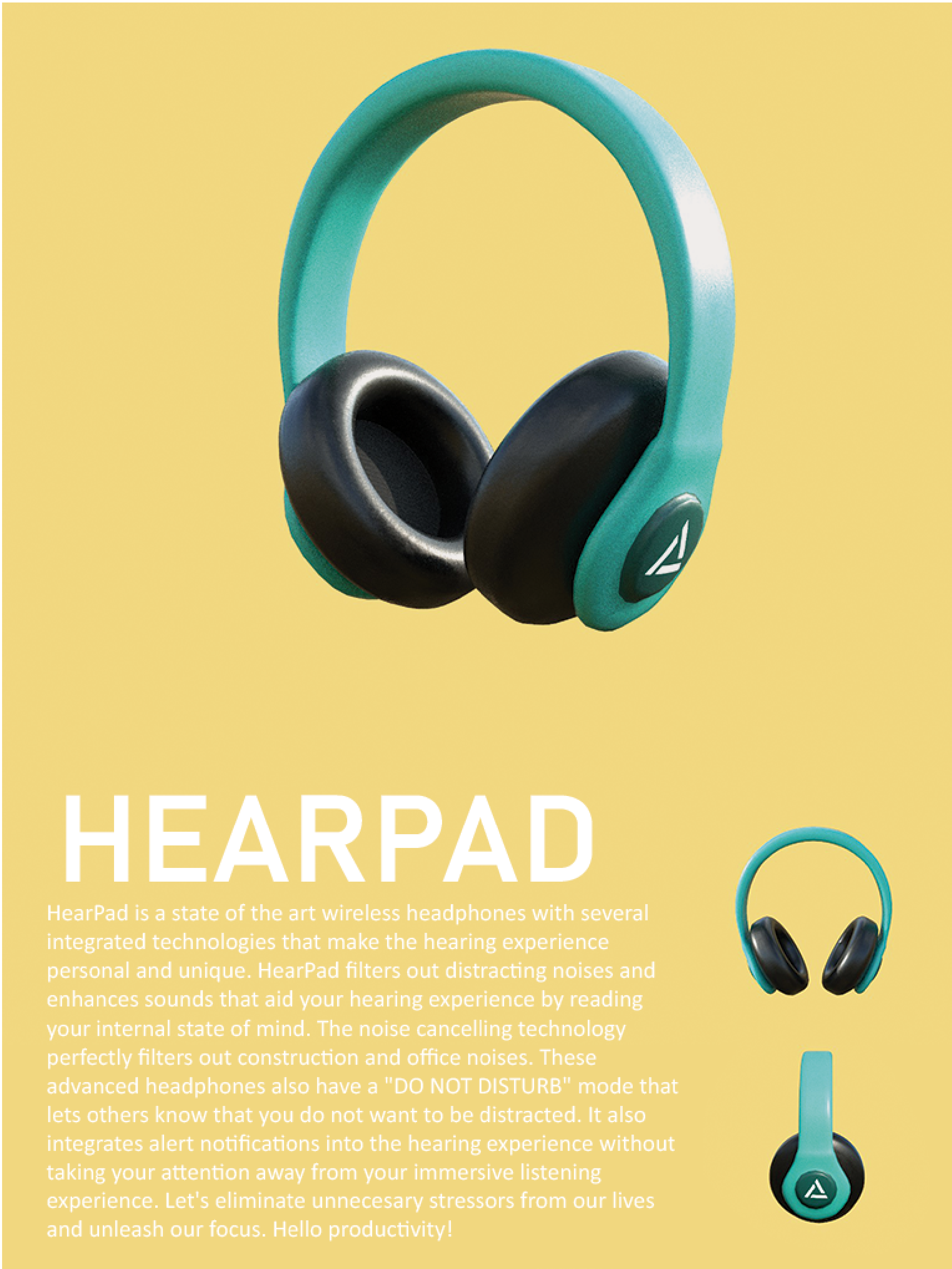


Figure B-3: One of the poster shown during the remote gallery tour.

Figure B-4: One of the poster shown during the remote gallery tour.

Figure B-5: One of the poster shown during the remote gallery tour.

Appendix C

Jupyter Notebook Data Analysis

```
In [ ]: import pandas as pd
        import matplotlib.pyplot as plt
```

```
In [ ]: # Import raw data
        d= pd.read_csv("./CommentsOnly.csv", index_col="Subject")
```

```
In [ ]: d
```

```
In [ ]: ### Drop the bad subjects from the raw dataframe
        d.drop(["S11", "S12","S18", "S23", "S25", "S26", "S27","S37", "S40","S48","S5
        0","S36","S108"], inplace=True)
```

```
In [ ]: test=d[d['Group']=='A'][d["Any additional thoughts?"].notnull()]
        test["Any additional thoughts?"]
```

```
In [ ]: def comment (group,comment):
            dn=d[d['Group']==group][d[comment].notnull()]
            [print(line) for line in (dn['Group'] + " & " + dn[comment] + '\\\\\\hlin
        e')]
```

```
In [ ]: comment("A", "Any additional thoughts?");
        comment("B", "Any additional thoughts?");
        comment("C", "Any additional thoughts?");
```

```
In [ ]: comment("B", "Did you see the guiding agent? If yes, please describe what you
         saw and what it resembles. If you did not see anything and just hear voice, w
        rite N/A.");
        comment("C", "Did you see the guiding agent? If yes, please describe what you
         saw and what it resembles. If you did not see anything and just hear voice, w
        rite N/A.");
```

```
In [ ]: comment("A", "Describe why you couldnt pay attention")
        comment("B", "Describe why you couldnt pay attention")
        comment("C", "Describe why you couldnt pay attention")
```

```
In [ ]: comment("A", "Were you able to understand the tour guiding agent? [Yes, Not al
        ways, No] If so, Which project did you struggle with?")
        comment("B", "Were you able to understand the tour guiding agent? [Yes, Not al
        ways, No] If so, Which project did you struggle with?")
        comment("C", "Were you able to understand the tour guiding agent? [Yes, Not al
        ways, No] If so, Which project did you struggle with?")
```

```
In [ ]:
```

```
In [ ]:
```


Figure C-1: Jupyter Noteboook with code for printing the comments

In [ ]:
```
import pandas as pd
import matplotlib.pyplot as plt

#TODO - do data analaysis filtered by:
#       filter by learner type
#       filter by country
#       filter by gender
#       filter by age
#       filter by used VR before
#       filter by English Fluency
#       filter by Project total
#       filter by Day
#       filter by Total Memory Retention
# MEAN of answers immediate, 24, and 72 in a chart
# Rate of memory decline between the 3 groups
# Comparison of rate of memory decline between
```

In [ ]:
```
# Import raw data
# d= pd.read_csv("./surveyraw.csv", index_col="Subject")
d= pd.read_csv("./PandasReady.csv", index_col="Subject")
```

In [ ]:
```
# Import a test csv
dc = pd.read_csv("./totalRetentionGroups.csv")
```

In [ ]:
```
### Drop the bad subjects from the raw dataframe
d.drop(["S11", "S12","S18", "S23", "S25", "S26", "S27","S37", "S40","S48","S5
0","S36","S108"], inplace=True)
```

In [ ]:
```
# Create a new dataframe and store all info by Group
dA = d[d["Group"]=="A"];
dB = d[d["Group"]=="B"];
dC = d[d["Group"]=="C"];
```

In [ ]:
```
# d
```

In [ ]:
```
project = lambda name: pd.DataFrame({"Group A": dA[name],"Group B": dB[name],
"Group C": dC[name]})
```

In [ ]:
```
project2 = lambda name,filt,filtval: pd.DataFrame({"Group A": d[name][(d["Grou
p"]=="A") & (d[filt]==filtval)],
                                            "Group B": d[name][(d["Group"]=="B"
) & (d[filt]==filtval)],
                                            "Group C": d[name][(d["Group"]=="C"
) & (d[filt]==filtval)]})
```


Figure C-2: Jupyter Noteboook with code for Data Analysis

In [ ]:
```
#### BOXPLOT FILTERED DATA
plt.figure(num=None, figsize=(12, 12))
plt.title("Boxplot of Total Memory Rentention")
plt.subplot(321)
project2('Total Memory Retention','Gender','Female').boxplot(column=['Group A'
, 'Group B', 'Group C'], grid=False)
plt.title("Boxplot of Total Memory Rentention \nFiltered by Gender = Female")
plt.subplot(322)
project2('Total Memory Retention','Gender','Male').boxplot(column=['Group A',
'Group B', 'Group C'], grid=False)
plt.title("Boxplot of Total Memory Rentention \nFiltered by Gender = Male")
plt.subplot(323)
project2('Total Memory Retention','English Fluency','Yes').boxplot(column=['Gr
oup A', 'Group B', 'Group C'], grid=False)
plt.title("Filtered by English Fluency = Yes")
plt.subplot(324)
project2('Total Memory Retention','Learner Type',4).boxplot(column=['Group A',
'Group B', 'Group C'], grid=False)
plt.title("Filtered by Learner Type = Mostly Visual")

plt.subplot(325)
project2('Total Memory Retention','Country','USA').boxplot(column=['Group A',
'Group B', 'Group C'], grid=False)
plt.title("Filtered by Country = USA")

plt.subplot(326)
project2('Total Memory Retention','Country','PA').boxplot(column=['Group A',
'Group B', 'Group C'], grid=False)
plt.title("Filtered by Country = Panama")
```


Figure C-3: Jupyter Noteboook with code for Data Analysis

In [ ]:
```
##### BOXPLOT PROJECTS

plt.figure(num=None, figsize=(12, 12))
plt.subplot(321)
project('HiGloves S1+S2+S3').boxplot(column=['Group A', 'Group B', 'Group C'],
grid=False)
plt.title("Boxplot of HiGloves S1+S2+S3")
plt.subplot(322)
project('BikePack S1+S2+S3').boxplot(column=['Group A', 'Group B', 'Group C'],
grid=False)
plt.title("Boxplot of BikePack S1+S2+S3")
plt.subplot(323)
project('Dr.Potty S1+S2+S3').boxplot(column=['Group A', 'Group B', 'Group C'],
grid=False)
plt.title("Boxplot of Dr.Potty S1+S2+S3")
plt.subplot(324)
project('HearPad S1+S2+S3').boxplot(column=['Group A', 'Group B', 'Group C'],
grid=False)
plt.title("Boxplot of Hearpad S1+S2+S3")
plt.subplot(325)
project('Brella S1+S2+S3').boxplot(column=['Group A', 'Group B', 'Group C'], g
rid=False)
plt.title("Boxplot of Brella S1+S2+S3")
plt.subplot(326)
project('Total Memory Retention').boxplot(column=['Group A', 'Group B', 'Group
C'], grid=False)
plt.title("Boxplot of Total Memory Retention (S1+S2+S3)")
```

