ACTIVE LISTENING TRAINING AGENT

Author: Ososeliase Inegbedion

Student ID: 230387748

MSc Artificial intelligence with Business Strategy

Supervisor: Dr Ulysses Bernardet

School of Engineering and Physical Sciences

Aston University

September 2024

Declaration

I declare that I have personally prepared this assignment. The work is my own, carried out personally by me unless otherwise stated and has not been generated using paid for assessment writing services or Artificial Intelligence tools unless specified as a clearly stated approved component of the assessment brief. All sources of information, including quotations, are acknowledged by means of the appropriate citations and references. I declare that this work has not gained credit previously for another module at this or another University, save for permitted elements which formed part of an associated proposal linked directly to this submission.

I understand that plagiarism, collusion, copying another student and commissioning (which for the avoidance of doubt includes the use of essay mills and other paid for assessment writing services, as well as unattributed use of work generated by Artificial Intelligence tools) are regarded as offences against the University's Assessment Regulations and may result in formal disciplinary proceedings.

I understand that by submitting this assessment, I declare myself fit to be able to undertake the assessment and accept the outcome of the assessment as valid.

Student signature: Ososeliase Inegbedion

Date: 30th September 2024

Acknowledgements

I would like to express my deepest gratitude to the following individuals and institutions who have been instrumental in the successful completion of this project.

First and foremost, I wish to extend my heartfelt thanks to Dr. Ulysses Bernardet, my esteemed supervisor, for his unwavering support, expert guidance, encouragement, and patience throughout the course of this research. His invaluable insights have been crucial to this work.

I am also sincerely grateful to Aston University for providing the necessary resources, laboratory facilities, and an enriching academic environment, which greatly contributed to the achievement of this project's objectives.

Lastly, I would like to extend my appreciation to the participants who generously gave their time and cooperation. Their contributions were indispensable in the successful execution of the experimental trials.

To each of these individuals and organizations, I am profoundly thankful for their support in bringing this project to fruition.

Contents

List of Figures and Tables

List of abbreviations

Mental Health First Aid – MHFA

Human-Computer Interaction – HCI

Augmented reality – AR

Head-mounted displays – HMDs

User interface – UI

User experience – UX

Mixed Reality Toolkit – MRTK

Inverse kinematics – IK

Inertial Measurement Units – IMUs

1. Introduction

In today's fast-paced and interconnected world, effective communication is an essential skill that underpins success in both personal and professional domains. Among the core components of effective communication, active listening stands out as a critical ability that fosters deeper understanding, empathy, and meaningful engagement between individuals. Whether in workplace settings, interpersonal relationships, or educational environments, the capacity to actively listen can significantly influence outcomes and interpersonal dynamics. However, despite its importance, many individuals struggle to develop and refine their active listening skills, often due to the limitations of traditional training methods. This gap highlights the need for innovative and accessible solutions to support the development of this vital competency.

Active listening involves more than just hearing words; it requires attentiveness to verbal and non-verbal cues, empathy, and the ability to process and respond meaningfully to what is being communicated. Existing approaches to active listening training often fall short in engaging participants in interactive, practical, and tailored experiences that can meaningfully enhance their skills. Traditional training methods, such as lectures, workshops, or role-playing exercises, may not provide the immersive and dynamic environments necessary to foster genuine improvement. Consequently, there is a growing interest in exploring technology-driven approaches that can provide more personalized, scalable, and effective active listening training.

This dissertation explores the development of an **Active Listening Training Agent**—an innovative solution designed to address the shortcomings of conventional training by leveraging augmented reality (AR) and immersive technology. By utilizing the Unity game engine and Microsoft HoloLens 2, this project aims to create a virtual human agent capable of simulating diverse and realistic interaction scenarios. The agent will serve as a training tool, enabling users to practice and enhance their active listening abilities in an engaging and interactive environment. Through the use of implicit feedback mechanisms during the interactions and explicit feedback at the end of each session, users will receive real-time insights and targeted guidance to improve their listening skills.

A key focus of this research is the design of interaction dynamics of the virtual agents. These dynamics are critical for ensuring that the training experience is both realistic and challenging, reflecting the complexities of real-world communication. Additionally, the project will investigate the effectiveness of implicit feedback mechanisms—such as tracking eye movements, head positioning, and verbal cues—to provide users with subtle, real-time feedback on their listening behaviours. The ability to vary the virtual agents' demographic characteristics and affective states (e.g., gender, emotional tone) further enhances the training by exposing users to a range of communication styles and contexts, which is especially relevant for applications in areas such as Mental Health First Aid (MHFA).

To assess the effectiveness of the Active Listening Training Agent, the project will incorporate empirical research methods, including the collection and analysis of user performance metrics. These metrics will capture the extent to which participants improve their active listening skills, as well as their ability to integrate feedback into subsequent interactions. Moreover, the dissertation will evaluate user satisfaction with the training system, considering factors such as engagement, realism, and the perceived effectiveness of the AR-based environment.

By introducing a novel, technology-driven approach to active listening training, this research aims to contribute to the advancement of communication skill development in both personal and professional contexts. The Active Listening Training Agent represents a forward-thinking solution that has the potential to revolutionize how individuals' approach and cultivate active listening skills, addressing current training limitations through the integration of immersive AR experiences and real-time feedback mechanisms.

Ultimately, this dissertation seeks to enhance the understanding of how virtual environments, combined with innovative feedback techniques, can effectively foster the development of one of the most crucial yet underdeveloped skills in human communication.

2. Literature review

Active listening is a crucial component of effective communication, characterized by the listener's ability to fully concentrate, understand, respond, and remember what is being said. It plays a fundamental role in various contexts, including interpersonal relationships, professional settings, and conflict resolution. According to scholars like (Argyle, 2013; Rogers, 1975), active listening is essential for building rapport, fostering empathy, and promoting mutual understanding between individuals.

Traditional methods for training active listening skills typically involve classroom-based instruction, workshops, role-playing exercises, and reflective practices. While these approaches provide valuable theoretical knowledge and opportunities for practice, they often lack interactivity, personalization, and real-world applicability. Research by (Gudykunst, 2004), suggests that traditional training methods may not adequately address the complexities of real-life communication situations, leading to limited skill transfer and retention.

Examining the underlying theories that support the development of communication skills, particularly active listening, is crucial in order to comprehend how technological solutions might enhance active listening.

2.1. Theoretical framework

Active listening draws upon a blend of psychological, cognitive, and educational principles. Rooted in theories like Carl Rogers' Person-Centred Theory, which emphasizes empathy and unconditional positive regard (Rogers, 1975), active listening embodies the essence of understanding and empathizing with the speaker's viewpoint. Bandura's Social Learning Theory further elucidates how active listening is a learned behaviour, reinforced by positive responses, thus encouraging its frequent use (Grusec, 1992). From a cognitive-behavioural perspective, active listening entails intricate cognitive processes such as attending to, interpreting, and responding to the speaker's message, highlighting the cognitive dimension of communication (Bodie et al., 2015). Research also suggests that attention to sounds enhances auditory processing, allowing for better detection of sounds and pitch differences (Chun, Golomb and Turk-Browne, 2011), and studies have explored the impact of training on improving listening in noise conditions (Song et al., 2011).

Numerous studies (Tri Yunianika, 2022; Milliner and Dimoski, 2019; Ali and M, 2022 ; Gilakjani and Ahmadi, 2011; Bozorgian and Fakhri Alamdari, 2017; Drollinger, Comer and Warrington, 2006) have demonstrated the efficiency of active listening training by highlighting the value of teaching people active listening techniques, such as using listening methods, developing metacognitive awareness, and improving listening comprehension.

Transactional Analysis is a psychological theory developed by Eric Berne that explores how individuals interact with one another through "transactions," or social exchanges. It identifies three ego states (Parent, Adult, and Child) that influence how people communicate and respond. It sheds light on interpersonal dynamics, emphasizing how active listening fosters positive interactions by validating the speaker's thoughts and feelings (Bodie et al., 2015). Constructivist Communication Theory proposes that communication is not simply the exchange of information but a collaborative process where individuals co-construct meaning through interaction. It emphasizes that understanding is built through active engagement, dialogue, and the exchange of perspectives. In this context, active listening plays a critical role by enabling individuals to clarify and negotiate meanings, fostering a shared understanding. By attentively listening and engaging in open dialogue, participants in communication can collaboratively shape and refine their interpretations of the conversation (Palincsar, 1998).

Practical models such as Egan's SOLER Model (Sit facing the speaker, Open posture, Leaning slightly forward, Eye contact, Relaxed) (Stickley, 2011) and John Keller's ARCS Model (Attention, Relevance, Confidence, Satisfaction) (Reynolds, Roberts and Hauck, 2017) offer structured approaches to understanding and training in active listening, focusing on non-verbal cues and key attention factors. However, it's important to acknowledge their limitations. The SOLER Model's emphasis on nonverbal communication may not fully account for cultural differences in interpreting body language, and its rigid guidelines may feel unnatural in certain contexts (Stickley, 2011). Similarly, while the ARCS Model addresses motivational factors in learning design, its complexity and focus on motivation may overshadow other crucial aspects of instructional design, such as content relevance and instructional methods. These limitations are important to consider when designing training programs for active listening, especially in technology-driven environments like augmented reality (AR).

The GROW Model (Goal, Reality, Options, Will/Way Forward), originally designed for coaching, offers a structured framework for setting goals, understanding current realities, exploring options, and committing to action (Herd, 2015). While the model can be adapted to guide learners in setting listening goals and exploring improvement options, it may overemphasize goal orientation and linear progression, potentially overlooking the complexity of emotional factors and non-linear processes involved in personal development. Additionally, its effectiveness depends on the individual's willingness to self-disclose and engage in open dialogue, which may vary depending on the context and the participant's comfort level (Herd, 2015).

In the context of language learning, research has explored the importance of cognitive strategies and metacognitive awareness in improving listening comprehension (Vandergrift and Tafaghodtari, 2010; Graham and Macaro, 2008). Explicit listening strategy training has been found to enhance learners' self-efficacy and skills (Milliner and Dimoski, 2019). The Four-Stage Listening Process offers structured approaches to understanding and training in active listening by receiving the auditory stimuli or information being communicated, understanding, evaluating, and responding (Anchal Malh, Wood and Chung, 2023). However, this model presents listening as a linear process, potentially oversimplifying the complex and dynamic nature of real-life listening experiences (Anchal Malh, Wood and Chung, 2023). It also predominantly focuses on cognitive aspects, possibly overlooking the influence of affective and behavioural factors.

In light of these limitations, a conceptual framework for the Active Listening Training Agent project can be proposed by integrating elements of active listening training, metacognitive awareness, cognitive strategies, and neural mechanisms underlying effective listening. The framework should address both cognitive processes and practical models while incorporating the latest technological advancements, especially in augmented reality (AR).

2.2. Augmented reality (AR) in Training and Education

Augmented Reality (AR) is transforming training and education by enhancing learning experiences through interactive digital content overlaid on the physical world. AR enables learners to engage with complex concepts in a tangible way, increasing comprehension and engagement. It is particularly effective in subjects requiring spatial understanding, like geometry or anatomy, by allowing students to manipulate 3D models in real-time. AR also promotes active learning by simulating real-world scenarios, fostering problem-solving and collaboration. Additionally, it supports personalized learning, adapting content to individual needs for more effective instruction.

A significant advantage of AR in education is its ability to enhance student engagement and understanding. For instance, AR applications allow for interactive learning experiences that can make complex subjects more accessible. Studies indicate that AR can significantly improve students' spatial abilities and visualization skills, particularly in subjects like geometry and science, where understanding three-dimensional relationships is crucial (Sun & Chen, 2019; Eg Su et al., 2022). AR has also been effectively utilized in medical education, where it aids in the training of surgical techniques by providing real-time overlays of anatomical structures during practice sessions, thereby enhancing clinical competence among students (Dhar et al., 2021; Dandıl, Serin and Şenol, 2022).

AR facilitates remote and distance learning, which has become increasingly important after the COVID-19 pandemic. The portability of AR applications on mobile devices allows students to engage with learning materials anytime and anywhere, thus overcoming geographical barriers (Alzahrani, 2020; Taha et al., 2023). This flexibility is particularly beneficial in fields such as medical training, where students can practice procedures in an augmented reality environment before performing them on actual patients (Dhar et al., 2021). The use of AR in e-learning contexts has been associated with positive learning outcomes, as it allows for the creation of immersive and interactive educational content that can cater to diverse learning styles (Rebollo et al., 2021; Sharmin & Chow, 2020).

Despite its benefits, the adoption of AR in educational settings is not without challenges. Factors such as poor internet connection, lack of training for educators, and limited access to AR resources can hinder the effective implementation of AR technologies in learning spaces (Taha et al., 2023; Alkhattabi, 2017). Educators need to be adequately trained to integrate AR into their teaching practices effectively. Research has shown that teacher competence in utilizing AR-based media is crucial for maximizing its educational potential (Karlimah Karlimah et al., 2023). Addressing these barriers is essential for the widespread adoption of AR in education, ensuring that both educators and students can fully leverage its capabilities.

2.3. Multimedia learning

The incorporation of multimedia components is crucial for augmenting the learning process, especially when paired with interactive technologies.

Multimedia learning, a dynamic approach to enhancing learning outcomes, integrates various forms of media such as text, images, audio, video, and interactive elements to optimize learning outcomes. At its core lies the Cognitive Theory of Multimedia Learning by Richard Mayer, which illuminates the mechanisms through which multimedia influences cognition and engagement (Mayer, 2014). Mayer's theory posits several principles that govern effective multimedia learning.

The Dual Coding Theory highlights the significance of presenting information through both verbal and visual channels simultaneously (Mayer, 2014). Building upon this, the Multimodal Principle suggests that employing multiple modalities, such as text paired with images, can further augment learning compared to utilizing a single modality.

The Contiguity Principle emphasizes the importance of temporal and spatial proximity between related words and pictures. When information is presented in close association, learners can more readily integrate verbal and visual representations, leading to enhanced comprehension and synthesis (Mayer, 2009).

The Modality Principle advocates for selecting appropriate multimedia modalities to convey information effectively. For instance, complex visual concepts may be best communicated through animations or videos, providing dynamic representations that aid in conceptualization (Mayer, 2001).

Conversely, the Redundancy Principle warns against unnecessary redundancy in multimedia presentations, as it can overload the cognitive system and impede learning (Mayer, 2009). Thus, it's imperative to streamline content and avoid extraneous information that detracts from the learning goals.

The Coherence Principle underscores the importance of organizing multimedia content in a logical and cohesive manner (Mayer, 2009). By minimizing extraneous information and maintaining a coherent structure, learners can navigate the material more efficiently, reducing cognitive load and enhancing understanding.

Moreover, the Personalization Principle highlights the benefits of tailoring multimedia content to match learners' interests and prior knowledge (Mayer, 2020). By personalizing the learning experience, educators can foster greater engagement and motivation, ultimately leading to improved learning outcomes.

Through thoughtful application of these principles, educators can harness the power of multimedia to optimize engagement, facilitate deeper understanding, and empower learners on their educational journey.

Users' interactions with the information are primarily based on multimedia learning, but the idea of Human-Computer Interaction (HCI) and the usage of virtual agents give the experience a more customized, interactive element (Knörzer, Brünken and Park, 2016; Plass et al., 2014; Lawson et al., 2021).

Multimodal feedback systems can be integrated with aural cues to further enhance the effectiveness of AR in training. Backchanneling and flexible turn-taking systems that foster more organic and interesting interactions in training contexts (Lala et al., 2017). These systems can aid in the development of a more sophisticated understanding of active listening among trainees by simulating human-like replies, including non-verbal cues and the nuances of conversational dynamics. This all-encompassing method of instruction guarantees that students are honing their comprehension and responding abilities to a wide range of communicative cues in addition to honing their listening skills.

The application of AR in training contexts has been shown to enhance motivation and engagement among learners. AR-based applications significantly improve training outcomes in various fields, including occupational safety and health (Kamal et al., 2021). This evidence suggests that similar methodologies could be applied to active listening training, where the immersive nature of AR can lead to higher levels of participant engagement and retention of skills. By creating a more interactive and stimulating learning environment, AR can facilitate deeper learning experiences that traditional training methods may not achieve.

2.4. Human-Computer Interaction (HCI) and Virtual Agents

Virtual human agents are computer-generated characters that simulate human-like behaviours and interactions. These agents simulate human behaviours and interactions, utilized in training, customer service, and therapeutic interventions (Sukthankar, Mandel and Sycara, 2007).

Performance optimization guides efforts to minimize latency and ensure smooth interaction responsiveness. And adaptivity and personalization are integral, enables virtual environments to tailor experiences to individual user preferences and characteristics (Dongas, 2018).

While Human-Computer Interaction (HCI) and Virtual Agents principles provide valuable frameworks for designing interactive digital systems, they are not without limitations. HCI principles often prioritize designing for the average user, potentially overlooking the diverse needs and preferences of individuals. Virtual agents, aiming to replicate human-like interactions, may also struggle to adapt to the complexities of human behaviour, leading to challenges in accurately modelling real-world interactions (Burgoon et al., 2000).

Ethical considerations in virtual agent development require addressing specific challenges such as ensuring user privacy by safeguarding sensitive data collected during interactions, minimizing bias in agent behaviour through diverse and representative datasets, and fostering trust by making the agent's decision-making processes transparent and explainable to users. These issues are critical to ensuring that virtual agents not only perform effectively but also align with ethical standards in real-world applications (Burgoon et al., 2000). Additionally, technical limitations, including accuracy and scalability issues in artificial intelligence and natural language processing, can hinder the effectiveness of virtual agents. Cultural sensitivity and long-term engagement also present challenges, as HCI and Virtual Agents principles may not adequately address cultural differences or sustain user interest over time.

In parallel, the exploration of virtual agents within HCI research opens up new avenues for enhancing user experience and interaction. These intelligent virtual entities can serve a multitude of roles, from tutoring and guidance in educational settings to providing emotional support and companionship in healthcare or therapy contexts (King et al., 2018). By incorporating insights from UX design and leveraging the capabilities of virtual agents, we can create digital experiences that not only meet technological requirements but also resonate deeply with human users, ultimately enhancing learning, interaction, and overall user satisfaction.

2.5. Application of AR and Virtual Agents in Active Listening Training

The integration of augmented reality (AR) and virtual agents in communication training, particularly for active listening, is a relatively novel yet promising field. While traditional methods are widely used to teach active listening, the approaches often fail to replicate the complexity of real-world interpersonal interactions (Tennant, Butler and Long, 2023). In contrast, AR and virtual agents offer dynamic, immersive environments that can simulate a wide variety of communication scenarios, enabling learners to engage in interactive and personalized experiences.

AR enhances the training experience by overlaying virtual elements onto the real world, which can include simulated conversations with virtual agents. This method has been shown to improve engagement and retention of skills among learners. For instance, auditory augmented reality (AAR) can create virtual sound sources from the agent that enrich the listener's environment, thereby enhancing the training experience by providing realistic auditory cues that are essential for effective active listening (Neidhardt and Zerlik, 2021). This immersive approach allows trainees to practice their listening skills in a context that closely resembles real-life situations, which is crucial for developing competence in active listening.

The use of head-mounted displays (HMDs) in AR applications has been shown to impact localization accuracy of sound sources, which is vital for active listening training. It highlights that the human brain utilizes interaural differences to determine the location of sounds, and any alterations in auditory perception due to HMDs can affect this ability (Poirier-Quinot and Lawless, 2023). This finding highlights the importance of designing AR training tools that account for these auditory localization challenges, ensuring that learners can accurately perceive and respond to auditory cues during training sessions.

Training in active listening can be improved even more by integrating virtual agents. Active listening strategies can be used by conversational user interfaces (UIs) like virtual assistants to promote user self-disclosure and enhance the user experience (Cho et al., 2022). These virtual agents can give trainees instant feedback by acting as a responsive listener, which reinforces good listening habits and enables them to make real-time adjustments to their communication tactics. In addition to facilitating the development of skills, this contact creates a sense of engagement and connection, both of which are necessary for successful learning.

In addition to the technological advancements, the psychological aspects of active listening training must also be considered. discuss the importance of active listening in informal helping conversations, noting that trained listeners are perceived as more helpful and supportive (Bodie et al., 2015). This highlights the need for training programs to not only focus on the mechanics of listening but also on the emotional and relational dimensions of communication. By integrating AR technologies that simulate real-life emotional scenarios, trainees can develop a more profound understanding of the impact of their listening behaviours on interpersonal relationships.

Moreover, the role of music and auditory experiences in enhancing active listening skills cannot be overlooked. emphasize that engagement with music can improve emotional recognition and auditory processing abilities, which are critical components of active listening (Calvin Kai-Ching Yu and Samuel Po-Shing Wong, 2023). By incorporating musical elements into AR training modules, educators can leverage these benefits to enhance learners' auditory skills and emotional intelligence, further enriching the training experience.

