USING VIRTUAL REALITY AND ELECTROENCEPHALOGRAPHY TO INVESTIGATE THE EFFECT OF EYE-GAZE IN JOINT ATTENTION.

Clíona Leatrice Kelly

Doctor of Philosophy

ASTON UNIVERSITY

May 2023

Aston University

Clíona Leatrice Kelly

Doctor of Philosophy, 2023

Abstract

This thesis presents the use of virtual reality combined with EEG techniques to investigate joint attention. When two individuals can successfully exchange information with each other, it can aid the development and strength of that relationship. However, some individuals struggle to engage in this exchange and as a result, this can lead to difficulties in developing social and romantic relationships.

Until recently, most of the experiments that have investigated social cognition have used either (i) on-screen computer experiments or, (ii) observation experiments. These methods are limited in either (i) ecological validity or, (ii) internal validity. One solution that was used to combat this limitation is by combining virtual reality and electroencephalogram (EEG) research methods to create a 'Neuro-VR' approach. This facilitated (i) the control the ecological validity by using a virtual human that participants interacted and collaborated with and, (ii) maintained internal validity by displaying a controlled and consistent environment and behaviour from the virtual human.

Studies one and two refined the gaze sequence and made changes to develop the paradigm. When participant-facing research could resume, I ran the refined paradigm with the Neuro-VR research method. The paradigm consisted of two levels: Collaboration; which dictated whether the virtual human would present informative gaze shifts to the target puzzle piece (Collaborative) or the virtual human will provide gaze shifts that are not informative (Non-Collaborative). The second level is Gaze Type; where the virtual human would engage in eye-contact with the participant (Eye) or not (No Eye).

We found that an immersive environment such as a virtual reality head-mounted display had an effect on Collaboration and Gaze Type when compared to a 2D screen. On a neural level, this resulted in significant alpha-band decreases that were predominantly observed in the effect of Collaboration and significant theta-band increases within Gaze Type comparisons. Overall, the findings suggest an effect of gaze on joint attention where direct eye-contact may facilitate faster cognitive processing (observed through faster response times) that possibly indicates an influence of social gaze. Further investigations using dynamic stimuli should be used to explore the effect.

Keywords: Electroencephalography; Joint Attention; Virtual Reality; Virtual Humans.

Statement of Originality

Declaration of Originality

The work in this thesis was carried out in the Insititute of Health and Neurodevelopment at Aston University under the supervision of Professor Tim Meese, Professor Klaus Kessler, Dr Ulysses Bernardet and Dr Johanna Zumer. I confirm that the work presented forthwith is my own. The material contained within this thesis is original and where information has been derived from external sources, this has been clearly indicated.

Clíona Leatrice Kelly

Publications

Publications that have arisen from work detailed in this thesis:

Gregory, S E, **Kelly, C L** and Kessler, K.

"Look into my "virtual" eyes: What dynamic virtual agents add to the realistic study of joint attention". Frontiers in Virtual Reality, Virtual Reality and Human Behaviour.**2021**

Kelly, C L, Bernardet, U, and Kessler, K.

"A Neuro-VR toolbox for assessment and intervention in Autism: Brain responses, to non-verbal, gaze and proxemics behaviour in Virtual Humans". 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), pp. 565-566. **2020**

Dedication

Dedicated to the community around me built from friends and family but, most importantly to God, who kept and guided me throughout.

Proverbs 2:5-6: **Then shalt thou understand the fear of the Lord, and find the knowledge of God. For the Lord giveth wisdom: out of his mouth *cometh* knowledge and understanding.**

Dedicated to **Ashan Corrick**, while the research I conduct does not directly relate to glioblastoma brain tumours, your diagnosis and journey remind me of why we continue to push to understand the brain. I pray that your story continues to spread awareness and that you may rest in everlasting peace.

To my grandparents, your journey to the UK from Jamaica as part of the Windrush generation was not in vain. I love you, Rest In Peace **Nanny Kelly** and **Charles Kelly**. To my Irish grandparents and family who continually supported me and helped to raise me as the person I am today, **Aunty Margaret** and **Uncle John**, thank you. Rest In Peace **Mary McDonagh** and **William Raymond McDonagh**.

Dad, your endless support does not go unnoticed and I thank you for guiding and always supporting me in everything I do.

Mum, you've made mighty sacrifices for us and I will never be able to repay your selflessness. Thank you for reminding me of the importance of life and encouraging me to live and explore.

Brothers, Andrew and Seani Pembroke. You have been my biggest supporters from day one and have always encouraged and believed in me. Thank you for helping to raise me and teach me all you know. A paragraph will never be enough to express my gratitude for the unconditional love you show me.

Sisters, Naomi, Rebecca and Jessica Kelly. Thank you for being there whenever I needed you, I truly appreciate your support and all the love shared.

De-Shaine, you are my best friend and the love of my life. I'm so grateful for all you are. Your love, support and patience have been my comfort and strength since we met in this journey.

Nephew, Nieces and God Children, Kymani, Jayda, Ramaea, Namiyah, Rome, Khali, Naturi, Kayson, Kayan and Amias.

Acknowledgements

I would like to express my deepest thanks to my village which has raised and supported me through this stage in my development. There are too many names to list but the collective support of so many has allowed me to achieve this milestone. Thank you all for your encouragement and prayers. In particular, I would like to thank:

Professor Klaus Kessler - You pushed me past what I thought I was capable of at the end of my BSc. I am grateful for your belief in me as a researcher and all that you have taught me both professionally and personally.

Professor Tim Meese - Thank you for joining the project in the last stretch and supporting me through to the end.

Supervisory Team - Dr Ulysses Bernardet, Dr Johanna Zumer and Professor Adrian Burgess, thank you for your support in various aspects throughout this project.

Members of the **ALIVE Lab** and **CHL lab**, colleagues and friends at Aston University throughout this journey. And a special thanks to the support and guidance from Nicholas Alexander, Nicola Ngombe, Sam Gregory, Rachel Green, Ulrike Senftleben, Ellen Black, Ben McNulty, Foyzul Rahman, Laura Shapiro, Niteen Mulji and Jake Diggins. Each of you have supported me in unique ways and have helped guide me through this process. Thank you for making my time at Aston beautiful and enjoyable.

Family - Thank you to my mum, dad, brothers, sisters and extended family who have always shown me patience, love and support. To Maxine, Robert and the entire Murray family, thank you for treating me like your own. You've encouraged and supported me with more love than I could imagine.

Friends - Those who are like family; Monifah Kondwani, Sherayne Allison and Aleyah Downie, this journey wouldn't have been the same without you, thank you for all the laughs, love and encouragement. Close friends who have been supportive and patient with me throughout this time; Kimberley Pam, Letifah Clarke, Kyle Giles-Brown, Fatmata Daramy and Flavia Ojok, thank you. The **African-Caribbean Research Collective**, a community that provided a safe space and laughter whenever I needed it, thank you. **Black In Neuro**, my online family that this wouldn't be possible without; we came together with a mission for change and became a family along the way.

External support - Thank you to all of the external grants, training, additional funding and support that have been provided to me throughout this project. And a special thank you to Niyo Enterprise and Taran3D, for their XR Bootcamp that supports, uplifts and empowers Black women.

List of Abbreviations

2D Two Dimensional

3D Three Dimensional

ASD Autism Spectrum Disorder

ALIVE Aston Labs for Immersive Virtual Environments

ANOVA Analysis of Variance

AQ Autism Quotient

BOLD Blood Oxygen Level Dependant

CAVE Cave Automated Virtual Environment

HMD Head Mounted Display

JA Joint Attention

Neuro-VR Neuroimaging and Virtual Reality

ms Milliseconds

kΩ Kiloohm

OBS Open Broadcaster Software

OPM Optically-pumped magnetometers

MEG Magnetoencephalography

EEG Electroencephalography

fNIRS Functional near-infrared spectroscopy

fMRI Funcitonal magnetic resonance imaging

LSL Lab Streaming Layer

SNR Signal-to-noise ratio

SD Standard Deviation

SDK Software Development Kit **s** Seconds

τ Tau

TBI Traumatic Brain Injury

M Mean

MAD Mean Absolute Deviation

JASP Jasper's Amazing Software Program

μ Mu

IJA Initiating Joint Attention

RJA Responding to Joint Attention

IPT Immersive project technology

ADHD Attention Deficit Hyperactive Disorder

VH Virtual Human

VR Virtual Reality

SQUID Superconducting Quantum Interference Devices

AI Artificial Intelligence

ToM Theory of Mind

IVE Immersive Virtual Environments

IPTs Immersive Project Technology Systems

EEG-VR Electroencephalography-Virtual Reality

PC Personal Computer

OLED Organic Light Emitting Diode

JSON JavaScript Object Notation

IPD Interpupillary distance

ROI Region of Interest

TFR Time Frequency Representation

Contents

List of Figures

Chapter 1

General Introduction

This chapter provides a general introduction to the contents of the thesis and the project motivation. It begins with an overview of social communication, where joint attention is defined as a process and the benefits of its study are discussed. This chapter also reports the current literature on eye-gaze and how virtual reality may be a beneficial tool to be used in experiments that explore social interaction.

1.1 Introduction

As human beings, we are naturally social creatures who develop and build together within tribes, communities and societies. The ability to communicate effectively with one another is crucial, helping to cement relationships, build rapport and initiate cooperation. Conversely, an inability to do this can lead to frustration, lasting consequences for inclusion and difficulties in progressing in various environments e.g. schooling. Communication can range from verbal to non-verbal cues, which describe behaviours that are signals to indicate an action. Historically speaking, it has been a tool, that on a fundamental level, has allowed us to indicate a potential threat or alternatively, to signal the location of food sources (Zeiträg, Jensen, & Osvath, 2022). At a foundational level, these cues are important to us, and they have become increasingly intricate as we have become more advanced as a species. This is supported by some anthropological literature that suggests an evolutionary concept of 'survival of the friendliest' which dictates a preference for prosocial homo-sapiens (Hare, 2017). While potential dangers are evident in our lives, their likelihood and frequency have decreased. As a result, gestures, spoken language and non-verbal expressions are now being used to communicate a much more sophisticated range of information.

When discussing communication, language is the most common and obvious reference that is assumed. Language, whether it is written or spoken can be ambiguous and its complexities can lead to different interpretations between the 'initiator' (the person opening the communication) and the 'responder' (the person being communicated to). In the written case, we can proofread, gain opinions from others and edit the language until it is interpreted in the way we wish, but in the spoken case, there are added layers of information and ambiguities are more easily overlooked. Intonations, context and expression can each affect how we perceive certain words or sentences. For example, intonations can distinguish between two sentences that use the same phrase i.e. whether the pitch of the sentence uses a falling intonation (an assertion) or a rising one (asking a question). Pitch and intonation have even been shown to positively affect word recognition in toddlers (Singh & Chee, 2016). Spoken language, therefore, provides the initiator with more freedom to accurately express their communicative intent. This is particularly true when communicating complex information that can have multiple meanings depending on the context or situation. Such communication can be used to build social relationships, rapport and a shared understanding (Salazar et al., 2021).

However, misinterpretation or incorrect use of language can lead to (i) conflict, (ii) a breakdown of communication and (iii) social faux pas, which can all lead to the dissolution of a relationship. Each of these can make it increasingly difficult to develop, maintain and progress social relationships. In addition, there are many other signals that are shared in social communication, some that can even override or change the interpretation of the speech. For example, non-verbal signals such as body language and eye-gaze can inform an observer how someone may be feeling (Vuilleumier & Pourtois, 2007) or where they are focusing their attention (Frith & Frith, 2007), respectively. As well as gaining information from non-verbal signals, we also mirror them. When a person displays facial expressions such as pain (Botvinick et al., 2005) or fear (Adolphs, 2002), we mirror and experience the same emotion ourselves (Gallese, Keysers, & Rizzolatti, 2004). Overall, we use other communicators' behaviours as a reference and guide to understanding. Ultimately, there are a number of cognitive processes that are recruited to decipher social cues. Researching the cognitive processes associated with particular cues will be beneficial in extending our knowledge of brain areas responsible for social communication. Specifically, I aimed to develop a paradigm that could be used to investigate and accurately quantify the role of eye-gaze in joint attention and the corresponding neural networks. This would be achieved by making use of immersive virtual reality for higher ecological validity and flexibility in the social partner used. Consequently, this will increase our grasp on disorders that are characterised by difficulties in social communication, such as Autism Spectrum Disorder (ASD), a neurodevelopmental disorder that affects the way a person communicates and interacts with the world. This is discussed in relation to joint attention in Subsection 1.1.2.3.

Technological advancements have also had and continue to have an influence on our communication, with interactions now occurring in new domains. The internet age has brought about the addition of online interactions, where those that engage can text, voice call or video call each other without physically sharing the same space. However, such modes of interaction are also prone to miscommunication and some of the common mishaps highlight the important role of particular non-verbal signals in our communication. During the COVID-19 pandemic, the lockdowns forced the world to make use of online meetings more frequently and many reported problems with the natural cueing and signalling that we rely on in-person for turn-taking. These difficulties were centred on the mismatch of latency by either (i) overlapping talk (Seuren, Wherton, Greenhalgh, & Shaw, 2021), (ii) silence where speech should occur (Boland, Fonseca, Mermelstein, & Williamson, 2022) or (iii) the difficulty of following gaze shifts with dislocated cameras and screens (Hills, Clavin, Tufft, Gobel, & Richardson, 2022). Online video-calling platforms demonstrated that it is not just the obvious auditory signals from speech and language that we rely on to judge our place in a communicative exchange. Being able to effectively engage in dialogue through a 'call and response' style, may be more successful using behaviours such as body posture and eye-gaze. Therefore, I posit that non-verbal behaviour is underestimated in these interactions and further research is needed. Specifically, I will interrogate the importance of eye gaze in facilitating successful joint attention.

1.1.1 The study of social interaction

Typically, the investigation of social interaction has involved either the observation of real humans (Lachat, Hugeville, Lemarechal, Conty, & George, 2012) or, measuring the effects of a static social image (Driver et al., 1999). These two methods of investigating social interaction sit on opposite ends of the spectrum, where the former leads to an unlimited number of confounding variables and conversely, the latter is deemed too simple to ensure ecological validity. Irrespective of their limitations, the earlier studies of social interaction have been crucial in informing subsequent research. Often, researchers have utilised naturally occurring communication and social phenomena. These observations have encouraged researchers to develop psychological theories and tests that investigate the onset and the (delayed) development of various social behaviours. By building on these foundations with modern technology, researchers can design experiments that are more realistic and relevant to the natural environment.

According to the American Psychological Association, social interaction involves the reciprocal stimulation or response between two or more individuals. It also includes the development of cooperation and competition, the influence of status and social roles and the dynamics of group behaviour with leadership and conformity, that may lead to the establishment of social relationships (VandenBos, 2007). Within an experimental setting, two agents in a dyadic exchange can both be real humans, virtual humans (discussed in detail in subsection 1.1.4) or one of each. In this thesis, social interaction describes the way that two (or more) of these agents exchange signals to each other by using verbal and/or non-verbal behaviours (Frith & Frith, 2007). Current technology (see Figure 1.1), provides tools that allow more ecologically valid experiment setups for each of these exchanges e.g. dual neuroimaging or "hyperscanning" (Salazar et al., 2021; Dravida, Noah, Zhang, & Hirsch, 2020) and virtual reality (Caruana, Brock, & Woolgar, 2015; Jyoti & Lahiri, 2020). The former gives researchers the ability to scan two participants concurrently. This tool can provide neuroimaging data for both initiators and responders during live human-to-human interactions. On the other hand, virtual reality uses immersive technologies to display the experimental paradigms and potentially increase a user's presence within the scene.

As humans, we can often judge when we are having a good conversation and building rapport with another agent(s). However, it can be difficult to define what makes a good conversation and what makes two agents feel connected after interacting. Studies have aimed to identify aspects of social interactions that contribute to building a strong connection and have identified a phenomenon called "social connectedness", defined as behaving jointly or contingently (Leong et al., 2017), as a core feature. Researchers have used neuroimaging to assess the neural synchronisation between two individuals, to demonstrate that patterns such as synchronised behaviour (Marsh, Richardson, & Schmidt, 2009), conversational patterns/turn-taking (Stevanovic & Peräkylä, 2015) and direct eye contact (Leong et al., 2017) are actions that can affect social connectedness. However, these behaviours rarely occur in isolation and are often demonstrated in conjunction with a verbal exchange. Behaviours that create a successful communicative exchange can be complex and diverse; ranging from overt, clear actions to subtle microexpressions (McDonald, Newby-Clark, Walker, & Henselwood, 2018). It is extremely difficult to control and measure these subtle changes when human confederates are used in experiments

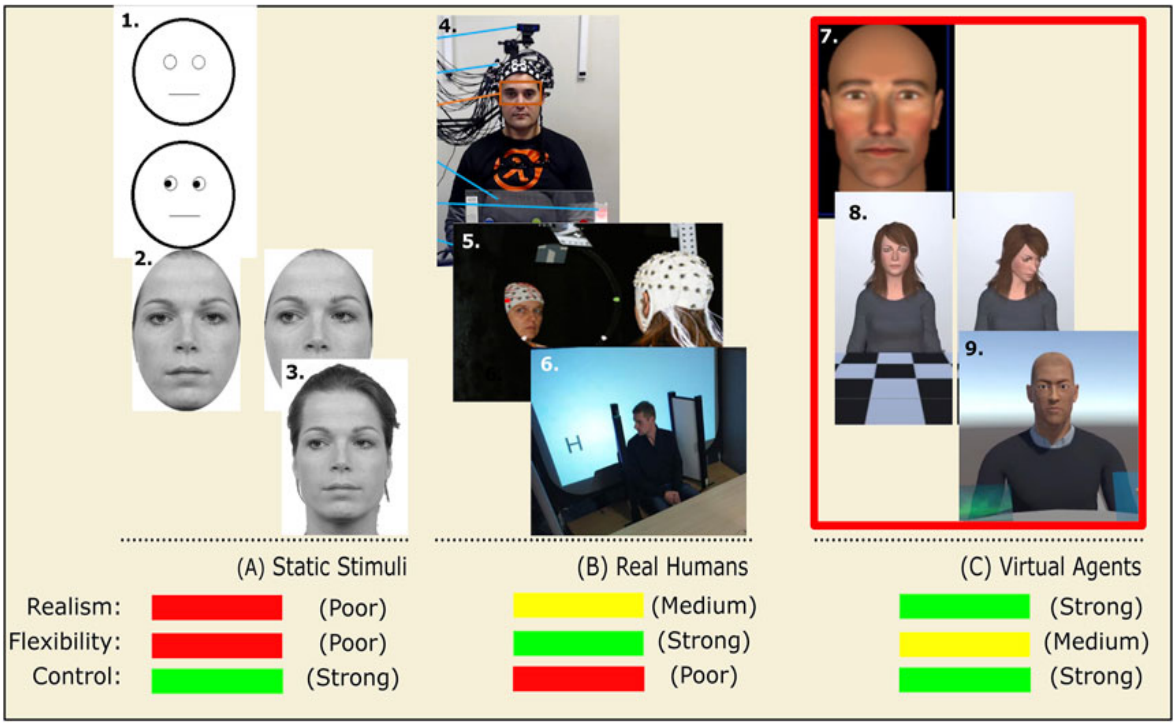

Figure 1.1: The figure is taken from (Gregory et al., 2021) and provides examples of the stimuli presented in interaction studies. Column A displays static stimuli with poor realism, poor flexibility and strong control. Column B is the use of real humans with medium realism, strong flexibility and poor control. Neuroimaging techniques display fNIRS (4) and EEG (5). Lastly, column C is the use of virtual humans with strong realism, medium flexibility and strong control.