In [ ]:
```
### BOXPLOT TOTAL CORRECT ANSWER OF EACH SURVEY
d.boxplot(by="Group", column=['TOTAL CORRECT S1', 'TOTAL CORRECT S2', 'TOTAL C
ORRECT S3'], grid=True, figsize=(10,8))
```

In [ ]:
```
# Testing for normal distribution
my_data=project('Total Memory Retention').hist()
```

In [ ]:
```
### CALCULATE MEAN AND SD AND FORMAT DATA FOR LATEX
pd.set_option("display.max_columns", 100)
d.groupby("Group").mean();
```

In [ ]:
```
d.groupby("Group").std();
```

In [ ]:
```
#### CORRECT ANSWER PER DAY means and sd
d.groupby("Group").mean()[['TOTAL CORRECT S1','TOTAL CORRECT S2',
                         'TOTAL CORRECT S3']].plot.bar(figsize=(8, 5), grid=Tr
ue)
plt.legend(bbox_to_anchor=(1.35, 1), loc='upper right', ncol=1)
```

In [ ]:
```
means = d.groupby("Group").mean().applymap(lambda n: 'M=%3.3f' % n)
stds = d.groupby("Group").std().applymap(lambda n: 'SD=%3.3f' % n)
```

In [ ]:
```
means+', '+ stds
```


Figure C-4: Jupyter Noteboook with code for Data Analysis

In [ ]:
```
test=(means+', '+stds).apply(lambda row: ' & '.join(row.values),axis=0)
test[0]
```

In [ ]:
```
test2= ( (test.index) + ' & ' + test + ' \\\\\\hline' )
```

In [ ]:
```
### FORMATTING FOR LATEX
[print(line) for line in test2];
```

In [ ]:
```
######## MORE STATS ANALYSIS FOR NORMAL DISTRIBUTION

import statsmodels.api as sm
from scipy.stats import norm
import pylab
### QQPLOT
# https://towardsdatascience.com/6-ways-to-test-for-a-normal-distribution-whic
h-one-to-use-9dcf47d8fa93
sm.qqplot(d["Total Memory Retention"], line='45')
pylab.show()
```

In [ ]:
```
from scipy.stats import kstest, norm
ks_statistic, p_value = kstest(dA["TOTAL CORRECT S3"], 'norm')
print(ks_statistic, p_value)
```

In [ ]:
```
# SHAPIRO-WILK TEST TO CHECK FOR NORMAL DISTRIBUTION. P>0.05 = NORMAL DISTRIBU
TION
import scipy.stats as stats
shapiro_test = stats.shapiro(d["Total Memory Retention"])
shapiro_test
```

In [ ]:
```
# BARTLETT'S TEST TO CHECK FOR HOMOGENEITY OF VARIANCES
import scipy.stats as stats
w,pvalue=stats.bartlett(dA["Total Memory Retention"],dB["Total Memory Retentio
n"],dC["Total Memory Retention"])
print( 'w=%3.3f, p-value=%3.3f' % (w,pvalue))
```


Figure C-5: Jupyter Noteboook with code for Data Analysis

```
In [ ]: ## One way ANOVA on Total Memory Retention.
fvalue,pvalue =stats.f_oneway(dA["Total Memory Retention"],dB["Total Memory Re
tention"],dC["Total Memory Retention"])
print( 'F=%3.3f, p-value=%3.3f' % (fvalue,pvalue))

## One way ANOVA on Total Survey 1.
fvalue,pvalue =stats.f_oneway(dA["TOTAL CORRECT S1"],dB["TOTAL CORRECT S1"],dC
["TOTAL CORRECT S1"])
print( 'F=%3.3f, p-value=%3.3f' % (fvalue,pvalue))

## One way ANOVA on Total Survey 2.
fvalue,pvalue =stats.f_oneway(dA["TOTAL CORRECT S2"],dB["TOTAL CORRECT S2"],dC
["TOTAL CORRECT S2"])
print( 'F=%3.3f, p-value=%3.3f' % (fvalue,pvalue))

## One way ANOVA on Total Survey 3.
fvalue,pvalue =stats.f_oneway(dA["TOTAL CORRECT S3"],dB["TOTAL CORRECT S3"],dC
["TOTAL CORRECT S3"])
print( 'F=%3.3f, p-value=%3.3f' % (fvalue,pvalue))


## One way ANOVA on Cognitive Dissonance.
fvalue,pvalue =stats.f_oneway(dA["Cognitive Dissonance"],dB["Cognitive Dissona
nce"],dC["Cognitive Dissonance"])
print( 'F=%3.3f, p-value=%3.3f' % (fvalue,pvalue))


fvalue,pvalue =stats.f_oneway(dA["Social Presence: [I would have liked the exp
erience to continue]"],dB["Social Presence: [I would have liked the experience
to continue]"],dC["Social Presence: [I would have liked the experience to cont
inue]"])
print( 'F=%3.3f, p-value=%3.3f' % (fvalue,pvalue))


fvalue,pvalue =stats.f_oneway(dA["Social Presence: [I felt REINA provided a me
morable experience]"],dB["Social Presence: [I felt REINA provided a memorable
 experience]"],dC["Social Presence: [I felt REINA provided a memorable experie
nce]"])
print( 'F=%3.3f, p-value=%3.3f' % (fvalue,pvalue))


fvalue,pvalue =stats.f_oneway(dA["Social Presence: [I felt as though I was in
 the same space as REINA]"],dB["Social Presence: [I felt as though I was in th
e same space as REINA]"],dC["Social Presence: [I felt as though I was in the s
ame space as REINA]"])
print( 'F=%3.3f, p-value=%3.3f' % (fvalue,pvalue))
```

```
In [ ]: A=project2('Total Memory Retention','Gender','Female')['Group A']
B=project2('Total Memory Retention','Gender','Female')['Group B']
C=project2('Total Memory Retention','Gender','Female')['Group C']

fvalue,pvalue =stats.f_oneway(A[A.notnull()],B[B.notnull()],C[C.notnull()])
print( 'F=%3.3f, p-value=%3.3f' % (fvalue,pvalue))
```


Figure C-6: Jupyter Noteboook with code for Data Analysis

In [ ]:
```
### FUNCTION TO GET ONE WAY ANOVA FOR FILTERED CONDITION
def filteredAnova (data,filt,val):
    A=project2(data,filt,val)['Group A']
    B=project2(data,filt,val)['Group B']
    C=project2(data,filt,val)['Group C']
    A=A[A.notnull()]
    B=B[B.notnull()]
    C=C[C.notnull()]
    fvalue,pvalue =stats.f_oneway(A,B,C)
    print(data+ ' &' + filt + '='+ str(val)+ ' & ' + 'A=%d, B=%d, C=%d & %3.3f
& %3.3f' % (A.shape[0],B.shape[0],C.shape[0],fvalue,pvalue))
#     print('A=%d, B=%d, C=%d' % (A.shape[0],B.shape[0],C.shape[0]))
```

In [ ]:
```
filteredAnova('TOTAL CORRECT S1', 'Gender','Female')
filteredAnova('TOTAL CORRECT S1', 'Gender','Male')
filteredAnova('TOTAL CORRECT S1', 'Country','USA')
filteredAnova('TOTAL CORRECT S1', 'Country','PA')
filteredAnova('TOTAL CORRECT S1', 'English Fluency','Yes')
filteredAnova('TOTAL CORRECT S1', 'Learner Type',4)
```

In [ ]:
```
filteredAnova('TOTAL CORRECT S2', 'Gender','Female')
filteredAnova('TOTAL CORRECT S2', 'Gender','Male')
filteredAnova('TOTAL CORRECT S2', 'Country','USA')
filteredAnova('TOTAL CORRECT S2', 'Country','PA')
filteredAnova('TOTAL CORRECT S2', 'English Fluency','Yes')
filteredAnova('TOTAL CORRECT S2', 'Learner Type',4)
```

In [ ]:
```
filteredAnova('TOTAL CORRECT S3', 'Gender','Female')
filteredAnova('TOTAL CORRECT S3', 'Gender','Male')
filteredAnova('TOTAL CORRECT S3', 'Country','USA')
filteredAnova('TOTAL CORRECT S3', 'Country','PA')
filteredAnova('TOTAL CORRECT S3', 'English Fluency','Yes')
filteredAnova('TOTAL CORRECT S3', 'Learner Type',4)
```

In [ ]:
```
filteredAnova('Total Memory Retention', 'Gender','Female')
filteredAnova('Total Memory Retention', 'Gender','Male')
filteredAnova('Total Memory Retention', 'Country','USA')
filteredAnova('Total Memory Retention', 'Country','PA')
filteredAnova('Total Memory Retention', 'English Fluency','Yes')
filteredAnova('Total Memory Retention', 'Learner Type',4)
```

In [ ]:
```
### Social Presence ANOVA
filteredAnova('Social Presence: [I would have liked the experience to continu
e]', 'AR/VR Usage','Yes')
filteredAnova('Social Presence: [I felt REINA provided a memorable experienc
e]', 'AR/VR Usage','Yes')
filteredAnova('Social Presence: [I felt as though I was in the same space as R
EINA]', 'AR/VR Usage','Yes')
```

In [ ]:
```
d
```


Figure C-7: Jupyter Noteboook with code for Data Analysis

In [ ]:
```
###### ANOTHER ONE-WAY ANOTHER
import statsmodels.api as sm
from statsmodels.formula.api import ols
```

In [ ]:
```
#data=d
#data
# mod=ols(d['Total Memory Retention'],d['Group'], data=data).fit()
# table = sm.stats.anova_lm(mod, typ=2) # Type 2 ANOVA DataFrame
```

In [ ]:
```
d.mean();
```

In [ ]:
```
d.std();
```

In [ ]:
```
dA.shape[0]
```

In [ ]:
```
dB.shape[0]
```

In [ ]:
```
dC.shape[0]
```


Figure C-8: Jupyter Noteboook with code for Data Analysis

Appendix D

Surveys

8/5/2020 Qualtrics Survey Software

Introduction

Introduction

Thank you for participating in this remote user-study to help assess what makes a more memorable tour guiding experience. In this study, you will go through a simple gallery tour with five different projects. In each project, you will see a poster with images and text descriptions, and an audio description provided by the tour guiding AI agent through a mobile app. All participants will be randomly selected into three different groups. Each group will experience a different feature of the tour guiding AI agent. During the call, you should have already downloaded one of the three applications called, "Gallery Tour <A/B/C>" and tested that it works correctly. Please, read and follow the instructions on this survey closely while you navigate through this experiment. **You DO NOT need to take any notes.**

This survey takes approximately 15 minutes to complete and should be taken in one single sitting.