The implications of these findings extend to various fields, including education, healthcare, and corporate training. For instance, in nursing education, AR has been shown to improve biomedical education by providing enriched sensory inputs that facilitate learning (Chun-Wai Ma, 2019). This approach can be adapted for active listening training in healthcare settings, where effective communication is paramount for patient care. By utilizing AR technologies, nursing students can practice their listening skills in simulated patient interactions, thereby improving their ability to provide empathetic and effective care.

2.6. Gaps and Opportunities

The integration of Augmented Reality (AR) and virtual agents into training programs presents both significant challenges and promising opportunities. These technologies have the potential to enhance learning experiences across various domains, including healthcare, education, and emergency response. However, their implementation is not without obstacles, such as technological limitations, user acceptance, and the need for effective design.

One of the primary challenges in utilizing AR and virtual agents for training is the technological complexity involved in creating realistic and effective simulations. For instance, while virtual reality (VR) has shown effectiveness in medical education, particularly in surgical training, there remains a gap in the evidence supporting the use of certain technologies like mobile VR and virtual dissection tables (Jiang et al., 2022). Additionally, the fidelity of the virtual environment can significantly affect user engagement and learning outcomes. High-fidelity simulations may inadvertently distract users or create unrealistic expectations regarding the capabilities of virtual agents, complicating the training process (Bruijnes, Linssen and Heylen, 2019). Thus, achieving the right balance between realism and usability is crucial for effective training outcomes.

Another obstacle is the high cost of developing and maintaining AR systems, as well as the need for advanced hardware such as HoloLens or AR-enabled mobile devices (Palmarini et al., 2018). These costs can limit accessibility, particularly in smaller educational or training institutions.

Another challenge is the learning curve associated with using AR technologies. Many educators and trainers may lack the technical expertise required to effectively integrate AR into their curricula (Karlimah Karlimah et al., 2023). Without proper training, there is a risk that AR-based training programs may not be utilized to their full potential, leading to suboptimal learning outcomes. Additionally, technical issues such as system lag, poor connectivity, or hardware malfunctions can disrupt the immersive experience, diminishing the effectiveness of AR-based learning environments.

In addition to technological hurdles, while virtual agents can simulate a wide range of human behaviours, they are still limited in their ability to fully replicate the nuanced and unpredictable nature of human interactions (Karlimah Karlimah et al., 2023). Virtual agents may struggle to adapt to the spontaneous and context-specific behaviours of human users, which can result in less authentic interactions. This limitation may hinder the development of certain soft skills, such as empathy, that are crucial for active listening.

User acceptance of AR and virtual agents is a critical factor influencing their effectiveness. Studies have indicated that the design of virtual agents, including their appearance and behaviour, can significantly impact user engagement and learning (Tanaka & Nakamura, 2022). For example, research has shown that users' preferences for virtual agent design can vary based on individual characteristics such as gender and age, which suggests that personalized approaches may enhance training effectiveness (Tanaka & Nakamura, 2021). The embodiment of virtual agents, whether they are presented as fully animated characters or as speech-only interfaces can influence user rapport and the effectiveness of communication in training scenarios (Wang et al., 2019). Therefore, understanding user preferences and behaviours is essential for the successful integration of these technologies into training programs.

Despite these challenges, AR and virtual agents also present exciting opportunities for innovation in training. The ability to customize virtual agents to reflect different demographic characteristics, communication styles, and emotional states provides a unique opportunity to create highly tailored training experiences. This personalization can be particularly beneficial in sectors where diversity and inclusivity in communication are critical, such as in mental health support or global business settings.

The potential for scalability in AR-based training systems is a significant advantage. Unlike traditional training methods, which may require in-person facilitators and limited participant numbers, AR-based systems can be deployed to a large number of users simultaneously. This scalability makes AR an ideal solution for organizations looking to implement cost-effective, widely accessible training programs (Johnson et al., 2022).

The data-driven insights provided by AR systems, such as real-time feedback on user performance and detailed analytics on listening behaviours, offer valuable opportunities for refining and improving training programs. By capturing metrics such as eye movements, head position, and response accuracy, AR-based systems can provide trainers and learners with actionable insights into their communication skills, leading to more targeted and effective skill development.

2.7. Research Questions:

1. How do different user interaction dynamics affect the effectiveness of active listening training in virtual environments?

2. What are the most effective implicit feedback mechanisms for enhancing active listening skills during training sessions?

3. How well can active listening level be measured in an augmented reality environment?

3. Methodology

In this chapter, our primary objective is to design and implement an active listening training agent using the Unity game engine and HoloLens 2. This project seeks to seamlessly integrate various interactive elements, including lifelike gestures, natural speech, realistic animations, advanced tracking systems, and synchronized dynamics. The goal is to craft an immersive and engaging user experience where virtual agents simulate real-life conversations within the user's physical environment. Simultaneously, the system monitors and analyses the user’s listening behaviours in real-world scenarios, providing valuable insights into their active listening skills.

3.1. Participants

Participants for this study were recruited through email or in-person messages, targeting postgraduate students at Aston University, particularly in the AI department, as well as other interested individuals. The study is open to those aged 18 or older, fluent in English, with normal or corrected-to-normal vision, and no history of epilepsy, mobility issues, breathing difficulties, or severe motion sickness. Participants sign a consent form confirming voluntary involvement and understanding of the study criteria. These criteria were communicated to potential participants through a Participant Information Sheet (PIS) distributed via email. With the option to withdraw at any time without consequences.

To achieve the necessary level of social presence with the Virtual Agent, the experiments employed a high-resolution immersive display using the HoloLens 2 AR Head-Mounted Display (HMD). The decision was made to conduct in-person experiments in a controlled, well-equipped laboratory. A comprehensive risk assessment, essential for obtaining ethics approval, was completed beforehand, and all recommended health, safety, and ethical protocols were strictly followed throughout the study to mitigate potential risks associated with AR research.

3.2. Equipment

The development of the system relied on several advanced hardware and software components, which enabled a robust and immersive training environment.

Unity is a highly flexible and comprehensive platform for building 2D, 3D contents and interactive environments. It was used to design the virtual agents and construct the interactive training scenarios. Unity supports a variety of animation systems, including state machines and event-driven mechanisms, which allowed for the realistic behaviour of the virtual agents. The engine also facilitated the integration of multiple input systems, enabling the real-time collection of user data, such as eye movements, head position, and voice input. Unity's extensive libraries, plug-ins, and customizable scripting capabilities (using C#) were essential for tailoring the application to the specific requirements of active listening training in augmented reality (AR).

The HoloLens 2, an advanced AR headset, provided the hardware platform for delivering the interactive experience. It features a combination of sensors and display technology, creating a mixed reality environment where users can interact with virtual objects overlaid on the physical world. The key hardware components of the HoloLens 2 used in this system include:

Eye-Tracking Sensors:

The HoloLens 2 is equipped with infrared-based eye-tracking technology, which captures the user's gaze direction with high accuracy. This data is invaluable for understanding user engagement by monitoring where and how long the user is looking at the agent and surrounding environment. The sensors operate at low latency, ensuring smooth tracking that does not disrupt the immersive experience.

Head-Tracking System:

The HoloLens 2 employs a sophisticated tracking system that combines inertial measurement units (IMUs) with external environment mapping sensors. This advanced technology precisely monitors the user's head position and orientation in three-dimensional space, allowing for accurate recording of head movements. Such precision is crucial for analysing how users physically interact with virtual agents and their surroundings. Moreover, the head-tracking system plays a vital role in maintaining immersion by ensuring that virtual content remains perfectly aligned with the user's field of view, creating a seamless blend of digital and physical worlds.

Spatial Mapping and Gesture Recognition:

The HoloLens 2 contains depth-sensing cameras that enable real-time spatial mapping, allowing the system to understand and interact with the physical environment. This technology supports spatial awareness, meaning virtual objects can be placed and interacted with in relation to real-world surfaces and spaces.

Voice Detection and Microphone Array:

The HoloLens 2 features a microphone array that captures high-quality audio. Custom algorithms were used to detect volume spikes in the audio input, which served as indicators of verbal interruptions or active engagement. The low-latency audio capture, combined with noise cancellation, ensured accurate detection of these spikes, even in noisy environments.

Data Logging Systems:

To complement the immersive experience and track user behaviour, several custom data logging systems were built within the Unity engine. These systems captured data, including gaze direction, head movements (both position and orientation), and microphone input. These inputs were logged and stored in real-time, leveraging the HoloLens 2's internal sensors and Unity's ability to handle high-frequency data processing.

The integration of HoloLens 2's advanced sensors with Unity's development environment enabled a seamless, interactive training system in augmented reality, enhancing the realism and effectiveness of the active listening training agent.

3.3. Experiment design

When evaluating the effectiveness and impact of an Active Listening Training Agent developed using Unity and HoloLens 2, it is important to consider the structure of the training environment, the mechanisms for capturing user behaviours, and how users interact with the system. These components are critical in determining how well the system fosters and measures active listening skills in an augmented reality (AR) environment. This section will detail the session flow, data collection methods, and user experience evaluation in the experiment, ensuring a comprehensive approach to measuring the effectiveness of the AR agent.

The session flow of the AR system defines the sequence of user interactions with the training agent and the content delivered during the training. The system comprises seven scenes that follow a specific order: menu, instruction, topic 1 with session, or topic 2 with session 2 (randomly generated), break, and the remaining topic and session. This structure ensures that users are introduced to the system and provided with two main content sessions, each followed by a reflective break or transition, while allowing for variation in the order of the topics to enhance engagement and adaptability.

The system incorporates two main types of transitions:

1. **Button-based Transitions:** In scenes such as the menu, instruction, topic 1, break, and topic 2, users advance to the next scene by pressing buttons presented on a canvas. This method gives users control over when they are ready to proceed and provides them with a sense of agency during the training.
2. **Timer-based Transitions:** In session and session 2, where the focus is on listening to characters, the system uses timers (set for approximately 5 minutes) to transition automatically to the next scene. This design forces users to stay focused during the speaking sessions without the option

to manually skip or move forward. By logging users' behaviour throughout this timed interval, we can examine whether they stay engaged during the entire duration.

The two session scenes present different types of content and are randomly ordered in each trial to minimize order bias. One session involves a character discussing how artificial intelligence will impact the job market, featuring minimal animation, while the other session involves a character narrating a trip to Universal Studios with dynamic and expressive gestures. This randomization ensures that the order of content delivery does not unduly influence user performance or engagement levels.

Evaluating session flow is crucial, as it determines the user's journey through the training environment and the balance between interactive and passive elements. Users' experiences transitioning between scenes, particularly between the more interactive scenes (menu, topic selection) and the timed listening sessions, are key factors in understanding how well the flow supports sustained focus on the content.

Data Collection

Data collection is fundamental to understanding how users engage with the Active Listening Training Agent. The system uses three primary data streams: eye tracking, head tracking, and voice detection. These allow for a comprehensive analysis of user behaviour and attentiveness during the training sessions.

Eye Tracking

The system uses colliders in Unity to monitor where the user is looking at every 20 milliseconds interval. Specifically, it tracks whether the user's gaze is directed at the character's head, upper body, or elsewhere (away from the character). Eye tracking is crucial for measuring visual attention, as maintaining eye contact or focusing on the speaker's face or body is an indicator of active listening. A higher proportion of time spent looking at the character's head is considered a strong signal of attentiveness.

The colliders are strategically placed on the virtual character to ensure precise tracking, with data logs detailing every change in gaze position. This high precision allows the system to detect even subtle deviations in focus, helping to differentiate between highly engaged users and those who may be distracted or looking away.

Head Tracking

The system continuously monitors the user's head movements, capturing both position and rotation in space relative to the character. This is logged with the same high frequency (every 20 milliseconds) to ensure accurate tracking of the user's orientation during the session.

The head tracking system detects whether the user is physically oriented toward the virtual speaker. It measures how frequently and by how much the user moves their head away from the character. Large or frequent deviations from the speaker may indicate a lack of focus or engagement, while steady head orientation toward the character could suggest sustained attentiveness.

By combining head tracking data with eye tracking data, we can more accurately measure the user's overall body language and postural engagement, providing a richer understanding of how attentively they are listening.

Voice Detection

The system includes a voice detection feature that logs spikes in microphone input volume to detect when users speak or make audible sounds during the sessions. This functionality captures any verbal interruptions, which are then recorded as potential breaks in attentiveness.

Voice detection works by sampling the microphone input continuously, measuring the average volume and flagging instances where the volume exceeds a set threshold. These flagged instances are logged as interruptions, which could reflect either an active attempt to respond to the content or a breakdown in focus (e.g., talking to oneself or commenting out loud).

This data stream provides insights into the verbal engagement of the user and their tendency to interrupt the speaker, which can be analysed alongside other behavioural measures (eye and head tracking) to determine overall listening effectiveness.

The combination of these data collection methods allows for a nuanced understanding of how users interact with the AR training agent. By capturing real-time data on visual attention, physical orientation, and verbal responses, we can quantify levels of active listening and identify patterns in user behaviour that correlate with higher or lower listening effectiveness.

User Experience

To complement the objective data collected through eye, head, and voice tracking, we will also gather subjective data through three questionnaires administered at different points in the experiment. These questionnaires assess users' recall, perceptions of the training system, and overall experience:

- **Session 1 Recall Questionnaire:** After the first session, participants are given a questionnaire that tests their recall of the content discussed by the character. This assesses how well users retained information on the topic of artificial intelligence or Universal Studios, depending on which session was presented first.
- **Session 2 Recall Questionnaire:** After a break and the completion of the second session, users are presented with another questionnaire focusing on the second session's content. This recall test serves a similar function, assessing how well the user remembers the key points discussed during the second session.
- **User Experience Questionnaire:** After completing both sessions, participants are asked to fill out a third questionnaire that incorporates a Likert scale to assess their overall experience with the AR training system and the different agents. This includes questions about ease of use, engagement with the virtual characters, and perceived effectiveness of the training in improving their active listening skills.

These questionnaires are essential for evaluating the subjective experience of the users. By comparing their self-reported perceptions with the behavioural data collected during the sessions, we can draw comprehensive conclusions about the effectiveness of the training system in fostering active listening and its overall usability.

By structuring the experiment around session flow, data collection, and user experience, we create a robust framework for evaluating the effectiveness of the Active Listening Training Agent in AR. The combination of real-time tracking data with user feedback provides a comprehensive assessment of how the AR environment influences active listening behaviours, allowing for meaningful insights into the system’s potential for training in augmented reality environments.

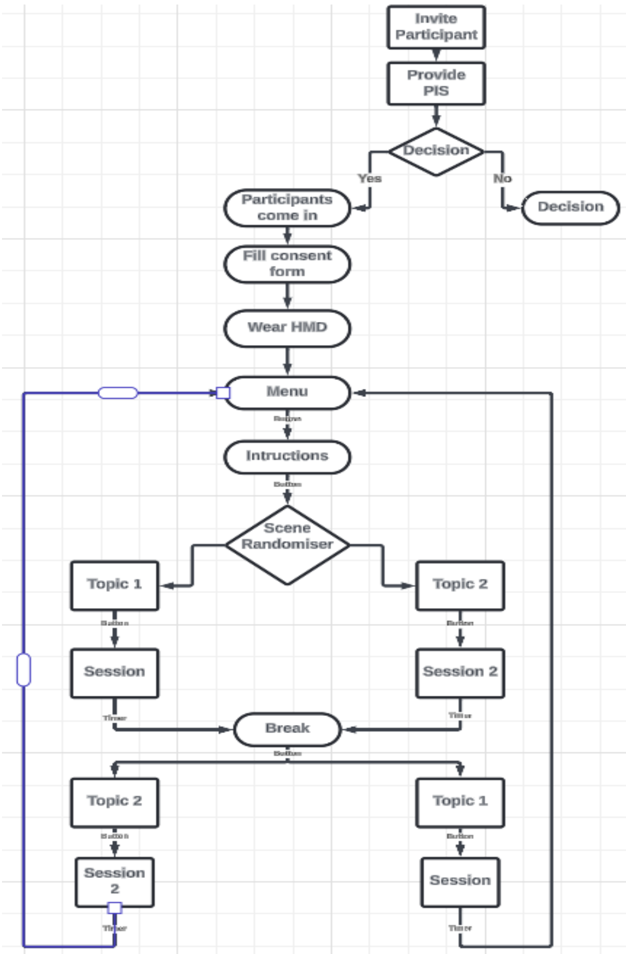

Fig 1 – Experiment workflow

3.4. Implementation

3.4.1. Scenes

The session flow of the AR system defines the sequence of user interactions with the training agent and how the content is delivered throughout the training. Built on Unity, the system comprises seven distinct scenes, each guiding the user through various phases of the training experience.

3.4.1.1. Scene transitions

Button-based transitions are used in scenes such as the menu, instruction, break, and both topic information sections. These transitions allow users to control their progression through the training by interacting with buttons that are displayed on a Screen Space Canvas. This canvas is anchored directly to the screen, ensuring that it remains responsive regardless of changes to the camera or game environment. The UI is built within a Menu Scene using Unity's UI components, where elements like titles and buttons are created. For example, the play button in the menu scene is linked to a function called ButtonHandlerPlay() in the _NextPageALT.cs script. When clicked, this button triggers a transition to the next scene, calculated using a SceneRandomizer script. Scene transitions are handled asynchronously through Unity's SceneManager, ensuring smooth changes between scenes without interruptions.

In contrast, Timer-based transitions sessions are time-constrained, typically lasting five minutes, and the user cannot manually skip ahead or exit early. A countdown timer is displayed on the screen, updated in real-time through the _GameTimerALT script. This script manages the countdown, updating the remaining time every second and displaying it in a Text element. Once the time runs out, the system automatically transitions to the next scene using the SceneRandomizer. The StartCountdown() coroutine handles the timed transitions, and the user's ability to skip these sessions is removed to maintain focus and consistency in the experience.

Fig 2 – Screen space canvas with UI button

3.4.1.2. Scene Randomization

The SceneRandomizer script is attached to a SceneManager GameObject, which controls the overall flow of scenes. Once a randomized sequence is generated at the start of the session, it remains consistent throughout the training to ensure that all users experience both topics, either beginning with topic1 -> session or topic2 -> session2, followed by the other pair after the break.

The SceneRandomizer script follows a straightforward yet effective approach for shuffling the topic and session scenes. The ShuffleScenes method ensures that the two topic-session pairs are shuffled randomly. The scenes are identified by their build indices in Unity's build settings: topic1 (index 2), session (index 3), topic2 (index 5), and session2 (index 6). The script creates a list containing these pairs and then shuffles them using the Fisher-Yates shuffle algorithm to ensure randomness. Once shuffled, the script assigns the order of the scenes to an internal array called sceneOrder, which stores the shuffled sequence. This array is used later to determine the next scene after the introduction and break scenes.

The **GetNextScene** method uses the current scene index to determine the next scene. After the introduction, it checks the sceneOrder array to decide whether to load **topic1** or **topic2** first.

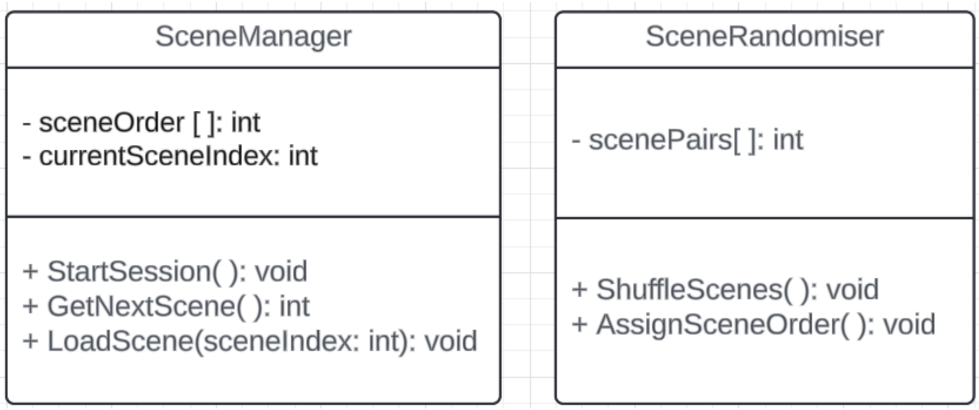

Fig 3 – Screen randomiser UML diagram

3.4.2. Developing augmented reality Agents

During implementation, the goal was to create an augmented reality environment for real-time interaction between virtual agents and users. This was achieved by integrating Unity with HoloLens 2, using the Microsoft Mixed Reality Toolkit (MRTK) and the Universal Render Pipeline (URP) to enhance system performance and visual quality for immersive experiences.