(Gregory et al., 2021). The term 'confederate' refers to a human agent who has been trained and paid to behave in a certain way during the experiment, without the participant knowing (Leavitt, Qiu, & Shapiro, 2021; Asch, 1951). Intensity, timing and subtle actions are all interesting measurements and potential manipulations that are of interest to psychologists but, are difficult or impossible to control in natural settings that use a confederate.

As a result, researchers are revisiting the classical approaches and reevaluating their credibility, with the aid of modern technology. One example of this is the use of neuroimaging techniques. Some of the main techniques include functional magnetic resonance imaging (fMRI), function near-infrared spectroscopy (fNIRS), magnetoencephalography (MEG) and optically pumped magnetometers (OPM). fMRI is a non-invasive, imaging technique that pairs a structural MRI scan with the measurement of blood oxygen level dependant (BOLD) responses. It provides excellent spatial resolution, down to approximately millimeter precision but poor temporal resolution, around 3 to 6 seconds, which is much slower than the underlying neural processes. A mobile alternative to fMRI is fNIRS, which is also a hemodynamic-based approach but uses infrared light to measure changes (Scarapicchia, Brown, Mayo, & Gawryluk, 2017). Another recent mobile imaging technique is OPMs that, similar to the substitution of fMRI for fNIRS, provides an alternative to MEG. Both are non-invasive techniques that measure the changes in magnetic fields created from the electrical activity within the cortex. As a result, the techniques provide a high spatial and temporal resolution that is particularly beneficial for epilepsy localisation (Tierney et al., 2019). However, MEG uses superconducting quantum interference devices (SQUIDs) that are placed within a static helmet, in a magnetically shield room. They are fixed within this environment and have a one-size-fits-all helmet. Conversely, OPMs utilise non-cryogenic field sensors that are lightweight and can be placed into mobile helmets (Brookes et al., 2022). By pairing neuroimaging and classical experimental approaches, ecological validity is maintained whilst gaining the benefit of objective neural data. Additionally, some of the new, dual neuroimaging techniques are also mobile, not only allowing for a two-person recording but also providing flexibility for the location of the experiment and the movement of the participants e.g. (Sun et al., 2023). Whereas classical approaches often observe the interaction of two humans, modern technology also allows the use of a substitute e.g. a robot, a virtual human or a photograph. This can be placed directly in front of the participant in a physical form or with the use of a computer. When displaying agents on a computer, the lack of reliability seen in observational and field studies can be better controlled and this control has benefits for experiment repetition and precision. Also, computers allow for precise and small changes to the independent variable that cannot be as easily achieved when using real humans and will provide more confidence in replicated effects.

However, these approaches to investigating social interaction have limitations. Generally speaking, observational experiments are an approach that most closely resembles real-life interactions. However, as these real-life interactions are composed of many different (confounding) variables it is difficult to confidently conclude that it was the experimenter's independent variable that solely affected their dependent variable. Agents may also unintentionally be over-expressive or inexpressive, potentially changing the meaning of the interaction that is being conveyed in the experiment. Conversely, computer-based experiments have typically used photographs or other static stimuli which have a higher grade of control but these experiments are so far removed from reality, it is difficult to generalise and apply their findings to real-life situations. One method that falls somewhere in between these two approaches is virtual reality, as it offers an aspect of realism in its high-resolution, field-of-view immersion and dynamic abilities (see Figure 1.1). Virtual reality combines traditional computer-based experiments with these technological advances but also, shares the ability to provide interactive environments and stimuli that are comparable to the real world. Together, this makes virtual reality the most appropriate, available method to investigate social interaction and is discussed in detail later in this thesis (Subsection 1.1.4).

1.1.2 Joint attention

Generally speaking, joint attention can be defined as the coordination of orienting two people toward an object or event (Shaw, Bryant, Malle, Povinelli, & Pruett, 2017; Kristen, Sodian, Thoermer, & Perst, 2011; Moore, Dunham, & Dunham, 2014; Mundy & Newell, 2007; Scaife & Bruner, 1975). Successfully coordinating attention helps to align mental states and representations across agents and is key to building social communication and relationships. This general definition of joint attention is broadly accepted throughout the literature. However, inconsistencies in experimental methods remain, for example, the inclusion criteria of actions that make up joint attention. In part, this is due to the different ways we can communicate and whether it involves speech, a physical gesture or non-verbal cues. However, the position of an individual's mental state i.e. a person's awareness of other people's awareness can also have an effect. The different levels of information that a person is equipped with, affects how they will interact in a given scenario. It is therefore important to design a study where the participant's mental state is clear and in turn, explicitly report this information.

Previous work suggests that the cognitive foundations of successful information exchange develop before the onset of language (Bretherton & Bates, 1979; Bruner, 1974). Non-verbal sharing has been observed in infants (6-9 months) where the visual gaze of partners follows and directs attention (Bakeman & Adamson, 1984; Bretherton & Bates, 1979; Scaife & Bruner, 1975). Bruner (1974) observed these pre-verbal referential behaviours and labelled them as early cognitive processes that are specific to social reference, but distinct from those for language development (Mundy & Jarrold, 2010). Paradigms of social communication have focused on stages of development and investigate the typical behaviours/abilities that present at certain ages (Baron-Cohen, 1995; Carpenter, Nagell, Tomasello, Butterworth, & Moore, 1998; Meltzoff & Brooks, 2008). These findings have been crucial to the development of the field and have provided strong foundations from which to work. However, these observational studies lack the experimental control needed to investigate the nuanced behaviours of joint attention and the roles and sequences of each agent in these dyadic social interactions.

When referring to joint attention, one could argue that the individual components (behaviours) that comprise it are often overlooked and instead it is the cluster of behaviours that are investigated by researchers. One example of this is the inclusion of physical gestures and the effect that it has on a person who is responding to a joint attention bid. Jyoti and Lahiri (2020) developed a virtual reality-based, joint attention task to investigate the benefits of virtual reality interventions in individuals with ASD. They looked at a hierarchical prompt protocol using (1) eye-gaze, (2) head turns, (3) finger pointing and (4) visually highlighting cues by applying an effect that makes them sparkle to the user. They found that typically developing participants were able to pick up the initial eye-gaze cues in the absence of the other cues. However, ASD participants showed impairments in following the isolated eye-gaze cues. Although we have learned a range of ways to engage in joint attention, eye-gaze remains a salient component in understanding others (Schilbach, 2015). Shepherd (2010) referred to eye-gaze as the 'window into social cognition' and argued that the ability to follow eye-gaze is the foundational behaviour in understanding another person's perspective. However, to demonstrate this point, observations were made with methods that incorporated different additional cues e.g. eye-gaze and a head turn compared to eye-gaze and speech. To overcome this problem, joint attention should only be investigated by either (i) focusing on one component of joint attention or (ii) acknowledging that the implications and the weight that each cue holds are not fully understood and therefore any conclusions that are drawn should be tentative.

1.1.2.1 Levels of Joint Attention

Another complexity when researching joint attention is cementing a definitive definition of the action. Siposova and Carpenter (2019) aimed to categorise different forms of 'social attention'. Their categorisation of social attention is largely based on each individual's perspective and the type of knowledge that they have in these attention states. They argue that joint attention is not characterised by a single process but a collection of various factors specific to the context. Social attention was once thought to be a single and simple behaviour under the umbrella term 'joint attention' but is now recognised to contain several levels. Sipsova and Carpenter's review led them to refine how joint attention is defined by identifying four different types and a necessary pre-condition of the ability to engage in individual attention (see Figure 1.2). These are monitoring attention, common attention, mutual attention and shared attention. This work replaced the traditional dichotomy of 'being engaged or not', with a 'scale of jointness' where these distinct levels are classified by the following factors: the agent's perspective, the agent's reliance on one another, the experience shared between them, and the information in their attention states. They also detailed eight factors that can shift the scale of jointness that are (i) the stimuli, (ii) the shared goals between the agents, (iii) the common ground shared between agents,

(iv) the shifts in contingency and timing, (v) the perceptual space around agents, (vi) behaviours are contingent to attending an object or event, (vii) the individual differences between agents and (viii) the relationship closeness of agents. These are explored in relation to the levels, later.

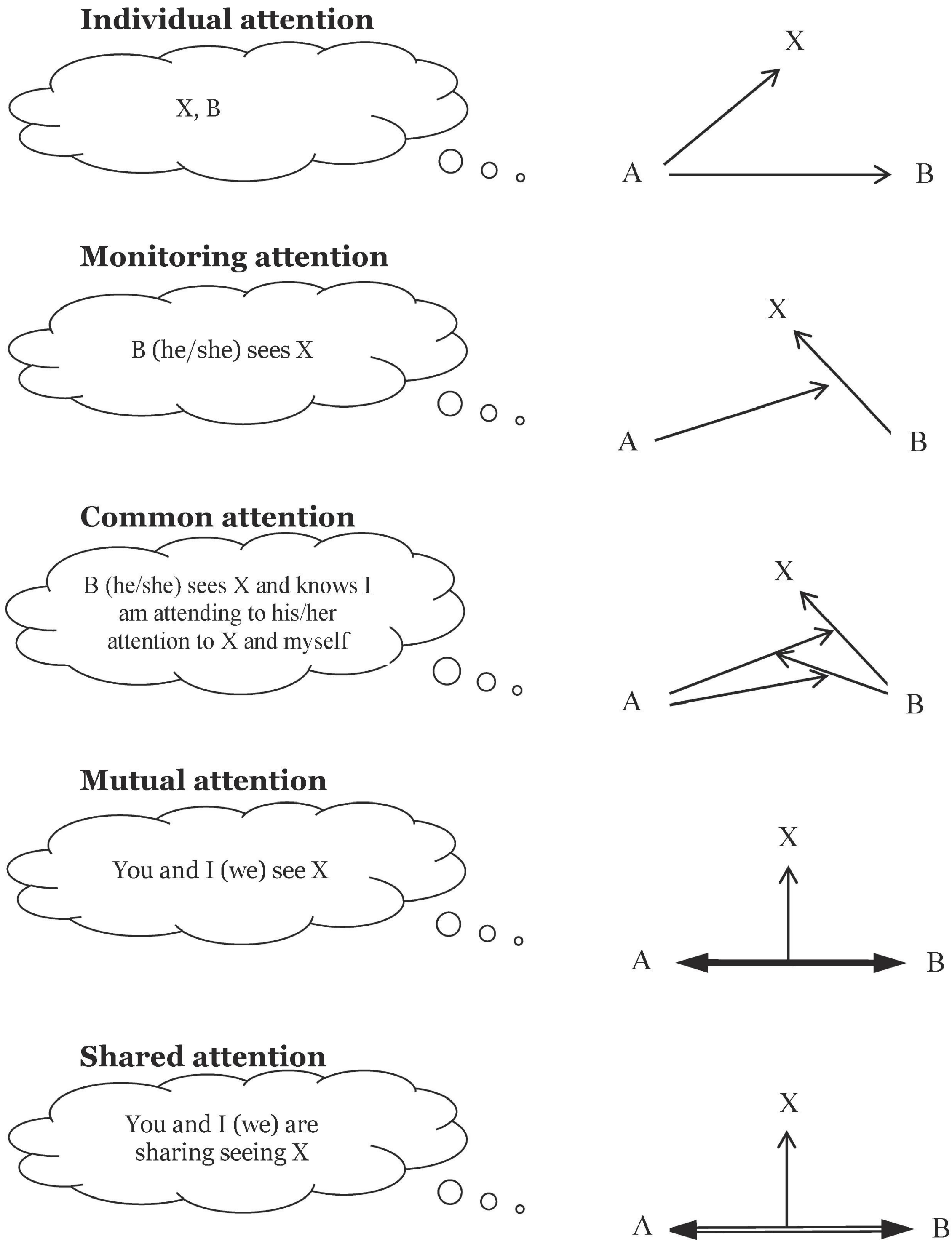

Figure 1.2: The figure is from the perspective of target individual A only. B is a second individual. X is the object of attention. Arrows represent what A and B are attending to. In individual attention, A attends to X and B. In monitoring attention, A attends to B's attention to X. In common attention, A attends to B's attention to X and herself. In mutual attention, A's attention to B's attention to X and herself is depicted via the solid two-way arrow, which represents non-communicative eye contact. In shared attention, the special two-way arrow represents communicative eye contact (and/or other bidirectional communication). Thought bubbles represent what A has in mind (this figure depicts a visual example). **Adapted from (Siposova & Carpenter, 2019)**

First, to start with, their levels of joint attention, Siposova and Carpenter began with their pre-condition of *individual attention.* This is a first-person perspective where only individual knowledge is gained. The experience of this is a dyadic interaction where the individual understands that they have attended to an object or event.

When describing each of the resulting levels of joint attention a workplace/office scenario can be used to illustrate the different levels of attention in a naturalistic setting.The first level *monitoring attention*, is one of the earliest examples of joint attention. This is when an agent takes an observer's perspective (third person) on a second agent and pays attention to what that second agent is attending to.

When Anna notices Ray looking out the window, Anna follows Ray's gaze and stares out the window while actively paying attention to what Ray is looking at.

Within monitoring attention, Siposova and Carpenter (2019) argue it can be taken further by manipulating attention. This occurs in a third-person perspective on a second agent who is attending to something else (see Figure 1.2). The third-person agent then behaves in a particular way to get the other agent to attend to something else. This type of attention involves the agent's knowledge and experiences. Occurring through unidirectional communication, manipulating attention relies on increased salience, thus the agent's attention is grabbed.

Sticking with the workplace example, Anna sees Ray collecting his empty wrappers together to go and put in the bin, without checking if Anna has any to get rid of. Anna then collects her rubbish together creating noise to grab Ray's attention, which he will likely offer to get rid of at the same time.

Common attention is held when two agents take a third-person perspective on each other, while also attending to what the other agent is attending to. In this scenario, both agents are aware that they have both attended to the same object or event, simultaneously. In addition to that, they are also attending to each other's attention to the given object or event.

For example, in the office, Anna and Ray find the room dark, Anna perceives the darkness and attends to Ray's perception of the darkness, Ray is also undergoing the same attention process.

Further, *mutual attention* is when two agents engage in a second-person relationship while concurrently attending to the same thing. As a result, both agents are immediately feeling each other's attention as well as their own. Mutual attention necessitates awareness without the need for deliberate communication.

In the office example, if the lights are suddenly turned off and Anna and Ray automatically hold hands without communicating, this would be mutual attention.

Following on from mutual attention, is *shared attention.* However, in shared attention, both agents need to communicate intentionally, either about the topic of attention or on the fact that they are both paying attention to the item.

For example, Ray notices another colleague outside the window and points it out to Anna, indicating his bid to communicate. He then makes eye contact with Anna and this is followed by dialogue about the other colleague. Following and guiding each other's attention, as well as producing bidirectional-shared focal attention are examples of shared attention.

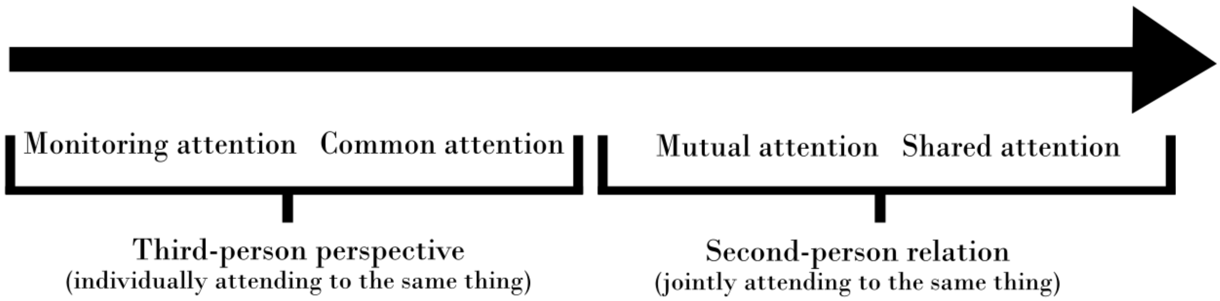

Figure 1.3: A schematic depicting Siposova and Carpenter's (2019) levels of jointness on a scale.