What is your Subject Number? (provided by researcher during the call)

Default Question Block

Before we start, we will like to collect some basic information to help us analyze the results. Your response is encrypted, and we will not be collecting information that will lead to your identification.

Age

- 18-29
- 30-39
- 40-49
- 50-59
- 60+

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_3KmaoeE07SxyqPP&ContextLibraryID=UR_1... 1/21

8/5/2020 Qualtrics Survey Software

Gender

- Male
- Female
- Other

Current Location (Country and State/Province)

Do you consider yourself fluent in conversational English?

- Yes
- No
- So so (Not a regular English speaker)

What is your occupation?

What is the highest level of education you have completed or are currently completing?

On a weekly basis, how many classes, tours, or meetings do you attend?

After a class, meeting, or tour, how much of the material do you remember after 24 hours without reviewing notes?

- A great deal
- A lot
- A moderate amount
- A little
- None at all

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_3KmaoeE07SxyqPP&ContextLibraryID=UR_1... 2/21

8/5/2020 Qualtrics Survey Software

How frequently do you talk with your conversational agent (e.g. Alexa, Cortana, Siri)?

- Several times a day
- Once a day
- Several times a week
- Once a week
- Rarely
- Never used before

What mobile device are you currently using?

- iphone
- android phone
- tablet
- ipad

Have you ever used an augmented reality or virtual reality applications/devices before?

- Yes
- No

Tutorial

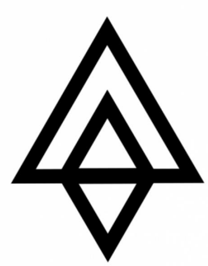
Tutorial

This is a tutorial project. Do not skip.

Instructions: Take a look at the poster/image below (read title, text, images if available). In this tutorial, you will only see an image with a symbol on it. The projects will be richer in information.

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_3KmaoeE07SxyqPP&ContextLibraryID=UR_1... 3/21

8/5/2020 Qualtrics Survey Software



Instructions: Open up the application and summon R.E.I.N.A. by placing your mobile camera in front of the image above. Try to hold your phone as steady as you can. You could place your elbow on the desk to provide support. When the audio stop playing, you can put your phone down, take some time to process the presented information, and move to the next page.

PROJECT 5

Instructions: Take a look at the poster below, make sure you read the text description just like you would touring a gallery.

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_3KmaoeE07SxyqPP&ContextLibraryID=UR_1... 4/21

Figure D-1: Survey 1

8/5/2020 Qualtrics Survey Software

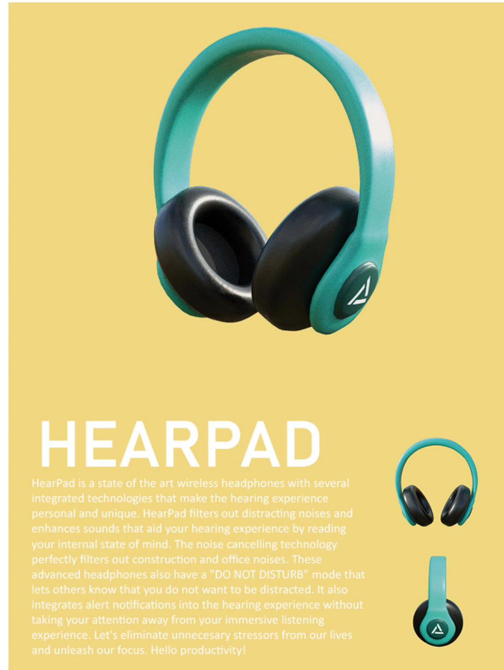
Instructions: Open up the application and summon R.E.I.N.A. by placing your mobile camera in front of the image above. Try to hold your phone as steady as you can. You could place your elbow on the desk to provide support. When the audio stop playing, you can put your phone down, take some time to process the presented information, and move to the next page.

PROJECT 3

Instructions: Take a look at the poster below, make sure you read the text description just like you would touring a gallery.

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_3KmaoeE07SxyqPP&ContextLibraryID=UR_1... 9/21

8/5/2020 Qualtrics Survey Software



https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_3KmaoeE07SxyqPP&ContextLibraryID=UR_... 10/21

8/5/2020 Qualtrics Survey Software

Instructions: Open up the application and summon R.E.I.N.A. by placing your mobile camera in front of the image above. Try to hold your phone as steady as you can. You could place your elbow on the desk to provide support. When the audio stop playing, you can put your phone down, take some time to process the presented information, and move to the next page.

PROJECT 2

Instructions: Take a look at the poster below, make sure you read the text description just like you would touring a gallery.

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_3KmaoeE07SxyqPP&ContextLibraryID=UR_... 11/21

8/5/2020 Qualtrics Survey Software

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_3KmaoeE07SxyqPP&ContextLibraryID=UR_... 12/21

Figure D-2: Survey 1

8/5/2020 Qualtrics Survey Software

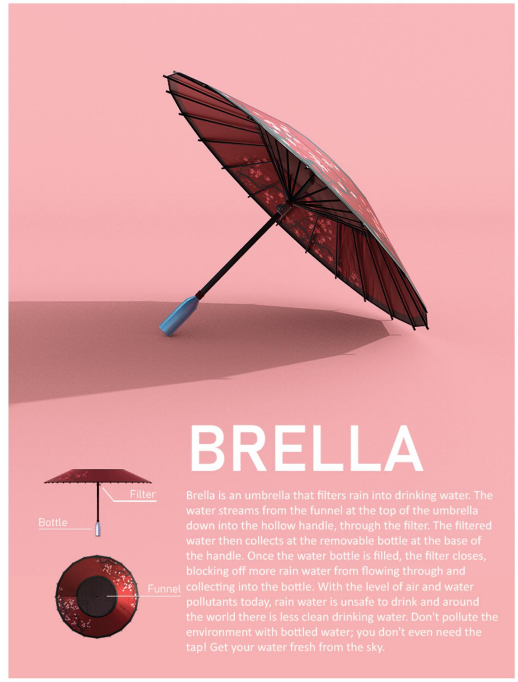
Instructions: Open up the application and summon R.E.I.N.A. by placing your mobile camera in front of the image above. Try to hold your phone as steady as you can. You could place your elbow on the desk to provide support. When the audio stop playing, you can put your phone down, take some time to process the presented information, and move to the next page.

PROJECT 1

Instructions: Take a look at the poster below, make sure you read the text description just like you would touring a gallery.

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_3KmaoeE07SxyqPP&ContextLibraryID=UR_… 13/21

8/5/2020 Qualtrics Survey Software



https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_3KmaoeE07SxyqPP&ContextLibraryID=UR_… 14/21

8/5/2020 Qualtrics Survey Software

Instructions: Open up the application and summon R.E.I.N.A. by placing your mobile camera in front of the image above. Try to hold your phone as steady as you can. You could place your elbow on the desk to provide support. When the audio stop playing, you can put your phone down, take some time to process the presented information, and move to the next page.

Before Questionnaire

You have reached the end of the gallery. **Please DELETE or UNINSTALL the application now before you continue.**

ASSESS 5

Read the question(s) carefully before you answer.

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_3KmaoeE07SxyqPP&ContextLibraryID=UR_… 15/21

8/5/2020 Qualtrics Survey Software

Select the correct statement about this project:

- ○ This is cut-resistant and can regulates temperature automatically to ensure that the user is comfortable.
- ○ This is a virtual reality haptic gloves that can make you feel and touch digital objects.
- ○ This can control compatible IoT devices by performing hand gestures

What material is this made out of?

- ○ Aluminized fabrics
- ○ Para-aramid fibers
- ○ Carbon fibers

ASSESS 4

Read the question(s) carefully before you answer.

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_3KmaoeE07SxyqPP&ContextLibraryID=UR_… 16/21

Figure D-3: Survey 1

8/5/2020 Qualtrics Survey Software

How much does it weight when unloaded?

- 12 lbs
- 15 lbs
- 13 lbs
- 20 lbs

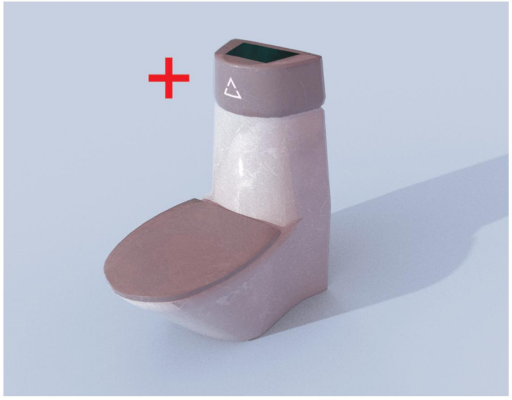
What is the material used to make the body frame?

Assess 3

Read the question(s) carefully before you answer.

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_3KmaoeE07SxyqPP&ContextLibraryID=UR_... 17/21

8/5/2020 Qualtrics Survey Software



How long it takes to get the result displayed?

- Within 60 seconds
- Within 25 seconds
- Within 42 seconds
- Within 30 seconds

What are three mentioned medical conditions it can detect?

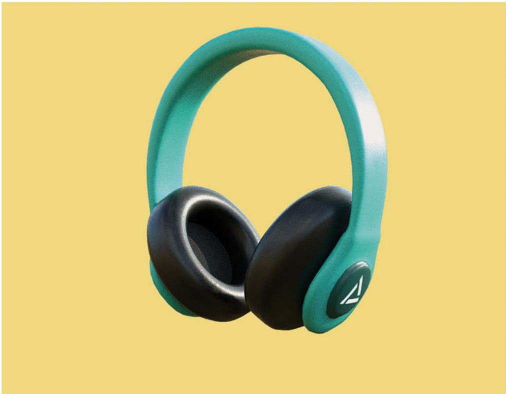
- Helicobacter Pilori, kidney disease, HIV
- Liver disease. kidney disease, bladder cancer
- Helicobacter Pilori, Liver disease, Diabetes

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_3KmaoeE07SxyqPP&ContextLibraryID=UR_... 18/21

8/5/2020 Qualtrics Survey Software

ASSES 2

Read the question(s) carefully before you answer.