3.4.2.1. HoloLens 2 connection

The connection between Unity and HoloLens 2 was established using MRTK, which is designed specifically for building mixed-reality applications. By incorporating MRTK into the Unity project, the system was tailored to HoloLens 2 features, enabling functionalities such as spatial mapping, hand and voice input controls, and holographic rendering. These configurations allowed for real-time interaction between users and virtual agents, ensuring that the augmented reality experience was both responsive and immersive. URP was employed to optimize rendering performance, striking a balance between visual fidelity and real-time responsiveness, which is essential for a smooth and engaging AR training environment.

3.4.2.2. Character

To develop realistic and responsive virtual agents, the CHL Character Library, a set of pre-built models suited for AR training scenarios, was integrated. Additionally, the Reallusion Auto Setup for Unity package was installed to ensure 3D models created in Reallusion tools like Character Creator and iClone were properly configured for Unity's URP shaders, preventing visual inconsistencies. Once integrated, these characters were incorporated into the AR environment. The project was then connected to HoloLens 2 for testing, where virtual agents were confirmed to appear realistically in the physical environment.

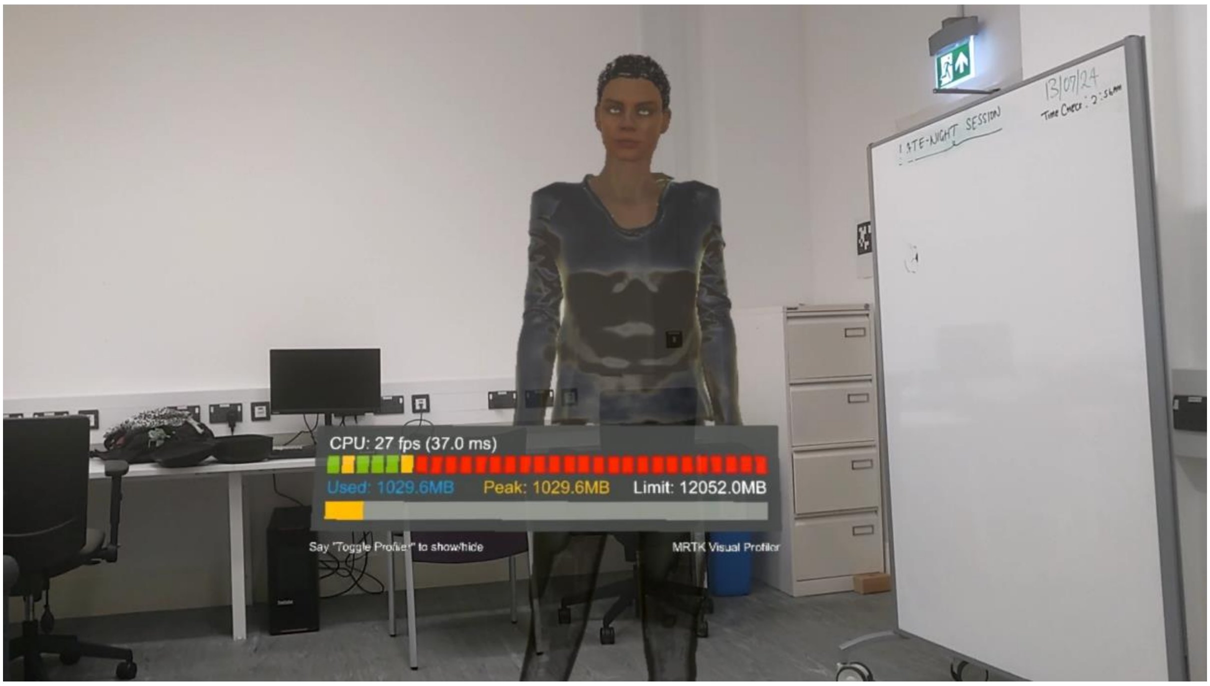

Fig 4 – Idle state of 1st virtual character

3.4.3. Audio

The implementation of audio was a crucial component in delivering an immersive and interactive experience for the virtual characters. The audio not only provided the spoken content but also synchronized with the characters' lip movements, enhancing the realism of the sessions. This section outlines the process of generating the audio content and executing it within the Unity environment to ensure smooth interaction.

3.4.3.1. Audio generation

The process of audio generation began with drafting two scripts aligned with the virtual characters' dialogue in the training sessions. These scripts were then converted into audio files using play.ht, an AI-based text-to-speech tool that produces high-quality, natural-sounding speech. The generated audio files were organized in an "Audio" folder within Unity, acting as the central repository for audio assets.

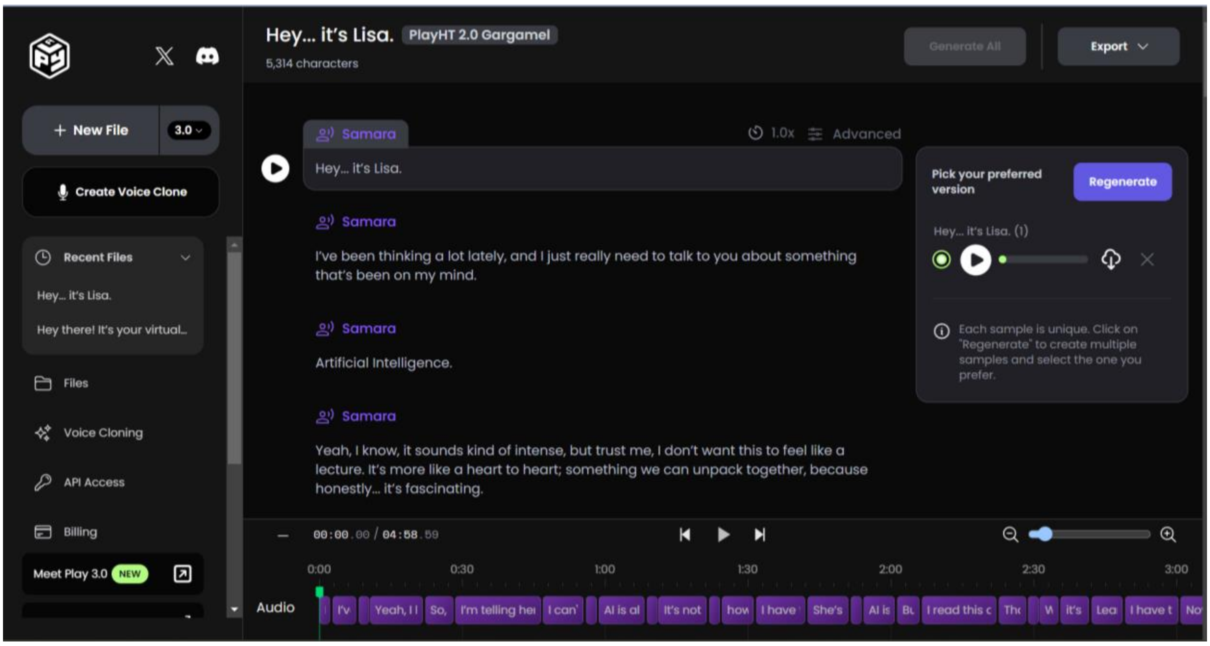

Fig 5 – Play.ht audio generator

To integrate the audio with the characters, the SALSA LipSync plugin was used. This plugin automatically synchronizes the lip movements of virtual characters with audio playback. The audio files were assigned to the characters' AudioSource components, allowing the SALSA LipSync system to manage real-time synchronization of mouth movements during playback.

3.4.3.2. Audio execution and delay

To enhance user experience, a delay mechanism was implemented to provide users with time to orient themselves within the virtual environment before the audio began. A custom C# script, named AudioDelay, was developed to introduce a 4-second delay before the audio started. This was achieved by referencing each character's AudioSource component and using the PlayDelayed() method to control the timing of the playback. This approach allowed users to locate the speaking agent before the dialogue commenced, contributing to a smoother, more engaging experience.

3.4.3.3. lip sync

Lip synchronization was achieved through the integration of the SALSA LipSync Suite and Character Library models, which ensured realistic and synchronized speech. This was a critical element in making the virtual characters' speech appear natural.

The implementation involved importing essential Unity packages, including SALSA LipSync Suite, oneclickbase, and oneclickcc, which provided the necessary assets and scripts for synchronizing lip movements with audio.

Character model preparation involved selecting the CC_Base_Body to apply lip sync to the mouth and face. The SALSA One-Click setup for Reallusion Character Creator (CC4) models was then implemented, automatically synchronizing lip movements with audio for natural speaking animations.

3.4.4. Gestures

The implementation of gestures in the system is a critical component for enhancing the realism and engagement of the virtual agents. The gestures include animation, gaze control, and facial expressions to ensure the character delivers non-verbal cues that are vital for active listening training. These components work together to create a more immersive and dynamic interaction experience.

3.4.4.1. Animation

The animation component plays a critical role in synchronizing character movements with audio to create an immersive and realistic experience. The implementation involves integrating Unity's Animator Controller with the SALSA LipSync system, allowing animations to trigger at precise moments during audio playback.

Initially, both 3D character models were equipped with Animator components to ensure smooth transitions. Animator Controllers (Animation1Controller and Animation2Controller) were then created and assigned to the characters for effective animation management. Various animation clips, such as Idle, Waving, and Talking, were imported in supported formats like ".anim" and ".FBX" and assigned to states controlled by Boolean parameters (e.g., isIdle, isWaving, isTalking). This setup enabled dynamic interaction, with animations transitioning smoothly based on the defined parameters.

The Animator scripts synchronize animation flow with audio by referencing Unity's Animator and SALSA LipSync components. In the Start method, the SALSA component is detected and initialized, while the audio begins playing, and the character's animation is set to idle. Synchronization is handled through the SyncAnimationsWithAudio() coroutine, which triggers animations at specific audio timestamps. For example, the waving animation is activated at four seconds by setting the isWaving parameter to true.

To maintain proper animation flow, the ResetAllBooleans() method resets other parameters before triggering each new animation, ensuring no overlap. After each animation, the character returns to idle through the SetIdleState() method, restoring a neutral position.

Testing and fine-tuning ensured the alignment of animations with audio playback. The SALSA component provided precise timing, while Invoke calls controlled the duration of each animation before transitioning back to idle.

Fig 6 – Animated state of 2nd virtual character

3.4.4.2. gaze controller

Gaze control is crucial for creating immersive experiences, as it allows the virtual character to follow the head movements of HoloLens 2 users, enhancing interaction realism. This behavior is implemented through a custom GazeController script, which directs the character's eyes and head to track the user's camera position within the virtual environment.

The main camera, representing the user's viewpoint on the HoloLens 2, serves as the gaze target, enabling the character to maintain eye contact and simulate natural human interaction. The GazeController script manages these mechanics by assigning the HoloLens 2 camera's transform to the character, ensuring smooth tracking of user movements.

Real-time updates to the gaze target enhance responsiveness. The OnAnimatorIK method is pivotal for implementing inverse kinematics (IK), controlling the character's head and eye movements to create a natural gaze. The parameters within SetLookAtWeight adjust the distribution of movement between the body, head, and eyes, allowing for fine-tuning of how the character follows the user's gaze.

This system balances realism and performance by enabling adjustable weights for head and eye movements. Fine-tuning these parameters allows for varying levels of engagement, ensuring that the virtual agent maintains natural eye contact that responds effectively to user movements.

3.4.4.3. facial expression

Facial expressions play a key role in the realism and effectiveness of an immersive Agent, enhancing the virtual character's ability to convey emotions. In this system, facial expressions are managed through the FacialExpressionControl.unitypackage and are controlled using the FaceController class, allowing for a wide range of emotional expressions that are crucial for implementing nonverbal ques. The character's facial expressions are designed to align with various interaction moments, ensuring that their emotions complement their verbal communication and other gestures.

The implementation leverages two main approaches to facial expression control. First, expressions can be set via predefined categorical emotions such as "Happiness," "Sadness," or "Surprise." These emotions are triggered based on the scenario and evolve naturally as the interaction progresses. Second, the system uses a more nuanced dimensional model, the Pleasure-Arousal-Dominance (PAD) framework, to allow for continuous, dynamic expressions. This model offers flexibility by adjusting how characters' emotions change over time, adding depth to their responses. Both approaches provide fine-grained control over the timing and intensity of each expression.

3.4.5. Eye tracking

The system was precisely calibrated to track users' gaze with high accuracy, focusing on specific areas of the virtual character, particularly the upper body, and logging this data at 20-millisecond intervals. The implementation process involved configuring MRTK's eye tracking capabilities, preparing the character model with appropriate colliders, developing custom scripts for gaze detection and data logging, and deploying the entire system on the HoloLens 2 for seamless functionality.

The initial step in implementing eye tracking involved configuring the system within MRTK. The project was set up ensuring that MRTK was properly imported and configured for HoloLens 2. Eye tracking was enabled by accessing the "Input" section within the MRTK configuration profile and cloning the DefaultMixedRealityInputSystemProfile for necessary modifications. The appropriate eye gaze data provider, OpenXREyeGazeDataProvider, was selected to ensure accurate tracking.

Following this, the "Use Eye Tracking Data" feature was activated under the "Pointers" section, enabling the system to use eye gaze for interaction. For testing purposes in the Unity Editor, eye tracking was simulated by enabling the "Input Simulation Service" in MRTK. The default head gaze cursor was replaced with an invisible prefab by adjusting the "Gaze Cursor Prefab" setting. Finally, it was ensured that the Universal Windows Platform (UWP) capability for Gaze Input was set through MRTK's tooling.

To track the user's gaze, colliders were added to key upper body parts of the character model. This process involved assigning colliders to specific body parts, including the head, chest, and shoulders. Different collider types were used: a SphereCollider for the head, a BoxCollider for the chest and shoulders. These colliders were precisely adjusted to ensure accurate gaze detection.

A custom script was then developed to detect when the user was looking at the character. The script, EyeTrackingLogger.cs, utilized the IMixedRealityEyeGazeProvider from MRTK to access the user's eye gaze data. A ray was continuously cast from the user's eye direction to determine if it intersected with the colliders on the character's head or upper body, allowing the system to identify specific areas being looked at. Gaze data was logged every 20 milliseconds using Unity's Debug.Log feature, recording whether the user was "Looking at character's head," "Looking at character's body," or "Not looking at character."

The data collection was managed through a coroutine running at regular intervals to ensure consistent logging. Lifecycle management was handled by starting the coroutine in the Start() method and stopping it in OnDisable() to avoid issues when the script was disabled.

3.4.6. Head movement

This system tracked the user's head position and rotation in real time. In MRTK, the camera represents the user's head, allowing for accurate monitoring of head movements during interactions with the virtual agent.

Within the Unity scene, a character representing the agent was placed at the world origin, positioned directly in front of the user. This character served as a key reference point, providing the baseline from which all head movements were measured. At the start of the scene, the user's head position and rotation were initialized, which captured the head's current position and orientation relative to the character. These initial values were stored and later used to compute the movement and rotation of the head over time as the user interacted with the virtual agent.

To systematically capture this data, a custom C# script, HeadMovementLogger.cs, was created to log the user's head movement every 20 milliseconds. The script continuously monitored the user's head position and rotation, logging both the current state and the changes from the initial reference point. Every 20 milliseconds, the script calculated how far the head had moved in 3D space and how much it had rotated, comparing the current position and orientation to the original state. This allowed for consistent tracking of the user's engagement with the character, providing valuable insights into how the user's focus shifted or remained fixed during the interaction. The script also included a stop function that deactivated logging when the scene changed, preventing unnecessary data collection after the user transitioned to another part of the application.

The logged data captured both absolute and relative movements of the head. For instance, when the user initially started the interaction, the head position might have been recorded as (0.5, 1.8, 0.2), with a head rotation of (Pitch: 0, Yaw: 90, Roll: 0). This indicated that the user's head was located slightly to the right, looking directly at the virtual agent, with no tilt or roll. After 20 milliseconds, the next log might show that the head had shifted slightly, with a new position of (0.7, 1.9, 0.3) and a rotation of (Pitch: 10, Yaw: 85, Roll: 5), indicating the user had moved slightly forward and upward while tilting their head upwards and to the left. The difference between the initial and current positions and rotations was also calculated and logged, providing a clear picture of how the user's head moved over time.

3.4.7. Voice detection for interruption

This section details the implementation of a voice detection system designed to track verbal interruptions when the virtual agent is speaking. The detection system employs volume spike detection to monitor interruptions. The system is designed to run alongside the character's audio without interference and detects external voice spikes that might interrupt the scene.

The implementation involves capturing audio samples from the HoloLens 2 microphone in real-time. These samples are continuously monitored, and their amplitude is analysed to identify potential interruptions. Specifically, the system looks for sudden increases in volume that exceed a predefined threshold, which might indicate someone speaking over the virtual character. The method relies on amplitude spikes in the recorded audio as markers of interruptions. The system further filters background noise to reduce the likelihood of false positives caused by environmental sounds or ambient noise.

The process starts with initializing the microphone in Unity to capture real-time audio input, which is continuously recorded into an AudioClip object. The script processes these audio samples by calculating their root mean square (RMS) value, which represents the average volume level. In each frame, the system compares the current average volume against a set threshold. If the volume exceeds this threshold, the system detects a spike and logs it as a potential interruption. The detection is handled non-intrusively to avoid disrupting the audio of the virtual character.

The system is further enhanced by recording each detected interruption. These logs, which capture the time and volume of the interruption. Allowing for further examination of when and how often interruptions occur throughout the character's speech.

In addition to detecting interruptions, the implementation includes effective resource management. The script ensures that the microphone is properly stopped, and its resources are released when the application is closed, preventing unnecessary performance bottlenecks.

3.4.8. Recording Data

In the system, the LogManager class serves as the primary tool for recording tracking data. The LogManager employs the Singleton design pattern to ensure that only one instance of the logger is active throughout the Unity application. This centralized instance logs data across multiple scenes or sessions. Specifically, tracking events are recorded in real time, capturing details such as the exact timestamp, and any contextual information at the time of the event.

During initialization, the LogManager sets up a uniquely named file for logging based on the date and time, avoiding overwrites. As the sessions progress, the LogManager continuously monitors for tracking data. The WriteTimeStampedEntry() method logs the exact moment, while additional context, then the system uses the WriteCSV() method to log structured data in a CSV format, facilitating easier analysis. To address real-time considerations, the LogManager prevents file access permission errors by opening and closing the log file with each entry. This approach ensures smooth logging, even if the application loses focus or switches scenes.

3.5. Experiment Procedure

Participants for the study were recruited from Aston University's postgraduate student body, as well as from external individuals who met the inclusion criteria. Those who qualified and consented after reviewing the Participant Information Sheet were invited to participate in the study, which took place at the AR/VR laboratory (MB212) at Aston University.

Upon arrival, participants signed a consent form, confirming their voluntary participation and understanding of the study. They were then fitted with a HoloLens 2 augmented reality (AR) headset. The study session lasted approximately 25 minutes.

Participants were guided through a series of scenes, starting with a menu, followed by an instructional segment explaining the tasks. This was followed by one of two randomized Topics, each with its associated Session. After a brief break, participants continued with the second Topic and its corresponding Session. The order of "Topic 1" and "Topic 2" was randomized, so either topic could be presented first, with its respective Session lasting 5 minutes. Participants had the option to sit or stand during the sessions, depending on their preference.

After the first session, participants completed a recall questionnaire about the content of that session, which was sent via Microsoft Teams. Following a short break, they proceeded to the second session and completed a similar recall questionnaire. To conclude, participants filled out a final questionnaire evaluating their overall experience with the system.

3.6. Results

The experiment was conducted with a total of 5 participants, each engaging with the active listening agent. The participants were randomly assigned different scene orders to minimize order bias. Two primary sessions were included: one featuring minimal animation on the topic of artificial intelligence's impact on the job market, and another more dynamic session where the virtual character described a visit to Universal Studios.

In this section, we provide an analysis of the collected data, including visual and behavioural metrics, as well as subjective feedback. The results have been divided categories: user interaction dynamics, and the measurement of active listening in AR. The specific findings for each research question are presented below.

3.6.1. Eye tracking and session analysis

Participants' gaze behaviour was recorded across both sessions, tracking how long their attention was focused on the character's head, upper body, or elsewhere. The data provides insights into how different session orders and content types affect user engagement.

Table below shows the percentage of time participants-maintained eye contact with the character's head and upper body across different sessions and scenes.