As previously mentioned, there are eight factors that can shift a scenario within the levels of joint attention. Firstly, the *salience* of a stimuli in a scene can provide clarity (high salience) or ambiguity (low salience) to an agent's interpretation and therefore, the level of joint attention attended to. For example, large, bright and dominating stimuli provide a clear object of attention, potentially providing higher salience and shifting attention to the right of the scale. Compared to small, dull and discrete stimuli that may not grab an agent's attention as easily (low salience) and therefore pushes attention to the left of the scale. Additionally, if there are relevant *shared goals* between agents, the strength and the existence of the goals have the ability to push scenarios to the right on the scale of jointness (see Figure 1.3). This is also true of *common ground* between agents and can include examples such as cultural and personally-shared common ground (Clark & Marshall, 1981; Foxe & Snyder, 2011). Stronger and more meaningful common ground can push a scenario to the right of the scale. Next, *contingency and timing* of joint attention can shift the level. If joint attention is established quickly after an exogenous and overt event, it is likely that this behaviour will increase the feeling of jointness, pushing the attention towards the right of the scale. For example, if something funny happened in front of two agents and they immediately look at each other, this would shift attention towards the right of the scale. On the other hand, if the two agents look at each other 30 minutes later when the object or event of attention is now less clear, this would shift attention to the left of the scale. *Perceptual space* can increase the certainty that another agent looked at an object i.e. closer spaces with clear boundaries increase an agent's ability to judge whether an object is being looked at and can push scenarios to the right of the scale. Another factor that can affect the level of joint attention is the observable *behaviours* of an agent. For example, sighing and irritable behaviour may signal boredom to other agents. *Individual differences* are also considered in these factors as some agents are more likely to be conscious of others and are more like to appear on the right-hand side of the scale. Lastly, the scale considers *relationship closeness*, which can include family, friends and partners. Agents are naturally more understanding and caring of those closest to them and would likely shift the scale to the right during those interactions in comparison to someone unfamiliar that would shift attention to the left of the scale. Overall, these factors provide relevant and strong arguments to be considered when investigating joint attention. Understanding the potential level of joint attention that can and possibly will be established as a result of the following factors is important when designing paradigms and interpreting data from joint attention studies. Interrogating these levels with ecologically valid paradigms will help us further elucidate the intricacies of this process.

Siposova and Carpenter's (2019) review of attention demonstrates the complexities of initiating and responding to another agent's bid for attention. Initiating and responding to joint attention probably both recruit similar skills (Caruana et al., 2015), and they can also require knowledge of another agent's mental state. This knowledge may dictate how an agent bids for another agent's attention and the behaviours necessary to do so. With the clarification of these definitions, the research aims and paradigms in this area can be refined to better understand the processes involved in joint attention. Based on these categories, I believe that the process studied in this thesis would be defined as *shared attention*, which is reflective of the traditional interpretation of joint attention. Without prompts, what is defined as intentional communication in a solely non-verbal exchange can be difficult to identify. However, in this set of experiments, participants were informed that the information would be conveyed to them through eye-gaze and they were aware that intentional communication was a part of the experiment.

Experimental paradigms designed to investigate social communication have largely focused on the stages of development and the typical behaviours and abilities that are present at certain ages (Baron-Cohen, 1995; Carpenter et al., 1998; Meltzoff & Brooks, 2008). This literature suggests these non-verbal sharing experiences are the cognitive foundations of successful information exchange (Bretherton & Bates, 1979; Bruner, 1974) and later consolidation of belief reasoning in typically developing children (Shaw et al., 2017; Charman et al., 2000a; Kristen et al., 2011). For example, Bruner (1974) characterised these pre-verbal referential behaviours as early cognitive processes specific to social reference, but distinct from those for language development (Mundy & Jarrold, 2010). Charman and colleagues (1997) suggested that impairments in joint attention can hinder the development of representational abilities that provide the foundation for pretend play. These developmental reports have been crucial in providing insight into how the basis of joint attention is first established and providing a frame of reference for future research. These frameworks have been fundamental in the development of cognitive neuroscience research, particularly when exploring the differences in neurodevelopmental disorders. Nevertheless, more research is needed to develop these frameworks and pinpoint the key behaviours and actions required for successful joint attention in social interactions.

The successful end result of joint attention can be observed in many different forms, from physical gestures such as pointing, which is seen in young children (Mundy & Jarrold, 2010), to very subtle executions like eye-gaze (Edwards, Seibert, & Bayliss, 2020). As mentioned at the beginning of this section, joint attention provides the opportunity to share representations of the world with one another. By ascertaining that both agents are focused on the same external object or event, joint attention allows for deeper and more insightful communication. However, both agents can in some cases, jointly attend to the same object from two different perspectives, which sometimes may cause problems. Using the Director's Task as an example (Krauss & Glucksberg, 1977; Rubio-Fernández, 2017), which is aimed at investigating perspective-taking and requires participants to recruit the process of joint attention, we can see a clear demonstration of communicative issues.

During the task, a confederate directs the participant to move objects in a vertical grid. Importantly, some of the cells are occluded for the confederate and so participants must understand the confederate's point of view. Being intentionally ignorant, the confederate would ask the participant to "move the small candle to the right", knowing there might be smaller candles in occluded cells. Here, it is crucial for the participant to recognise that the confederate cannot see these smaller candles. Nonetheless, participants show a tendency to consider moving the smallest candle in their privileged view before moving the smallest candle that the confederate can see (Barr, 2008; Converse, Lin, Keysar, & Epley, 2008). The task demonstrates the need to first, jointly attend to the same object and then, recruit higher-order processes such as perspective-taking. The completion of this task relies on the combination of these two processes. Similarly, other higher-order processes, such as mentalising and empathising, can also be accessed by building upon the foundation of joint attention (Stephenson, Edwards, & Bayliss, 2021). Therefore, joint attention can be used as a critical first step in enhancing our cohesion with other agents.

In more recent work, researchers have investigated joint attention during live face-to-face interaction, where both agents are participants (Dravida et al., 2020). While concurrently measuring cortical activity, Dravida and colleagues (2020) recorded spontaneous eye contact that was established between the two participants. The eye contact shared by the participants were then categorised as either “high eye contact” or “low eye contact” dyads. The researchers found that the levels of eye contact mediated responses in brain regions associated with joint attention, supporting the previous findings from some of the less naturalistic experiments.

Using real humans as interactive partners has a number of limitations which can influence the design of a study and the scope of manipulations that can be achieved (Gregory et al., 2021). First, the studies are more resource-heavy, as they typically require the same confederate to be used in all of the experiments which can lead to fatigue. This can be logistically difficult as it requires the scheduling of a minimum of three people (the participant, the confederate and the experimenter) and would also require payment and therefore acts as an additional cost (Leavitt et al., 2021). Second, the flexibility of questions and the design of the paradigm is limited, as it reduces the ability to probe time-dependent effects. Third, there are issues in the experimental control that threaten the validity of the experiment (Kuhlen & Brennan, 2013). For example, it is difficult for confederates to perform a repeated action in the same way within every trial and between participants. In these experiments, the participant’s experience needs to be considered. While using real humans may seem like a more natural approach, the experience is still unnatural by comparison to typical social experiences. For example, in gaze cueing tasks, participants are placed in an environment where (i) a stranger is often staring at them and does not communicate in a natural way, (ii) their social partner may be wearing unusual headgear such as a brain imaging cap (e.g. (Dravida et al., 2020)) and (iii) the task itself is not a typical social exchange, nor is it particularly engaging. In essence, participants will hold a preconceived schema that contains the expectations of how another real human can and might behave in the real world. When the real human agent is permitted to deviate from the prescribed schema, this can result in the participant perceiving the situation as strange for a number of reasons. This is often a subsequent effect of poorly trained confederates, who had minimal practice of the role and have not been trained as actors. These potential negative outcomes are often underappreciated (Sanko, Shekhter, Kyle Jr, Di Benedetto, & Birnbach, 2013).

Conversely, the expectations of virtual humans are very different from those of a real humans. Most people have come into contact with some form of technology that attempts to interact with them, whether a device like an Amazon Alexa, a virtual web assistant or an automated phone agent. With these rudimentary experiences, most participants hold a schema with limited expectations of how, broadly speaking, an interaction with technology will be. In summary, our expectations for the quality of interactions with artificial intelligence-type (AI) technologies (e.g.virtual humans) are low by comparison to those for real humans. This may begin to change with recent AI developments such as ChatGPT, an AI chatbot that mimics human conversation (Biswas, 2023). In turn, this may increase the need to include questionnaires or self-reports in experiments to test participants' knowledge, understanding and interpretation of current technological capabilities.

Intuitively, I feel there is relevance in meaningful social interactions that take place in a shared or common environment with an agent. Some of the questions that arise when researching social interaction can only be addressed effectively by using real humans, for example, when observing the interactions between familiar or socially bonded agents i.e. two or more friends. Whereas other research questions may require the participant to feel immersed in a virtual world interacting with a virtual human. The expectations held in this scenario would invoke the right neural circuitry that represents an anticipated interaction. In regards to this project, it is reasonable to assume that a completely non-verbal interaction with a virtual human agent is less likely to be deemed as a strange scenario to a participant, in comparison to a real human agent where certain social conventions may be expected. An example of this is an informal greeting upon the participant's arrival. Immersive virtual reality could potentially provide a consistent environment that circumvents these issues and allows participants to socially engage without distraction. In turn, the experimental control allows researchers to more readily measure the direct effects of the orchestrated social engagement. The rationale behind the experiments in this thesis was to use virtual reality to prime participants for preparing for one of two key roles: the initiator or the responder.

1.1.2.2 Responding to and initiating joint attention (RJA and IJA)

The process of joint attention involves one agent initiating joint attention (initiator) and another responding to joint attention (responder). The first part of the interaction begins with the initiatior executing the first cue (IJA) that bids for and guides the responder's attention. The second part of the interaction involves the requesting of the responder's attention. The responder must first process the cue accurately and then respond by following the gaze of the initiator, to the given object or event (RJA). Typically, the initiator's gaze will then alternate between the object and the responder to indicate that the object of attention is shared (Bruinsma, Koegel, & Koegel, 2004). Once complete, both agents will have generated a shared representation of the desired information and/or experience (Caruana et al., 2017a).

When joint attention is not successful, either agent (or both) might be at fault. For example, the initiator may not be clear enough with their bid for attention or the responder may struggle to understand whether their attention is being requested or not. This can lead to frustration for both agents and inhibit the development of a stronger social and communicative relationship (Rubio-Fernández, 2017). Developmental difficulties have been reported in IJA and RJA capabilities. Such difficulties have been consistently observed in individuals with ASD, Attention Deficit Hyperactive Disorder (ADHD), schizophrenia, bipolar disorder and depression (Lindner et al., 2019). One aim of this thesis was motivated to locate these difficulties and understand the finer behaviours that contribute to joint attention. The project began by investigating the core behaviours of the diagnosis of ASD and how a better understanding of joint attention might provide insight into the differences observed. Throughout the rest of this thesis, specific aspects of joint attention will be explored in relation to (the development of) ASD and the behavioural differences that are recorded in the literature. However, an experiment on behavioural differences in ASD was not conducted.

While this project is striving to better understand the role of eye-gaze in joint attention, the findings from the experiments can also be used to compare with neurodevelopmentally atypical populations (e.g. ASD). Particularly for ASD, these findings may reveal whether the differences manifest in eye-gaze and if so, how these differences are exhibited in comparison to typical behavioural and neurological recordings. Around 25-35% of individuals diagnosed with ASD are minimally- or non-verbal (Rose, Trembath, Keen, & Paynter, 2016). Focusing on non-verbal behaviours in the process of joint attention is more inclusive and importantly captures autistic individuals across the spectrum. Non-verbal behaviours are crucial and make up a large portion of what comprises joint attention. Eye-gaze in particular does not only benefit our own communication, but also our understanding of others (Morgan, Smith, & Freeth, 2021; Driver et al., 1999). While both roles in joint attention (IJA and RJA) can manifest in multiple physical forms (e.g. speech or pointing), in the explanations below I am referring to eye-gaze only.

Many studies have focused on RJA (Gredebäck, Fikke, & Melinder, 2010; Mundy, Kasari, Sigman, & Ruskin, 1995), particularly in young infants, as it typically precedes the development of IJA. This early development of RJA follows the basic principles of non-verbal joint attention as described by Siposova and Carpenter (2019), a bi-directional shared focal attention that is manifested through alternating eye-gaze between the object and the agent. In earlier reports and experiments, it is often referred to as "joint visual attention" or "visual coordination" (Murray et al., 2008). Studies focusing on RJA have reported its predictive capabilities for later language development (Mundy & Newell, 2007) and self-regulatory abilities (Vaughan Van Hecke et al., 2012). For example, following the direction of a head or eye-gaze is one of the earliest behaviours of joint attention that is observed (Murray et al., 2008). The literature has reported development of RJA from as early as six months (Bakeman & Adamson, 1984; Newson & Newson, 1975) and most have reported typical eye-gaze alternation within the first year (Hofsten, Dahlström, & Fredriksson, 2005; Jones & Carr, 2004). In typically developing children, pointing and gazing at an object that is out of their sight will prompt them to take an interest (Cilia, Touchet, Vandromme, & Driant, 2020; Deák, Flom, & Pick, 2000; Morales et al., 2000). After 12 months, this pointing is no longer necessary and infants become engaged through eye-gaze alone (Theuring, Gredebäck, & Hauf, 2007). By around 18 months, infants can accurately follow an agent's gaze toward an object outside their visual field (Carpenter et al., 1998). However, for children with ASD, these milestones often develop later than expected and are met with difficulty.

Initiating joint attention has been highlighted as an important distinction between humans and primates (Mundy & Newell, 2007). A few studies have observed the physical behaviours of IJA in non-human primates (Santos & Hauser, 1999; Tomasello, Carpenter, Call, Behne, & Moll, 2005). These researchers (Povinelli & Vonk, 2003) argue that while the primates they observed can understand the behaviour, they do not understand the mental states. For example, primates may understand the physical representation of gaze in terms of space and timing but may not understand the implications of gaze i.e. the experience of 'seeing' (Povinelli & Vonk, 2003). Although the literature is limited, and perhaps inconclusive, there is a view that the understanding of the intention to share attention, which is the basis of what we observe in IJA, may be unique to humans (Mundy & Newell, 2007). The ability to initiate joint attention in humans typically develops around twelve months of age (Bretherton & Bates, 1979). For individuals with ASD, IJA development is often later and difficulties continue until adulthood in some cases (Dawson et al., 2004; Lord et al., 2000). The late onsets of IJA and the ability to read another agent's intentions have been associated with a later development of Theory of Mind (ToM) (Brandone & Stout, 2022). ToM is the ability to interpret an agent's behaviour by taking their mental states into account i.e. beliefs, desires, intentions and knowledge (Bartsch & Wellman, 1995; Hughes & Dunn, 1998). False-belief tasks are used to test the development or absence of ToM around the age of 4-5 years old. The absence of ToM is typically associated with developmental disorders, ASD being one (Baron-Cohen, Leslie, & Frith, 1985; Tager-Flusberg, 2007).

1.1.2.3 Autism Spectrum Disorder (ASD) and Joint Attention

The current literature on ASD and neuroscience agree that there are notable differences when comparing an atypical and typical brain. However, the details of these differences and the validity of generalising them remain a matter of debate. Utilising a behaviour and/or neural signature during the diagnosis process would benefit clinicians as well as those seeking a diagnosis. Currently, key traits for the diagnosis of ASD are focused on the difficulties that the individual has with social communication (Bruinsma et al., 2004). These include deficits in joint attention (Billeci et al., 2017a), atypical gaze behaviour, such as avoiding eye contact (Griffin & Scherf, 2020; Spezio, Adolphs, Hurley, & Piven, 2007; Swanson & Siller, 2013) and, in children, possibly deficits in language development (Charman et al., 2003). When diagnosing ASD, a deficit in joint attention distinguishes ASD from other developmental disabilities for 80 to 90% of children (Lewy & Dawson, 1992) and is included as a measure in infant screening and in tools for diagnosis (Jones & Carr, 2004; Khowaja & Robins, 2013). It has been suggested that impairments in joint attention may have a direct relation to the differences in social behaviour observed in autistic people (Volkmar, Lord, Bailey, Schultz, & Klin, 2004). Studies that have focused on developing and increasing joint attention skills for autism have reported favourable outcomes (Meindl & Cannella-Malone, 2011). For example, Murray et al (2008) found that individuals with ASD would engage in conversation for longer and Whalen et al (2006) reported increases in social interactions and spontaneous speech. Therefore, behavioural and neural investigations into the process of joint attention in the typical population, have the potential to increase our understanding of how joint attention is expressed in the atypical population at both a cortical level and behaviorally (e.g.

eye-gaze sequences). Although traditional behavioural measures, such as response times and accuracy have held precedent in the investigation of joint attention, another aspect of interest is eye-movements. Measurements of eye-tracking have proven fruitful to the field, particularly when comparing typical and atypical populations (Mei, Zahed, Mason, & Quarles, 2018).