What is the name of this project?

What is one of the mode presented in this project?

- "STAY CALM"
- "BEING PRODUCTIVE"
- "DO NOT DISTURB"

ASSESS 1

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_3KmaoeE07SxyqPP&ContextLibraryID=UR_... 19/21

8/5/2020 Qualtrics Survey Software

Read the question(s) carefully before you answer.

What is this project about and what can it do?

- This umbrella reflects UV light off and protect you from skin cancer.
- This umbrella filters rain into drinking water.
- This umbrella hides you from been seeing by drones or satellites.

What is the filter composed of?

- A Brita filter.
- Nano-mesh
- A layer of sand, gravel, and activated carbon.

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_3KmaoeE07SxyqPP&ContextLibraryID=UR_... 20/21

Figure D-4: Survey 1

CONCLUSION

Thank you for your participation. Another questionnaire will be sent in 24hr and 72hr. Please make sure you look out for them in your email. Thanks!

Powered by Qualtrics

Figure D-5: Survey 1

8/5/2020 Qualtrics Survey Software

Introduction

Welcome back to the second survey on what makes a memorable augmented reality tour guiding experience. Similar to the first survey, you will be asked several questions about each of the presented projects. Read each question carefully before you answer.

This survey should take less than 10 minutes to complete.

What is your subject number?

ASSESS 4

Read the question(s) carefully before you answer.

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_aaXK8rMEHAczS4Z&ContextLibraryID=UR_1A... 1/7

8/5/2020 Qualtrics Survey Software

How long does it take to fold?

- ○ in under 3 minutes
- ○ in under five minutes
- ○ in under 30 seconds
- ○ in under a minute

Select the correct statement about this project

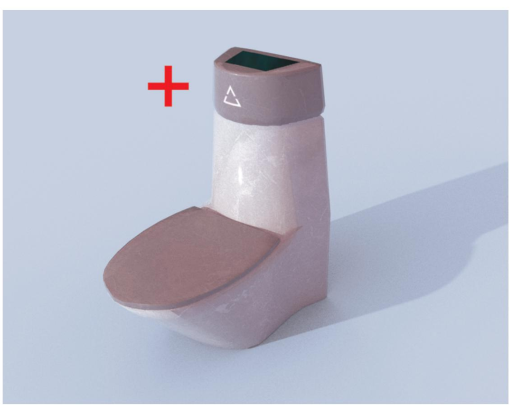
- ○ A bike that has a backpack as the seat for storage
- ○ A modular backpack that can be converted into a bike by attaching pedals and wheels on it.
- ○ A sleek, strong, and light bike that can be folded into a backpack

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_aaXK8rMEHAczS4Z&ContextLibraryID=UR_1A... 2/7

8/5/2020 Qualtrics Survey Software

Assess 3

Read the question(s) carefully before you answer.



How can users keep their results?

- ○ By providing their fingerprint, it will automatically send the encrypted information to the registered contact matching it.
- ○ By providing their email address via voice command, it will automatically send the encrypted information.
- ○ Through facial recognition, it will automatically send the encrypted information to the registered contact matching the facial structure.

What is the name of this project?

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_aaXK8rMEHAczS4Z&ContextLibraryID=UR_1A... 3/7

8/5/2020 Qualtrics Survey Software

ASSESS 5

Read the question(s) carefully before you answer.

What is the name of the technology mentioned in this project?

- ○ GoodSense
- ○ HiTemp
- ○ StrongSkin
- ○ HiSense

How does the technology mentioned work?

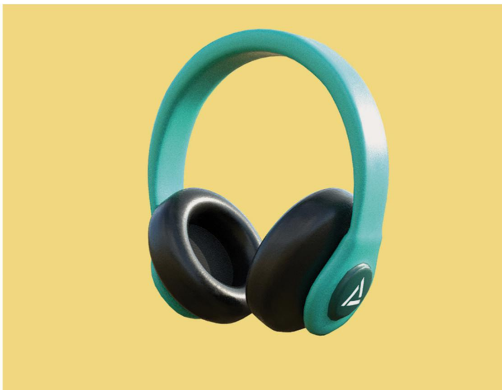
https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_aaXK8rMEHAczS4Z&ContextLibraryID=UR_1A... 4/7

Figure D-6: Survey 2 sent 24 hours after the experiment

- It identifies the material of the surface in contact and replicate the texture on the inner layer of the glove.
- It simulates the texture of the digital object to provide haptic feedback to the user.
- It provides vibration warning to keep user away from hazardous material.

ASSES 2

Read the question(s) carefully before you answer.



What are the four main features mentioned in this project?

- Adaptive mic system, Noise cancelling technology, Augmented reality technology, Calm notification technology
- Augmented reality technology, Super X-Fi technology , Noise cancelling technology, Selective sound augmentation

- Noise cancelling technology, Selective sound augmentation, Calm notification technology, and Context-awareness capabilities

What does it use to identify the user's mental state?

- Heart Rate
- Galvanic Skin Response
- Targeted Electroencephalography

ASSESS 1

Read the question(s) carefully before you answer.

Where is the filter located at?

- At the detachable bottle
- At the porous fabric
- At the hollow handle

What is the name of this project?

Block 6

Look out for the last survey which will be sent to your email in 48 hours. Thanks!

Powered by Qualtrics

Figure D-7: Survey 2 sent 24 hours after the experiment

8/5/2020 Qualtrics Survey Software

Introduction

Welcome back to the third and final survey on what makes a memorable augmented reality tour guiding experience. Similar to the first two surveys, you will be asked several questions about each of the presented projects. Read each question carefully before you answer.

This survey should take less than 10 minutes to complete.

What is your subject number?

ASSESS 5

Read the question(s) carefully before you answer.

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_bQ2CVvBpNJuZM0t&ContextLibraryID=UR_1A... 1/8

8/5/2020 Qualtrics Survey Software

What is the name of this project?

What is this resistant to?

- Tear-resistant
- Weather-resistant
- Heat-resistant
- Cut-resistant

ASSESS 4

Read the question(s) carefully before you answer.

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_bQ2CVvBpNJuZM0t&ContextLibraryID=UR_1A... 2/8

8/5/2020 Qualtrics Survey Software

What is the name of this project?

What is the carrying capacity of the backpack?

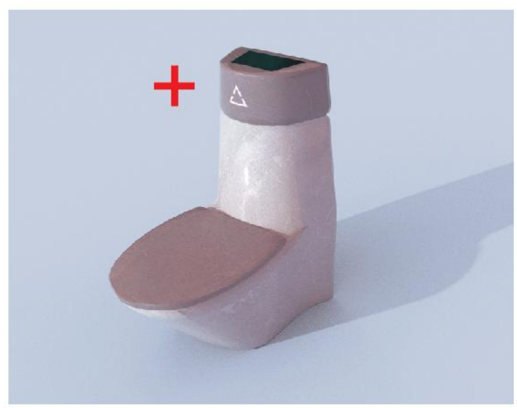
- 15L
- 20L
- 10L
- 25L

Assess 3

Read the question(s) carefully before you answer.

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_bQ2CVvBpNJuZM0t&ContextLibraryID=UR_1A... 3/8

8/5/2020 Qualtrics Survey Software



Select the correct statement(s) about this project

- A smart toilet system that sends samples of your urine and stool to your family doctor for analysis.
- A smart toilet system that provides dietary recommendations based on your urine and stool samples.
- A smart toilet system that performs medical analyses on urine and stool samples in real-time.
- A smart toilet system that automatically does a self-clean after every use.

How does it ID the user?

- Facial recognition
- Voice recognition

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_bQ2CVvBpNJuZM0t&ContextLibraryID=UR_1A... 4/8

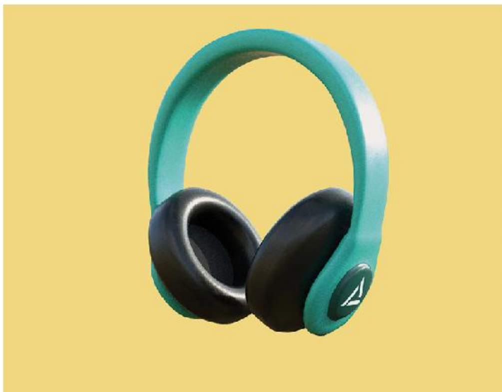
Figure D-8: Survey 3 sent 72 hours after the experiment

8/5/2020 Qualtrics Survey Software

- DNA profiling
- Fingerprint

ASSES 2

Read the question(s) carefully before you answer.



Select the correct statement described about this project.

- The augmented reality technology enhances the essential sound of the real world environment to make user more aware of their surrounding.
- The vibrating notification system warns user of their surrounding.
- The calming notification provides awareness of the user's surrounding by blending notification alerts into the listening experience.

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_bQ2CVVBpNJuZM0t&ContextLibraryID=UR_1A... 5/8

8/5/2020 Qualtrics Survey Software

What is the name of this project?

- HearThere
- HeadBand
- HearPad

ASSESS 1

Read the question(s) carefully before you answer.

What inspired the modeling of this project?

- Japanese tea ceremony umbrella called Nodatekasa
- Chinese oil-paper umbrella called Youzhisan

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_bQ2CVVBpNJuZM0t&ContextLibraryID=UR_1A... 6/8

8/5/2020 Qualtrics Survey Software

- Japanese oil-paper umbrella called Wagasa

How much filtered water can it hold?

- 32oz
- 20oz
- 12oz
- 16oz

Block 7

Did you see the guiding agent? If yes, please describe what you saw and what it resembles. If you did not see anything and just hear voice, write N/A.

Did the guiding agent ask you a question after explaining about the project?

- Yes
- No

What do you think about the tour guiding agent?

What do you think about the experience overall? Did it distract you or enhance your memory? Feel free to provide your perceived experience and any technical issue you have encountered.

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_bQ2CVVBpNJuZM0t&ContextLibraryID=UR_1A... 7/8

8/5/2020 Qualtrics Survey Software

Were you able to understand the tour guiding agent? [Yes, Not always, No] If so, Which project did you struggle with?