Table 1 – Participants eye tracking data

Participant	1st Session	% Time Focused on Head	% Time Focused on Body	2nd Session	% Time Focused on Head	% Time Focused on Body
1	AI Jobs	5.7	51.9	Universal Studios	11	54.2
2	AI Jobs	38.8	41.5	Universal Studios	59.3	25.5
3	Universal Studios	51.3	27.7	AI Jobs	42.5	44.3
4	Universal Studios	5.4	44.8	AI Jobs	14.9	32.4
5	Universal Studios	37.5	17.3	AI Jobs	50	27.7

Results based on sessions

Session 1 (Mean) - Head= 27.74, Body= 36.64

Session 2 (Mean) – Head= 35.54, Body= 36.82

Session 1 (Standard deviation) - Head= 20.96, Body= 13.937

Session 2 (Standard deviation) – Head= 21.5, Body= 12.13

Results based on Topic

AI jobs (Mean) - Head= 30.38, Body= 39.56

Universal Studios (Mean) – Head= 32.9, Body= 33.9

AI jobs (Standard deviation) – Head= 19.04, Body= 9.62

Universal Studios (Standard deviation) – Head= 23.9, Body=15.12

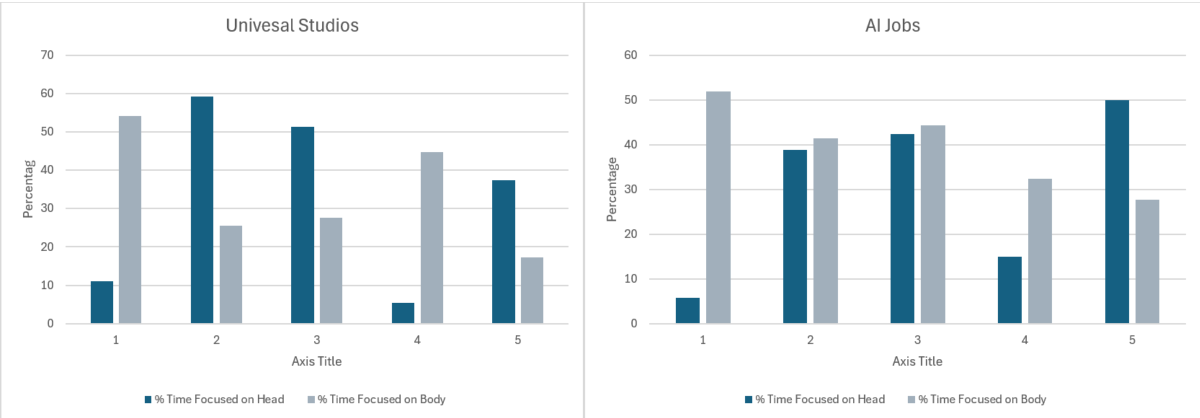

Fig 7 – Graph of eye tracking data based on topics

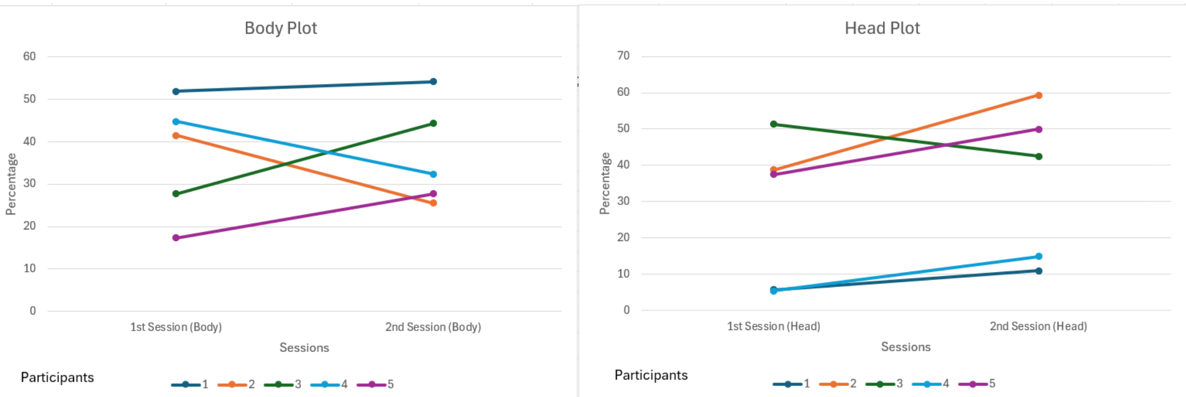

Fig 8 - Graph of eye tracking data based on sessions

Across both sessions, participants generally spent more time focusing on the body than the head, although the degree of focus varied. In Session 1, participants directed 27.74% of their attention to the head and 36.64% to the body. In Session 2, attention to the head increased to 35.54%, while focus on the body remained nearly constant at 36.82%. This slight shift in engagement toward the head in the second session may indicate greater attention to facial expressions or gestures as the session progressed. Another possible explanation is that participants, having completed the first questionnaire, felt the need to listen more carefully in anticipation of the second questionnaire. Despite this increase in head focus, the body continued to capture a significant portion of participants' gaze throughout both sessions, possibly due to the seated position of participants, making the character's body more convenient to look at.

When the data is analysed by session topic, a notable difference emerges. In the "AI Jobs" session, participants focused 30.38% of the time on the head and 39.56% on the body, on average. In contrast, during the "Universal Studios" session, attention to the head increased slightly to 32.9%, while focus on the body dropped to 33.9%. These differences may be attributed to the more animated gestures in the "Universal Studios" session, which likely drew more attention to the character's head and facial expressions.

The variability in participants' focus is reflected in the standard deviation values. In Session 1, the standard deviation for head focus was 20.96%, while for body focus, it was 13.94%, indicating diverse patterns of attention, particularly toward the head. In Session 2, the standard deviation for head focus increased slightly to 21.5%, while for body focus, it was lower at 12.13%, suggesting a more consistent focus on the body among participants in the second session.

When broken down by topic, the "AI Jobs" session exhibited less variability in body focus (SD = 9.62%) compared to the "Universal Studios" session (SD = 15.12%). Additionally, head focus during the "Universal Studios" session displayed greater variability (SD = 23.9%) than in the "AI Jobs" session (SD = 19.04%). This suggests that participants were more consistent in focusing on the body during the "AI Jobs" session, whereas the more animated "Universal Studios" session led to varied patterns of attention toward both the head and body.

Overall, participants tended to focus more on the body than the head, although this trend shifted slightly in the second session, where head focus increased. This shift could reflect a growing engagement with the virtual agent's facial expressions and gestures over time. The higher variability in head focus across sessions suggests that participants differed in how they engaged with the agent's facial cues and body language. The more dynamic "Universal Studios" session likely prompted greater attention to facial expressions, while the less animated "AI Jobs" session led to a more consistent focus on the body.

3.6.2. Head Tracking

This data captured, tracks the user’s head orientation over time, enabling an analysis of their focus during interactions. By measuring vertical head movements, represented as pitch angles. The system captures how often users look up, down, or maintain a steady gaze, which provides valuable insights into attentiveness.

Data collected includes five participants measurements, each representing a specific angle of the user’s head relative to the character over the 5-minute session. To ensure the analysis remained accurate, the data was reconstructed and rows where all pitch values were missing were excluded. To normalize the data and avoid skew from extreme values (such as those around 350 degrees), a transformation was applied. Pitch values exceeding 180 degrees were normalized to a range between -180 and 180, providing a more meaningful representation of the head movement.

The accompanying diagrams below visualizes the pitch data over time, showing the trends in head movements throughout the training sessions. The graphs provide a visual representation of how users’ focus shifted during their interactions with the agents. By grouping the data into specific intervals that correspond to different sessions, we can identify how head tracking behaviour varied between different parts of the sessions. The diagram clearly highlights moments of significant head movement and periods where users maintained a fixed gaze.

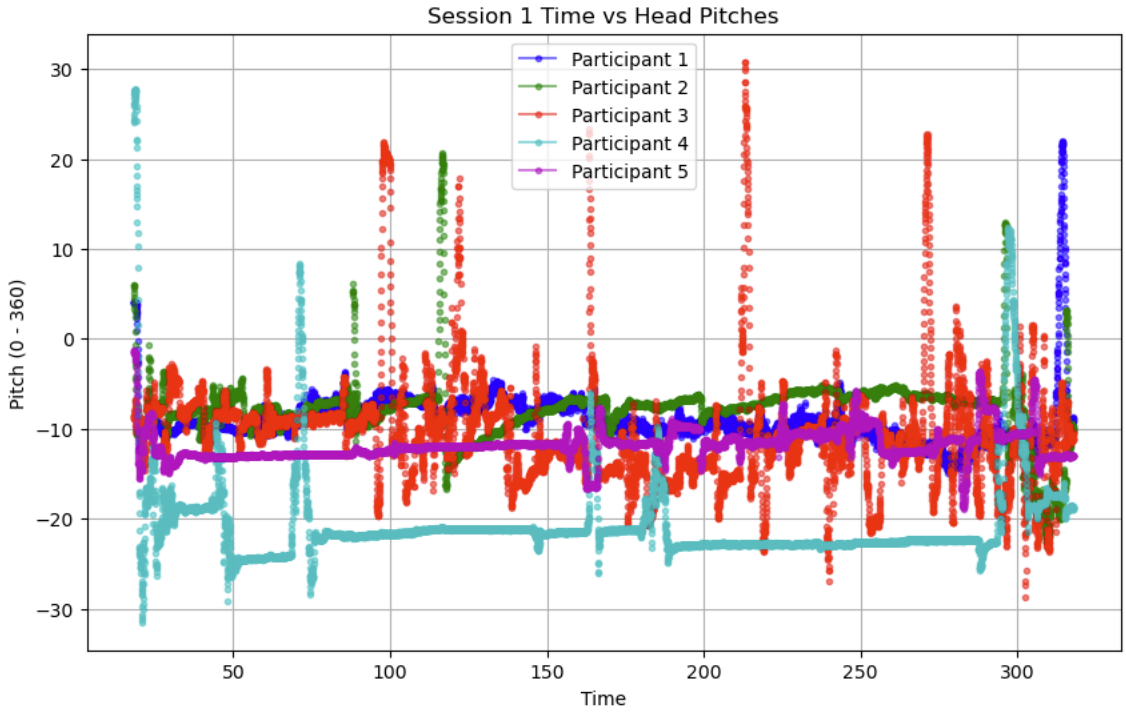

Fig 9 – Session 1 head tracking plot

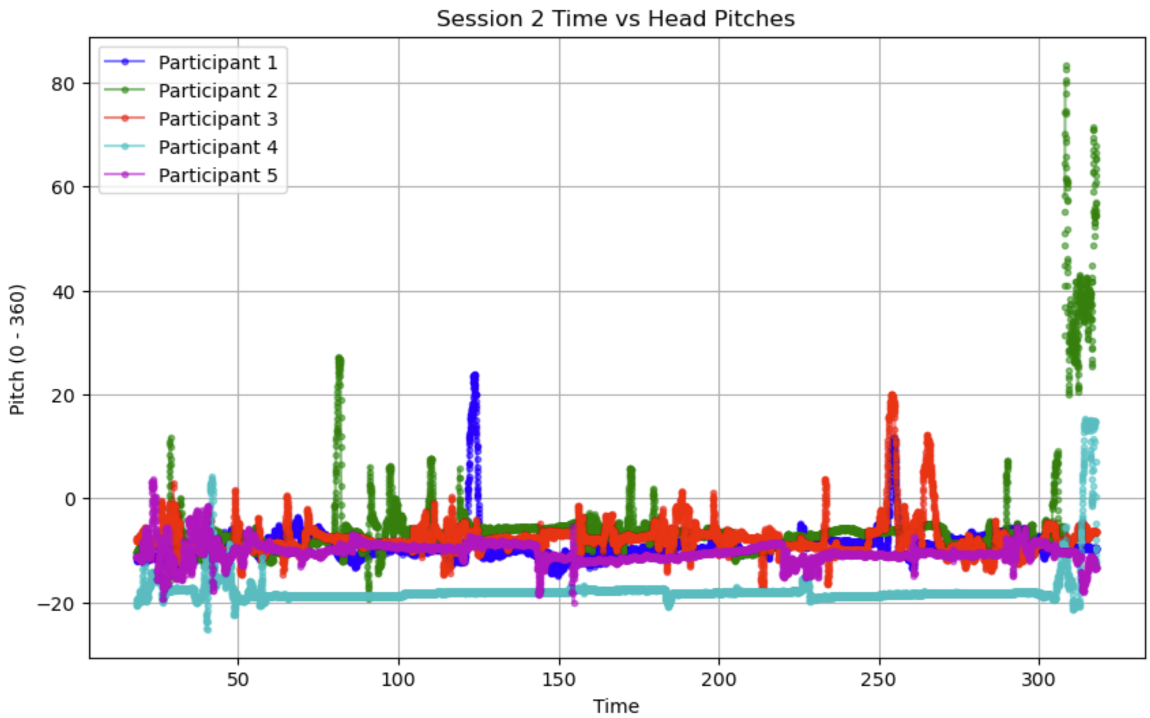

Fig 10 - Session 2 head tracking plot

Several insights were derived from the head tracking data. First, users maintained relatively stable head positions when the agent was directly speaking to them, indicating attentiveness. Sudden shifts in pitch angles often correlated with moments when users reoriented themselves or adjusted their gaze. In contrast, periods of increased head movement typically occurred during moments of heightened engagement, such as when the agents introduced animated gestures. These shifts suggest that users became more physically active when their attention was more intensely drawn to the agent.

Moreover, a comparison between the two sessions, reveals interesting interaction dynamics. Users demonstrated more consistent head orientation towards Session 2, which may be attributed to better engagement due to visual characteristics or the need to listen more carefully in anticipation of the questionnaire.

3.6.3. Verbal Interruptions

Verbal interruptions were tracked during the sessions by monitoring spikes in microphone input. This method allowed for the detection of instances where participants spoke over or interrupted the virtual character. Each spike in volume above a predefined threshold was logged as a potential interruption, capturing the frequency of these events.

Figure 11 below presents the frequency of verbal interruptions recorded across both sessions. Participants exhibited varying levels of interruptions depending on the session position, with some sessions showing no interruptions, while others experienced multiple.

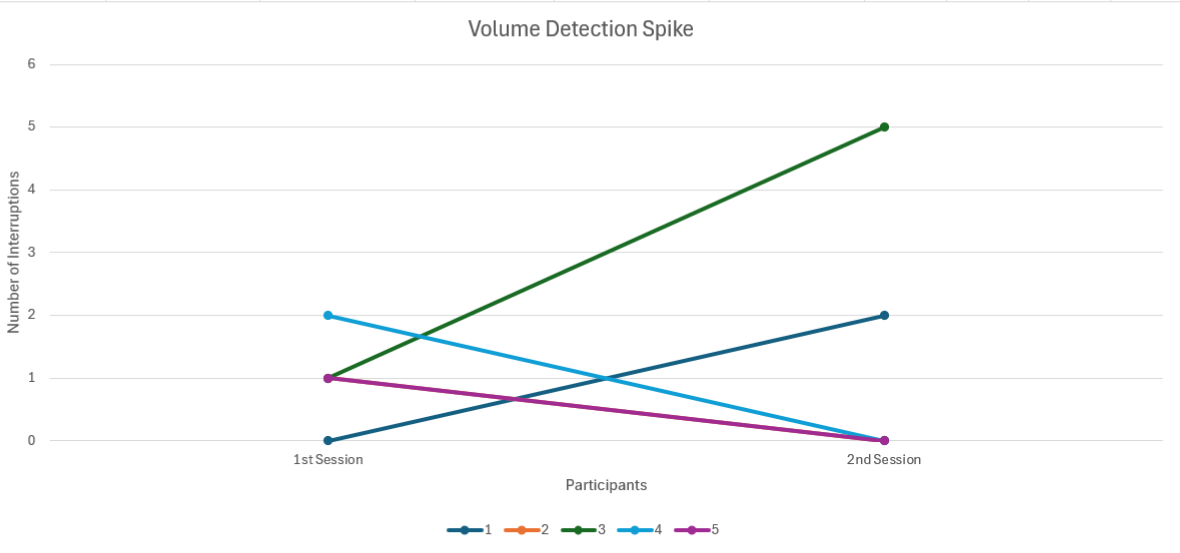

Fig 11 – Verbal interruption graph

The data reveals distinct differences between participants' behaviour across sessions. For example, Participant 3 interrupted frequently during the AI Jobs session, logging five interruptions, while Participant 4, who participated in the same session, recorded none. These variations suggest that the content of the session may influence the likelihood of interruptions. In contrast, Universal Studios sessions generally saw fewer interruptions, though Participant 1 recorded two interruptions during this session.

3.6.4. Measuring Active Listening in AR

To evaluate the effectiveness of feedback mechanisms for measuring active listening in augmented reality (AR), we analysed a combination of eye-tracking data and recall questionnaire scores from each session. Eye tracking measured the percentage of time participants focused on the virtual agent, while recall questionnaires evaluated how much information they retained. By correlating these metrics, we identified trends between attention and recall performance.

The table below summarizes participants' focus and recall across two sessions on different topics. Focus was measured by tracking the time participants looked at the agent's head and body, paired with their corresponding recall scores.

Table 2 – Participant's eye tracking vs recall score

Participant	1st Session	% Time Focused	% Recall Score	2nd Session	% Time Focused	% Recall Score
1	AI Jobs	57.6	76.9	Universal Studios	65.2	80
2	AI Jobs	80.3	76.9	Universal Studios	84.8	93.3
3	Universal Studios	79	100	AI Jobs	86.8	76.9
4	Universal Studios	50.2	40	AI Jobs	47.3	69.2
5	Universal Studios	54.8	73.3	AI Jobs	77.7	84.6

There appears to be a general positive relationship between the percentage of time participants remained focused and their recall performance. For instance, participant 2 demonstrated the highest focus across both sessions (80.3% and 84.8%) and consistently scored higher on recall (76.9% and 93.3%). In contrast, participant 4, who displayed lower focus percentages (50.2% and 47.3%), also showed significantly lower recall scores (40.0% and 69.2%).

It is notable that topic familiarity or interest might also have influenced performance. For example, participant 3 performed significantly better on recall for the "Universal Studios" topic (100% recall) compared to "AI Jobs" (76.9% recall), despite focusing less during the second session. This suggests that factors beyond attention, such as prior knowledge or intrinsic interest in the subject matter, may play a role in recall performance.

These findings suggest that eye tracking can serve as a valuable metric for measuring active listening, especially when coupled with recall data. Participants who remained more focused on the agent during the session tended to achieve higher recall scores, indicating better engagement and retention. However, the variability seen in the relationship between focus and recall across different topics implies that cognitive load, content complexity, and participant interest for a more comprehensive measure of active listening in AR.

3.6.5. User Experience Evaluation

Participants completed a post-experiment questionnaire to assess their overall experience with the system. This evaluation aimed to capture their perceptions in several key areas, including ease of use, emotional engagement, and the behaviour and visual realism of the augmented reality agents. A Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree) was used to measure user satisfaction across different aspects of the system. Table 3 summarizes the average user satisfaction ratings in each category.

Table 3 – Users experience average ratings

Area	Average Rating
Ease of Use and Comfort	5.25
Emotion and Engagement	5.266667
Agent 1 Behaviour and Mannerisms	5.177778
Agent 1 Visual Realism and Appearance	5.342857
Agent 2 Behaviour and Mannerisms	5.622222
Agent 2 Behaviour and Mannerisms	5.514286

The system was generally well-received, with participants reporting moderate to high satisfaction regarding usability and comfort. Users found the interface intuitive and easy to navigate, especially when interacting with the virtual humans in the augmented reality environment. This positive response reflects the system's ability to provide a smooth user experience, crucial for fostering engagement in immersive settings.

In terms of engagement, participants felt emotionally connected during their interactions with the virtual agents, showing that the system effectively maintained their attention. The participants' emotional responses were consistent, indicating that the design successfully captured their interest, a critical factor when evaluating active listening in augmented reality.

Agent 1, which featured little to no animations, received lower ratings for its behaviour and mannerisms. While participants found its actions somewhat coherent, the lack of fluid movement diminished the overall realism and effectiveness of its interactions. Without synchronized animations to match its audio, Agent 1 felt less natural, making its behaviour seem more static and less engaging.

In contrast, Agent 2, which had more animations synced to its audio, was rated much higher in terms of behaviour and engagement. The animations provided a more natural flow to its interactions, aligning well with the spoken dialogue and contributing to a stronger sense of realism. Participants found that Agent 2's movements felt more coherent and lifelike, making the interaction more immersive and satisfying. This significant difference between the two agents highlights the importance of synchronized animations in creating convincing virtual agents that can hold user attention and foster engagement.

Regarding visual realism, Agent 1's static nature also affected perceptions of its appearance. While the agent was generally seen as visually realistic, the lack of animation limited how immersive and lifelike it felt to users. Participants appreciated the detailed appearance but noted that the overall experience could be improved with animations to match the audio and enhance visual realism.

Agent 2, with its more fluid animations, scored higher in visual realism as well. Participants noted that the synchronization between Agent 2's audio and movements contributed to a more cohesive and immersive experience. The animations helped bring the agent to life, making the interactions more convincing and enjoyable.

Overall, participants found the system engaging and user-friendly, but Agent 2 outperformed Agent 1 significantly in both behaviour and visual realism due to its use of animations. These findings underscore the importance of dynamic, synchronized animations in virtual agents to create more engaging and realistic interactions.