1.1.3 Eye Gaze

The foundation of eye-gaze literature is supported by gaze studies using static images of faces, such as those used in Driver et al (1999). Tasks have included direct or averted gaze (Adams & Kleck, 2005), the familiarity of the face (Sterling et al., 2008) and gaze cueing to targets (Friesen & Kingstone, 1998). This research has been fundamental to the general advancement of our understanding of joint attention at a behavioural level, by providing insight into gaze sensitivity and how this might inform interactions in a social setting (Edwards et al., 2020). However, traditional experiments that use static images lack a true reflection of the dynamic processes in everyday social interactions (Schilbach, 2015), with evidence showing that outcomes do not always reflect those that may be seen with real human interaction e.g. (Risko, Laidlaw, Freeth, Foulsham, & Kingstone, 2012; Risko, Richardson, & Kingstone, 2016; Gobel, Kim, & Richardson, 2015). Laboratory research has found that participants tend to direct their gaze to the eye-region of face stimuli (T. Smith & Mital, 2013; Võ, Smith, Mital, & Henderson, 2012), whereas real-life studies report that participants avoid direct eye contact (Laidlaw, Foulsham, Kuhn, & Kingstone, 2011; Foulsham, Walker, & Kingstone, 2011; Gallup et al., 2012). The lack of generalisability is particularly evident when considering how dynamic eye-gaze can be in terms of (i) speed, (ii) the path taken to a target of attention and (iii) dwell time or time spent looking at a target. Therefore, it is important to find alternative approaches that better reflect real-life social interactions, while maintaining experimental control.

In this project, we originally set out to measure eye movements by making use of the eye-tracker that is embedded within the virtual reality head-mounted display (HTC Vive Pro Eye). The initial intention was to use the recorded eye movements in a closed-loop system where both RJA and IJA were investigated. IJA trials would use each participant's gaze location to determine the virtual human's response, in turn, developing a response-specific paradigm and thereby increasing the feeling of realism. RJA trials would record human eye movements to capture repetitive sequences or gaze shifts in response to a bid for attention. The experiments in Chapters 2, 3 and 4 were designed with this in mind but owing to COVID-19, plans were curtailed, meaning the eye-gaze data that was recorded were less valuable than initially hoped. This will be discussed in detail in Subsection 4.4.3.

1.1.3.1 Eye gaze and neuroscience

Researchers regularly report behavioural effects in reaction and fixation times in response to eye-gaze manipulations e.g. (Senju, Hasegawa, & Tojo, 2005; Bayliss et al., 2013). More recently, research has steered towards investigating the appearance of these effects at a neural level. This has largely focused on event-related potentials (ERPs) that have been investigated for differences between averted vs. direct gaze (Itier, Alain, Kovacevic, & McIntosh, 2007; Driver et al., 1999), social vs. non-social gaze (Gregory, 2021; Wahl, Michel, Pauen, & Hoehl, 2013; Barry, Graf Estes, & Rivera, 2015) and gaze fixation (McPartland, Cheung, Perszyk, & Mayes, 2010). However, neuroimaging-based investigations of eye-gaze are on the increase, particularly those combining mobile neuroimaging equipment for human-to-human studies of gaze (Dravida et al., 2020) and more recently the pairing with virtual reality (Gregory, Wang, & Kessler, 2022).

The overall understanding of joint attention, on a neural level, is largely based on static agents that are displayed on a 2D screen. However, even at this level, researchers have observed the effects of eye gaze on higher-order cognitive processing, that are not present in the equivalent non-social tasks. For example, faster gaze cueing in response to head turns and eye-gaze than a non-social cue such as an arrow or a car (Wahl et al., 2013; Barry et al., 2015). While both cues may induce similar behavioural responses they are processed completely differently (Ristic, Friesen, & Kingstone, 2002). This default social effect that eye-gaze has is considered as an altercentric (other-centred) processing (Kampis & Southgate, 2020). This describes the behaviour of intentionally processing the gaze of others. When watching another person gaze at an object, the object acquires properties that would not have been attributed to it if were not gazed at (Shteynberg, 2010; Becchio, Bertone, & Castiello, 2008). While these fundamental studies have evidenced a social implication of eye-gaze, it is important to explore these data further using more dynamic and ecologically valid stimuli.

1.1.4 Virtual Reality

1.1.4.1 Introduction to virtual reality

For consistency, I will use the term 'virtual humans' throughout this thesis. Note that in some of the published work mentioned, virtual humans may be referred to as 'virtual agents', 'virtual characters' or 'electronic confederates'. All of these terms refer to the same thing i.e. a computer-generated human that has been programmed by the experimenter.

Virtual reality refers to the "realistic" computer-generated interactive environment that incorporates high-speed, three-dimensional graphics, audio feedback and peripheral devices (Limniou, Roberts, & Papadopoulos, 2008). The virtual environment can be displayed with traditional inputs such as a screen, keyboard and mouse on a desktop system. However, this technology does not move users into a virtual world. Conversely, *immersive* virtual environments (IVEs) provide displays and peripheral inputs to enhance a user's perception they are actually in the virtual world.

These involve stereoscopic displays, such as head-mounted displays, that provide users with the experience or perception of an ecologically realistic environment (Bailenson, Blascovich, Beall, &

Figure 1.4: A photograph of a participant wearing the HTC Vive Pro Eye head-mounted display, with a displayed virtual environment that fills their field of view.

Loomis, 2003). While viewing the virtual environments, users are presented with a slightly different image for each eye, which is similar to that of the natural eyes (Limniou et al., 2008) and a tracked hand-held device can control interaction and movement (Lee & Lyons, 2004). IVEs and virtual reality more generally are now beginning to be implemented in mental health research, and virtual reality is an emerging tool for intervention in disorders like ASD, depression and anxiety (Lindner et al., 2019). Using self-reported measures, Bouchard and colleagues (2017) reported virtual reality exposure therapy as a superior tool to in vivo exposure-based therapies for social anxiety disorder. As the technology continues to develop, the equipment is becoming cheaper, more readily available and more easily accessed outside of the main virtual reality community. This allows for a greater uptake from the general public and, with the right programs available, more affordable options for public healthcare systems like the National Health Service (NHS). The use of screen-based virtual reality (including both computer and mobile) has increased in many public sectors. In particular, there has been an increase in the use of virtual humans in these applications, for example, presenters for TV and web applications (Bickmore et al., 2021), simulation-based training and for teaching (García et al., 2022).

1.1.4.2 Virtual humans

Virtual humans are defined as computer-generated virtual reality characters with human-like appearances, in contrast to an avatar, which is a humanoid representation of a user in a virtual world (De Paolis & Bourdot, 2018). Virtual humans have been used across experimental platforms to investigate social interaction (Gregory et al., 2022; Mei et al., 2018). Importantly, similar social behaviours have been found during interactions between real-to-virtual humans, as are observed with real-to-real humans (for a review, see (Bombari, Schmid Mast, Canadas, & Bachmann, 2015)). In our lab we have used virtual humans in traditional screen-based studies (Green, Shaw, & Kessler, 2023), as well as in immersive virtual reality (Alexander, 2023 (in prep.)). They can also be used with technology such as immersive project technology systems (IPTs) which projects images onto surface(s) around the user (Garau et al., 2003).

There are several important benefits of using virtual humans instead of real humans, that go beyond the factors of experimental control that have been outlined already. First, it makes it easier to use age-matched stimuli in research studies. While it might be difficult to have an age-matched confederate when investigating child development or ageing, an age-matched virtual human is more readily available. Further, it is possible to conduct studies making use of virtual humans in different ways. For example, investigating the effects of multi-agent joint attention, would be easier in a virtual environment using an avatar (to represent the participant) and virtual humans (as social partners), than it would be using multiple human confederates. Further, the experience of an interaction would be much more realistic than using a screen-based set-up with traditional tools. This is also more cost-effective, as ready-made virtual humans are available on several platforms (e.g. Maximo and SketchFab), often free or cheap to purchase. They can also be programmed for experimental studies using freeware such as Unity, Unreal Engine and Godot.

Furthermore, virtual humans can mimic the dynamic behaviour of human eye-gaze with great precision by design. For example, it is possible to map real eye-gaze onto virtual humans and create a very naturalistic interaction, but this can also be computer-controlled by defining specific locations to be looked at. Crucially, processes can be applied to the action such as 'tweening', which aims to create a smooth animation between two specified frames. Within this, transitions can be applied to reflect real-life eye-gaze movements i.e. slow to begin and slow to end. The use of virtual humans holds many benefits, one being their performative nature; virtual humans do not need to function properly (Ruhland et al., 2015), as long as the implied behaviour reflects human-likeness. For example, a virtual human doesn't need to look directly at an object for the user to believe it is being looked at. It does not affect a participant's ability to correctly perform tasks (i.e. to localise the virtual human's attention through their gaze (Ruhland et al., 2015)). This is often corroborated by the context given to participants before taking part in an experiment. If instructions to the participant state that the virtual human will direct their gaze to an object, it creates an expectation that a shift in eye-gaze will inform a communicative exchange of information. This is particularly evident in joint attention when one agent has been clearly designated as the initiator or as the responder.

There are several counter-arguments to using virtual humans as social partners. The most important question is whether virtual humans can replicate reality enough, to show effects that can be generalised to real-world social interactions. Importantly, it can be argued that because the participants know that virtual humans are not real humans, this knowledge could affect the interaction. This is particularly evident when discussing the context of an interaction. I consider this with the following comparisons and further discuss the importance of expectations: (i) virtual world behaviours versus real-world behaviours, and (ii) real-world context-specific versus real-world everyday life (general). For example, in comparison-1: *virtual world versus real-world*, people will happily shoot and kill virtual characters in computer games, but would not engage in such behaviour in real life. Similarly, people will engage in a physical altercation to gain rewards in a game but would be less likely to behave the same way in reality. However, it is important to note the differences between a computer game scenario and a research experiment. In real life, people behave differently depending on the context and the task at hand (Richesin, Baldwin, & Wicks, 2022). In comparison-2: *real-world context-specific versus real-world everyday life*, people will shoot others in paintball when engaged with willing participants but would be less likely to take a paintball gun out on the street and use it there. Similarly, people will bump into cars on the fairground ride but would be less likely to drive like that on the road. I propose that virtual humans can be seen in the same way.

Depending on the contextual set-up of an experiment, which is typically provided by the researcher, participants will hold specific, context-dependent expectations. A common scenario given to participants is that they will be working against or in competition with an agent. Consequently, the participant will adjust their expectations and behave in response to the given context i.e. behaving competitively against the agent. Whereas, if the participants are informed that the agent is a collaborative, social partner who they need to work with to complete the task, the participants are likely to treat them as such. Participants have been observed adjusting their behaviours in virtual reality experiments. Research on ethics using virtual reality has demonstrated how the use of virtual humans, in a context-specific scenario, can even influence moral-decision making dilemmas.

A commonly known dilemma, created by Foot (1967), is described as the doctrine of double effect. This is the distinction between what an individual perceives to be a consequence of their voluntary action and what they had intended. One popular example of this is the 'Trolley problem', where participants are told that there is a train coming down the tracks and the brakes do not work. If the train continues, it will kill five people who are on the tracks. However, the participant has control of a lever that can divert the train onto a different track, changing the course of the train from killing five people to an alternative track where only one person would be killed. Participants are then asked whether they would pull the lever. In virtual reality experiments, participants are required to pull a lever (sometimes virtual, sometimes physical). In a study by (Niforatos, Palma, Gluszny, Vourvopoulos, & Liarokapis, 2020), the authors compared responses between paper-based and virtual reality-based scenarios. Results were explained in relation to two main views (i) deontology; where moral principles are considered above everything else (e.g. not killing an intruder who intends to harm your family and others because killing is wrong) and (ii) utilitarianism; where the best course of action is based on the consequences (e.g. killing an intruder because they plan to harm your family and potentially others). Results showed that while both the paper-based and the virtual reality scenarios employed utilitarianism, virtual reality invoked a significantly higher percentage of utilitarian decision-making. It is a common finding that participants do not always accurately report their subjective opinions on questionnaires, this is sometimes a result of social desirability bias and demand characteristics (Nichols & Maner, 2008). Therefore, I propose that the virtual reality scenario was comparatively more representative of real-life behaviours. In conclusion, while the issue of gameplay not being reflective of real life is argued (Kiili, 2007), I suggest that the context given mitigates this and affords the use of virtual humans as a more naturalistic method of data collection.

1.1.4.3 Practicalities of virtual reality

The use of virtual reality in academia and industry has increased considerably over recent years (S. Kim & Kim, 2020; Wilson & Soranzo, 2015). Concurrently, the utility and functionality of virtual reality tools have also improved. The latest developments, particularly for head-mounted displays, offer a better spatial resolution, wider visual field, lighter hardware, and some include a in-built eye-tracker. In turn, these hardware improvements have encouraged the continued advancement of software capabilities and propelled creativity in the development of virtual spaces and scenarios. As a result of this greater usability, there has also been an increase in virtual reality applications in the fields of engineering (di Lanzo et al., 2020), rehabilitation (Holper et al., 2010) and social neuroscience (G. Kim, Buntain, Hirshfield, Costa, & Chock, 2019).

In general, there are studies that have used virtual reality and have reported it as such, irrespective of the equipment used and its subsequent level of immersion. Here, *immersion* is defined as the objective level of sensory fidelity a virtual reality system provides. Comparatively, *presence* describes the subjective psychological response a user has to a virtual reality system (Bowman & McMahan, 2007; Slater, 2003). Therefore, differences in immersion account for the equipment used. For example, IPTs, such as a Cave Automatic Virtual Environment (CAVE), where participants stand in a three- to six-wall rectilinear partial enclosure and use projectors to display the virtual environment onto the walls. CAVEs require participants to wear stereoscopic shutter glasses, that allow a dynamic viewpoint with perspective projection, that updates the virtual scene consistent with the participant's location in the space. Another example is a scene that has been created in a 3D software engine and has been displayed on a computer screen. Lastly, the use of a head-mounted display, which is a headset that fills the user's field of view, by displaying an environment that has been created in a 3D software engine.

Levels of immersion also consider many other factors including software abilities, such as realism of light, frame rate and refresh rate (Bowman & McMahan, 2007; Slater, 2003).

With these various possibilities, it is important to clearly define and refer to the different levels of immersion that have been used in an experiment. Many studies claim to have used virtual reality but often fail to make a clear reference to the equipment used and therefore, identify the potential presence of the user. A critical distinction is needed to ascertain whether the experiment involved dynamic presentations either (i) on displays that are perceived to be external to the participant or, (ii) on displays that are intended to be unnoticed by the participant with the aim of feeling present in that scene. Development of an environment using 3D development software like Unity or Unreal engine, is classed as virtual reality by some journals and so it is important to differentiate the use of *immersive* virtual reality. For example, in this following set of experiments, the scene was built using 3D software but was displayed online on a 2D computer screen (Studies 1 and 2). The tools used to develop the paradigm are clearly stated but 'virtual reality' is not used as a term to describe the experiment as it has connotations of immersion. Alternatively in Study 3, a head-mounted display was used and therefore, *immersive* virtual reality is an appropriate and justified term to describe the set-up. Consistency within the literature using these terms is essential to accurately depict the experiment set-up and cross-reference comparative studies.

Within social neuroscience and psychology, virtual reality offers a research tool that can be used in combination with more traditional methods, such as questionnaires and behavioural responses. It has offered psychologists the opportunity to validate classical paradigms and replicate results in an environment with higher realism (Rey & Alcañiz, 2010). Bridging the gap between the classical on-screen/paper paradigms and more sophisticated technologies will enhance the investigation of social behaviour further, with the use of dynamic stimuli. To truly understand how people engage in social interactions, we can make use of immersive equipment and virtual humans that can be manipulated and have agency in a task. This will empower future experiments bringing greater insight into how human social interaction affects our social cognition.

1.2 Summary of aims and rationale

The aim of this PhD project was to investigate non-verbal social cues, specifically eye-gaze, when working with a virtual human. This project positioned eye-gaze as the fundamental behaviour in joint attention and explored the differences in both behaviour and neurophysiological measures. As a result of the COVID-19 pandemic, I began by using a series of online studies and later developed one participant-facing experiment. The online experiments utilised the virtual environment/world to display the scene and virtual human. Response times and accuracy were used to measure the behavioural effects of the manipulations. The online studies served as experiments that allowed us to enhance and adjust the sequences as needed. In the participant-facing experiment, scenarios were presented in an HTC Vive Pro Eye, which provided an immersive environment that encapsulated participants' field-of-view. This also gave an opportunity to record participants' eye movements in an environment that doesn't provide a clear border or guideline to where the display starts and ends. Therefore, allowing participants to gaze more naturally around the environment, as they might in real life. Building upon the behavioural differences found in the online studies, EEG was concurrently recorded in the participant-facing experiment. This provided an objective measure of the potential neural changes that relate to joint attention and specifically, the role of eye-gaze in this process.

Based on the theoretical background and tools available, I have developed a basic joint attention scenario that focused on responding to joint attention. This allowed me to focus on developing a paradigm that engaged the participants with the virtual human and created an experimental foundation that could later incorporate initiating joint attention. The paradigm was centred around completing puzzles. This provided a task that could require agents to work together by aiming to complete both puzzle boards (the participants and the virtual humans) as quickly as possible, or in competition with each other by completing their own puzzle board faster than the agent. In the current project, participants are tasked with working with the virtual human with the aim to complete both puzzles, together.

In all of the studies, participants began the experiment seated at a table with the virtual human opposite them. On the table were two puzzle boards that need to be completed, each board contained a label that dictated whether it belonged to the participant or the virtual human. Participants were tasked with completing the two puzzle boards as quickly and as accurately as possible. To do so, they must work collaboratively with the virtual human who engaged with the participant using eye-gaze only. In each trial, the virtual human was shown one puzzle piece (hidden from the participant) and performed an eye-gaze sequence that the participants were told to attend. The gaze sequence performed by the virtual human was designed around two levels: (i) Collaboration and (ii) Gaze Type where the virtual human's gaze movements varied. Within these levels, the factors reflected either

(i) Collaborative gaze and directed the participant's attention to the corresponding puzzle boards the piece belonged or Non-Collaborative gaze and was uninformative (the virtual human gazed only at the puzzle piece), and (ii) Eye which either involved direct eye contact or NoEye which did not. Participants were asked to use this gaze sequence to determine where the puzzle piece belonged and use a keyboard press to indicate their answers. In Collaborative conditions, there was a correct response that the participant could give however, in Non-Collaborative conditions, the trials were designed to have no correct response and therefore participants were required to guess their response. As a result, we could not compare accuracy but were interested in the effect on response time and potential cognitive load due to either clear (Collaborative) or unclear (Non-Collaborative) information given.