Do you think you are more of a visual or auditory learner?

- Visual
- Auditory

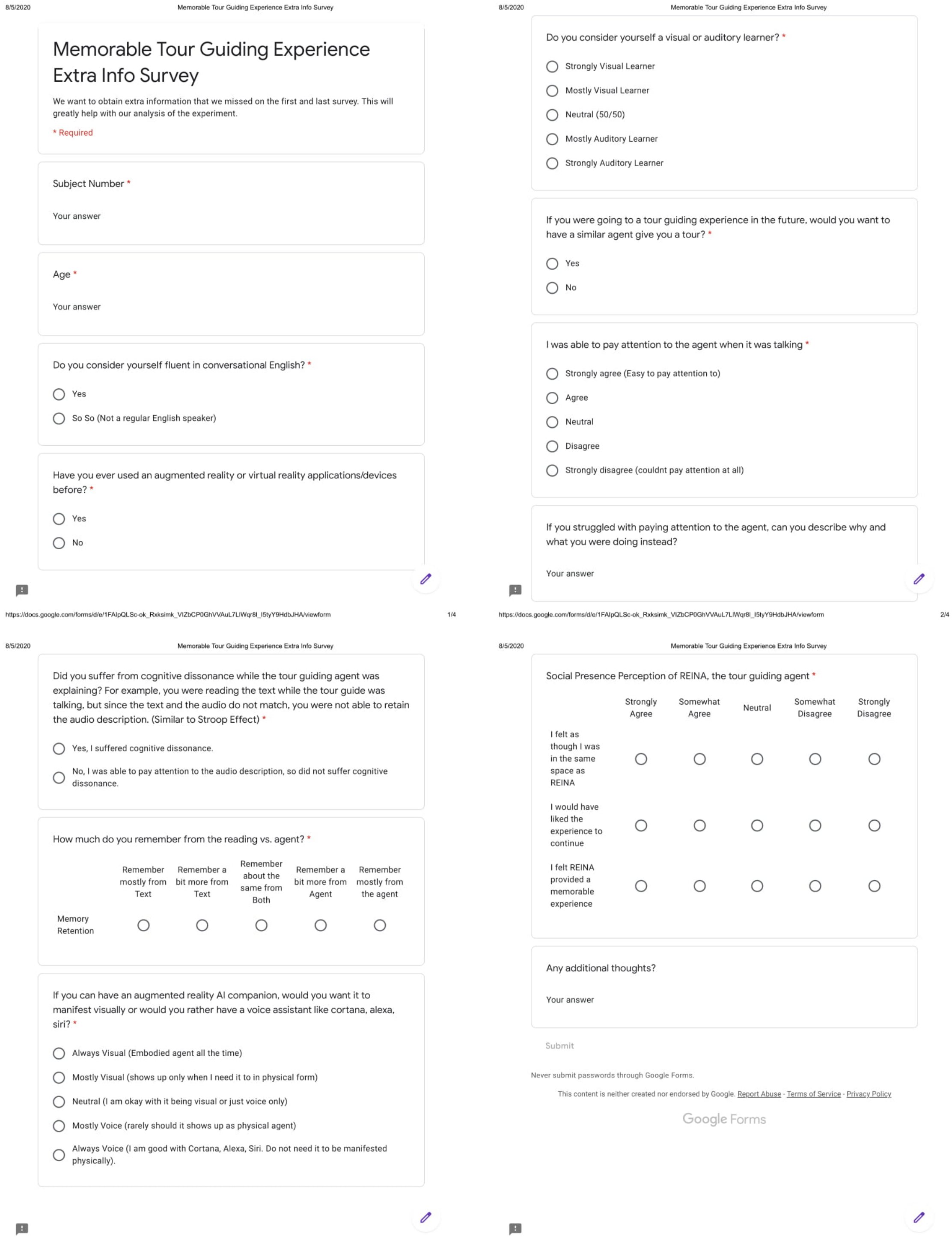
Block 6

You made it! You have completed all the required surveys. We want to thank you for your contribution to this work. We will be sending you a $15 Amazon gift card to the email you provided during our call. If you don't receive it by the end of July, please contact elenack@mit.edu.

Powered by Qualtrics

https://mit.co1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview?ContextSurveyID=SV_bQ2CVVBpNJuZM0t&ContextLibraryID=UR_1A... 8/8

Figure D-9: Survey 3 sent 72 hours after the experiment



8/5/2020 Memorable Tour Guiding Experience Extra Info Survey

Memorable Tour Guiding Experience Extra Info Survey

We want to obtain extra information that we missed on the first and last survey. This will greatly help with our analysis of the experiment.

* Required

Subject Number *

Your answer

Age *

Your answer

Do you consider yourself fluent in conversational English? *

- Yes
- So So (Not a regular English speaker)

Have you ever used an augmented reality or virtual reality applications/devices before? *

- Yes
- No

https://docs.google.com/forms/d/e/1FAIpQLSc-ok_Rxksimk_VIZbCP0GhVVAuL7LIWqr8I_I5tyY9HdbJHA/viewform 1/4

8/5/2020 Memorable Tour Guiding Experience Extra Info Survey

Do you consider yourself a visual or auditory learner? *

- Strongly Visual Learner
- Mostly Visual Learner
- Neutral (50/50)
- Mostly Auditory Learner
- Strongly Auditory Learner

If you were going to a tour guiding experience in the future, would you want to have a similar agent give you a tour? *

- Yes
- No

I was able to pay attention to the agent when it was talking *

- Strongly agree (Easy to pay attention to)
- Agree
- Neutral
- Disagree
- Strongly disagree (couldnt pay attention at all)

If you struggled with paying attention to the agent, can you describe why and what you were doing instead?

Your answer

https://docs.google.com/forms/d/e/1FAIpQLSc-ok_Rxksimk_VIZbCP0GhVVAuL7LIWqr8I_I5tyY9HdbJHA/viewform 2/4

8/5/2020 Memorable Tour Guiding Experience Extra Info Survey

Did you suffer from cognitive dissonance while the tour guiding agent was explaining? For example, you were reading the text while the tour guide was talking, but since the text and the audio do not match, you were not able to retain the audio description. (Similar to Stroop Effect) *

- Yes, I suffered cognitive dissonance.
- No, I was able to pay attention to the audio description, so did not suffer cognitive dissonance.

How much do you remember from the reading vs. agent? *

	Remember mostly from Text	Remember a bit more from Text	Remember about the same from Both	Remember a bit more from Agent	Remember mostly from the agent
Memory Retention	○	○	○	○	○

If you can have an augmented reality AI companion, would you want it to manifest visually or would you rather have a voice assistant like cortana, alexa, siri? *

- Always Visual (Embodied agent all the time)
- Mostly Visual (shows up only when I need it to in physical form)
- Neutral (I am okay with it being visual or just voice only)
- Mostly Voice (rarely should it shows up as physical agent)
- Always Voice (I am good with Cortana, Alexa, Siri. Do not need it to be manifested physically).

https://docs.google.com/forms/d/e/1FAIpQLSc-ok_Rxksimk_VIZbCP0GhVVAuL7LIWqr8I_I5tyY9HdbJHA/viewform 3/4

8/5/2020 Memorable Tour Guiding Experience Extra Info Survey

Social Presence Perception of REINA, the tour guiding agent *

	Strongly Agree	Somewhat Agree	Neutral	Somewhat Disagree	Strongly Disagree
I felt as though I was in the same space as REINA	○	○	○	○	○
I would have liked the experience to continue	○	○	○	○	○
I felt REINA provided a memorable experience	○	○	○	○	○

Any additional thoughts?

Your answer

Submit

Never submit passwords through Google Forms.

This content is neither created nor endorsed by Google. Report Abuse - Terms of Service - Privacy Policy

Google Forms

https://docs.google.com/forms/d/e/1FAIpQLSc-ok_Rxksimk_VIZbCP0GhVVAuL7LIWqr8I_I5tyY9HdbJHA/viewform 4/4

Figure D-10: Survey 4 sent after all participants were done with the first three.

Appendix E

Survey Responses: Comments on Attention

Table E.1: Responses to "Describe why you could not pay attention".

Group	"Describe why you could not pay attention"
A	It's tough reading and listening at the same time for normal museum behavior. One or the other learning method is nice but not both.
A	It could help to allow pause it (and rewind it)
A	I was looking at the picture and reading the texts
A	The only conflicts I had were the inconsistencies between the audio and the text, and the fact that everything was happening at the same time.
A	Bright poster and image had more of my attention
A	I think because I had already read most of what the agent was saying.
A	I think I got the version where no one showed up, which made me have to look away from the posters so that I didn't read while the voice talked. It's hard to listen to someone/something when there is similar text in screen, UNLESS there's a person to look at. I was aware of this so I was able to pay attention to the audio but it was not frictionless as I had to actively not look at the text.
A	Skimming the text again, or getting distracted by other notifications on my screen
A	I had to listen and read at the same time so it was kind of hard to understand
A	There were occasions when it was difficult to pay attention to the agent because I was trying to read the text which was different in certain parts from what the agent was saying and so it became a little distracting. Lol I just saw this is the question below
A	I think i did because Im not so used to just listen and remember, i need visuals
A	I struggle just when cognitive dissonance happened
A	Visualizing what the agent was saying, but couldn't keep up cause it was too fast.

Table E.2: Responses to "Describe why you could not pay attention".

Group	"Describe why you could not pay attention"
B	If I history payed attention to the agent, it was totally fine. But when I tried to also read, that was distracting.
B	I just zoned out and thought about other things. Mostly about a similar project I had worked on.
B	With the agent, it felt quite monotone thus made it a little hard to follow along.
B	Didn't know if I should be reading or listening and got distracted
B	It was easy to pay attention to but sometimes it felt slow and got monotonous. It was also annoying that it only played once and could not be replayed or you could not control where to pause resume it from
B	I was trying to read the poster
B	I also like to explore the spaces visually and physically myself (sense of wonder), while the agent is speaking beside me
B	I don't remember. I think I was trying to pay attention as much as possible.
B	well part of it was probably because I was hung over, I was focusing on how the orb moved along with the words rather the words themselves
B	I was actually trying to focus on what it was saying because I could not read the lips and found that it was different from the text.
B	there was quite a bit of content to remember overall, and my eyes kept wanting to look at text
B	Looking at R.E.I.N.A or trying to process the words
B	I was too distracted by the animation of the agent
B	I was listen to the agent but i think that the same voice in all the projects make me bored.
B	One time, because the agent suddenly disappear when it moved to one side and in the moment I was searching for it that is why I wasn't paying attention.
B	I was trying to look closely at the photo for the described features/parts.
B	Having to hold my arm up to look at the dot was distracting when I could have essentially the same experience just listening and reading the poster
B	My mind sometimes wandered
B	I do not struggled paying attention
B	I thought that there was a lot of repetition (from the text descriptions) so I would tend to tune out the assistant.