3.7. Ethical Considerations

When conducting research involving human participants, it is essential to ensure that the study adheres to ethical guidelines that protect participants' rights, safety, and well-being. The Active Listening Training Agent experiment was designed and conducted following the ethical standards outlined by Aston University's Ethics Committee, ensuring that all participants were treated with respect and integrity throughout the study. Key ethical considerations for the experiment include informed consent, participant confidentiality, risk minimization, and the right to withdraw.

Confidentiality was strictly maintained. All personal data, including eye tracking, head tracking, and voice detection logs, were anonymized and stored securely. Data were assigned unique codes to protect participants' identities, and access was restricted to authorized researchers. The study also complied with GDPR regulations, ensuring that participants' data privacy was safeguarded at all times.

To minimize risks, participants were allowed to sit or stand during the sessions to ensure comfort while using the HoloLens 2 headset. Regular breaks were provided, and participants were instructed that they could pause or stop the experiment at any time if they experienced discomfort. The study posed minimal physical risks, but steps were taken to address any potential issues immediately.

Finally, participants were informed of their right to withdraw from the study at any point without any negative consequences. They were also given the option to request the removal of their data after the study if they chose. These measures ensured that participants felt comfortable and in control throughout the experiment, maintaining the ethical integrity of the research.

4. Project Management

Effective project management was essential for the success of this research project, ensuring that the objectives were met within the designated timeline and that high-quality deliverables were produced. This section outlines the core project management strategies used throughout the project and their importance.

4.1. Development Methodology

The Agile method is a flexible, iterative approach to software development and project management, chosen for its adaptability to changing requirements. It enabled the project to be responsive to real-time feedback, essential for refining the Active Listening Training Agent. Utilizing the Scrum framework, work was organized into sprints focused on delivering specific features, facilitating regular reviews and feedback integration. Sprint planning prioritized high-impact tasks, while continuous testing helped identify issues early in the process.

4.2. Timeline and Milestones

The project was divided into clear phases to ensure steady progress and maintain focus. It spanned for 20 weeks, with the learning, research, and Unity setup phases concluding by Week 8. These early weeks focused on understanding augmented reality, Unity development, and setting up the project environment, including GitHub for version control and collaboration.

By Week 8, the development phase commenced, which included system design, implementation, and iterating through multiple feedback mechanisms. Testing and iterations took place continuously to refine the system. These development cycles extended until Week 16, with project demos occurring around the same time to present progress and gather initial feedback. Ethics approval was also obtained during this phase to ensure compliance with research standards.

Experimentation and data collection occurred during Weeks 18 and 19, when human trials were conducted. This phase involved user testing, gathering real-time data on attentiveness and interaction dynamics, and using three different questionnaires to assess participants' experiences.

Report writing was an ongoing process, beginning in Week 1 with the literature review, followed by methodology, data analysis, and reflections as the project progressed. This early start allowed me to document and reflect on the project in real time, easing the intensive writing period that occurred from Week 17 onwards. By Week 20, the report was finalized for submission, reflecting on the entire project life cycle, from development to user testing and evaluation.

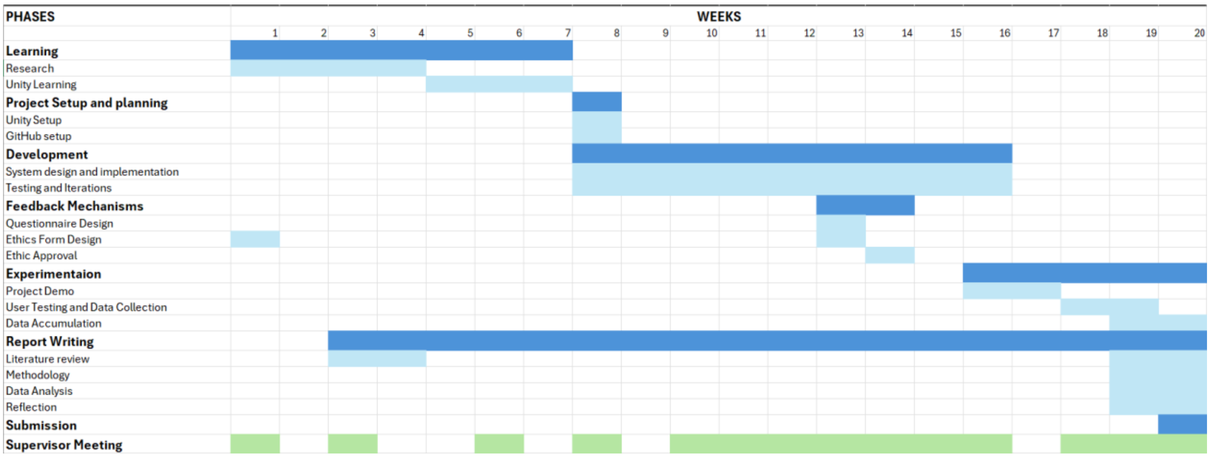

Fig 12 – Project plan in Gantt chart

This structured timeline was critical for maintaining focus at each stage of the project and ensuring that research, development, testing, and reporting were completed in a timely manner. Dividing the project into distinct phases allowed for efficient management of tasks and resources, ensuring that each phase received appropriate attention without rushing or delays.

4.3. Version Control and Collaboration

At the start of the project, I integrated GitHub with the Unity project for version control, a crucial step in managing code changes. Each feature or task was developed on its own branch to maintain the integrity of the main branch, minimizing the risk of introducing errors. This version control allowed smooth collaboration with my supervisor, who was added as a monitor on the GitHub repository. My supervisor reviewed code, providing feedback at key phases of development.

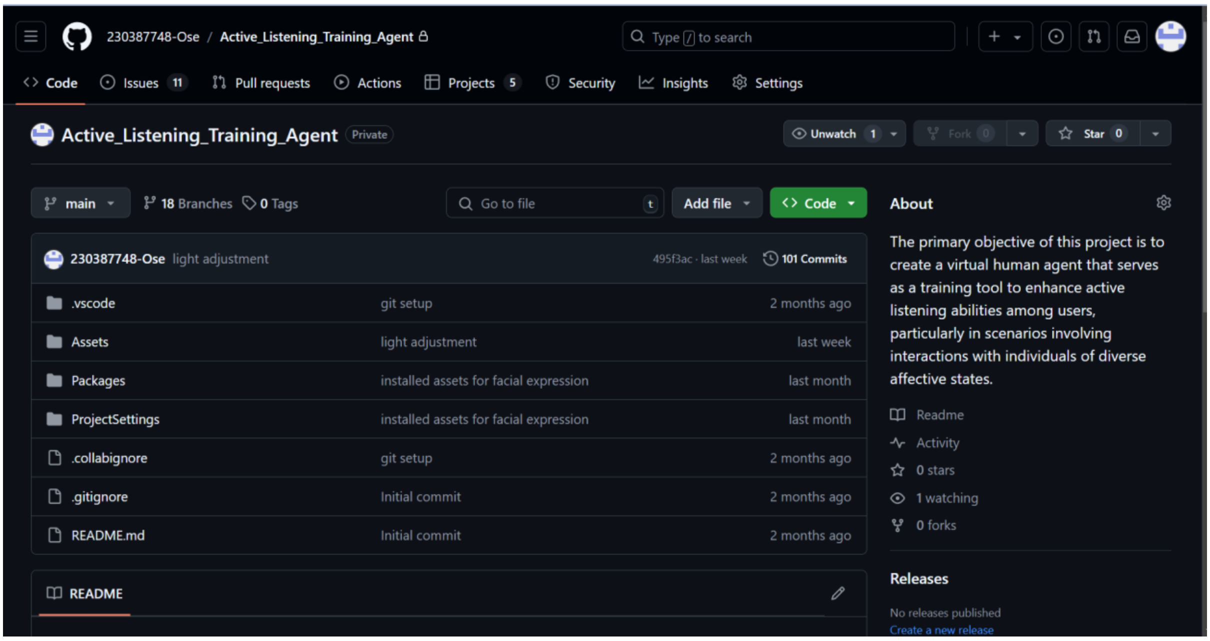

Fig 13 – GitHub collaboration

This structured method ensured the transparency of the development process and enabled easy reversion to previous versions if any major issues were encountered. It contributed significantly to code quality and the overall success of the project, ensuring a smooth and collaborative workflow.

4.4. Task Breakdown and User Stories

Project deliverables were divided into smaller, actionable tasks, represented as user stories, and managed using GitHub Issues. This approach allowed me to work on one manageable aspect of the project at a time, ensuring each task had a clear focus. Whenever I was ready to work on a new feature, I would create a branch from the issues board, allowing each task to be tracked and completed independently.

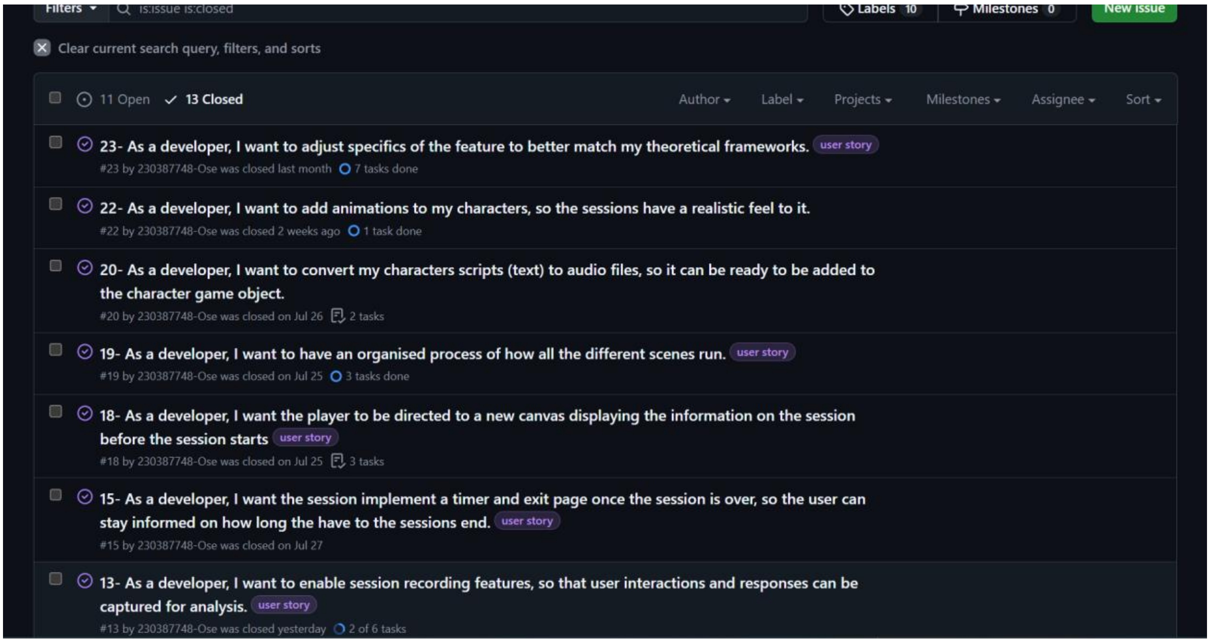

Fig 14 – GitHub issues

Breaking the tasks into user stories was key for maintaining clarity and making the overall project manageable. It allowed for focused work on individual components, while still contributing to the larger goals, preventing the process from becoming overwhelming.

4.5. Workflow Management with Kanban Boards

GitHub's project functionality, specifically through Kanban boards, was used to track and manage the workflow. Tasks were categorized as "Backlog" (upcoming tasks for that sprint), "Ready" (tasks ready to begin), "In Progress" (tasks being worked on), "In Review" (tasks awaiting supervisor feedback), and "Done" (completed tasks that had been merged into the main branch). This sprint-based workflow helped keep development on track, ensuring smooth transitions between tasks and prioritizing work efficiently.

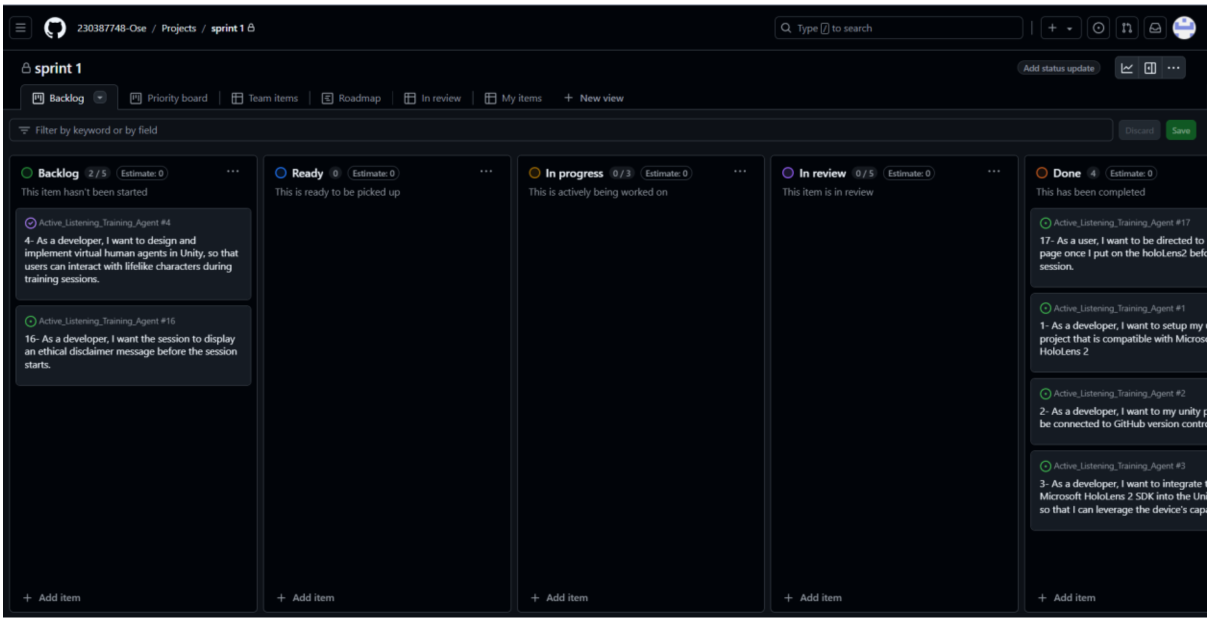

Fig 15 – GitHub kanban board

The Kanban board was particularly helpful during periods of testing and iteration, ensuring smooth transitions between tasks, prioritizing work effectively, and managing feedback efficiently.

4.6. Monitoring and Review

Progress was continuously monitored through GitHub's built-in project tracking features, allowing for ongoing assessments and any necessary adjustments to the project plan. The system facilitated timely reviews and approvals from my supervisor, ensuring that all work met the project's quality standards before being integrated into the main branch.

This review process ensured that each feature was properly implemented and aligned with the project's overall objectives. Regular monitoring was essential for identifying potential issues early, leading to more polished and reliable outcomes.

4.7. Importance of Project Management

Structured project management practices, such as version control, task breakdown, and workflow tracking, were essential to keeping the project organized and on schedule. These strategies enabled smooth transitions between phases, improved collaboration, and allowed the development process to remain flexible and responsive to challenges. As a result, the system was successfully completed, with both technical implementation and research components delivered on time and to a high standard.

5. Evaluation and Reflection

5.1. Result discussion

The study highlights the importance of different user interaction dynamics in enhancing the effectiveness listening in an augmented environment. Participants who maintained focus on the characters' heads and aligned their head orientation accordingly demonstrated better recall and comprehension. This suggests that non-verbal cues, such as consistent eye contact and head orientation, positively influence active listening. Participants who also minimized verbal interruptions showed improved information retention, reinforcing the role of controlled interactions in effective listening.

Regarding implicit feedback mechanisms, eye-tracking and voice detection emerged as the most effective tools for enhancing listening skills. Participants who kept their gaze focused on the character's head performed better in recall tasks, indicating that eye-tracking feedback can guide attention and improve engagement. Similarly, voice detection, which identified interruptions, revealed that less frequent interruptions correlated with better listening outcomes.

The study also demonstrates that active listening levels can be adequately measured in an AR environment using eye-tracking, head-tracking, and voice detection. Eye-tracking data showed a strong link between focused attention and improved recall, while head-tracking data indicated that consistent orientation toward the speaker correlated with greater attentiveness. Additionally, voice detection effectively captured interruptions, with fewer interruptions leading to better comprehension. These results show that AR environments can moderately measure and enhance active listening by capturing key behavioural indicators.

5.2. Limitation

This study faced several limitations that impact the interpretation and generalizability of its findings. First, the sample size was relatively small, with only 5 participants, which restricts the ability to generalize the results to a larger population. Additionally, the participants shared similar demographic characteristics, being aged between 22-26 and having backgrounds in computer science. This lack of diversity may not represent the experiences and learning behaviours of a broader audience, limiting the study's relevance to a wider population.

In terms of experimental design, although the session order was randomized, there remains the possibility of order bias. Participants' prior experiences, expectations, or differences in the content of the sessions may have influenced their engagement or recall. The standardized session duration of 5 minutes might not have been sufficient for participants to fully exhibit their active listening skills. For some, the session length may have been too short to engage deeply, while for others, it may have been too long, leading to a decline in attention due to individual differences in attention span. The positioning of the virtual character relative to participants might also have introduced bias, as variations in viewpoint could influence how attentively participants.

Technological limitations also played a role in shaping the study's outcomes. The augmented reality environment and the HoloLens 2 had limitations regarding visual clarity, tracking accuracy, and system responsiveness. These factors may have affected how participants interacted with the virtual agents, potentially compromising the reliability of the eye and head-tracking data. The sensors might not have accurately captured subtle or rapid movements, leading to incomplete or less precise data. The equipment also did not have the capability to track additional features such as participant emotions, dynamic verbal responses, biometric data, or body language, further limiting the comprehensiveness of the data collected.

The content and agent design presented another limitation. The two session topics may not have equally engaged all participants. Interest and prior knowledge of the subject matter likely varied, influencing how participants interacted with the content and their recall performance. Additionally, the complexity of each topic could have affected the participants' engagement levels, further contributing to variability in the results.

Environmental factors also posed challenges. The controlled setting of the study may have influenced participant comfort levels, potentially affecting how they interacted with the virtual agents. External distractions such as noise or movement in the physical space could have affected participant focus and performance during the sessions.

Human factors, including cognitive fatigue, might have influenced participants' attentiveness, particularly in later sessions. Repeated exposure to similar tasks could lead to a decrease in focus and engagement. Additionally, individual differences, such as personality traits, mood, and familiarity with AR technology, likely contributed to variations in how each participant engaged.

Finally, the use of subjective measures through self-reported questionnaires introduced potential biases. Participants may not have accurately conveyed their true feelings or engagement levels, as responses could be influenced by social desirability, misunderstanding of the questions, or difficulties in recalling details accurately. The recall questionnaires, in particular, may have been influenced by factors such as topic complexity, attention span, and familiarity with the content, further complicating the interpretation of the results

5.3. Implications and future directions

The data emphasize the importance of dynamic interactions in boosting user engagement, with fluid animations, like those of Agent 2, improving the training experience and effectiveness in developing listening skills. Future studies should expand to larger, more diverse participants and explore varied session topics to broaden the system's applicability. Incorporating advanced feedback mechanisms, such as real-time prompts or adaptive virtual agent responses, could enhance interactivity and personalization. Additionally, integrating machine learning to adjust session difficulty based on engagement levels would provide a more tailored and effective training experience.

5.4. Business Opportunity

The development of the Active Listening Training Agent in augmented reality (AR) offers significant business potential, as analysed through PESTEL, Porter's Five Forces, and the Ansoff Matrix.

Macro-Environment (PESTEL)

The political landscape increasingly supports digital transformation, with governments and organizations funding immersive technologies for skill development. Economically, businesses are seeking cost-effective training methods, and AR-based solutions address this need by reducing in-person training costs while catering to the shift toward remote work. Social trends emphasize the importance of soft skills like active listening, and AR's immersive nature aligns with growing demands for more interactive communication training. Technologically, AR advancements, particularly in eye-tracking and voice detection, make the system scalable. Environmental concerns also drive adoption, as virtual training minimizes resource usage. Legally, compliance with data privacy regulations, especially around tracking, will be critical as the technology scales.

Industry and Sector Analysis (Porter's Five Forces)

The threat of new entrants in the AR training space is moderate, as developing sophisticated systems requires substantial expertise. Supplier power is high, given the limited number of AR hardware providers, but this may decrease as more options enter the market. Buyers have moderate bargaining power due to the availability of alternative training platforms, though the unique features of AR-based active listening systems provide differentiation. The threat of substitutes, such as traditional e-learning or in-person workshops, exists but lacks the real-time, interactive feedback AR provides. Industry rivalry is still low, as the AR training sector is emerging, allowing early entrants to secure a competitive edge by offering advanced features and strong evidence of effectiveness.