As previously mentioned, communication needs to be intentional in order to develop shared attention. One example of doing this is by engaging in direct eye contact and directing the agent's attention i.e. bidirectional-shared focal attention. When two or more agents engage in eye contact while paying attention to an item, this behaviour can provide an intentional indication that both agents are sharing attention. In this paradigm, we aimed to achieve this within the Eye conditions where a trial would begin with the virtual human engaging in direct eye contact, then performing either a Collaborative or Non-Collaborative gaze shift followed by the virtual human's gaze returning to being engaged in eye contact. Both the initial eye contact made and later eye contact was critical for the virtual human to create and develop shared attention with the participant and possible social influence. Direct eye contact in this paradigm is of particular interest in the studies and its effect will be explored throughout.

This paradigm was developed with the following aims:

1. Investigate the fundamental behaviours that create successful joint attention.
   (a) By restricting communication to one behaviour, eye-gaze.
   (b) By manipulating the behaviour of the virtual human's eye-gaze and observing the behavioural effects.
2. Investigate the neural underpinnings of this basic joint attention behaviour.
   (a) To understand the electrophysiological frequencies that are associated with the gaze manipulations.
   (b) To understand the neural signatures that are associated with the gaze manipulations.

These aims are recurring themes throughout the thesis and form the foundation of the specific hypothesis highlighted in each of the experimental chapters.

1.3 Overview of thesis

The aim of this thesis was to take a reductionist method to investigate joint attention. As previously mentioned, joint attention in everyday situations is often observed as a combination of actions and behaviours. Within these actions, eye-gaze remained a consistent behaviour that was considered when investigating joint attention. Therefore, I believe that investigating eye-gaze as the primary factor in joint attention will extend the boundaries of scientific knowledge in understanding what makes successful joint attention.

The literature review supports the view that non-verbal interactions are key in neurodevelopment, often acting as an indicator for future outcomes: such as spoken language (Kasari, Gulsrud, Freeman, Paparella, & Hellemann, 2012) and emotion regulation (Morales, Mundy, Crowson, Neal, & Delgado, 2005). While the behavioural expressions are well reported, more research is needed around the neurophysiological differences in response to particular behaviours involved in joint attention, such as eye-gaze. This motivated the second general aim of this thesis which was to understand the neural correlates of communicative eye-gaze in the process of joint attention. Targeting non-verbal behaviours increases the scope of the research to include those with reduced communicative abilities, for example, individuals with low-functioning ASD. In addition, virtual reality has proven useful in previous social neuroscience research in both neurotypicals and atypical populations (X. Wang, Xing, & Laffey, 2018; Mei et al., 2018), namely through the employment of virtual humans (Bailenson et al., 2003). This is further extended by later making use of a head-mounted display to provide a great level of immersion and presence for participants.

Following on from this introductory chapter, **Chapter 2** introduces the first set of online experiments, where I develop the structure of the paradigm by testing parameters such as the speed of gaze, the effects of a disembodied head and the subsequent incorporation of a body. **Chapter 3** extends such work homing in on the key features of the gaze sequence and refining the trials within the conditions. A Neuro-VR investigation comes to fruition in **Chapter 4** where the finalised paradigm is interrogated with in-person participants, electroencephalography (EEG) and virtual reality. This thesis then concludes with **Chapter 5** which presents a critical evaluation of my work, looking at ways to improve the investigation of joint attention within our field of cognitive neuroscience, whilst also being inclusive of a wider range of participants. I will also look at the future outlook of my work and the potential next steps for researchers building on Neuro-VR paradigms.

1.4 Covid Note

While completing this PhD, the U.K. went into lockdown and participant-facing research was paused part way through the first year until halfway through the second year of this project. The original plan for this project involved face-to-face testing and use of the head-mounted display in all experiments. As this was no longer feasible, multiple contingency plans were created, one of which involved transforming the paradigm into an online experiment. For this reason, these initial online studies have focused on refining the paradigm in an online format and the final revision was used in the in-person experiment.

Chapter 2

Investigating non-verbal bids of attention with a virtual human: online studies

This chapter discusses the first set of experiments that I conceived, that work to refine the experimental paradigm; a randomly interleaved 2x2 factorial design that manipulated either (i) eye-gaze, by the virtual human engaging in eye contact or not and (ii) collaboration, which dictates whether the virtual human would provide informative gaze shifts or not. All of these studies have been created in Unity but displayed on a computer screen without the use of a virtual reality headset. Experiments were collected through online platforms and then collated and analysed.

2.1 Introduction

A key to building communication and social relationships is the alignment of mental representations through joint attention. Generally speaking, joint attention presents a combination of behaviours that have a common goal of directing both agents' attention to an object or event. However, the finer details of this process are poorly understood. Research that has investigated joint attention has often focused on static stimuli, with only a few examples that have used dynamic gaze. Some of the research that has made use of more dynamic stimuli have continued to use either unrealistic settings (e.g. in an fMRI scanner) (Caruana et al., 2015) or have used 'simple' virtual humans as agents (Mei et al., 2018) with jerky movements. This first study serves as a close comparison to such dynamic paradigms, with a view to later (Chapter 4) increase immersion and potentially the presence of the user by incorporating a head-mounted display that covers the participants' field of view and therefore provides a common environment for both agents to interact.

Building a paradigm that focuses on the fundamental behaviour of joint attention is an important first step to investigating the process as a whole. Understanding these phenomena is not only informative for further elucidating the inner workings of successful communication between humans but, can also help researchers better understand instances where this communication may break down. This research could be useful to understand differences seen in 'atypical' individuals, such as those diagnosed with ASD and ADHD, where differences in neural signatures have been reported. As I continue to explore and better understand how these processes are expressed in the typical population, our findings may help us identify if and where there are any statistically significant differences between typical and atypical populations.

I began my investigation by constructing a physical model of the proposed gamified paradigm, which was created on pieces of cardboard and played between Aston Laboratory for Immersive Virtual Environments (ALIVE) lab members. Using this physical model, each agent had their own puzzle board and cut out puzzle pieces. One agent would perform IJA gaze sequences and the other would respond (RJA) by indicating where they believed the agent was directing them (either their own puzzle board or the agent's puzzle board). This physical enactment was used to get an idea of how I could translate this model into a virtual paradigm and begin to identify critical parameters that could be controlled. In the first instance, human agents worked collaboratively and were instructed to only use their eyes to communicate with the other agent. This gave me an understanding of how human agents would naturally perform and use gaze in these scenarios. I built on this by designing another scenario, where one agent would be told to work collaboratively on 50% of their IJA turns. Observations and self-reports were made and used to inform the development of the first version of the paradigm. As expected when using human agents, there are many subtle behaviours that were picked up by the agents that either aided or hindered the communication between them e.g. a widening of eyes to emphasise. This was both observed and reported by lab members, which further supported the need to use a virtual human within an immersive environment.

I translated these observations into Study 1 (1.1-1.5), a collection of experiments which aimed to understand the contribution that eye-gaze had in the combination of behaviours observed in joint

Figure 2.1: A photograph of the physical cardboard game. This was created to play the proposed game in person and observe the natural gaze behaviour in response to the task.

attention. To do this, I restricted communication to only eye-gaze and manipulated aspects of the sequence and measured the effects. The eye-gaze sequences were translated and developed onto a virtual human that was created as the participant's partner. As alluded to in Chapter 1, virtual humans were chosen as they act as an appropriate agent for social interaction, by displaying a human-like appearance and behaviour. They also provide experimental control, as each animation is directly programmed and thus editable by the researcher. Unlike human confederates, virtual humans can consistently and repetitively perform the same behaviour, a critical feature needed for the reproducibility of experiments. In this paradigm, this was achieved by restricting the virtual humans' communication solely to changes in eye movements that mimicked human eye-gaze. The desired eye movements were performed by the researcher in the head-mounted display and were mapped onto the virtual human in real time. Predetermined eye-gaze sequences for each condition were performed by the researcher and screen recorded to later edit and use in the online experiments. Focusing on one behaviour, eye-gaze in this instance, was crucial to understanding the basics of joint attention. Therefore, Study 1 would measure whether participants could (i) complete the task using only eye-gaze and if so, (ii) measure participants' behavioural responses (response time and accuracy) to the eye-gaze manipulations.

Using the eye-gaze manipulations, the paradigm focused on two factors, Collaboration and Gaze Type.

These were created with the specific goal of understanding the role of collaborative gaze and eye contact, respectively. The task required participants to work with a virtual human to complete two puzzles. On each trial, a puzzle piece, hidden from the participant, was presented to the virtual human. The virtual human then performed a gaze sequence that had been preprogrammed and derived from human eye-gaze recordings. The gaze mimicry was developed using C# and the native software development kit (SDK) for the HTC Vive Pro Eye. Making use of the in-built eye-tracker, the experimenter's eye-gaze was tracked and mapped onto the virtual human, using the SDK and later, in Chapters 2 and 4, I incorporated C# code to enhance the application of the tool.

In recent literature, dynamic eye-gaze has been used with virtual humans to further understand the underlying mechanisms of joint attention. Caruana and colleagues (2017) investigated 'intention monitoring' with a virtual human as an agent, in a collaborative game. Participants were placed in a functional magnetic resonance imaging (fMRI) scanner and viewed a screen with a virtual human's head in the centre, and a row of three houses above and below the virtual human's head (an example of stimuli is presented in Figure 2.2. In the task, a burglar was present in one of the houses. The target for participants was to identify the house that the burglar was occupying. This information would be communicated to the participant, by the virtual human using eye-gaze. In each trial, participants engaged with the virtual human who would display either a 'Search' or 'NoSearch' condition which dictated the eye-gaze sequence. In the Search phase, the virtual human was programmed to gaze around the environment in search of the target, a behaviour described as a 'noncommunicative gaze shift'. The virtual human would then display a gaze shift towards the target, a 'communicative gaze shift', to invoke a response from the participant. In the NoSearch conditions, the virtual human would immediately perform a communicative gaze shift towards the target and, therefore, the information would be immediately conveyed without noncommunicative gaze shifts.

Caruana and colleagues (2017) then replicated the experiment with a non-social condition, that replaced the virtual human with an arrow. Their results revealed that the noncommunicative gaze shifts in the Search phase resulted in slower gaze following in comparison to the NoSearch phase. This revealed a measurable effect from intention monitoring when responding to joint attention behaviour. They argued that intention monitoring recruits additional cognitive resources that are required to monitor an agent's intentions i.e. determine whether or not an agent's gaze is an informative cue that should be followed. In the first set of experiments in Study 1 (1.1 - 1.3), I programmed a comparable Search gaze sequence at the beginning of each trial. This is reported as the 'initial sequence' (see section C

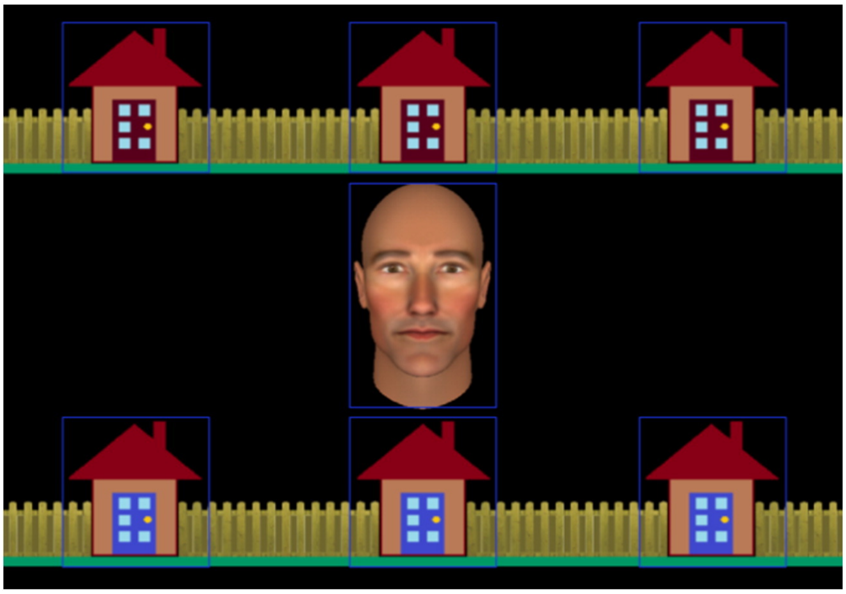

Figure 2.2: Image adapted from (Caruana et al., 2017a), displaying the participant's view of the stimuli with blue rectangles representing the gaze areas of interest (the blue rectangles were not visible to the participants).

of Figure 2.3). In the following experiments, the role of the initial sequence was explored within my novel paradigm. This was to determine whether the initial sequence was a necessary section of the trial needed to study the role of eye-gaze in the process of joint attention.

In these experiments, the scenario recruits a virtual human that is used as the participants' collaborative partner. Each trial begins with the virtual human's eye-gaze following the initial sequence. Firstly, the virtual human looks at the puzzle piece presented and then laterally gazes at both puzzle boards (2.3; counterbalanced). This is followed by a randomly selected gaze sequence from one of the conditions: Collaboration (Collaborative vs. Non-Collaborative, communicative or noncommunicative gaze shifts, respectively) and Gaze Type (Eye vs. No Eye, direct eye contact or not, respectively). In the conditions where the virtual human is providing collaborative eye-gaze and engaging in eye contact (Collaborative, Eye conditions), the sequence illustrates the prototypical behaviour of joint attention or alternatively, shared attention, as described by Siposova and Carpenter (2019). In Collaborative, Eye conditions, the virtual human is programmed to look directly at the participant which engages the participant in eye contact and directly requests their attention (as displayed in Figure 2.4). Next, the virtual human directs their eye-gaze towards the correct puzzle board and then reverts back to direct eye contact with the participant before the trial ends. As described by Siposova and Carpenter (2019), the Collaborative, Eye condition includes bi-directional shared gaze behaviour that alternates between the initiator (the virtual human) and the responder (the participant). The inverse is also true in Non-Collaborative,

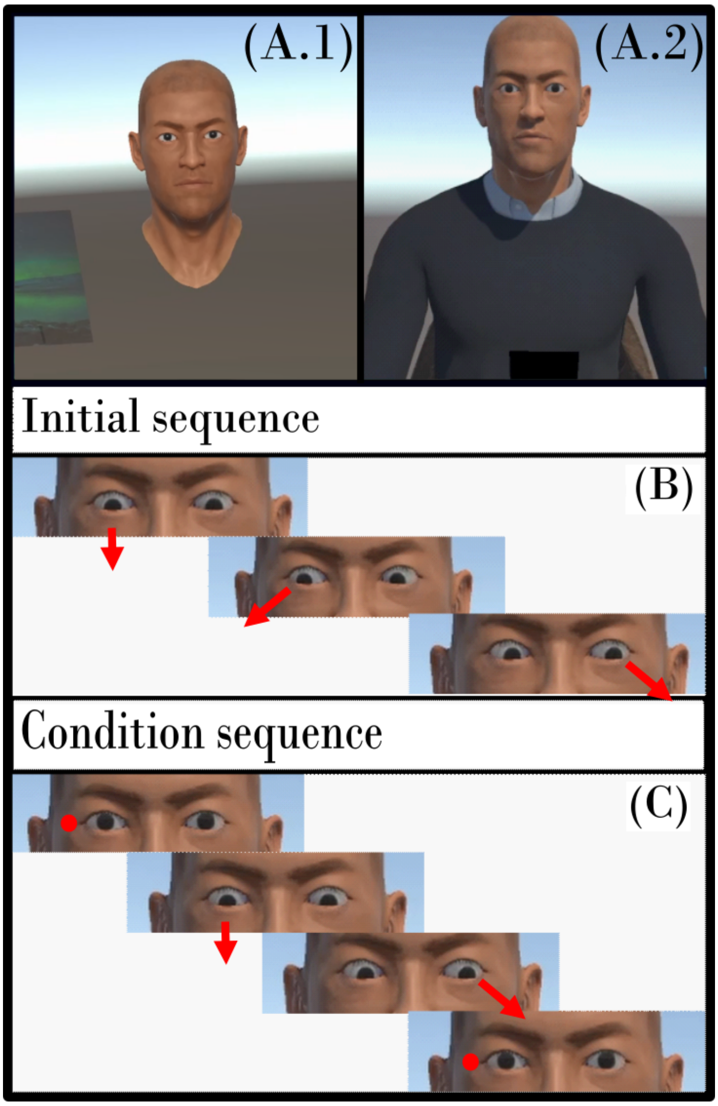

Figure 2.3: A schematic displaying the sequence used in the online studies 1.1-1.5. **Section A.1**, is a still from the experiment which used only a decapitated head (1.1) and **A.2** is a still from the experiments that used a head and body (1.2-1.5). **Section B**, demonstrates the initial sequence that is performed at the beginning of every trial in Experiments 1.1-1.3. **Section C**, is an example of the Collaborative, Eye condition. The red arrows and dots denote where the eye gaze is directed. The red dots represent facing straight forward and arrows pointing towards the gaze direction.