Table E.3: Responses to "Describe why you couldnt pay attention".

Group	"Describe why you could not pay attention"
C	I usually like to rewind and listen when it's providing information e.g. audiobooks and podcasts. Since I had to close and reopen the app everytime to scan the image, I wasn't as enthusiastic about doing the same in order to listen multiple times like I usually would. I do think I was looking at the image while listening, though.
C	I was also attempting to read and was reading faster than the speed which the agent was speaking in.
C	I was reading the text.
C	I was paying attention
C	Looking at the product and thinking what can it be used for or what are some flaws
C	Looking at the poster on the screen
C	Focusing on repositioning the camera to the correct location.
C	I was trying to read the text at the same time, but it seemed like the text differed from the audio.
C	I struggled to pay attention because I was also trying to read the text at the same time.
C	Needing to pay attention to a virtual agent is harder than a physical human being. I suppose having someone physical that is conscious of my attention and focus makes me pay attention more.
C	I'm a more visual learner. For all the posters, I got how the projects work mainly from the text. Sometimes it's a little difficult for me to understand how the projects work in mind just through the description from the agent. But maybe because the processing speed of my brain is not fast enough...
C	I was able to attend to its speech, but not to its visual elements. I was looking at the actual poster (on my laptop screen).
C	I was looking at the shape of the agent and the little particle that were leaving its body. I was also trying to read at the same time to match up the information I was receiving.
C	The agent wouldn't always read from the text. While I listened to the agent, I also tried reading the text, which caused me to not give either the agent or text 100% of my attention.
C	The voice was very soft and somewhat put me to sleep. But also they weren't saying the same as the written text entirely.

Appendix F

Survey Responses: Comments on Agent Visualization

Table F.1: Description of the tour guiding agent by some of the participants in Group B and C.

Group	"Did you see the guiding agent? If yes, please describe what you saw and what it resembles."
B	A circle with particles floating out if the top as the agent spoke.
B	It looked like a sphere. It was shining mix of blue-white colors
B	Yes, it looked like an orb, kinda looked like a jellyfish. Particles floated out of it consistently when it spoke.
B	Yes, blue sphere with voice command
B	The guiding agent looked like a round orb with electricity sparks
B	Yes, it looked like a blue ball of energy
B	Yes, it is a glowing blue orb that releases electric "stem". It just looks like some kind of sci-fi ball.
B	An orb
B	Yes, I saw a sky-blue circle with swirling white contrails. It was called Reina(?).
B	I saw a blue orb moving with the voice emphasis
B	Yes, I saw an orb that shone whenever it was talking. In some occasions, I looked away while it was speaking to focus on it.
B	yes, reminds me of a water droplet or Navi from Legend of Zelda
B	R.E.I.N.A. Circular shaped, with a galactic feel in it. Remind me of Siri. When it is talking, the middle will start lighting up.
B	I saw a blue orb that pulsed while it talked.
B	Yes, i could see my agent and could also hear it. It was like an circle like blue color.
B	Yes, I saw the guiding agent. I saw it like a big circle with blue particles around it.
B	Yes, it was like a circular ball of energy, and inside of it I saw somethings like thunders. In one of the Product, the guiding agent first appear and then it moved to the right side and seems like it disappears.
B	Yes, is a blue energy like sphere that glow while talking
B	I saw R.E.I.N.A. i saw her like a blue floating orbe that reminds me a little bit of Jarvis from "Iron Man".
B	Yes. Blue ethereal sphere, makes movement with the speech.
B	I saw a blueish glowing orb with lightning-like effects coming from it.
B	An orb that grows and emits electricity with respect to the voice
B	Yes, it was a blue orb emitting particles and with lightning inside.
B	I saw what looked like a floating blue orb with some particles orbiting around it. I think it lit up a little as it was speaking.
B	I saw the guiding agent. It resembled a plasma globe.
B	Yes it look like a fairy of the legend of zelda
B	Yes I saw the guiding agente, and I heared it. I saw a electric bubble and it resembles me like a talking bubble.

Table F.2: Description of the tour guiding agent by some of the participants in Group B and C.

B	Yes, blue circle full of rays. Look like a stormy rain.
B	Yes! It was a bubble with "energy" coming out of it very fast
C	It looks like a floating 3D Cortana. Or from Maplestory, the "Bonus XP" that can be obtained.
C	Yes, it was a little blue orb that would float and vibrate as it spoke. It flowed and had blue squiggly lines coming off it. It resembled a guiding light, kinda like one of those Disney movies (Frozen I think), where the lights guide the character.
C	Yes. It looked like a blue crystal sphere with graphics/coding emerging from it.
C	I saw a white orb of light along with her voice.
C	Yes, a sphere
C	Yes. I saw a blue bubble-like sphere. It emited somewhat of energy that went upward.
C	Yes. The agent looked like a blue lit bulb.
C	Yes. It was a translucent sphere that looked like water flowing a bit and it lit up as it spoke.
C	Yes, I saw an orb that emit some blue smoke. It was most similar to a talking snow globe.
C	Yes, I was able to see and hear the agent. The agent resembles a bubble.
C	To be really honest, I don't remember. I don't remember seeing anything particular except for the screen, but I can't remember for sure.
C	Yes, it was a floating orb that talked. I don't think it resembled anything specific to me beyond that.
C	Electric glowing vaporous ball
C	Yes, I did. It's on the top left conner of the posters. It's like a transparent glass ball with blue electric spark inside.
C	Yes. A blue glowing orb.
C	Yes, it resembled an oracle. It had particles flowing out of it and it floated on the top left corner of each product. It was purple.

Table F.3: Description of the tour guiding agent by some of the participants in Group B and C.

C	Yes, it was an orb that spoke with a female voice. Every time it would speak, the orb would flash.
C	Yes, I saw a sphere shooting electricity and waves of color.
C	A blue ball
C	Looked like a bubble
C	Yes, the guiding agent looked like a floating and vibrating energy orb.
C	Yes, a floating circle with dynamic flow
C	I say a translucent floating sphere. In side the sphere there are animated light patterns.
C	Yes, the guide agent was like a transparent clouds with electricity and colors blue and white.
C	Yes, the agent was like a blue purple - ish sphere and it resembles a magic ball
C	I saw like a bubble in space or something like an orb.
C	yes, i saw a sphere blue with lightning bolts , floating in my pc , like a cortana from halo
C	yes, like navi(from Zelda) sparkling.
C	Yes. The guilding agent was a blue orb with a white center, and pulsates when it speaks.
C	Yes a blue orb with effervescence

Appendix G

Survey Response: Additional Comments

Table G.1: Additional comments from the participants

A	Multimodal for the win, great idea!
A	Everything is great
A	This was fun! :)
A	It was hard to tell when REINA was done talking. It felt like the agent stopped abruptly.
A	REINAs voice and ways of addressing us (greetings) and explaining each project made her sound somewhat robotic and monotonous. For me, a more memorable experience would be obtained if the information provided by her would have been more casually introduced, rather than sound as someone/something who reads literally from a source; it sounded with lack of emotion. Maybe some inflections in her voice (and if possibly, she interacting with some features of the project) could make me feel more driven and eager towards her. However, it was good that the information she conveyed was different to the one provided in the texts, since it was an interesting info, it was memorable. The memorable trait came from the quality of the info rather than from REINA's voice/ways itself.
A	Could have been better if the avatar showed up talking or interacting with the product.
A	Loved the projects and Reina
A	I definitely think a visual presence would have been helpful for me personally. Just the voice felt like a listening tour in a museum
A	I remember better the projects that matched with my needs (I.e. Brella)
A	Prefer to see how each of those items(bikepack, headdra, brella) work in a video with an audio explanation background. This will help me remember better. It's like going to a movie, and you remember the scene that has an interesting story. Instead of a jpeg image with text form and an audio description. That was harder to remember.
B	I think the agent and the way it physically manifest was a cool idea but maybe improving upon it would make it more appealing. In reference to the question above about retaining from audio vs text, I am not sure I remember where the information I remembered was from, as in, in retrospect its hard to recall whether it came from audio or text. It did not mess up the information when it was different for audio and text but it was annoying when after listening to the audio I was reading the text and did not find similar information and had no way to go back to the audio to get that
B	I would rather just voice unless the visual manifestation resembled a human
B	I thought the voice acting was excellent, but trying to focus both on the text and what Reina was saying was hard. Maybe adding a feature for repetition or question answering may have helped.

Table G.2: Additional Comments from the participants.

B	Having a physical guide to me is a must, but much like a real guide, the AI should consume a bulk of my attention. I think if the graphic only showed the product and REINA narrated while the words appeared like subtitles, it would've stuck more
B	Any other thoughts, was a nice experience
B	I'd really hope one day i can have a companion like R.E.I.N.A.
B	I believe in the instructions, I was told to not actively try to memorize what was being said but more so actively listen. Because of this, I think I took this more as a casual 'strolling through an exhibit' type of experience leading me to focus more on the general gist of each product rather than the specific details.
B	Great study case, hope you all can go forward with augmented reality
B	It was a different experience, but now you mention it in one of the questions before. I think I would like it better if what REINA says was similar to the description text, so I can follow the reading. But if is not equal, just similar is ok.
C	I loved it!!
C	Keep going. The project has great potential.
C	If I can use REINA I would get it for my phone.
C	If the guiding agent has a face, it might make it easier for people to remember what they are saying.
C	Although I could pay attention easily to the agent, I don't think I could for a long duration.
C	sorry don't remember my number. pls double hecknif it's correct. very good research topic!!
C	I think it would be better if the REINA can use animation to show how the projects work while she's introducing them. This is more intuitive for me and easier to understand and remember. Actually when I looked at posters about some new projects, I'm more interested in the big idea, i.e. what the project can do and how it solve the problem. I will not pay too much attention to the details, e.g. the BRELLA project, I will not pay attention to the capacity of the bottle. Some details may not be worthy for me to remember.
C	I would have liked to see Reina's embodiment playing more role in the experience. It would have been helpful for Reina to use its visual elements and animacy (movement and position) that synchronizes to the information being presented.
C	It was a really good experience and I would like to do more of them in the future!
C	The experiment was very cool, I never did anything like this before.