Internal Factors (Ansoff Matrix)

In terms of growth, a market penetration strategy would focus on deepening the use of the Active Listening Training Agent within corporate training programs, emphasizing its immersive and interactive benefits. Market development opportunities exist in sectors like healthcare and education, where communication skills are essential, as well as through geographic expansion. For product development, additional features like multi-lingual support or AI-driven real-time feedback could enhance the system's appeal. Diversification could extend the platform into other soft skills training, such as conflict resolution or empathy, broadening its application and market reach.

By leveraging macro-environmental trends and competitive advantages in technology, the system holds significant business potential. Targeting industries focused on communication, developing scalable solutions, and expanding into related soft skills training positions the system for success in both virtual and real-world training environments.

5.5. Conclusion

In conclusion, this dissertation advances the understanding of active listening training through AR, demonstrating how dynamic, interactive environments enhance skill development. The system, built using Unity and Microsoft HoloLens 2, enables users to improve their active listening abilities by engaging in realistic scenarios with real-time feedback. Implicit feedback mechanisms, such as visual prompts and audio cues reinforces learning, while diverse virtual agents simulate a wide range of communication contexts.

This research shows that user interaction dynamics significantly impact training effectiveness, with participants demonstrating measurable improvements in focus, responsiveness, and overall listening skills. The integration of quantitative metrics and qualitative assessments offers a comprehensive evaluation of active listening in AR. These findings underscore AR's potential for creating immersive, impactful training experiences, and suggest that future studies should refine these methods and validate the approach across broader populations.

References

Ali and M, D.O. (2022). The Effect of an Instructional Listening Program on Secondary Students' Listening Achievement. *Deleted Journal*, 2019(1), pp.474–485. doi:https://doi.org/10.37653/juah.2022.171792.

Alkhattabi, M. (2017). Augmented Reality as E-learning Tool in Primary Schools' Education: Barriers to Teachers' Adoption. *International Journal of Emerging Technologies in Learning (iJET)*, 12(02), p.91. doi:https://doi.org/10.3991/ijet.v12i02.6158.

Alzahrani, N.M. (2020). Augmented Reality: A Systematic Review of Its Benefits and Challenges in E-learning Contexts. *Applied Sciences*, [online] 10(16), p.5660. doi:https://doi.org/10.3390/app10165660.

Anchal Malh, Wood, S. and Chung, K.C. (2023). The Art of Listening. *The Art of Listening. Plastic and Reconstructive Surgery*, 151(5), pp.921–926. doi:https://doi.org/10.1097/prs.0000000000010065.

Argyle, M. (2013). *Social Skills and Health (Psychology Revivals)*. Routledge.

Bodie, G.D., Vickery, A.J., Cannava, K. and Jones, S.M. (2015). The Role of 'Active Listening' in Informal Helping Conversations: Impact on Perceptions of Listener Helpfulness, Sensitivity, and Supportiveness and Discloser Emotional Improvement. *Western Journal of Communication*, [online] 79(2), pp.151–173. doi:https://doi.org/10.1080/10570314.2014.943429.

Bozorgian, H. and Fakhri Alamdari, E. (2017). Multimedia listening comprehension: Metacognitive instruction or metacognitive instruction through dialogic interaction. *ReCALL*, 30(1), pp.131–152. doi:https://doi.org/10.1017/s0958344016000240.

Bruijnes, M., Linssen, J. and Heylen, D. (2019). Special issue editorial: Virtual Agents for Social Skills Training. *Journal on Multimodal User Interfaces*, 13(1), pp.1–2. doi:https://doi.org/10.1007/s12193-018-00291-7.

Burgoon, J.K., Bonito, J.A., Bengtsson, B., Cederberg, C., Lundeberg, M. and Allspach, L. (2000). Interactivity in human–computer interaction: a study of credibility, understanding, and influence. *Computers in Human Behavior*, 16(6), pp.553–574. doi:https://doi.org/10.1016/s0747-5632(00)00029-7.

Calvin Kai-Ching Yu and Samuel Po-Shing Wong (2023). A moderated mediation study of music engagement, adaptive music listening, and dreaming. *Psychology of Music*, 52(1). doi:https://doi.org/10.1177/03057356231159199.

Cho, E., Nasim Motalebi, S. Shyam Sundar and Abdullah, S. (2022). Alexa as an Active Listener: How Backchanneling Can Elicit Self-Disclosure and Promote User Experience. 6(CSCW2), pp.1–23. doi:https://doi.org/10.1145/3555164.

Chun, M.M., Golomb, J.D. and Turk-Browne, N.B. (2011). A Taxonomy of External and Internal Attention. *Annual Review of Psychology*, 62(1), pp.73–101. doi:https://doi.org/10.1146/annurev.psych.093008.100427.

Chun-Wai Ma (2019). Nursing Students’ Perceptions of Biomedical Education with Augmented Reality. *Integrative Biomedical Sciences*. doi:https://doi.org/10.18314/gjbs.v4i1.1851.

Dandıl, E., Serin, Z. and Şenol, Y. (2022). SABAS: A Smartphone-Aided Training Simulator based on Virtual and Augmented Reality for Brain Anatomy Assessment. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, [online] 13(3), pp.252–276. doi:https://doi.org/10.18662/brain/13.3/366.

Dhar, P., Rocks, T., Samarasinghe, R.M., Stephenson, G. and Smith, C. (2021). Augmented reality in medical education: students’ experiences and learning outcomes. *Medical Education Online*, [online] 26(1), p.1953953. doi:https://doi.org/10.1080/10872981.2021.1953953.

Dick, E. (2021). *Current and Potential Uses of AR/VR for Equity and Inclusion*. [online] itif.org. Available at: https://itif.org/publications/2021/06/01/current-and-potential-uses-arvr-equity-and-inclusion/.

Dongas, R. (2018). Special Session—Designing Virtual Learning Experiences: Iterating Immersive Environments. *2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*, 11, pp.1225–1227. doi:https://doi.org/10.1109/tale.2018.8615328.

Drollinger, T., Comer, L.B. and Warrington, P.T. (2006). Development and validation of the active empathetic listening scale. *Psychology and Marketing*, 23(2), pp.161–180. doi:https://doi.org/10.1002/mar.20105.

Eckert, M.A., Teubner-Rhodes, S. and Vaden, K.I. (2016). Is Listening in Noise Worth It? The Neurobiology of Speech Recognition in Challenging Listening Conditions. *Ear and Hearing*, 37, pp.101S110S. doi:https://doi.org/10.1097/aud.0000000000000300.

Eg Su, G., Zubir, N.S., Zakaria, N.H. and Ahmad, J. (2022). Handheld Augmented Reality Application for 3D Fruits Learning. *International Journal of Innovative Computing*, 12(2), pp.69–75. doi:https://doi.org/10.11113/ijic.v12n2.378.

Gilakjani, A.P. and Ahmadi, M.R. (2011). A Study of Factors Affecting EFL Learners' English Listening Comprehension and the Strategies for Improvement. *Journal of Language Teaching and Research*, 2(5). doi:https://doi.org/10.4304/jltr.2.5.977-988.

Graham, S. and Macaro, E. (2008). Strategy Instruction in Listening for Lower-Intermediate Learners of French. *Language Learning*, 58(4), pp.747–783. doi:https://doi.org/10.1111/j.1467-9922.2008.00478.x.

Grusec, J.E. (1992). Social learning theory and developmental psychology: The legacies of Robert Sears and Albert Bandura. *Developmental Psychology*, 28(5), pp.776–786.

Gudykunst, W.B. (2004). *Bridging Differences: Effective Intergroup Communication*. [online] *Google Books*. SAGE. Available at: https://books.google.co.uk/books?hl=en&lr=&id=B9LReJA44WwC&oi=fnd&pg=PP1&dq=Gudykunst+and+Shapiro+(1996)+Traditional+methods+for+training+active+listening+skills&ots=SKQpGlKky6&sig=lI_Mb8hMM1jF4O7j2Krs4_sPNow#v=onepage&q&f=false.

Herd, A. (2015). The Case of 'Not Enough Time:' Using GROW and Motivational Interviewing Coaching Frameworks. *New Horizons in Adult Education and Human Resource Development*, 27(3), pp.46–49. doi:https://doi.org/10.1002/nha3.20110.

Jiang, H., Vimalesvaran, S., Wang, J.K., Lim, K.B., Mogali, S.R. and Car, L.T. (2021). Virtual Reality in Medical Students' Education: A Scoping Review (Preprint). *JMIR Medical Education*, 8(1). doi:https://doi.org/10.2196/34860.

Kamal, A.A., Junaini, S.N., Hashim, A.H., Sukor, F.S. and Said, M.F. (2021). The Enhancement of OSH Training with an Augmented Reality-Based App. *International Journal of Online and Biomedical Engineering (iJOE)*, 17(13), pp.120–134. doi:https://doi.org/10.3991/ijoe.v17i13.24517.

Karlimah Karlimah, Asep Nuryadin, Lilis Susilawati, Hartanto, B., Muiz, A., Muhammad, Sri Ariyani Desmawati and Irni Rachmawati Putri (2023). The Development of Teachers' Competences in Utilizing Augmented Reality-Based Media in Geometry Learning. *ABDIMAS Jurnal Pengabdian Masyarakat*, 6(4), pp.4555–4563. doi:https://doi.org/10.35568/abdimas.v6i4.4015.

Keller, J. (2014). *ARCS Model of Motivational Design Theories (Keller) - Learning Theories*. [online] Learning Theories. Available at: https://learning-theories.com/kellers-arcs-model-of-motivational-design.html.

King, D., Tee, S., Falconer, L., Angell, C., Holley, D. and Mills, A. (2018). Virtual health education: Scaling practice to transform student learning. *Nurse Education Today*, 71(71), pp.7–9. doi:https://doi.org/10.1016/j.nedt.2018.08.002.

Knörzer, L., Brünken, R. and Park, B. (2016). Emotions and multimedia learning: the moderating role of learner characteristics. *Journal of Computer Assisted Learning*, 32(6), pp.618–631. doi:https://doi.org/10.1111/jcal.12158.

Lala, D., Milhorat, P., Inoue, K., Ishida, M., Takanashi, K. and Kawahara, T. (2017). *Attentive listening system with backchanneling, response generation and flexible turn-taking*. [online] DOI.org (Crossref). doi:https://doi.org/10.18653/v1/W17-5516.

Lawson, A.P., Mayer, R.E., Adamo-Villani, N., Benes, B., Lei, X. and Cheng, J. (2021). The positivity principle: do positive instructors improve learning from video lectures? *Educational Technology Research and Development*. [online] doi:https://doi.org/10.1007/s11423-021-10057-w.

Mayer, R.E. (2001). Modality Principle. *Multimedia Learning*, pp.134–146. doi:https://doi.org/10.1017/cbo9781139164603.009.

Mayer, R.E. (2014). Cognitive Theory of Multimedia Learning. *The Cambridge Handbook of Multimedia Learning*, [online] pp.43–71. doi:https://doi.org/10.1017/cbo9781139547369.005.

Mayer, R.E. (2019). Coherence Principle. *Multimedia Learning*, [online] pp.89–107. doi:https://doi.org/10.1017/cbo9780511811678.007.

Mayer, R.E. (2020). Personalization Principle. *Cambridge University Press eBooks*, pp.305–321. doi:https://doi.org/10.1017/9781316941355.019.

Milliner, B. and Dimoski, B. (2019). Explicit Listening Strategy Training for ELF Learners. *The Journal of AsiaTEFL*, 16(3), pp.833–859. doi:https://doi.org/10.18823/asiatefl.2019.16.3.5.833.

Neidhardt, A. and Zerlik, A.M. (2021). The Availability of a Hidden Real Reference Affects the Plausibility of Position-Dynamic Auditory AR. *Frontiers in Virtual Reality*, 2. doi:https://doi.org/10.3389/frvir.2021.678875.

Palincsar, A.S. (1998). SOCIAL CONSTRUCTIVIST PERSPECTIVES ON TEACHING AND LEARNING. *Annual Review of Psychology*, [online] 49(1), pp.345–375. doi:https://doi.org/10.1146/annurev.psych.49.1.345.

Palmarini, R., Erkoyuncu, J.A., Roy, R. and Torabmostaedi, H. (2018). A systematic review of augmented reality applications in maintenance. *Robotics and Computer-Integrated Manufacturing*, 49, pp.215–228. doi:https://doi.org/10.1016/j.rcim.2017.06.002.

Plass, J.L., Heidig, S., Hayward, E.O., Homer, B.D. and Um, E. (2014). Emotional design in multimedia learning: Effects of shape and color on affect and learning. *Learning and Instruction*, [online] 29, pp.128–140. doi:https://doi.org/10.1016/j.learninstruc.2013.02.006.

Poirier-Quinot, D. and Lawless, M. (2023). Impact of wearing a head-mounted display on localization accuracy of real sound sources. *Acta Acustica*, 7, pp.3–3. doi:https://doi.org/10.1051/aacus/2022055.

Rebollo, C., Remolar, I., Rossano, V. and Lanzilotti, R. (2021). Multimedia augmented reality game for learning math. *Multimedia Tools and Applications*, 81. doi:https://doi.org/10.1007/s11042-021-10821-3.

Reynolds, K.M., Roberts, L.M. and Hauck, J. (2017). Exploring motivation: integrating the ARCS model with instruction. *Reference Services Review*, 45(2), pp.149–165. doi:https://doi.org/10.1108/rsr-10-2016-0057.

Rogers, C. (1975). Empathic: An Unappreciated Way of Being. *The Counseling Psychologist*, 5(2), pp.2–10. doi:https://doi.org/10.1177/001100007500500202.

Sharmin, N. and Chow, A.K. (2020). Augmented Reality Application to Develop a Learning Tool for Students: Transforming Cellphones into Flashcards. *Healthcare Informatics Research*, 26(3), pp.238–242. doi:https://doi.org/10.4258/hir.2020.26.3.238.

Song, J.H., Skoe, E., Banai, K. and Kraus, N. (2011). Training to Improve Hearing Speech in Noise: Biological Mechanisms. *Cerebral Cortex*, 22(5), pp.1180–1190. doi:https://doi.org/10.1093/cercor/bhr196.

Stickley, T. (2011). From SOLER to SURETY for Effective non-verbal Communication. *Nurse Education in Practice*, [online] 11(6), pp.395–398. doi:https://doi.org/10.1016/j.nepr.2011.03.021.

Sukthankar, G., Mandel, M. and Sycara, K. (2007). Creating physically embodied agents using realistic human motion models. *Simulation & Gaming*, 39(1), pp.64–82. doi:https://doi.org/10.1177/1046878107309686.

Sun, K.T. and Chen, M.H. (2019). Utilizing Free Augmented Reality App for Learning Geometry at Elementary School in Taiwan. *International Journal of Distance Education Technologies*, 17(4), pp.36–53. doi:https://doi.org/10.4018/ijdet.2019100103.

Taha, S., Enas Abulibdeh, Hanan El-Gendi, Hanan Shaher Almarashdi, Alghazo, E.M., Imad Ibraheem Mostafa, Abusini, B.S., Salem Khalil Al-Aqtash, Abed, Ayasrah, M. and Mohannad Salameh Alkhalaileh (2023). Exploring Media and Communication Students' Perception of Egyptian Universities' Use of Augmented Reality in Learning. *Studies in Media and Communication*, 11(4), pp.209–209. doi:https://doi.org/10.11114/smc.v11i4.6121.

Tanaka, H. and Nakamura, S. (2021). The Acceptability of Virtual Characters as Social Skills Trainers: Usability Study (Preprint). doi:https://doi.org/10.2196/preprints.35358.

Tanaka, H. and Nakamura, S. (2022). The Acceptability of Virtual Characters as Social Skills Trainers: Usability Study. *JMIR Human Factors*, 9(1), p.e35358. doi:https://doi.org/10.2196/35358.

Tennant, K., Butler, T.J.T. and Long, A. (2023). *Active Listening*. [online] Nih.gov. Available at: https://www.ncbi.nlm.nih.gov/books/NBK442015/.

Tri Yunianika, I. (2022). Evaluation of Teachers' Active Listening Training at Madrasah Ibtidaiyah Hidayatul Athfal. *Evaluasi: Jurnal Manajemen Pendidikan Islam*, 6(2), p.322. doi:https://doi.org/10.32478/evaluasi.v6i2.1106.

Vandergrift, L. and Tafaghodtari, M.H. (2010). Teaching L2 Learners How to Listen Does Make a Difference: An Empirical Study. *Language Learning*, [online] 60(2), pp.470–497. doi:https://doi.org/10.1111/j.1467-9922.2009.00559.x.

Wang, I., Smith, J.S. and Ruiz, J. (2019). Exploring Virtual Agents for Augmented Reality. *Human Factors in Computing Systems*. doi:https://doi.org/10.1145/3290605.3300511.

Appendices

Code

In [1]:
```
import matplotlib.pyplot as plt
import pandas as pd
```

In [2]:
```
data = pd.read_csv('C:/Users/osose/Documents/Dissertation/disso/n data/Session1.csv')
```

In [3]:
```
data.head()
```

Out[3]:

	Time	Pitch 1	Pitch 2	Pitch 3	Pitch 4	Pitch 5
0	18.2843	NaN	NaN	NaN	NaN	NaN
1	NaN	NaN	NaN	NaN	NaN	NaN
2	18.2843	3.9	6.0	352.1	24.1	358.4
3	NaN	NaN	NaN	NaN	NaN	NaN
4	18.2843	4.0	5.8	351.4	24.3	358.6

In [4]:
```
# Drop rows with missing pitch values
cleaned_data = data.dropna(subset=['Pitch 1']).copy()
cleaned_data = data.dropna(subset=['Pitch 2']).copy()
cleaned_data = data.dropna(subset=['Pitch 3']).copy()
cleaned_data = data.dropna(subset=['Pitch 4']).copy()
cleaned_data = data.dropna(subset=['Pitch 5']).copy()
cleaned_data = data.dropna(subset=['Time']).dropna(subset=['Pitch 1', 'Pitch 2', 'Pitch 3', 'Pitch 4', 'Pitch 5'], how='all').cop
```

In [5]:
```
# Normalize pitch values to a range between -180 and 180
def normalize_pitch(pitch):
    return ((pitch + 180) % 360) - 180

cleaned_data['Pitch 1'] = cleaned_data['Pitch 1'].apply(normalize_pitch)
cleaned_data['Pitch 2'] = cleaned_data['Pitch 2'].apply(normalize_pitch)
cleaned_data['Pitch 3'] = cleaned_data['Pitch 3'].apply(normalize_pitch)
cleaned_data['Pitch 4'] = cleaned_data['Pitch 4'].apply(normalize_pitch)
cleaned_data['Pitch 5'] = cleaned_data['Pitch 5'].apply(normalize_pitch)
```

In [6]:
```
cleaned_data.head()
```

Out[6]:

	Time	Pitch 1	Pitch 2	Pitch 3	Pitch 4	Pitch 5
2	18.28430	3.9	6.0	-7.9	24.1	-1.6
4	18.28430	4.0	5.8	-8.6	24.3	-1.4
6	18.28430	4.0	NaN	-8.6	24.5	-1.3
10	18.31252	4.0	6.0	NaN	NaN	-1.3
12	18.31250	NaN	5.9	-8.5	24.9	NaN

In [8]:
```
# Define the figure size
plt.figure(figsize=(10, 6))

# Plot each pitch column with a unique color
plt.plot(cleaned_data['Time'], cleaned_data['Pitch 1'], marker='.', linestyle='-', color='b', label='Participant 1', alpha=0.5)
plt.plot(cleaned_data['Time'], cleaned_data['Pitch 2'], marker='.', linestyle='-', color='g', label='Participant 2', alpha=0.5)
plt.plot(cleaned_data['Time'], cleaned_data['Pitch 3'], marker='.', linestyle='-', color='r', label='Participant 3', alpha=0.5)
plt.plot(cleaned_data['Time'], cleaned_data['Pitch 4'], marker='.', linestyle='-', color='c', label='Participant 4', alpha=0.5)
plt.plot(cleaned_data['Time'], cleaned_data['Pitch 5'], marker='.', linestyle='-', color='m', label='Participant 5', alpha=0.5)

# Add labels and title
plt.xlabel('Time')
plt.ylabel('Pitch (0 - 360)')
plt.title('Session 1 Time vs Head Pitches')

# Add a legend to differentiate the lines
plt.legend()

# Display grid lines for better readability
plt.grid(True)

# Show the plot
plt.show()
```

Questionnaires

Questionnaire1

Active Listening Training Agent Questionnaire

7 Aug 2024

You are invited!
Location of the event is AR/VR Lab (MB212) at Aston University.