NoEye conditions, where the virtual human does not engage in eye contact and does not direct their gaze to either of the puzzle boards (the virtual human stares at the puzzle piece only). The trial sequence for the experiment is relatively simple in terms of gaze patterns and the task required of the participant. However, with only a few dynamic gaze studies reported in the literature, I made an intentional decision to begin with a foundational paradigm and build complexity. This would allow me to better understand the effects of the chosen manipulations on dynamic gaze in joint attention.

A large proportion of the literature on joint attention has relied on gaze studies using static images of faces (Gregory et al., 2021) and paradigms where participants are merely observing others (Wilms

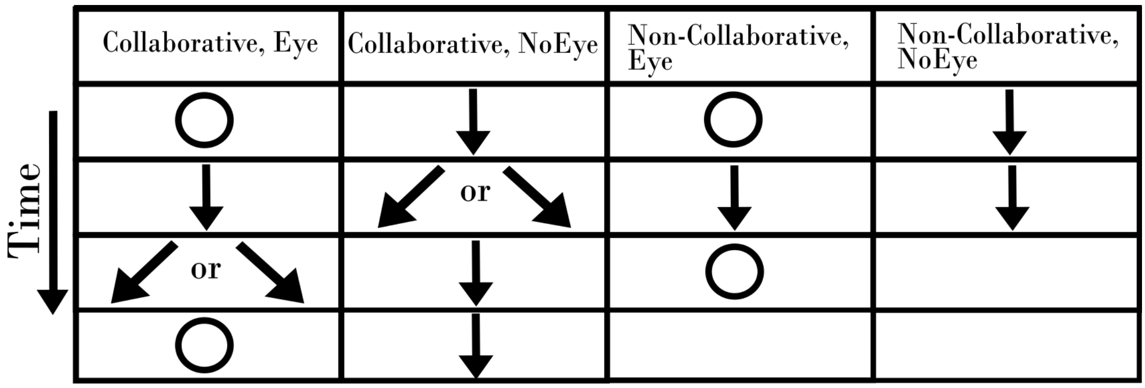

Figure 2.4: The schematic depicts the gaze locations for each condition in Experiments 1.1 - 1.5. The circle (**O**) represents the virtual human directly at the participant and the arrows point toward the direction that virtual human gazes. Downward arrows represent the virtual human looking down (looking in the direction of the puzzle piece) and diagonally pointed arrows represent gaze toward either of the puzzle boards. These directions are visualised in Figure 2.10.

et al., 2010). Tasks have included direct or averted gaze (Adams & Kleck, 2005), familiarity of the face (Sterling et al., 2008) and gaze cueing to targets (Friesen & Kingstone, 1998; Frischen, Bayliss, & Tipper, 2007). Many of these paradigms have been used in cognitive neuroscience to investigate the neural networks associated with these processes. However, few have used interactive platforms. Subsequently, there is a need to develop ecologically valid paradigms that offer real-time interaction with dynamic stimuli (Schilbach, 2015). In this chapter, I have utilised 3D software (Unity3D, 2019) for the development of virtual environments and recorded the scenarios to use as videos in an online experiment format. While these experiments do not take full advantage of the capabilities of the software and technology (due to the COVID-19 restrictions), they do provide an added element of realism by making use of dynamic gaze. Therefore, these experiments begin to unpack how participants interact with dynamic agents when relying solely on eye-gaze.

In line with the literature, I hypothesise that Collaborative conditions will yield faster responses from participants, as they will have been provided with a clear direction towards a given target puzzle board. Additionally, in Eye conditions, the virtual human will provide an alternating gaze between the target puzzle board and the participant that suggests an understanding of a shared mutual representation in joint attention. Therefore, I predict that this eye contact will result in faster response times and greater accuracy. From these expectations, I developed hypotheses.

2.2 Hypotheses

Our investigation of the role of eye contact and collaborative gaze between real and virtual humans led us to investigate two key hypotheses:

1. H1: Participants will respond faster in Collaborative conditions in comparison to Non-Collaborative conditions.
2. H2: When the virtual human engages in eye contact (Eye conditions), response times will be faster and more accurate (Collaborative trials only) than NoEye conditions.

2.3 Methodology

The main experiment was created using Unity Technologies (Unity3D, 2019), a development software used to create 2D and 3D scenes. Unity is free to download and allows real-time development of environments in virtual reality. Scenes can be developed using open-access assets, which include a catalogue of development kits, templates and 2D or 3D models. The current experiments made use of assets that were free from the Unity Store or already present in the software. The table and chair were downloaded from the Unity Store and the floor and background environment were a part of the software's default set-up. The virtual human's head was provided as part of the HTC Vive Pro Eye's SDK and the body was downloaded from the Maximo website, free of charge. Two puzzle board images were displayed on the table: an octopus and an aurora borealis. These two images were chosen as they were distinctly different, this aimed to reduce the ambiguity for the participant. They were saved from a Google images search and imported into Unity as 2D sprites to be manipulated in the scene: both images were tilted at a +/-45 degree angle and their opacity was reduced. See figure 2.7.

As mentioned in the introduction, the original plan for the experiment was to create a closed-loop system that considered the location of the user's eye-gaze as an input that would affect the virtual human's behaviour. Consequently, the scene was initially developed to have the virtual human mimic the eye-gaze of the user. I used Open Broadcaster Software (OBS) software to record the mimicked eye-gaze sequence for each of the conditions. The video recordings were edited to the same length and used as the trials in the experiment. PsychoPy3 software was used to develop the experiment and allowed me to control the main aspects including presentation timing, order of sequence and recording participant responses. Once completed, the experiment was uploaded to Pavlovia (pavlovia.org/), a website that allows researchers to upload, store and run experiments from.

The experiment was published with an accompanying questionnaire. I decided on using the Autism Quotient (AQ) questionnaire (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001) which I inputted and published on the Qualtrics website. Within the experiment structure, the questionnaire was placed at the beginning and once it was completed, participants were redirected to Pavlovia for the main task. The Autism Quotient is 50 item questionnaire (see Appendix G) that is used as a measure of autistic traits in the general public. This questionnaire was not developed for diagnostic purposes but can provide an indication as to whether a member of the public possesses traits that are associated with autism. This helped to ensure that my pool of participants and therefore results, solely reflected neurotypical behaviours. This step was taken with the long-term aim of using this data to compare with results from an autistic population. Once all sections of the experiment were collated and linked together, an advert was posted to the Prolific (Prolific.co/) website. Prolific provides an online platform to connect eligible participants with paid experiments and was a well developed website during the time of collection.

As part of the screening, participants who scored 26 and above in the Autism Quotient were excluded (Ashwood et al., 2016). This cut-off was chosen as it was one standard deviation above the mean score for neurotypicals and would ensure that the resulting data was only comprised of neurotypical behaviour.

2.3.1 Participants

Participants that were recruited through Prolific, were screened online through the criteria set by the researcher. These were firstly, age, which was set at 18-40 years as research has demonstrated a significant difference in cognitive abilities, including attention, between older (typically 45 years +) and younger adults (Hedden & Gabrieli, 2004; Murman, 2015; Fortenbaugh et al., 2015). As the last experiment would primarily recruit university students, I have used 40 years old as a cut-off to reduce any potential variability I would see as a result of ageing differences. Next, participants needed to be able to speak and be fluent in English, as the instructions for the task were in English and any help provided from the researcher would be in English. Third, participants were required to have normal-to-corrected vision, to ensure they could see the eye movements that were being made by the virtual human. With Experiment 3 (which uses EEG) in mind, I asked whether participants had ever sustained a traumatic brain injury (TBI). Many studies have reported both neurological and behavioural differences in working memory and information processing when a person has experienced a TBI (Christodoulou, 2001; Grafman, Jonas, & Salazar, 1990). I have used this exclusion criterion to reduce variability in the data that might fall outside of the typical behaviours. This allowed me to compare these results directly with data from Experiment 3 and keep the pool of participants uniform across all studies. Lastly, if a participant had already taken part in any of the experiments I had uploaded to the Prolific website (including pilots), they were excluded, as familiarity may have an effect on response times.

In Experiment 1.1 there were 80 participants, 42 were female and the mean age of all participants was 26 years. Participant's age ranged between 18-40 years with a standard deviation of 6 years. Experiment 1.2 consisted of 80 participants, 38 female and the total mean age of all participants was 27. Participant's age ranged between 18-40 years with a standard deviation of 7 years. Experiment 1.3 had 82 participants, 40 female and 27 years as the mean age. Participant's age ranged between 18-40 years with a standard deviation of 6 years. In Experiment 1.4 there were 38 participants, 17 were female, 21 were male and they had a mean age of 24 years. Participant's age ranged between 19-40 years with a standard deviation of 4 years. Finally, Experiment 1.5 had a total of 46 participants, 23 of them were female and they had a mean age of 23 years, collectively. Participant's age ranged between 19-33 years with a standard deviation of 4 years. The study was approved by Aston University Health and Life Sciences Ethics Committee.

2.3.2 Paradigms

2.3.2.1 Experiment 1.1

The experiment began with participants filling out the Autism Quotient questionnaire (Baron-Cohen et al., 2001) on the Qualtrics website (https://astonpsychology.eu.qualtrics.com/). Once completed, participants were re-directed to the online experiment on Pavlovia. Participants conducted the experiments at home, on their desktop PCs (experimental timings were therefore independent of variability in network connection speeds). The online experiment used a disembodied head (see Figure 2.5) as a virtual human in a total of 80 trials, with each trial (video) lasting 12 seconds (Collaboration (Collaborative vs. Non-Collaborative) and Gaze Type (Eye vs. NoEye)). The experiment began with written instructions on screen that read, "*Thank you for agreeing to take part in the study. You will now be presented with a succession of tasks (trials) and they will go as follows: The virtual human (VH) will appear on screen. Shortly after, they will be presented with a puzzle piece in front of them that is only visible to them. You will only see the back of the puzzle piece (a black square). The VH will perform an eye-gaze sequence that will suggest whether the puzzle piece is for yours or their own board. They will try to be truthful but might not always get it clearly across. It is your task to decide which puzzle board you think they are suggesting. Place your dominant index finger on the '↑' key and the other index finger on the '↓' key of your keyboard, as indicated in the images on the right. You will indicate whether you believe the puzzle piece is for yourself (using the '↓' key, pointing to you) or for the VH (using the '↑' key, pointing to them). You must do this as quickly and as accurately as possible. You will first go through some practice trials. Please press the space bar when you are ready.*" accompanied by two images to visualise their required hand placement (see figure 2.6). All conditions started with the virtual human on screen and a puzzle board on either side. Next, a puzzle piece was presented to the virtual human, hidden from the participant. The virtual human then began to communicate using eye-gaze and started with the initial sequence (Figure B of 2.3), where the virtual human looked at the puzzle piece and then at both puzzle boards (counterbalanced).

Figure 2.5: A still image of the experiment scenario for Experiments 1.1-1.2. The image shows a disembodied virtual human head in the centre of the screen with two puzzle boards at the bottom of the screen on either side.

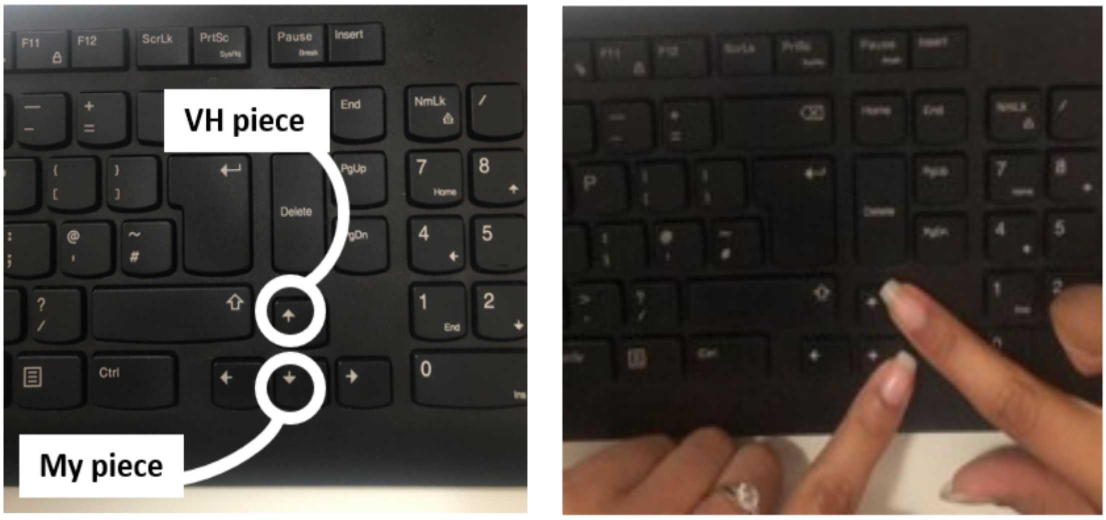

Figure 2.6: Figure displays the two images that were shown during together with the instructions on the screen. The image on the left had labels to visualise the button correspondence. And the image on the right demonstrates the required hand placement of participants.

This was then followed by the condition sequence where collaborative or non-collaborative gaze shifts were performed by the virtual human. The virtual human would either direct the participant's attention to the location of the target puzzle (Collaborative) or would gaze solely at the puzzle piece (Non-Collaborative). These gaze sequences either did (Eye) or did not (NoEye) involve eye contact. Participants were told to respond by pressing the arrow keys on the keyboard. The up arrow indicated the puzzle piece that belonged to the virtual human and the down arrow indicated the puzzle piece that belonged to the participant. Participants could only respond once the video had finished playing. An example of the eye gaze changes during a trial sequence can be seen in sections B and C of Figure 2.3.

2.3.2.2 Experiment 1.2

The second experiment followed the same procedure as Experiment 1.1, but the speed of the virtual human's eye-gaze during the initial sequence was increased to 1.5 times faster than the original version. The motivation for this change was to make the gaze resemble a glance, as if the virtual human was performing a check. Comparatively, the slower gaze length in Experiment 1.1 could be interpreted as an informative gaze shift that suggests the intention to communicate. This reduced the video time to 9 seconds per trial. I expected that an increase in speed to the initial sequence would lead participants to interpret and therefore, process the sequence as a "quick glance". This is opposed to participants interpreting the initial sequence as an informative gaze sequence that was to be understood and potentially recruiting additional cognitive resources.

2.3.2.3 Experiment 1.3

The third experiment was another replication of Experiment 1.1, with the inclusion of a body for the virtual human and a table (see Figure 2.7). As a body was not provided with the HTC Vive Pro Eye SDK, I used Mixamo, a website with free 3D characters. Previous research has suggested that a disembodied head may not provide the realism needed when investigating social questions (Gregory et al., 2021). By including a body, I was aiming to create a virtual human and virtual environment that was an appropriate replication of real life. I had intended to do this in a controlled manner that kept the focus/manipulation on eye-gaze behaviour.

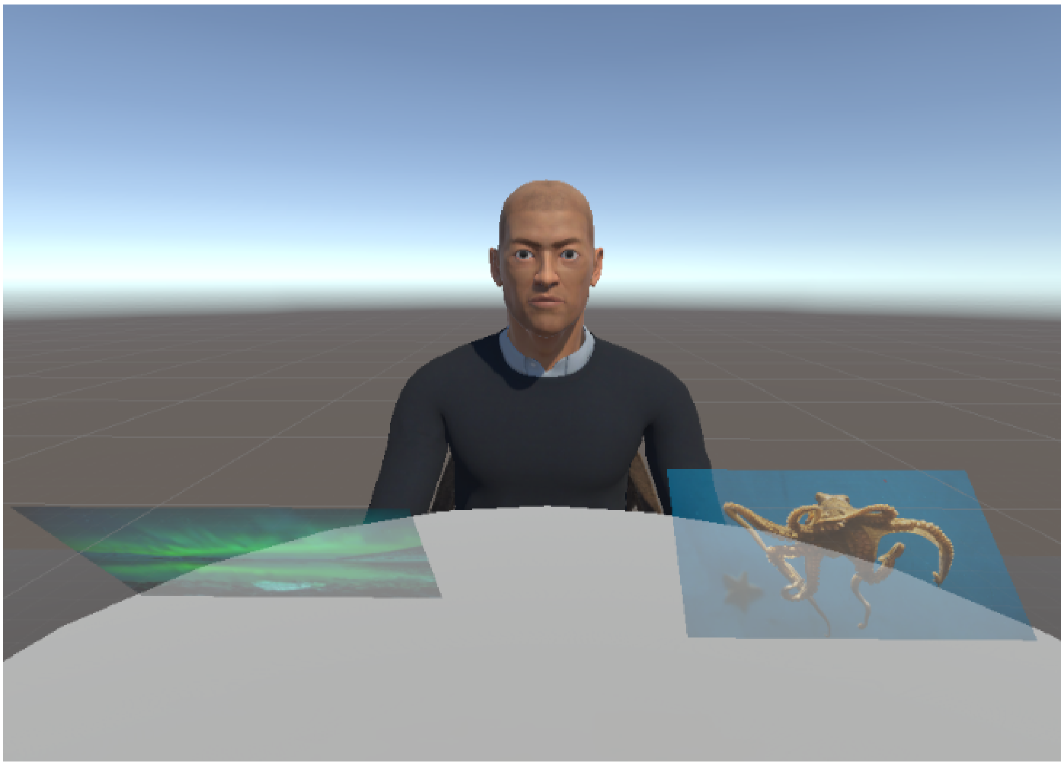

Figure 2.7: This still image displays the environment set-up that participants would interact with in the online experiments 1.3-1.5. The virtual human is positioned opposite the participant at a table, with the two puzzle boards present.

2.3.2.4 Experiment 1.4

The next experiment replicated Experiment 1.3, using a virtual human with a body. However, in this experiment I removed the initial sequence, leaving only the condition sequence to constitute the whole trial. The initial sequence was originally used to display a natural behaviour of familiarity. When I first developed the experiment, I re-enacted the experiment in real life with two humans and found the initial sequence as a checking behaviour before signalling to the other player. However, I did not initially consider that when a participant becomes familiar with the experiment, they will stop performing an initial sequence. As a result, I decided to investigate this in our study design with Experiment 1.4. The effect size from Experiment 1.1 was used to inform our power analysis.