Table G.3: Additional Comments from the participants.

C	I believe maybe the main problem was ... hmmm what should I pay attention (Text or Reina). As we mostly rely more on sight, rather than hearing for information about our environment, it was hard to make the transition while touring. At the end REINA gave way more info than the text but you dont know that unless you really pay attention to her.
C	I think what would be helpful with REINA would be to have the option of adding captions under her as an option for those who want to read what REINA is saying.
C	If I could use Reina sparingly and in a more directed way e.g. when I wanted more detailed information, I think the agent could have proved very useful. I am neurodiverse and struggle with being given information in a linear way. I would have liked to be able to jump around and explore the product features at my own pace. Overall it was very interesting and this sort of agent could certainly enhance my experience and retention of intervention if applied more flexibly. Thanks!

Bibliography

[1] PARC History: A LEGACY OF INVENTING THE FUTURE.

[2] Fabio Abbattista, Graziano Catucci, Giovanni Semeraro, and Fabio Zambetta. SAMIR: A Smart 3D Assistant on the Web. *PsychNology Journal*, 2(1):43–60, 2004.

[3] Jordan Abdi, Ahmed Al-Hindawi, Tiffany Ng, and Marcela P Vizcaychipi. Scoping review on the use of socially assistive robot technology in elderly care. *BMJ open*, 8(2):e018815, 2018.

[4] Henny Admoni and Brian Scassellati. Social eye gaze in human-robot interaction: a review. *Journal of Human-Robot Interaction*, 6(1):25–63, 2017.

[5] Susan B Barnes. Alan Kay: Transforming the computer into a communication medium. *IEEE Annals of the History of Computing*, 29(2):18–30, 2007.

[6] Christoph Bartneck and Jodi Forlizzi. A design-centred framework for social human-robot interaction. In *RO-MAN 2004. 13th IEEE international workshop on robot and human interactive communication (IEEE Catalog No. 04TH8759)*, pages 591–594. IEEE, 2004.

[7] Tony Belpaeme, James Kennedy, Aditi Ramachandran, Brian Scassellati, and Fumihide Tanaka. Social robots for education: A review. *Science robotics*, 3(21), 2018.

[8] Kirsten Bergmann and Manuela Macedonia. A Virtual Agent as Vocabulary Trainer: Iconic Gestures Help to Improve Learners' Memory Performance. In Ruth Aylett, Brigitte Krenn, Catherine Pelachaud, and Hiroshi Shimodaira, editors, *Intelligent Virtual Agents*, pages 139–148, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.

[9] Timothy Bickmore and Justine Cassell. Relational Agents: A Model and Implementation of Building User Trust. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '01, pages 396–403, New York, NY, USA, 2001. ACM.

[10] M Billinghurst and J Savage. Adding intelligence to the interface. In *Proceedings of the IEEE 1996 Virtual Reality Annual International Symposium*, pages 168–175, 3 1996.

[11] Frank Biocca, Chad Harms, and Judee K Burgoon. Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence: Teleoperators & virtual environments*, 12(5):456–480, 2003.

[12] Michael Bratman and others. *Intention, plans, and practical reason*, volume 10. Harvard University Press Cambridge, MA, 1987.

[13] Joost Broekens, Marcel Heerink, Henk Rosendal, and others. Assistive social robots in elderly care: a review. *Gerontechnology*, 8(2):94–103, 2009.

[14] Rodney A Brooks. Intelligence without reason, computers and thought lecture. *Proceedings of IJCAI-91, Sidney, Australia*, 32:37–38, 1991.

[15] Rodney A Brooks. Intelligence without representation. *Artificial intelligence*, 47(1-3):139–159, 1991.

[16] David J H Burden. Deploying Embodied AI into Virtual Worlds. In Tony Allen, Richard Ellis, and Miltos Petridis, editors, *Applications and Innovations in Intelligent Systems XVI*, pages 103–115, London, 2009. Springer London.

[17] Abraham G Campbell, John W Stafford, Thomas Holz, and Gregory M P O'Hare. Why, when and how to use augmented reality agents (AuRAs). *Virtual Reality*, 18(2):139–159, 2014.

[18] J Cassell, T Bickmore, M Billinghurst, L Campbell, K Chang, H Vilhjálmsson, and H Yan. Embodiment in Conversational Interfaces: Rea. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '99, pages 520–527, New York, NY, USA, 1999. ACM.

[19] Justine Cassell. More than just another pretty face: Embodied conversational interface agents. *Communications of the ACM*, 43(4):70–78, 2000.

[20] Justine Cassell, Catherine Pelachaud, Norman Badler, Mark Steedman, Brett Achorn, Tripp Becket, Brett Douville, Scott Prevost, and Matthew Stone. Animated Conversation: Rule-based Generation of Facial Expression, Gesture &Amp; Spoken Intonation for Multiple Conversational Agents. In *Proceedings of the 21st Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '94, pages 413–420, New York, NY, USA, 1994. ACM.

[21] Justine Cassell, Joseph Sullivan, Elizabeth Churchill, and Scott Prevost. *Embodied conversational agents*. MIT press, 2000.

[22] Lalith Chandralal and Fredy-Roberto Valenzuela. Exploring memorable tourism experiences: Antecedents and behavioural outcomes. *Journal of Economics, Business and Management*, 1(2):177–181, 2013.

[23] Kerstin Dautenhahn. Socially intelligent robots: dimensions of human–robot interaction. *Philosophical transactions of the royal society B: Biological sciences*, 362(1480):679–704, 2007.

[24] Maartje M A De Graaf and Somaya Ben Allouch. Exploring influencing variables for the acceptance of social robots. *Robotics and autonomous systems*, 61(12):1476–1486, 2013.

[25] Ewart J de Visser, Frank Krueger, Patrick McKnight, Steven Scheid, Melissa Smith, Stephanie Chalk, and Raja Parasuraman. The world is not enough: Trust in cognitive agents. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 56, pages 263–267. Sage Publications Sage CA: Los Angeles, CA, 2012.

[26] Ewart J de Visser, Samuel S Monfort, Kimberly Goodyear, Li Lu, Martin O'Hara, Mary R Lee, Raja Parasuraman, and Frank Krueger. A Little Anthropomorphism Goes a Long Way: Effects of Oxytocin on Trust, Compliance, and Team Performance With Automated Agents. *Human Factors*, 59(1):116–133, 2017.

[27] Ewart J de Visser, Samuel S Monfort, Ryan McKendrick, Melissa A B Smith, Patrick E McKnight, Frank Krueger, and Raja Parasuraman. Almost human: Anthropomorphism increases trust resilience in cognitive agents. *Journal of Experimental Psychology: Applied*, 22(3):331, 2016.

[28] Eric Deng, Bilge Mutlu, and Maja Mataric. Embodiment in socially interactive robots. *arXiv preprint arXiv:1912.00312*, 2019.

[29] Douglas C Engelbart. Augmenting human intellect: A conceptual framework. *Menlo Park, CA*, 1962.

[30] Eric Fassbender and W Heiden. The virtual memory palace. *Journal of Computational Information Systems*, 2(1):457–464, 2006.

[31] Brian J Fogg and Hsiang Tseng. The elements of computer credibility. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 80–87. ACM, 1999.

[32] Terrence Fong, Illah Nourbakhsh, and Kerstin Dautenhahn. A survey of socially interactive robots. *Robotics and autonomous systems*, 42(3-4):143–166, 2003.

[33] Jodi Forlizzi and Carl DiSalvo. Service robots in the domestic environment: a study of the roomba vacuum in the home. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, pages 258–265. ACM, 2006.

[34] Stan Franklin and Art Graesser. Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents. In *International Workshop on Agent Theories, Architectures, and Languages*, pages 21–35. Springer, 1996.

[35] Wilbert O Galitz. *The essential guide to user interface design: an introduction to GUI design principles and techniques*. John Wiley & Sons, 2007.

[36] Michael A Goodrich and Alan C Schultz. *Human-robot interaction: a survey.* Now Publishers Inc, 2008.

[37] D Gouaillier, V Hugel, P Blazevic, C Kilner, J Monceaux, P Lafourcade, B Marnier, J Serre, and B Maisonnier. Mechatronic design of NAO humanoid. In *2009 IEEE International Conference on Robotics and Automation*, pages 769–774, 5 2009.

[38] Heather M Gray, Kurt Gray, and Daniel M Wegner. Dimensions of Mind Perception. *Science*, 315(5812):619, 2007.

[39] Donatien Grolaux, Peter Van Roy, and Jean Vanderdonckt. Migratable user interfaces: beyond migratory interfaces. pages 422–430, 2004.

[40] Robert H Guttman, Alexandros G Moukas, and Pattie Maes. Agent-mediated electronic commerce: A survey. *Knowledge engineering review*, 13(2):147–159, 1998.

[41] Matej Hoffmann and Rolf Pfeifer. The implications of embodiment for behavior and cognition: animal and robotic case studies. *arXiv preprint arXiv:1202.0440*, 2012.

[42] Thomas Holz, Abraham G Campbell, Gregory M P O'Hare, John W Stafford, Alan Martin, and Mauro Dragone. MiRA—Mixed Reality Agents. *International Journal of Human-Computer Studies*, 69(4):251–268, 2011.

[43] Sooyeon Jeong and Cynthia Lynn Breazeal. Improving Smartphone Users' Affect and Wellbeing with Personalized Positive Psychology Interventions. In *Proceedings of the Fourth International Conference on Human Agent Interaction*, HAI '16, pages 131–137, New York, NY, USA, 2016. ACM.

[44] Alan C Kay. The early history of Smalltalk. In *History of programming languages—II*, pages 511–598. 1996.

[45] K Kim, L Boelling, S Haesler, J Bailenson, G Bruder, and G F Welch. Does a Digital Assistant Need a Body? The Influence of Visual Embodiment and Social Behavior on the Perception of Intelligent Virtual Agents in AR. In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 105–114, 10 2018.

[46] Tomoko Koda and Pattie Maes. Agents with faces: the effect of personification. In *Proceedings of HCI 1996*, pages 189–194, 1996.

[47] Robyn Kozierok and Pattie Maes. A learning interface agent for scheduling meetings. In *Proceedings of the 1st international conference on Intelligent user interfaces*, pages 81–88, 1993.