* Required

Section 1/6

1. Participant ID

2. What is your age?

3. What gender do you identify as?

- Male
- Female
- Prefer not to say
- Transgender
- Non-binary

Section 2/6

4. The virtual humans have the appearance of a human *

1	2	3	4	5	6	7

Strongly Disagree Strongly Agree

5. The virtual humans have a human-like manner *

1	2	3	4	5	6	7

Strongly Disagree Strongly Agree

6. The virtual humans seems natural from its outward appearance

*

1	2	3	4	5	6	7

Strongly Disagree Strongly Agree

7. The virtual humans acts like a living organism *

1	2	3	4	5	6	7

Strongly Disagree Strongly Agree

8. The virtual human's appearance is appropriate *

1	2	3	4	5	6	7

Strongly Disagree Strongly Agree

9. The virtual humans are easy to understand *

1	2	3	4	5	6	7

Strongly Disagree Strongly Agree

10. The virtual humans do their tasks well *

1	2	3	4	5	6	7

Strongly Disagree Strongly Agree

11. I like the virtual humans *

1	2	3	4	5	6	7

Strongly Disagree Strongly Agree

12. The virtual humans have a distinctive character *

1	2	3	4	5	6	7

Strongly Disagree — Strongly Agree

13. I will use the virtual humans again in the future *

1	2	3	4	5	6	7

Strongly Disagree — Strongly Agree

14. The first virtual human is boring *

1	2	3	4	5	6	7

Strongly Disagree — Strongly Agree

15. The interaction with the first virtual human captured my attention *

1	2	3	4	5	6	7

Strongly Disagree — Strongly Agree

16. The second human is boring *

1	2	3	4	5	6	7

Strongly Disagree — Strongly Agree

17. The interaction with the second virtual human captured my attention *

1	2	3	4	5	6	7

Strongly Disagree — Strongly Agree

18. The virtual human's behaviour does not make sense *

1	2	3	4	5	6	7

Strongly Disagree — Strongly Agree

19. I see the interaction with the first virtual human as something positive *

1	2	3	4	5	6	7

Strongly Disagree — Strongly Agree

20. I see the interaction with the second virtual human as something positive *

1	2	3	4	5	6	7

Strongly Disagree — Strongly Agree

21. The first virtual human is emotionless *

1	2	3	4	5	6	7

Strongly Disagree Strongly Agree

22. The second virtual human is emotionless *

1	2	3	4	5	6	7

Strongly Disagree Strongly Agree

23. The virtual humans are social entities *

1	2	3	4	5	6	7

Strongly Disagree Strongly Agree

24. The emotions I feel during the interactions are caused by the virtual humans *

1	2	3	4	5	6	7

Strongly Disagree Strongly Agree

Section 3/6

25. The appearance of the first virtual human is realistic *

1	2	3	4	5	6	7

Strongly Disagree Strongly Agree

26. The appearance of the second virtual human is realistic *

1	2	3	4	5	6	7

Strongly Disagree Strongly Agree

27. The first virtual human's voice is realistic *

1	2	3	4	5	6	7

Strongly Disagree Strongly Agree

28. The second virtual human's voice is realistic *

1	2	3	4	5	6	7

Strongly Disagree Strongly Agree

29. The first virtual human's gestures are realistic *

1	2	3	4	5	6	7

Strongly Disagree Strongly Agree

30. The second virtual human's gestures are realistic *

1	2	3	4	5	6	7

Strongly Disagree Strongly Agree

31. The first virtual human's actions are coherent with its speech *

1	2	3	4	5	6	7

Strongly Disagree Strongly Agree

32. The second virtual human's actions are coherent with its speech *

1	2	3	4	5	6	7

Strongly Disagree Strongly Agree

33. It was easy to follow the first virtual human's speech *

1	2	3	4	5	6	7

Strongly Disagree — Strongly Agree

34. It was easy to follow the second virtual human's speech *

1	2	3	4	5	6	7

Strongly Disagree — Strongly Agree

Section 4/6

35. I felt the character was immersed in the environment *

1	2	3	4	5	6	7

Strongly Disagree — Strongly Agree

36. The system ran without glitches *

1	2	3	4	5	6	7

Strongly Disagree — Strongly Agree

37. I felt comfortable wearing the HoloLens headset *

1	2	3	4	5	6	7

Strongly Disagree — Strongly Agree

38. I can imagine that in the future a virtual human can replace a real human teacher *

1	2	3	4	5	6	7

Strongly Disagree — Strongly Agree

39. Using the AI system is enjoyable. *

1	2	3	4	5	6	7

Strongly Disagree — Strongly Agree

40. Using the AI training system increases my attention *

1	2	3	4	5	6	7

Strongly Disagree — Strongly Agree

Questionnaire 2

Active Listening Training Agent Questionnaire (2)

9 Aug 2024

You are invited!
Location of the event.

* Required

Section 1/6

1. Participant ID

Section 2/6

Session 1

Based on Lisa's presentation, please answer the following questions.
Select the best answer for each question.

2. What was the primary topic discussed by Lisa, the virtual agent? *

- ◯ The impact of artificial intelligence on climate change
- ◯ The role of AI in transforming the job market over the next ten years
- ◯ The history of artificial intelligence development
- ◯ How to program artificial intelligence systems

3. Which industry did Lisa mention as already being significantly impacted by AI? *

- ◯ Healthcare
- ◯ Agriculture
- ◯ Manufacturing
- ◯ Retail

4. According to Lisa, what has AI allowed Sarah's marketing agency to do? *

- ◯ Hire more employees
- ◯ Focus on creative campaigns by automating routine tasks
- ◯ Reduce overall business expenses
- ◯ Develop AI-driven marketing strategies

5. Lisa believes that AI will only result in job losses without creating new opportunities. *

- ◯ True
- ◯ False

6. What example does Lisa provide to illustrate the potential negative impact of AI on jobs? *

- ◯ Automation in manufacturing plants
- ◯ AI in retail replacing customer service staff
- ◯ Self-driving cars reducing the need for truck drivers
- ◯ AI tools in healthcare diagnosing diseases

7. Which of the following is NOT a potential new job role mentioned by Lisa that could emerge due to AI advancements? *

- ◯ AI specialists
- ◯ Ethical tech developers
- ◯ Social media influencers
- ◯ Data analysts

8. What personal story does Lisa use to demonstrate the benefits of upskilling in response to AI? *

- ◯ Her mother learning about AI technologies
- ◯ Jake moving from retail to a tech division by learning data analysis
- ◯ Sarah's transition to an AI specialist
- ◯ Uncle Joe becoming a self-driving car engineer

9. What ethical concern does Lisa highlight regarding the development and use of AI? *

- ◯ AI's potential to become self-aware and autonomous
- ◯ Data privacy and bias in algorithms
- ◯ The environmental impact of AI technologies
- ◯ AI replacing human creativity in art and music

10. What analogy does Lisa use to explain AI to her mother? *

- ◯ Comparing AI to a personal assistant
- ◯ Joking about a talking toaster
- ◯ Relating AI to a smartphone app
- ◯ Describing AI as a new family member

11. How AI might create new job opportunities according to Lisa's presentation. *

- ◯ By completely automating all existing jobs, leading to the creation of new industries.
- ◯ By taking over all manual tasks, eliminating the need for human labour.
- ◯ By eliminating some roles while creating demand for new jobs such as AI specialists, data analysts, and ethical tech developers.
- ◯ By reducing job opportunities in all sectors, leading to economic decline.

12. What skills do you think are most important to develop for the future job market? *

- ◯ Technical skills related to AI and machine learning
- ◯ Soft skills like creativity and emotional intelligence
- ◯ Business and management skills
- ◯ All of the above

13. How can governments and businesses ensure that AI is used responsibly? Provide two strategies mentioned or implied in the talk.
Choose two of the following:
a) Develop comprehensive policies that address data privacy and algorithmic bias.
b) Allow AI to develop freely without regulations to encourage innovation.
c) Collaborate with international organizations to create global AI standards.
d) Create job displacement programs to mitigate the impact on affected workers.
e) Ensure that only private companies handle AI regulation and implementation. *

- ◯ a and b
- ◯ a and c
- ◯ b and d
- ◯ c and e

14. Which statement best captures the overall message of Lisa's presentation? *

- ◯ AI will replace most human jobs, and we should be prepared for a jobless future.
- ◯ AI presents both challenges and opportunities, and adaptation through learning is essential.
- ◯ AI will primarily benefit tech industries, leaving other sectors behind.
- ◯ AI's influence will be limited and won't significantly impact the job market.

Questionnaire 3

Active Listening Training Agent Questionnaire (3)

9 Aug 2024

You are invited!
Location of the event.

Section 1/6

1. Participant ID

Section 2/6

Session 2

Based on Desmond's presentation, please answer the following questions.
Select the best answer for each question.

2. What was the first place Desmond visited at Universal Studios, and why was it significant to him?

- ◯ The Simpsons Ride; because he loves the show and wanted a pink donut.
- ◯ The Jurassic Park area; because he loves dinosaurs.
- ◯ The Wizarding World of Harry Potter; because he is a Potterhead.
- ◯ The Simpsons Ride; because he watched the show.

3. What item did Desmond's friend Alex purchase, and what was Desmond's humorous observation about it?

- ◯ A Gryffindor scarf; Desmond joked about it being as magical as it looked.
- ◯ A Butterbeer mug; Desmond joked about drinking too much sugar.
- ◯ A dinosaur toy; Desmond joked about it being ferociously fun.
- ◯ A wand; Desmond joked about it being a stick of wood.

4. Which ride did Desmond describe as feeling like living out his childhood dreams, and what specific elements did he highlight about the experience?

- ◯ The Transformers ride; Desmond highlighted the 3D effects.
- ◯ The Forbidden Journey ride; Desmond highlighted flying on a broomstick and battling a dragon.
- ◯ The Jurassic World ride; Desmond highlighted the T-Rex chase.
- ◯ The Simpsons ride; Desmond highlighted the cartoonish fun.

5. What precaution did Desmond suggest for those planning to go on the Jurassic World ride?

- ◯ Bring a camera for the epic scenes.
- ◯ Wear comfortable shoes for the long wait.
- ◯ Bring a poncho or a change of clothes to avoid getting wet.
- ◯ Wear sunglasses for the outdoor portions.

6. Where did Desmond and Alex go to eat, and what iconic food item did they try there?

- ◯ Moe's Tavern; they tried a Duff beer.
- ◯ The Kwik-E-Mart; they tried a Squishee.
- ◯ Lard Lad; they tried a giant pink donut.
- ◯ Krusty Burger; they tried a Krusty Burger.

7. Why did Desmond likely compare waking up on the day of the trip to Christmas morning?

- ◯ Because he received gifts from his friends.
- ◯ Because he was filled with excitement and anticipation.
- ◯ Because it was a holiday for him.
- ◯ Because he decorated his house for the occasion.

8. How does Desmond's description of the Nighttime Lights at Hogwarts Castle contribute to the overall theme of the adventure?

- ◯ It emphasizes the technological advancements at Universal Studios.
- ◯ It highlights the park's ability to provide thrilling rides.
- ◯ It emphasizes the sense of magic and wonder throughout the experience.
- ◯ It shows the variety of dining options available at the park.

9. What can you infer about Desmond's personality based on his experiences and observations at Universal Studios?

- ◯ He is a serious and reserved individual.
- ◯ He is adventurous, nostalgic, and has a sense of humour.
- ◯ He prefers solitary activities.
- ◯ He is uninterested in fictional worlds.

10. Why might Desmond have chosen to share his experience at Universal Studios with you?

- ◯ To complain about the long lines and crowds.
- ◯ To convey his excitement and relive the memorable day.
- ◯ To share detailed technical aspects of the rides.
- ◯ To discuss the park's food pricing.

11. How did Desmond's relationship with his best friend, Alex, enhance his experience at the park?

- ◯ Alex's seriousness helped Desmond focus on the details.
- ◯ Their shared excitement and interests made the experience more enjoyable.
- ◯ Alex's planning skills allowed them to avoid crowds.
- ◯ Desmond felt responsible for showing Alex around the park.

12. What emotions do you think Desmond experienced during the Nighttime Lights at Hogwarts Castle, and why were they significant?

- ◯ Boredom and fatigue; the day was too long.
- ◯ Confusion and uncertainty; the lights were overwhelming.
- ◯ Awe and wonder; the experience encapsulated the day's magic.
- ◯ Indifference; he had seen similar light shows before.

13. How does Desmond's narration encourage others to visit Universal Studios?

- ◯ By emphasizing the park's technological innovations.
- ◯ By highlighting the fun, nostalgia, and variety of experiences available.
- ◯ By detailing the park's dining options.
- ◯ By describing the challenges of navigating large crowds.

14. How does Desmond use humour in his storytelling.

- ◯ He makes self-deprecating jokes about his excitement to lighten the mood, which makes the story more relatable and enjoyable.
- ◯ He makes jokes about the park's staff to highlight the importance of good service.
- ◯ He mocks the park's attractions to show their shortcomings, which adds a critical tone to his narrative.
- ◯ He uses humour sparingly, which makes his descriptions dry and factual.

15. What role does nostalgia play in Desmond's experience?

- ◯ Nostalgia makes him feel disconnected from the present, creating a negative tone in his narrative.
- ◯ Nostalgia enhances the story by connecting his childhood interests with current experiences, making it relatable.
- ◯ Nostalgia is not present in his narrative, as Desmond focuses on the future.
- ◯ Nostalgia creates a longing for more traditional theme park experiences, detracting from his enjoyment.

16. How does Desmond structure his story to maintain interest and engagement throughout the narrative.

- ◯ He focuses solely on technical details, which makes the story engaging for engineering enthusiasts.
- ◯ He mixes vivid descriptions, personal anecdotes, and humour, maintaining a lively and dynamic narrative.
- ◯ He provides a detailed itinerary of the day, which adds a monotonous tone to his storytelling.
- ◯ He only focuses on negative experiences, keeping the story engaging through criticism.

Ethics Approval

MARK
LAST MARKED ATTEMPT

Complete

ATTEMPT 2
10/08/24 02:35

Complete

Submission

In Person VR Study.zip

Comments

Feedback to Learner
12/08/24 00:16

Congratulation! Your ethics approval application has been approved. Please find the review form below for your reference.

Best wishes for your dissertation.

Joe

11/8/2024
244700-230387748a_Ososeliase Inegbedion_Ethics Application Review Form_AR_Studies_Survey .pdf

Aston University
BIRMINGHAM UK

College of Engineering & Applied Sciences Research Ethics Committee
Ethics Application Review Form

When completing this form please refer to the reviewer guidance notes

Ethics application number	**244700-230294664a**	**Date of review**	**3-July-23**
Study title	Active Listening Training Agent		
Name of applicant(s)	Ososeliase Inegbedion		
Name of reviewer	**Joe Yuen**		

STUDY PROTOCOL Please provide comments on whether the protocol for the study is clear and appropriate, including rationale, design, methods for recruitment and data analysis.
This study plans to conduct a VR study to explore the effectiveness of an Active Listening Training Agent within augmented reality environments. The experiment will be taken placed at Aston University Main building (MB212) with risk assessment. Participants interaction with the AR system will be observed and record. **Participants will be recruited via email.**

MATERIALS

Please note the following:

1. All materials should include the Aston University logo and have a date and version number on each page.
2. University pro-formas for the Participant Information Sheet and Consent form should be used. [Note: forms can be edited to suit the individual project.]
3. GDPR compliance statement (Appendix A) needs to be included in the PIS.

Please provide comments on the following:	
Participant Information Sheet	**Well prepared**
Consent form	**Participants will be informed they can withdraw at any time during the studies and their data could be deleted within 14 days after the studies have taken place.** **Personal information will be deleted after report submission.** **The pseudonymised data will be stored for at least 6 years after completion of the study**
Debrief (if applicable)	**N/A**

Aston University
BIRMINGHAM UK

Any other supplied materials (e.g. Poster, Advert for recruitment, questionnaire, interview prompts, etc)	**Recruitment message, Questionnaire, Risk Assessment Record**

RISK ASSESSMENT Has a risk assessment been carried out? Is the assessment of the level of risk appropriate and are actions proposed to mitigate risks?
A H&S risk assessment are well-prepared.

CONSENT AND WITHDRAWAL
Are processes for obtaining consent clear and appropriate?
The PIS and consent form will be sent along with the recruitment message. Participants must sign the consent form and attach it in their replies.
Is the process for withdrawal clearly explained to participants?
Yes, participant can be withdrawn at any time during the interview. Data submitted by participant will be anonymised after 14 days and withdrawal of submitted data is not possible.

RESEARCH GOVERNANCE	
Is the researcher competent to carry out the research, or are they being supervised by a competent researcher?	**Yes**
Have details been provided about how long data will be stored and where?	**Data with personal information will be deleted after submission.**
Have details been provided about who will access the data?	**Project supervisor**
Has indemnity and sponsorship been specified (to note all staff and student research is covered by Aston University and this should be stated unless there is another sponsor)	**Aston University is the only sponsor as stated**

ANY OTHER COMMENTS

Aston University
BIRMINGHAM UK

Please make sure the distribution list of recruitment message is taken from your personal / supervisor's network and the use of email address does not breach GDPR.

REVIEW OUTCOME: Please underline one

Approve (study can proceed)
~~**Conditional (minor clarifications, amendments)**~~
~~Not Approve (extensive revisions and resubmission required)~~
~~No opinion (expert advice needed)~~

Ethics Application

Ethics Declaration

Aston University

Ethics Declaration Form for CS4700/CS4705 Student Dissertation Projects

EPS Department: Computer Science

Project Title: Active Listening Training Agent

Supervisor Name and e-Mail: Dr Ulysses Bernardet name / u.bernardet@aston.ac.uk

Student Name, Number and e-Mail: Ososeliase Inegbedion/ 230387748/ 230387748@aston.ac.uk

Section 0: Ethics Declaration Questions

If you are conducting a study for your supervisor where your supervisor already has ethical approval for the work to be conducted, please consult with your supervisor to ensure that (s)he has applied for an amendment to the existing ethical approval in order to permit your involvement. A letter from the College Research Ethics Committee confirming that permission will need included with your project report. You do not need to complete the remainder of this form.

Please answer the following questions honestly and carefully:

Will your project involve <u>any</u> of the following?	*Delete as applicable*
The NHS – either patients selected via the NHS or clinical staff working for the NHS	Not Permitted
Participants under the age of 18	Not Permitted
Vulnerable groups – individuals with physical disabilities, mental health disabilities/ill health, individuals with learning difficulties, prisoners/detained persons, people over 65 years of age, and/or pregnant women	Not Permitted
Any of the following: i) clinical procedures or ii) physical intervention or iii) penetration of the participant's body or iv) prescription of compounds additional to normal diet or other dietary manipulation/supplementation or v) collection of bodily secretions or vi) involve human tissue which comes within the Human Tissue Act (e.g., surgical operations; taking body samples including blood and DNA; exposure to ionizing or other radiation; exposure to sound light or radio waves; psychophysiological procedures such as fMRI, MEG, TMS, EEG, ECG, exercise and stress procedures; administration of any chemical substances)	Not Permitted
Delving into topics that might be sensitive, embarrassing or upsetting or where it is possible that criminal or other disclosures requiring action could take place (e.g., during interviews) – including but not limited to projects focusing on mental health	Not Permitted
Human participants - *sessions wearing the Microsoft HoloLens 2 augmented reality headset.* - *consent form.* - *sessions active listening is monitored, and data is recorded.* - *questionnaires assessing the content and efficacy of the presentation experienced.* - *questionnaires gauging their subjective experience and impressions of the sessions.* - *the system database and administered questionnaires are collated and analysed.*	Yes

Testing of apparatus - *Developed using unity game engine* - *HoloLens 2 for virtual testing user experience*	Yes
Risk to you, including:	Yes/No
lone working during data collection	No
travel to areas where you may be at risk	No
risk of emotional distress	No
other: *please outline*	No
Any risk to the environment	No
Any conflict of interest	No
Work/research that could be considered controversial or be of reputational risk to Aston University	No
Social media data and/or data from internet sources that could be regarded as private	No
Any other ethical considerations *(please state here or contact the Research Ethics Officer via your College Ethics inbox if there are any substantial ethical considerations you are aware of and would like to flag for the REC.)*	No

If you answer 'yes' to any of the questions above, you will need to continue to complete the rest of this form as well as any required participant-facing documentation (see Ethics Guidance provided). You and your supervisor will need to sign the declarations section at the end of this form before submitting it, along with all other paperwork, for ethical approval. No data collection and use (including recruitment of participants) may take place before ethical approval has been granted. All ethics paperwork, including evidence of ethical approval, should be included in your final report submission.

If you answered 'no' to all of the questions above, you may skip to Section 2. You and your supervisor will need to sign the declarations section at the end of this form. The form will need included in your final report submission.

Section 1: Study Details

Please provide the following information about your study. Be as detailed as possible. Where a question is not relevant, please indicate 'Not Applicable' but also explain why you believe that to be so.