2.3.2.5 Experiment 1.5

In the final experiment, I replicated Experiment 1.4 but doubled the number of trials, resulting in 160 trials in total, 40 per condition. In Experiment 3, which ran concurrently with EEG and will be discussed fully in Chapter 4, a higher number of trials were needed in order to achieve a good signal-to-noise ratio (SNR). A good SNR is achieved by maximising the amount of recording (signal) to the "noise" or unwanted signal that is recorded with EEG. This is important when working with a new paradigm and potential effects that could be masked by disruptions to the signal for example, movement. Doubling the number of trials in Experiment 1.5 allowed me to check whether this manipulation would have any significant effect on the behavioural level and therefore, a potential neurological effect.

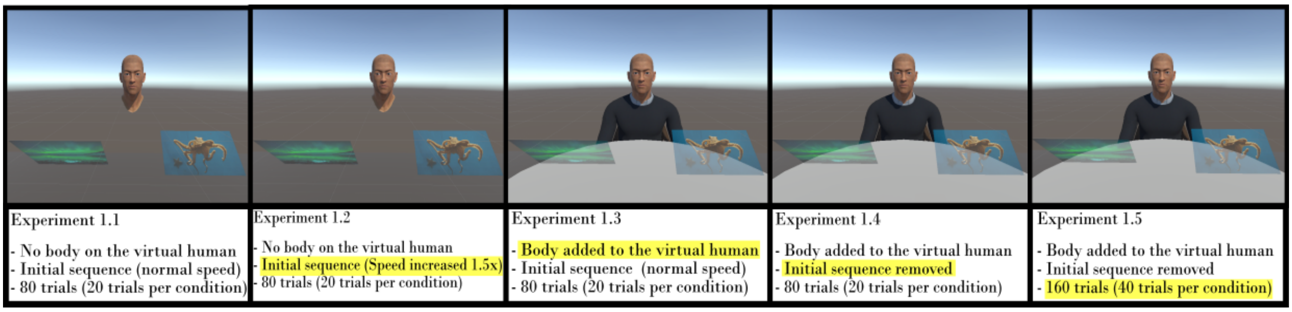

Figure 2.8: A figure that represents the changes made to each experiment with visual depictions accompanied by a written explanation of what was changed. Starting from the left, I displayed a still image of Experiment 1.1 and below the image, I have bullet points of the experimental structure. Sequentially, across the rest of the figure, still images of each experiment (Experiments 1.2 - Experiment 1.5) are display with the experiment structures included. In the written explanation of each of the experiments, one of the bullet points is highlighted for each experiment. The highlighted bullet point is the change that had been made for that experiment.

2.3.3 Results

2.3.3.1 Behavioural Analysis

To begin, the response times and accuracy data were collected via Pavlovia and then extracted using Matlab. Then at the individual level, the mean response times were calculated and responses above or below 2 standard deviations (SD) were removed. This data was then visualised and inspected. Upon visual inspection, I noticed the data was still heavily skewed with large outliers persisting in the data. To mitigate against this, I initially planned to continue to use SD to remove outliers at the group level. However, I found that the SD was influenced by the large outliers and was not a suitable measure of removal for this data set. As a result, I decided to use an alternative method that provided a robust measure of the variability in the data, this was the Mean Absolute Deviation (MAD). The MAD describes the average of the absolute deviations from the mean. It is a robust statistic that is more resilient to outliers than the SD. Based on a Gaussian (normal distribution), it has been reported that the MAD is ≈80 % of a SD (see Equation 2.1). Specifically, it is reported that $1SD$ is equal to 0.79788 MAD (Geary, 1935). I used this to calculate the MAD that was equivalent to $3SD$s for each experiment (see Equation 2.2).

$$MAD \div \ SD = 0.79788 \tag{2.1}$$

$$0.79788 \times 3 = 2.39364 \approx 3SD \tag{2.2}$$

This number was used for outlier rejection at the group level. For example, if a participant had three or more conditions that were above the (MADx2.39), their data was removed from the experiment. As a result of this method, nine participants were removed in Experiment 1.1, twelve participants were removed in Experiment 1.2, seven participants were removed in Experiment 1.3, three participants were removed from Experiment 1.4 and eight participants from Experiment 1.5. In this instance, I was aiming to be as conservative as possible with the data and so I had to use larger constraints than what is typically used in psychology experiments that are normally distributed data. Following this, the data was transferred into Jeffreys's Amazing Statistics Program (JASP), an open-source statistics program, to perform a one-way analysis of variance (ANOVA) on the response times of correct answers for Collaborative trials and all answers for Non-Collaborative trials.

Due to the nature of the Non-Collaborative gaze sequences, the virtual human does not look at either puzzle board. The virtual human either looks up (toward the ceiling) or down (at the puzzle piece). At these central gaze points, the virtual human's gaze is equidistant from both boards. Therefore, in the Non-Collaborative condition, the *correct* answer is not defined and it is meaningless to derive an accuracy measure. However, accuracy was established in Collaborative gaze sequences, where the virtual human looked at one of the puzzle boards. The puzzle board that the virtual human gazed at was assigned as correct. A t-test was then performed to determine the potential effect of Gaze Type (Eye versus NoEye) within the Collaborative condition only. The analysis procedure was identical for all experiments.

2.3.3.2 Behavioural Results

A repeated-measures ANOVA was conducted on mean response times, with two within-subject factors: Collaboration (Collaborative, Non-Collaborative) and Gaze Type (Eye, NoEye). Reaction times were **faster** (three experiments) and **more accurate** (one experiment) for Collaboration, consistent with trial information content, but **slower** for eye contact (one experiment), revealing an unexpected aftereffect that suggested a time-consuming cognitive resource for eye-gaze. Participants were instructed to respond to these behaviours as quickly and accurately as possible by indicating to which board they were being directed. Virtual human speed and construction (with or without body) were manipulated across experiments but had no effect on the general pattern of results.

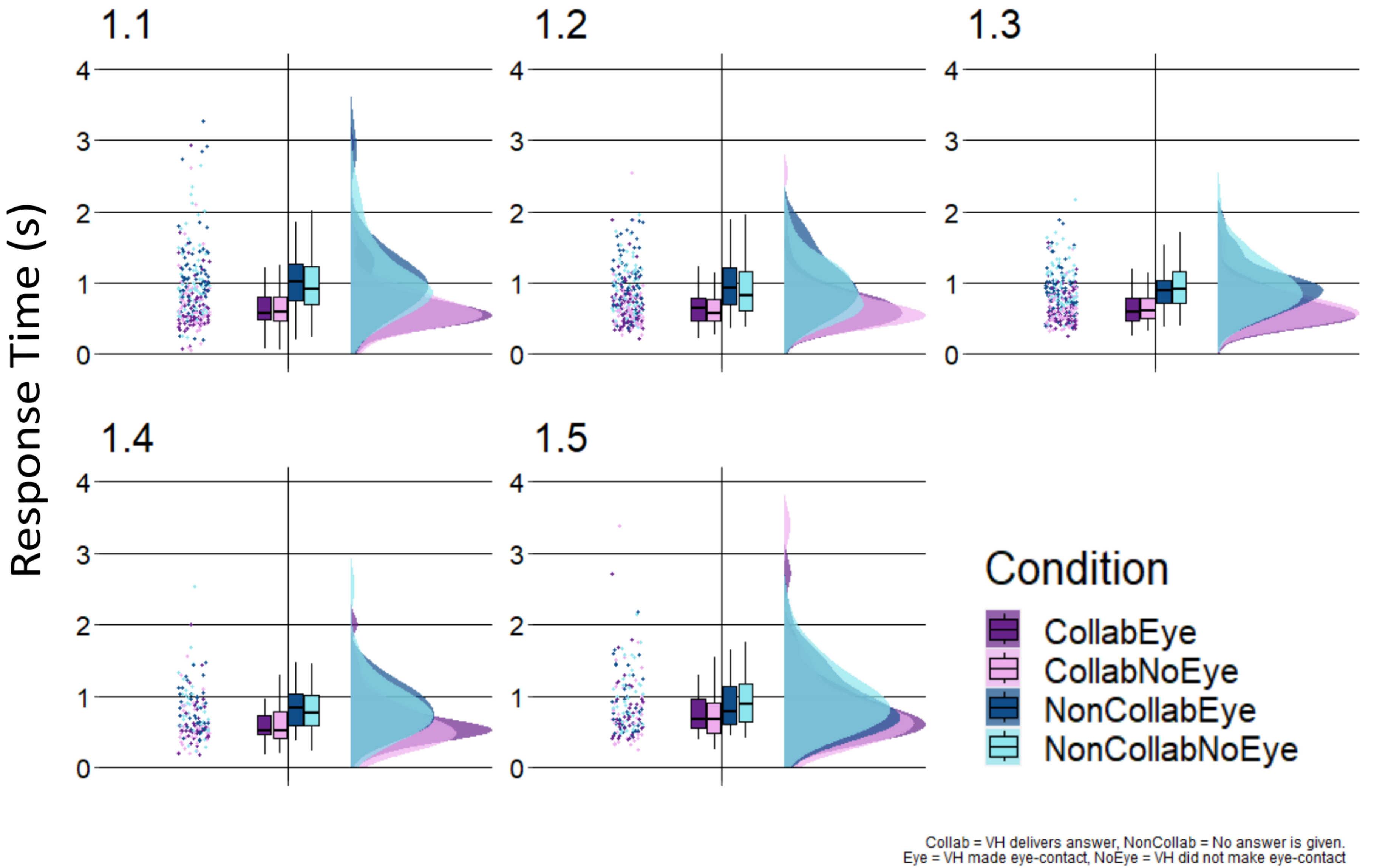

Figure 2.9: The response time results (milliseconds) of Experiments 1.1-1.5 visualised using raincloud plots. Collaborative conditions are in purple and Non Collaborative conditions in blue. Darker colours depict Eye conditions and lighter colours NoEye conditions. The individual points (the "rain") represent the raw data for each participant. The box plots the distribution of the data, with upper (75% quantile) and lower hinge (25% quantile). The middle horizontal line represents the median (50% quantile) and the whiskers display a ± 95% confidence interval. Lastly the "cloud" shows the spread of the data. The raincloud plots were created using R Statistics (R Core Team, 2021) and The Raincloud Plot package by (Allen et al., 2021).

The response time data for Experiments 1.1-1.5 are visualised as raincloud plots in Figure 2.9.

In Experiments 1.1 to 1.4, I observed a significant main effect of Collaboration on response time, with the Collaborative condition faster than Non-Collaborative (Experiment 1: $F(1,70)=77.56$, $p<.001$ ($d=.53$); Experiment 2: $F(1,68)=66.39$, $p<.001$, ($d=.49$); Experiment 3: $F(1,73)=95.74$, $p<.001$, ($d=.57$); and Experiment 4: $F(1,34)=39.18$, $p<.001$, ($d=.54$)). Experiment 5 revealed no significant main effect of Collaboration ($F(1,37)=.92$, $p=0.35$, ($d=.02$)).

In Experiments 1.1, 1.3 to 1.5, no significant main effect of Gaze Type on response time was found (Experiment 1.1: $F(1,70)=.20$, $p=.66$ ($d=.00$); Experiment 1.3: $F(1,73)=1.95$, $p=.17$, ($d=.03$); Experiment 1.4: $F(1,34)=.38$, $p=.54$, ($d=.01$); Experiment 1.5: $F(1,37)=.12$, $p=.73$, ($d=.00$)). Experiment 1.2 revealed a significant main effect of Gaze Type $F(1,68)=4.6$, $p=0.04$, ($d=.06$). These response time data has been visualised.

Similarly, there was no significant interaction between Collaboration-by-Gaze Type for any of the experiments (Experiment 1.1: $F(1,70)=.02$, $p=0.88$ ($d=3.48$); Experiment 1.2: $F(1,68)=1.73$, $p=.19$, ($d=.03$); Experiment 1.3: $F(1,73)=0.92$, $p=.34$, ($d=.01$); Experiment 1.4: $F(1,34)=.76$, $p=.39$, ($d=.02$); Experiment 1.5: $F(1,37)=.2.30$, $p=.14$, ($d=.06$)).

As mentioned above, accuracy was derived only for Collaborative conditions. Therefore, Gaze Type was examined within the accuracy data. In Experiment 1.1, there was a significant difference in accuracy for Collaborative conditions when Gaze Type was at 75% Eye (M=14.99, SD=3.10) versus 83% for NoEye (M=16.68, SD=3.26); $t(79)$=-5.96, $p<.001$, Cohens d=-.67. However, the accuracy data for Experiment 1.2 where Eye was 81% (M=16.13, SD=3.40) and NoEye was 81% (M=16.18, SD=3.72); $t(79)$=-.21, $p=0.84$, Cohens d=-.02, Experiment 1.3 where Eye was 83% (M=16.63, SD=4.01) and NoEye was 82% (M=16.35, SD=4.04); $t(81)=1.28$, $p=0.20$, Cohens $d=.14$, Experiment 1.4 where Eye was 84% (M=16.87, SD=3.44) and NoEye was 83% (M=16.53, SD=4.07); $t(38)=1.01$, $p=.32$, Cohens $d = .16$, and Experiment 1.5 where Eye was 82% (M=32.59, SD=6.45) and NoEye was 81% (M=32.24, SD=6.26); $t(46)=.72$, $p=.48$, Cohens $d = .11$ all found no significant difference between the conditions.

2.4 Discussion

2.4.1 Summary of findings

All five studies aimed to investigate how collaboration and eye contact with a virtual human affect participants' response times and accuracy (where applicable) when responding to a bid for joint attention. Previous studies have focused on the addition of deictic signals (i.e. pointing and verbal communication; (Bacos, 2020; Jyoti & Lahiri, 2020) and how that differs from a solitary action, for example, only pointing. Many studies have reported a benefit when using multiple deictic signals (Cilia et al., 2020; Deák et al., 2000; Morales et al., 2000). However, these studies ignore some of the smaller, more subtle, behaviours that one action can display and the resulting impact that a solitary action can have on the overall process of joint attention. In this study, I investigated eye-gaze and the differences in gaze behaviour that may increase or hinder the ability to understand and respond to a cue correctly.

Overall, the results support hypothesis one (H1), as I found that participants responded more quickly during Collaborative conditions in comparison to Non-Collaborative conditions. This was expected as these conditions provided clear signals as to where the puzzle piece should be placed.

In all experiments, I expected eye contact to facilitate performance by decreasing response times and increasing accuracy. However, I found that participants' responses were slower in Eye conditions in Experiments 1.1 and 1.2. One explanation for this unexpected increase in response times could be a result of the disembodied head and its validity when using virtual humans as social partners. Some studies argue that humans do not interact with virtual humans in the same way and phenomena such as the uncanny valley effect may serve as a reminder that the virtual human is not a real human being (Mori, MacDorman, & Kageki, 2012), thus limiting their utility for these experiments. Increasing the realism of the virtual human was considered during the development of the paradigm and a body was included from Experiment 1.3 onward. Additionally, in Experiment 1.2 I observed a significant effect of Gaze Type where Eye conditions were significantly slower than NoEye. The change made in Experiment 1.2 (increasing the speed of the initial sequence) was to create the appearance that the initial sequence was an informal "check" to see which puzzle board the puzzle piece belonged to before providing a communicative gaze shift. Based on the results from Experiment 1.1 and specifically 1.2, participants may be interpreting the initial sequence as intentional and communicative gaze shifts.