[48] Yamato Kuwahara, Hung-ya Tsai, Yuya Ieiri, and Reiko Hishiyama. Evaluation of a Campus Navigation Application Using an AR Character Guide. In *International Conference on Collaboration and Technology*, pages 242–250. Springer, 2019.

[49] Yezdi Lashkari, Max Metral, and Pattie Maes. Collaborative interface agents. *Readings in agents*, pages 111–116, 1997.

[50] Brenda Laurel. Interface agents: Metaphors with character. *Human Values and the design of Computer Technology*, pages 207–219, 1997.

[51] Kwan Min Lee, Younbo Jung, Jaywoo Kim, and Sang Ryong Kim. Are physically embodied social agents better than disembodied social agents?: The effects of physical embodiment, tactile interaction, and people's loneliness in human–robot interaction. *International journal of human-computer studies*, 64(10):962–973, 2006.

[52] J C R Licklider. Man-Computer Symbiosis. *IRE Transactions on Human Factors in Electronics*, HFE-1(1):4–11, 3 1960.

[53] Henry Lieberman. Autonomous interface agents. In *CHI*, volume 97, pages 67–74. Citeseer, 1997.

[54] Henry Lieberman and Ted Selker. Agents for the user interface. *Handbook of Agent Technology*, pages 1–21, 2003.

[55] Mary Margaret Lusk and Robert K Atkinson. Animated pedagogical agents: does their degree of embodiment impact learning from static or animated worked examples? *Applied Cognitive Psychology*, 21(6):747–764, 2007.

[56] Alexander Maedche, Stefan Morana, Silvia Schacht, Dirk Werth, and Julian Krumeich. Advanced user assistance systems. *Business & Information Systems Engineering*, 58(5):367–370, 2016.

[57] Pattie Maes. Agents that reduce work and information overload. In *Readings in human–computer interaction*, pages 811–821. Elsevier, 1995.

[58] Pattie Maes. Artificial Life Meets Entertainment: Lifelike Autonomous Agents. *Commun. ACM*, 38(11):108–114, 11 1995.

[59] Pattie Maes and Robyn Kozierok. Learning interface agents. In *AAAI*, volume 93, pages 459–465, 1993.

[60] Kohei Matsumura, Yasuyuki Sumi, and Takumi Gompei. Embodiment of guidance robot encourages conversation among visitors. *Journal of Information Processing*, 25:352–360, 2017.

[61] Jan Miksatko, Kerstin H Kipp, and Michael Kipp. The Persona Zero-Effect: Evaluating Virtual Character Benefits on a Learning Task with Repeated Interactions. In Jan Allbeck, Norman Badler, Timothy Bickmore, Catherine Pelachaud, and Alla Safonova, editors, *Intelligent Virtual Agents*, pages 475–481, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.

[62] Nelson Minar, Matthew Gray, Oliver Roup, Raffi Krikorian, and Pattie Maes. Hive: Distributed agents for networking things. In *Proceedings. First and Third International Symposium on Agent Systems Applications, and Mobile Agents*, pages 118–129. IEEE, 1999.

[63] Kenzaburo Miyawaki and Mutsuo Sano. A virtual agent for a cooking navigation system using augmented reality. In *International Workshop on Intelligent Virtual Agents*, pages 97–103. Springer, 2008.

[64] Elaine Mordoch, Angela Osterreicher, Lorna Guse, Kerstin Roger, and Genevieve Thompson. Use of social commitment robots in the care of elderly people with dementia: A literature review. *Maturitas*, 74(1):14–20, 2013.

[65] Omar Mubin, Catherine J Stevens, Suleman Shahid, Abdullah Al Mahmud, and Jian-Jie Dong. A review of the applicability of robots in education. *Journal of Technology in Education and Learning*, 1(209-0015):13, 2013.

[66] Clifford Nass, B J Fogg, and Youngme Moon. Can computers be teammates? *International Journal of Human-Computer Studies*, 45(6):669–678, 1996.

[67] Clifford Nass, Jonathan Steuer, and Ellen R Tauber. Computers are social actors. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 72–78. ACM, 1994.

[68] Kim K. Lee M. Schubert R. Erickson A. Bailenson J.N. Bruder G. & Welch G. Norouzi, N. Walking Your Virtual Dog: Analysis of Awareness and Proxemics with Simulated Support Animals in Augmented Reality. In *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 253–264, 2019.

[69] Hyacinth S Nwana and others. Software agents: An overview. *Knowledge engineering review*, 11(3):205–244, 1996.

[70] Mohammad Obaid, Radosław Niewiadomski, and Catherine Pelachaud. Perception of spatial relations and of coexistence with virtual agents. In *International Workshop on Intelligent Virtual Agents*, pages 363–369. Springer, 2011.

[71] Catherine S Oh, Jeremy N Bailenson, and Gregory F Welch. A systematic review of social presence: definition, antecedents, and implications. *Front. Robot. AI 5: 114. doi: 10.3389/frobt*, 2018.

[72] Richard Pak, Nicole Fink, Margaux Price, Brock Bass, and Lindsay Sturre. Decision support aids with anthropomorphic characteristics influence trust and performance in younger and older adults. *Ergonomics*, 55(9):1059–1072, 2012.

[73] Amit Kumar Pandey and Rodolphe Gelin. A mass-produced sociable humanoid robot: pepper: the first machine of its kind. *IEEE Robotics & Automation Magazine*, 25(3):40–48, 2018.

[74] Paola Pennisi, Alessandro Tonacci, Gennaro Tartarisco, Lucia Billeci, Liliana Ruta, Sebastiano Gangemi, and Giovanni Pioggia. Autism and social robotics: A systematic review. *Autism Research*, 9(2):165–183, 2016.

[75] Romain Pourchon, Pierre-Majorique Léger, Élise Labonté-LeMoyne, Sylvain Sénécal, François Bellavance, Marc Fredette, and François Courtemanche. Is augmented reality leading to more risky behaviors? An experiment with Pokemon Go. In *International Conference on HCI in Business, Government, and Organizations*, pages 354–361. Springer, 2017.

[76] Anand S Rao, Michael P Georgeff, and others. BDI agents: from theory to practice. In *Icmas*, volume 95, pages 312–319, 1995.

[77] Rebecca Reynoso. A Complete History of Artificial Intelligence, 2019.

[78] Oscar Rosello, Marc Exposito, and Pattie Maes. NeverMind: Using Augmented Reality for Memorization. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, UIST '16 Adjunct, pages 215–216, New York, NY, USA, 2016. ACM.

[79] Katie Salen Tekinba$\$cs. Afraid to roam: The unlevel playing field of Pokémon Go. *Mobile Media & Communication*, 5(1):34–37, 2017.

[80] Teri Schamp-Bjerede. What Clippy and Tux can teach us: incorporating affective aspects into pedagogy. *Högskolepedagogisk debatt*, (1):47–56, 2012.

[81] Susanne Schmidt, Gerd Bruder, and Frank Steinicke. Effects of virtual agent and object representation on experiencing exhibited artifacts. *Computers & Graphics*, 83:1–10, 2019.

[82] Ryan M Schuetzler, Justin Scott Giboney, G Mark Grimes, and Jay F Nunamaker Jr. The influence of conversational agent embodiment and conversational relevance on socially desirable responding. *Decision Support Systems*, 114:94–102, 2018.

[83] Byung-Kuk Seo, Kangsoo Kim, and Jong-Il Park. Augmented reality-based onsite tour guide: A study in Gyeongbokgung. pages 276–285, 2010.

[84] Stela H Seo, Denise Geiskkovitch, Masayuki Nakane, Corey King, and James E Young. Poor thing! Would you feel sorry for a simulated robot? A comparison of empathy toward a physical and a simulated robot. In *2015 10th ACM/IEEE*

International Conference on Human-Robot Interaction (HRI), pages 125–132. IEEE, 2015.

[85] Maeve Serino, Kyla Cordrey, Laura McLaughlin, and Ruth L Milanaik. Pokémon Go and augmented virtual reality games: a cautionary commentary for parents and pediatricians. *Current opinion in pediatrics*, 28(5):673–677, 2016.

[86] Claude Elwood Shannon. A mathematical theory of communication. *ACM SIGMOBILE mobile computing and communications review*, 5(1):3–55, 2001.

[87] Mikey Siegel, Cynthia Breazeal, and Michael I Norton. Persuasive robotics: The influence of robot gender on human behavior. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2563–2568. IEEE, 2009.

[88] Nattaon Techasarntikul, Photchara Ratsamee, Jason Orlosky, Tomohiro Mashita, Yuki Uranishi, Kiyoshi Kiyokawa, and Haruo Takemura. The Effect of the Presence of an Embodied Agent in an AR Guiding System. *Transactions of the Virtual Reality Society of Japan*, 25(1):68–77, 2020.

[89] Ravi Tejwani, Felipe Moreno, Sooyeon Jeong, Hae Won Park, and Cynthia Breazeal. Migratable AI. *arXiv preprint arXiv:2007.05801*, 2020.

[90] Sam Thellman, Annika Silvervarg, Agneta Gulz, and Tom Ziemke. Physical vs. virtual agent embodiment and effects on social interaction. In *International Conference on Intelligent Virtual Agents*, pages 412–415. Springer, 2016.

[91] Alan M Turing. Computing machinery and intelligence. In *Parsing the turing test*, pages 23–65. Springer, 2009.

[92] Suzanne Vossen, Jaap Ham, and Cees Midden. Social influence of a persuasive agent: the role of agent embodiment and evaluative feedback. In *Proceedings of the 4th International Conference on Persuasive Technology*, page 46. ACM, 2009.

[93] Weiquan Lu, Linh-Chi Nguyen, Teong Leong Chuah, and Ellen Yi-Luen Do. Effects of mobile AR-enabled interactions on retention and transfer for learning in art museum contexts. In *2014 IEEE International Symposium on Mixed and Augmented Reality - Media, Art, Social Science, Humanities and Design (ISMAR-MASH'D)*, pages 3–11, 2014.

[94] Bob G Witmer and Michael J Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7(3):225–240, 1998.

[95] Chia-Pin Yu, Wei-Chun Chang, and Jittakon Ramanpong. Assessing Visitors' Memorable Tourism Experiences (MTEs) in Forest Recreation Destination: A Case Study in Xitou Nature Education Area. *Forests*, 10(8):636, 2019.