Study Details	
Project Objectives *Please provide a brief outline of your overall project objectives*	The project aims to create an Active Listening Training Agent using virtual human agents in an AR environment, powered by Unity and Microsoft HoloLens 2. It seeks to enhance active listening through immersive scenarios and feedback mechanisms.
Study Objectives *Please explain how the study you are seeking ethical approval to conduct contributes to your overall project objectives*	The study aims to explore the development and effectiveness of an Active Listening Training Agent within augmented reality environments, focusing on immersive scenarios, real-time feedback mechanisms, and participant engagement.
Data Collection Method(s) to be Used *Please outline your proposed data collection methods – e.g., questionnaire/survey, interview, observational study, etc. Justify their use and explain how you will conduct the data collection in practice, including timeframe*	The data collection methods for the project encompass a blend of quantitative and qualitative approaches to comprehensively evaluate user interaction and training outcomes. Firstly, participants will engage in two informative sessions with virtual human agents wearing the Microsoft HoloLens 2 augmented reality headset, facilitated by the Unity game engine platform. During these sessions, both implicit feedback during interactions

	and explicit feedback at the session end will be recorded. To assess the effectiveness and user experience, participants will complete two distinct questionnaires following each session: one evaluating the content and efficacy of the presentation experienced, and another gauging subjective experiences and impressions throughout the sessions. The timeframe for data collection will involve successive 5-minute sessions during which active listening features will be monitored and data recorded. After each session, questionnaires will be administered promptly.
Data to be Collected *Please briefly outline the type of data to be collected*	The research design involves collecting diverse logs of audio and video recordings to prepare datasets ensuring the authenticity of the Active Listening Training Agent's learning experiences. During 5-minute sessions with virtual human agents, the system monitors active listening features and records data. including eye tracking (focusing on the speaker, frequency, and duration of eye contact), head movement (nods, and head tilts), and voice analysis (interruptions). AR headsets equipped with eye-tracking sensors and built-in accelerometers and gyroscopes capture these metrics, while this data is used to analyse user engagement. Post-sessions questionnaires assess content efficacy and participant experience.
Location of Data Collection *Please briefly outline where you will be collecting the data – e.g., where you will be conducting your study.*	Data collection for the project will occur in a controlled augmented reality environment within the Aston University building. The AR/VR laboratory (MB212) within Aston University.
Participant Recruitment *Please outline how you will recruit participants to your study and how you will ensure that participants are not coerced to participate.*	To recruit participants for the study, we'll start by clearly explaining the project's goals and procedures. Potential participants will be provided with a recruitment message either by email or in person, depending on access to individuals. This will entail to post graduate students at Aston University within the AI department or outside interested parties. Participants will then sign a consent form, confirming their voluntary involvement and understanding of the study. Which also ensure participants know they can withdraw at any time without consequences.
Data Storage *Please outline where you will store your data (ideally, on an encrypted server; USB drives are not permissible)*	In compliance with data security protocols, all data generated from the project will be securely stored on my personal laptop or an encrypted cloud database. Given the sensitivity of the information collected, USB drives are strictly prohibited to mitigate potential data breaches or loss. By utilizing an encrypted cloud database, we ensure that the data remains protected from unauthorized access while maintaining accessibility for analysis and research purposes. This approach aligns with best practices for safeguarding research data, prioritizing confidentiality and integrity throughout the data storage process.

Data Deletion *Please outline when you will erase the data you collected. For personally identifiable information, it should be deleted once an anonymised version is created, e.g., audio recordings should be deleted once the corresponding anonymised transcripts are created. For anonymised data, it should be kept until the final grade of the dissertation is released OR there is agreement from participants (via the consenting process) to retain and share data for/with future research(ers). Please provide justification if you have no plan to delete the collected data.*	Data collected for the project will be managed according to strict protocols to ensure privacy and ethical standards. Personally identifiable information will be promptly deleted once anonymized versions are created, such as age and gender being erased once anonymized versions are generated. Participants understand that they are able to withdraw their data up to 14 days after taking part in the study by contacting the study team; after this period, their data will be anonymized and no longer withdrawable. Anonymized data will be retained until the final grade of the dissertation is released or until there is agreement from participants, obtained through the consenting process, to retain and share data for or with future researchers.
Data Access *Please indicate who will have access to the data you collected - e.g., your supervisor and any other members of the research team.*	Access to the data collected in this study will be limited to authorized personnel involved in the research project, including the supervisor and possible markers. Participants' data privacy and confidentiality will be strictly maintained, adhering to ethical guidelines and consent agreements.
For Secondary Data/Dataset Use Only: Compliance with Terms & Conditions of Use *If you will be in receipt of secondary data OR will be using an online, publicly available dataset, please provide evidence that you are observing any terms and conditions associated with its use and have permission to use it. Be mindful that just because data is available online does not mean that you are ethically entitled to use it for your study; this needs proven. If you are being given data by, for example, a third party, you need to be sure that the individual has permission to share the data with you.*	N/A
Risk *Please outline any risks to either the participants in your study and/or yourself in the conduct of the study and what you have done to mitigate that risk*	Participants may experience discomfort or disorientation from prolonged AR use, potential motion sickness, or psychological distress from immersive scenarios. To mitigate these risks, the study adheres to ethical guidelines, ensuring informed consent, providing short 5-minute sessions with breaks in between. Additionally, the development process prioritizes user safety by thoroughly testing the AR environment for stability and usability and incorporating features to minimize motion sickness. Researchers also face risks such as technical glitches, data security breaches, or psychological strain from managing participant experiences. Mitigation strategies include regular system maintenance, data encryption protocols, and establishing a system for running the sessions.

Section 2: Declarations

The following declaration should be signed by both you, the student, and your supervisor

Student:

I confirm the following:

- The above is an accurate representation of my study activities;
- That I shall not commence participant recruitment and/or data collection without ethical approval to do so (where applicable);
- That I shall seek further ethical approval should I need to make any changes to my study after ethical approval has been granted (where applicable);
- That I shall conduct my study with integrity and in accordance with the ethical approval granted (where applicable);
- That, where necessary, I shall use existing or secondary data in accordance with terms and conditions of use or ethical approval, as applicable;
- That I understand that if I breach the terms of the approval granted I may not be able to use the data collected in my project report and may face disciplinary procedures; and
- That I shall respect my participants (where applicable) and the data I have collected and am using.

Supervisor:

I confirm the following:

- That I have reviewed the content of this form and all associated paperwork and am happy with its standard and accuracy;
- That I shall monitor the student's conduct of the study in accordance with the ethical approval granted (where applicable); and
- That I shall report to the person(s) granting ethical approval any breaches of approval and ensure that no data is included in the student's work that has been collected in breach of approval.

Both student and supervisor should sign* and date below:	
Signatures: Ososeliase Inegbedion	Date:06/08/2024

** note, typed/e-signatures are acceptable.*

Consent Form

Aston University
BIRMINGHAM UK

Active Listening Training Agent
Consent Form

Name of Student: Ososeliase Inegbedion

Please initial boxes

1.	I confirm that I have read and understand the Participant Information Sheet (Version PIS-IP VR 1.1 06-08-2024) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	
2.	I acknowledge the risks associated with participation in this study and confirm that I have **no known history of epilepsy or other condition which could be triggered by the use of virtual reality**, have **no mobility issues that would stop me walking (including on a treadmill)**, **have no breathing difficulties that could be exacerbated by light exercise**, and **do not suffer from severe motion sickness**.	
3.	I understand that my participation is voluntary and that I am free to withdraw at any time during the study, without giving a reason and without my legal rights being affected.	
4.	I understand that I am able to withdraw my data up to 14 days after taking part in the study by contacting the study team, after this time my data will be anonymised and I will no longer be able to withdraw.	
5.	I agree to my personal data and data relating to me collected during the study being processed as described in the Participant Information Sheet.	
6.	I agree to my anonymised data being used by research teams for future research.	
7.	I agree to take part in this study.	

Name of participant	Date	Signature
Name of Person receiving consent.	Date	Signature

Recruitment Message

Aston University
BIRMINGHAM UK

Active Listening Training Agent
Recruitment Message

Dear <name>,

My name is Ososeliase Inegbedion, and I am currently undertaking my dissertation project as part of my MSc degree at Aston University. The aim of my project is to study and improve active listening by immersing individuals in dynamic scenarios, offering real-time feedback, and integrating cutting-edge AI features.

This study is open to all participants aged 18+ years old, who are fluent in English, have normal or corrected to normal vision (via glasses or contact lenses) and who have **no history of epilepsy or other condition which could be triggered by the use of virtual reality**, have **no mobility issues that would stop you walking on a treadmill**, **have no breathing difficulties that could be exacerbated by light exercise,** and **do not suffer from severe motion sickness**. We would respectfully request that, in particular, for safety reasons if you do suffer from any of the issues highlighted above, you do not proceed to participate in this study.

To this end, if you meet the inclusion criteria noted above, I would like to invite you to take part in a research study at Aston University during which we will explore the effectiveness of virtual agents on Active Listening skills, focusing on immersive scenarios, real-time feedback mechanisms, and participant engagement within a VR environment. It is anticipated this session should take approximately 25 minutes of your time.

Before agreeing to take part in this interview, please would you read the attached **Participation Information Sheet**, which will detail further the purpose of the interview. If you have any questions, please email me and I will be only too happy to answer them. If, after having any questions answered to your satisfaction, you would like to participate in the interview, please read and sign the attached **Consent Form**, and send it back to 230387748@aston.ac.uk. Upon receipt, I shall be in touch to set up the session at a date/time that is convenient to you.

I look forward to hearing back from you soon.

Kind Regards,

Ososeliase Inegbedion

Participant Information Sheet

Aston University
BIRMINGHAM UK

Active Listening Training Agent

Participant Information Sheet

Invitation

I would like to invite you to take part in a research study.

Before you decide if you would like to participate, take time to read the following information carefully and, if you wish, discuss it with others such as your family, friends or colleagues.

Please ask me, my contact details can be found at the end of this information sheet, if there is anything that is not clear or if you would like more information before you make your decision.

What is the purpose of the study?

The aim of this project is to study and improve active listening by immersing individuals in dynamic scenarios, offering real-time feedback, and integrating cutting-edge AI features.

The aim of this specific study is to explore the development and effectiveness of Active Listening skills using virtual agents, focusing on immersive scenarios, real-time feedback mechanisms, and participant engagement.

Why have I been chosen?

You are being invited to take part in this study because you have responded to an advert to take part in this study. Your insights and participation are crucial in shaping the future of active listening training. Your unique perspective and engagement will contribute significantly to our understanding and refinement of effective training methods within augmented reality environments.

This study is open to all participants aged 18 years old or over, who are fluent in English, have normal or corrected to normal vision (via glasses or contact lenses) and who have **no history of epilepsy or other condition which could be triggered by the use of augmented reality**, have **no mobility issues that would stop you walking (including on a treadmill), have no breathing difficulties that could be exacerbated by light exercise**, and **do not suffer from severe motion sickness**. We would respectfully request that, in particular, for safety reasons if you do suffer from any of the issues highlighted above, you do not proceed to participate in this study.

What will happen to me if I take part?

If you decide to participate, you will be invited to attend a study session at Aston University which should take about 25 minutes to complete.

You will be required to wear an augmented Reality (AR) headset via which you will be 'entered' into an augmented reality environment comprising the VR laboratory within Aston University and virtual agents. You will be exposed to virtual agents designed to engage users in two 5 mins informative sessions with a break in between, where participants aim to absorb the content efficiently and answer questionnaires following their completion. Each session will last 5 minutes, during which participants will remain seated or standing in a comfortable environment based on preference. During the sessions, a participants active listening skills will be monitored in real-time.

Following completion of the tasks, will be asked to remove your AR. You will be asked to complete a survey designed to measure your experience in the AR environment, completing the tasks; these should take 5-10 minutes to complete in total.

All equipment – the AR headset and computer surfaces – will be sanitised between participants.

Do I have to take part?

No. It is up to you to decide whether or not you wish to take part.

If you do decide to participate, you will be asked to sign and date a consent form. You can halt your participation at any time just by letting the student know and any data collected up to that point will not be used. You would still be free to withdraw from the study at any time up to 2 weeks from the date of your interview without giving a reason by emailing the student.

Will my taking part in this study be kept confidential?

Yes. A code will be attached to all the data you provide to maintain anonymity. Analysis of your data will be undertaken using coded data.

If we need to collect personal data (such as a name and contact details) we will only use this for the purposes outlined in this participant information sheet – e.g., to contact you to arrange the interview.

The data we collect will be stored in a secure document store (paper records if applicable) or electronically on a secure encrypted mobile device, password protected computer server or secure cloud storage device.

To ensure the quality of the research, Aston University may need to access your data to check that the data has been recorded accurately, e.g., for the purposes of audit. If this is required, your personal data will be treated as confidential by the individuals accessing your data.

What are the possible benefits of taking part?

While there are no direct benefits to you of taking part in this study, the data gained will offer unique and engaging insights into AR technology, potentially improving digital literacy and comfort with virtual environments. Your involvement may ultimately contribute to the development of more effective training programs, educational tools, or virtual onboarding platforms.

What are the possible risks and burdens of taking part?

If you have any history of epilepsy (or related condition that could be trigged by the use of the virtual reality system) or any severe motion sickness, please DO NOT participate in this study.

While the risks have been mitigated to the best of our knowledge, we cannot guarantee there being no risk.

For some people, using a Augmented Reality headset can cause the feeling of nausea, similar to motion sickness. The movement system and environment has been designed and optimised to minimise the risk of this occurring. If at any point you do feel unwell, you will be free to stop immediately and allow yourself to recover; you will be able to withdraw completely at any point. **We recommend that if you suffer from severe motion sickness you DO NOT participate in this study.**

Please also take the time to make sure you feel safe with the equipment. You will be able to ask for help and take breaks at key points during your session.

You will be asked to sanitise your hands before taking part in the study; the researcher/student will do the same. All equipment will be thoroughly sanitised for each individual.

In terms of burden, the sesssion will require approximately 25 minutes of your time.

What will happen to the results of the study?

The results of this study will be published in the MSc dissertation report of the student, wherein your identity will remain confidential.

The results of this study may be published in scientific journals and/or presented at conferences. If the results of the study are published, your identity will remain confidential.

The anonymized results may be used for research by other research teams as described in Appendix A.

Expenses and payments

There are no expenses or payments being provided for participation.

Who is funding the research?

The study is being funded by Aston University.

Who is organising this study and acting as data controller for the study?

Aston University is organising this study and acting as data controller for the study. Research data will be used only for the purposes of the study or related uses identified in this Information Sheet or Appendix A.

Who has reviewed the study?

This study was given a favorable ethical opinion by delegated authority of the Engineering & Physical Sciences Research Ethics Committee.

What if I have a concern about my participation in the study?

If you have any concerns about your participation in this study, please speak to the student and s(he) will do his/her best to answer your questions. Contact details can be found at the end of this information sheet.

If the student is unable to address your concerns or you wish to make a complaint about how the study is being conducted you should contact the Aston University Research Integrity Office at research_governance@aston.ac.uk or telephone 0121 204 3000.

Research Team Details

Ososeliase Inegbedion, 230387748@aston.ac.uk , and +447307495063
Dr Ulysses Bernardet, u.bernardet@aston.ac.uk , and +44121 204 3893

Thank you for taking time to read this information sheet. If you have any questions regarding the study please don't hesitate to ask one of the research team.

Appendix A: Transparency Statement

Aston University takes its obligations under data and privacy law seriously and complies with the Data Protection Act 2018 ("DPA") and the General Data Protection Regulation (EU) 2016/679 as retained in UK law by the Data Protection, Privacy and Electronic Communications (Amendments etc) (EU Exit) Regulations 2019 ("the UK GDPR").

Aston University is the sponsor for this study based in the United Kingdom. We will be using information from you in order to undertake this study. Aston University will process your personal data in order to register you as a participant and to manage your participation in the study. It will process your personal data on the grounds that it is necessary for the performance of a task carried out in the public interest (GDPR Article 6(1)(e). Aston University may process special categories of data about you which includes details about your health. Aston University will process this data on the grounds that it is necessary for statistical or research purposes (GDPR Article 9(2)(j)). Aston University will keep identifiable information about you for 6 years after the study has finished. Your rights to access, change or move your information are limited, as we need to manage your information in specific ways in order for the research to be reliable and accurate. If you withdraw from the study, we will keep the information about you that we have already obtained. To safeguard your rights, we will use the minimum personally identifiable information possible.

You can find out more about how we use your information at https://www.aston.ac.uk/about/statutes-ordinances-regulations/publication-scheme/policies-regulations/data-protection or by contacting our Data Protection Officer at dp_officer@aston.ac.uk.

If you wish to raise a complaint on how we have handled your personal data, you can contact our Data Protection Officer who will investigate the matter. If you are not satisfied with our response or believe we are processing your personal data in a way that is not lawful you can complain to the Information Commissioner's Office (ICO).

When you agree to take part in a research study, the information about you may be provided to researchers running other research studies in this organisation and in other organisations. These organisations may be universities, NHS organisations or companies involved in health and care research in this country or abroad.

This information will not identify you and will not be combined with other information in a way that could identify you. The information will only be used for the purpose of research, and cannot be used to contact you.

Risk Assessment

Risk Assessment Record

Description of Activity: **Stationary VR/AR experiments** Location of Activity: **MB212/MB216 cubicle**

1. Hazards associated with the activity	**2.** Consequences of exposures to the hazards	**3.** Who or what at risk and numbers affected	**4.** What control measures are in place to prevent this from happening	**5.** With Initial	Initial	Initial	**6.** Risk Category	**7.** Action required to reduce the risk to an acceptable level	**8.** Responsibility for action and timescales to complete	**9.** With Revised	Revised	Revised
Participants invited on campus	General concerns about safety for people visiting the campus.	The participants.	The experiment will be conducted during normal university working hours with security being on campus. The instructor will supervise the experiment at all times. SafeZone is suggested. First aid will be available.	3	1	3	Acceptabl e	No action required.		3	1	3
Electricity: VR/AR machinery is mains and battery powered. The participant utilises a Head-mounted Display (HMD) and controllers.	Electrical shock from machinery leading to severe injury and/or death	Researcher Participant Cleaning Staff Visitors	VRAR machinery is supplied by Insight Manufacturer's instructions and safety information will be followed Machinery visually inspected before use for damage. Computer to which the Head-mounted Display is connected is grounded and on a circuit breaker.	4	1	4	Acceptable	Current risk level acceptable no further actions required		4	1	4
VR/AR leading to disorientation leading to giddiness and/or injury when using HMD (head-mounted Display)	Participant may feel ill and giddy when using HMD. Participant may be injured due to lack of orientation in 'real life' setting when using the HMD.	Participant	Time spent using HMD is less than 15 minutes Participant is under the instruction and supervision of researcher during the experiment and can withdraw from the experiment at any time. Clear concise instructions provided to participant Designated area for participant which has been cleared of trip and slip hazards	3	2	6	Acceptable	Current risk level acceptable no further actions required		3	2	6
Fire	Serious injury due to fire and/or smoke inhalation from building/equipment fire	Researcher Participant Cleaning staff Visitors	Aston University has fire detection and evacuation procedures in place. See electricity controls above	5	1	5	Accepta ble	Current risk level acceptable no further actions required		5	1	5
Slips, trips, falls	Bruises, abrasions, and possible broken bones could occur following a slip, trip or fall over trailing cables, spillages, etc.	Researcher Participant Cleaning staff Visitors	Good housekeeping in place Cables tied off and cleared from walkways Designated researcher and participant areas cleared of trip hazards Spillages cleared up immediately	3	2	6	Acceptable	Current risk level acceptable no further actions required		3	2	6
Manual Handling: handling, carrying, and wearing of VRAR equipment	Musculoskeletal injury from carrying and wearing VRAR equipment.	Researcher	VR/AR IT equipment is relatively lightweight (approximately less than 10kg) VR and AR HMD are relatively light (approximately 0.5kg) and designed to fit the user.	3	1	3	Acceptabl e	Current risk level acceptable no further actions required		3	1	3

10.

Assessor Name: Ulysses Bernardet / Amal Htait

Manager responsible for the assessment: Ulysses Bernardet

Peer Reviewer (where required)

Date: 08/08/2024

Review Due: 08/08/2025
(Max. 12 months)

1-3	Insignificant
7-6	Low
7 - 10	Moderate
12-19	High
>20	Critical
Greater than 12	Action Needed

Risk Assessment = Level of Risk	1. Negligible	2. Minor	3. Moderate	4. Major	5. Critical
1. Rare	**1**	**2**	**3**	**4**	**5**
2. Unlikely	**2**	**4**	**6**	**8**	**10**
3. Possible	**3**	**6**	**9**	**12**	**15**
4. Likely	**4**	**8**	**12**	**16**	**20**
5. Almost Certain	**5**	**10**	**15**	**20**	**25**