Additionally, as the study was run on remote, 2D computer screens, I believed it did not give the

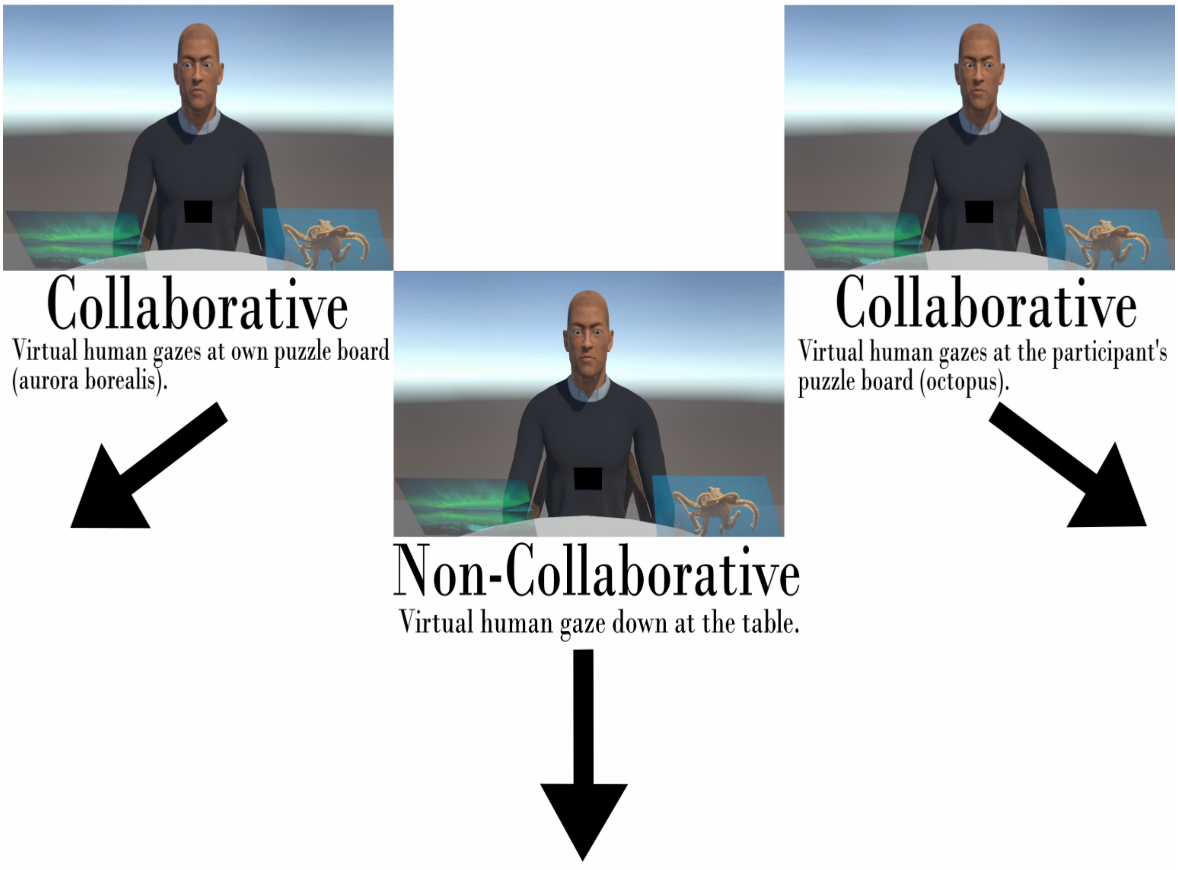

Figure 2.10: The figure shows a collage of still images from each of the conditions. On the left and right-hand side, the images display the Collaborative conditions and in the central, downward image is a display of a Non-Collaborative condition. Next to each image is the condition label ('Collaborative' or 'Non-Collaborative'), a written description of the gaze direction and an arrow pointing in the same direction as the gaze.

opportunity to increase the participant's immersion and potentially their presence by being in a shared environment with the virtual human. The use of a head-mounted display allows for one stable visual space to be presented that maintains consistency and continuity. In this preferred context, the virtual human's potential of being a social partner may become more favourable. Further improvements would include the use of animation. For example, an idle animation of the virtual human's body, small breathing movements and a slight posture sway that would reflect natural human behaviour. These behaviours in an immersive environment are not only important to reduce the potential of the uncanny valley effect but, can also increase the ecological validity of experimental paradigms when exploring social communication.

An alternative explanation to the observed increased response times in Eye conditions, particularly in the Non-Collaborative, Eye condition (e.g. Expt 1.2) may be due to the influence of direct gaze. In this condition, the virtual human is engaging in eye contact with the participant for the whole trial and does not provide any direction as to where the puzzle piece belongs. These increased response times suggest an effect of gaze but, not one that was anticipated: direct eye-gaze may imply communication intention. This explanation suggests that the participants may have believed that the virtual human had intended to communicate with them, indicated by engaging in eye contact, but did not follow through. This is supported by the aforementioned work of (Caruana et al., 2017a) who looked at the differences between a virtual human searching for a target and then communicating the location of the target to the participants (Search conditions), in comparison to the virtual human communicating the location of the target without searching (NoSearch conditions). They reported that the intention monitoring process, which would happen during the Search condition had a significant effect on RJA, which resulted in slower response times. In Caruana et al's (2017) study, participants identify a cue as being an intentional bid to initiate joint attention. In this study, the combination of the initial sequence and the Non-Collaborative, Eye condition may have engaged a similar process of intention monitoring. Eye contact suggests an intention to communicate and it is plausible that an agent may expect that some form of communication or information, will follow after (Frith & Frith, 2007). Therefore, when no information is conveyed after a trial displays the initial sequence and a Non-Collaborative, Eye condition, additional cognitive resources may be recruited to understand or account for the unexpected outcome.

One potential shortcoming on the experimental design is the systematic control between the condition sequences (see Figure 2.4). For example, differences in the length of the trial and the number of gaze shifts performed in each trial. This has been revised and restructured in Chapter 3 to balance the conditions on a number of factors. Overall, I have demonstrated that collaborative gaze improves participants' performance (response time) when working with a virtual human. I stipulate that the increased response times in Non-Collaborative, Eye conditions are a result of intention monitoring which leads the participant to recruit more cognitive resources.

2.4.2 Implications for subsequent experiments

Upon review of this first set of experimental results, I went back to the paradigm to assess its effectiveness and validity. As a result, I made changes to a range of aspects that included but was not limited to the length of trials, the time taken to perform a gaze shift and the removal of the initial sequence. In addition, I also assessed the effects of the initial sequence and the benefits it holds. After this process, I speculated that the initial sequence may feel repetitive and unnecessary after the first trial. The initial sequence involves the virtual human looking at the puzzle piece and then both puzzle boards. While looking at the puzzle piece makes sense for each trial, looking at both puzzle boards for every trial seemed redundant and behaviour that would not reflect real life in such a simple task. Secondly, I examined the sequence of conditions with the overall aim to balance (i) the number of gaze shifts made, (ii) the time the informative gaze was performed, (iii) counterbalancing of gaze locations in non-collaborative conditions, (iv) the length of a trial and (v) computer-controlled gaze. These are key confounding errors that I had thought about when revisiting the paradigm to interpret the results. In later experiments, I addressed these issues and made the appropriate changes to refine the paradigm and increase validity.

Chapter 3

Refining gaze sequences to investigate the role of eye-gaze in joint attention

In this chapter I outline a revised version of the previous sets of experiments detailed in Chapter 2. In the following experiments (2.1 and 2.2), I address the outcomes and limitations of Study 1 through the removal and addition of specific scenes within the paradigm. Adjustments were made to more accurately explore the role of eye-gaze in a social interaction. The manipulations focused on using computer-controlled gaze and re-defining the gaze sequences by adapting their number, timing and locations.

3.1 Introduction

After reviewing the first set of experiments, I worked to revise the paradigm. The main goal of Study 2 (2.1 and 2.2) was to develop a paradigm that was optimised to investigate the role of eye-gaze in joint attention, in addition to being compatible with a subsequent experiment that would include EEG and virtual reality (Chapter 4). Study 1 helped to refine the configuration of the experiment where I concluded that (i) the initial sequence should be removed (ii) the virtual human should be presented with a body, and (iii) the number of trials could be doubled for the EEG experiment without compromising the trend of results. Firstly, I investigated the validity of the initial sequence and whether it provided the paradigm with behaviours that influenced the participant's response or interpretation of the virtual human's eye-gaze. The results from the manipulations on the initial sequence presented findings that suggested an effect of intention monitoring. Although these results were interesting, the aim of this project was to not only investigate the role of eye-gaze in joint attention but to address what behaviours support successful joint attention. In addition to this, the initial sequence becomes redundant after familiarity with the experiment is reached. The location of the puzzle boards and their images requires little memory and so this familiarity is quickly learnt. I concluded that the removal of the initial sequence did not disturb the overall aim of the paradigm (as seen in Figure 2.9). The removal of the initial sequence also reduced the time taken for each individual trial, which allowed me to maximise the number of trials when testing using EEG. Secondly, a decision was made to keep the body that was added to the head of the virtual human in Experiments 1.3 - 1.5. It was my impression that a virtual human with a body as opposed to a disembodied head would provide an increased level of presence for the participant and the illusion that the virtual human had autonomy in preparation for Experiment 3, in Chapter 4. Lastly, the number of trials were not doubled in the following two experiments (2.1 and 2.2) as they were online experiments, and the increased number was specifically for Experiment 3 in order to record EEG with enough signal-to-noise ratio.

For Study 2, my goal was to create conditions that were equally balanced and had high validity for both of the factors: Gaze Type and Collaboration. This was completed by refining the gaze sequence (see Figure 3.1) and paradigm through computer-controlled gaze and optimally matching (i) the number of gaze shifts made, (ii) the time the informative gaze was performed, (iii) counterbalancing of gaze locations in non-collaborative conditions and (iv) the length of a trial.

In order to rectify this, with the later goal of using the paradigm while recording EEG, each condition contained two gaze shifts/transitions. The first was from the starting position to the second. At this second gaze position, the collaborative or non-collaborative gaze shift would take place (pictured in Figure 3.2). These were systematically placed in the second position of each trial, to maximise the quality of EEG data from Experiment 3. The second gaze shift/transition was from the second position to the third. Together, these two changes that have been detailed above, addressed the points i and ii, to ensure that the length of each of the conditions was matched. Lastly, at the time that these experiments were being developed, the COVID-19 restrictions had been lifted and access in to the laboratory was granted. With this, I experimented controlling the virtual human's eye-gaze with pre-programmed locations that were computer-controlled. The first human-controlled gaze recordings (Experiment 1.1 and 1.2) displayed raw gaze, where the shifts included blinks and less standardised or clean shifts in gaze. One could argue that this was better for ecological validity and displayed "true" human gaze behaviour. However, as I continued to develop the paradigm, the importance of systematic gaze shifts increased for experimental reliability. As a result, the human-derived gaze recordings became more systematic (observed in Experiments 1.3 - 1.5). When compared, the computer-controlled gaze shifts were indistinguishable from the human-controlled gaze shifts. Additionally, the computer-controlled version helped maintain the precision and consistency of each gaze shift and for these reasons, was adopted into Experiments 2.1 and 2.2.

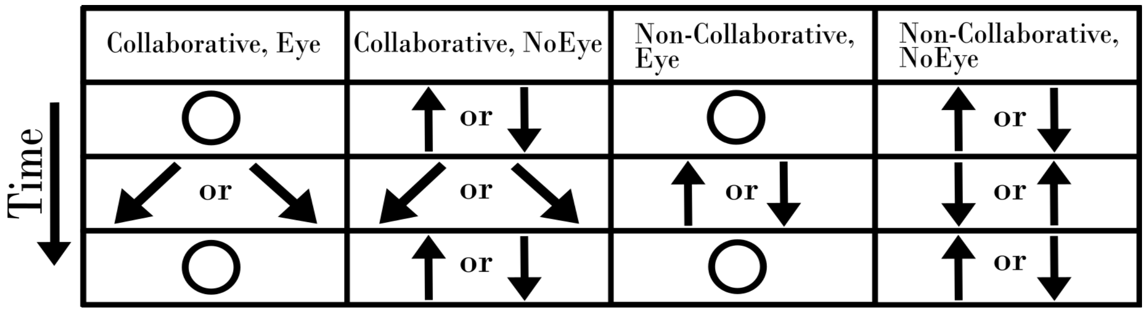

Figure 3.1: **Schematic of the new sequence steps.** The schematic characterises the refined sequence of gaze shifts that the virtual human performs in each of the conditions. The circle (**O**) represents direct eye contact with the participant and the arrows indicate where the virtual human was looking. The arrow pointing down represents the virtual human looking directly down in front. Similarly, the arrow up would be gaze upwards (e.g. the direction of where the lamp was in Experiment 2.2). Arrows pointing diagonally left and right represent eye-gaze towards either of the puzzle boards. These directions are visualised in Figure 3.2.

For Experiment 2.1 the scene environment remained the same. Conversely, in Experiment 2.2 a ceiling light was added to the environment (Figure 3.4), directly above the virtual human's head. In the new gaze sequences described above (Experiments 2.1 and 2.2), I have included an upwards gaze location to counterbalance the downward gaze shift. This new gaze shift motivated the inclusion of a ceiling light in the environment in order to provide an object that is being looked at, as the other gaze shifts have an object present in the line of sight. With these changes made, I predicted the following hypotheses:

3.2 Hypotheses

H1: The main effect of Collaboration will maintain and I will see a significant main effect of Gaze Type in comparison to Experiments 1.1-1.5.

H2: In Experiment 2.2, I expect Non-Collaborative conditions to have a difference in the response time pattern in comparison to Experiments 1.1-1.5.

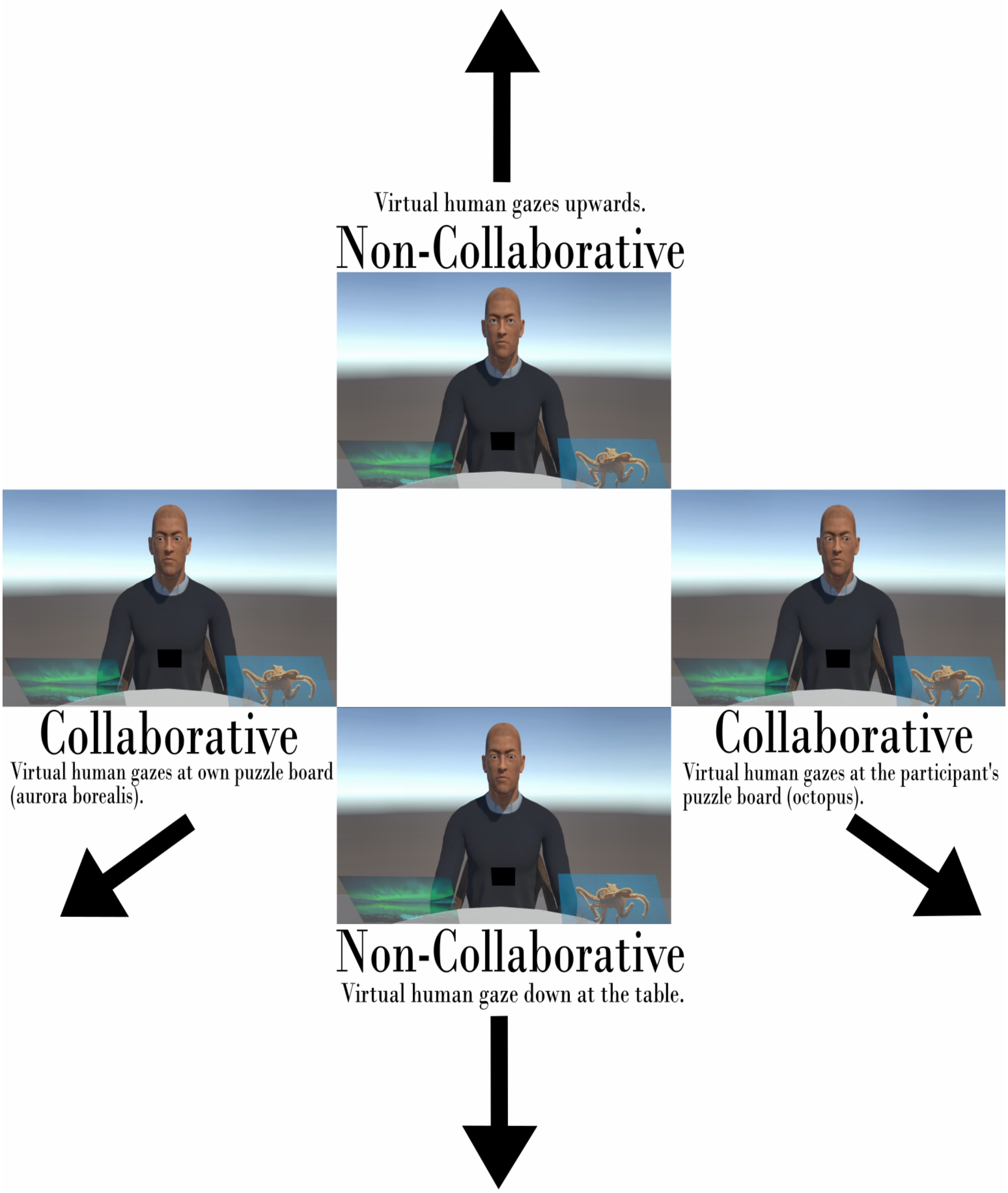

Figure 3.2: The figure pictures a collage of still images from each of the conditions. Horizontally, the images display the Collaborative conditions and vertically, the images display Non-Collaborative conditions. Next to each image is the condition label ('Collaborative' or 'Non-Collaborative'), a written description of the gaze direction and an arrow pointing in the same direction as the gaze.

3.3 Methodology

Similar to the previous chapter, the experiments in this chapter were also created within Unity technologies (Unity3D, 2019) and ran online. The physical build and presented environment of the experiment remained the same which involved the two puzzle board images: an octopus and an aurora borealis. However, in the current experiment, the virtual human's eye-gaze shifts were programmed and animated to rotate towards predetermined location points in the Unity software. These location points were created by placing cubes (made visible as black cubes in Figure 3.3, 1a) in the areas where the gaze was going to be directed. These cubes were placed as markers for the sphere to move towards and therefore, the virtual human's gaze to follow.

Using Visual Studio software, I used C# to code the virtual human's eye-gaze to follow a sphere (made visible in Figure 3.3, 1b and 1c) that was positioned in front of the virtual human. The cube locations were determined by the intuition of the experimenter; whether the virtual human's gaze implied that the object (puzzle board or puzzle piece) was being *looked at*. Again, these sequences were recorded using OBS recording software and the videos formed as part of the online experiment set up that was created in PsychoPy3. These experiments followed the same upload and distribution procedure as before; (i) uploading to Pavlovia and (ii) advertising to registered participants on the Prolific website. Again, the Autism Quotient questionnaire was utilised to ensure that only neurotypical behaviours comprised the data.

3.3.1 Participants

We recruited participants through an online recruitment platform (Prolific.co/). Participants were screened through the website's criteria: by age (18-40yrs), the ability to speak and understand fluent English, having normal or corrected vision, having scored below 26 on the AQ(Baron-Cohen et al., 2001; Ashwood et al., 2016), having never experienced traumatic brain injury and were excluded if they had already taken part in any of our other experiments in this study. The justifications for such criteria are identical to those explained in section 2.3.1. In the first experiment (Experiment 2.1) with the new paradigm, I analysed 80 participants, 40 were female and the whole group had an average age of 27 years. Participants age ranged between 19-40 years with a standard deviation of 5 years. In the second experiment (Experiment 2.2), I analysed 80 participants, also with 40 females and had a total mean age of 28 years. Participants age ranged between 19-39 years with a standard deviation of 6 years.

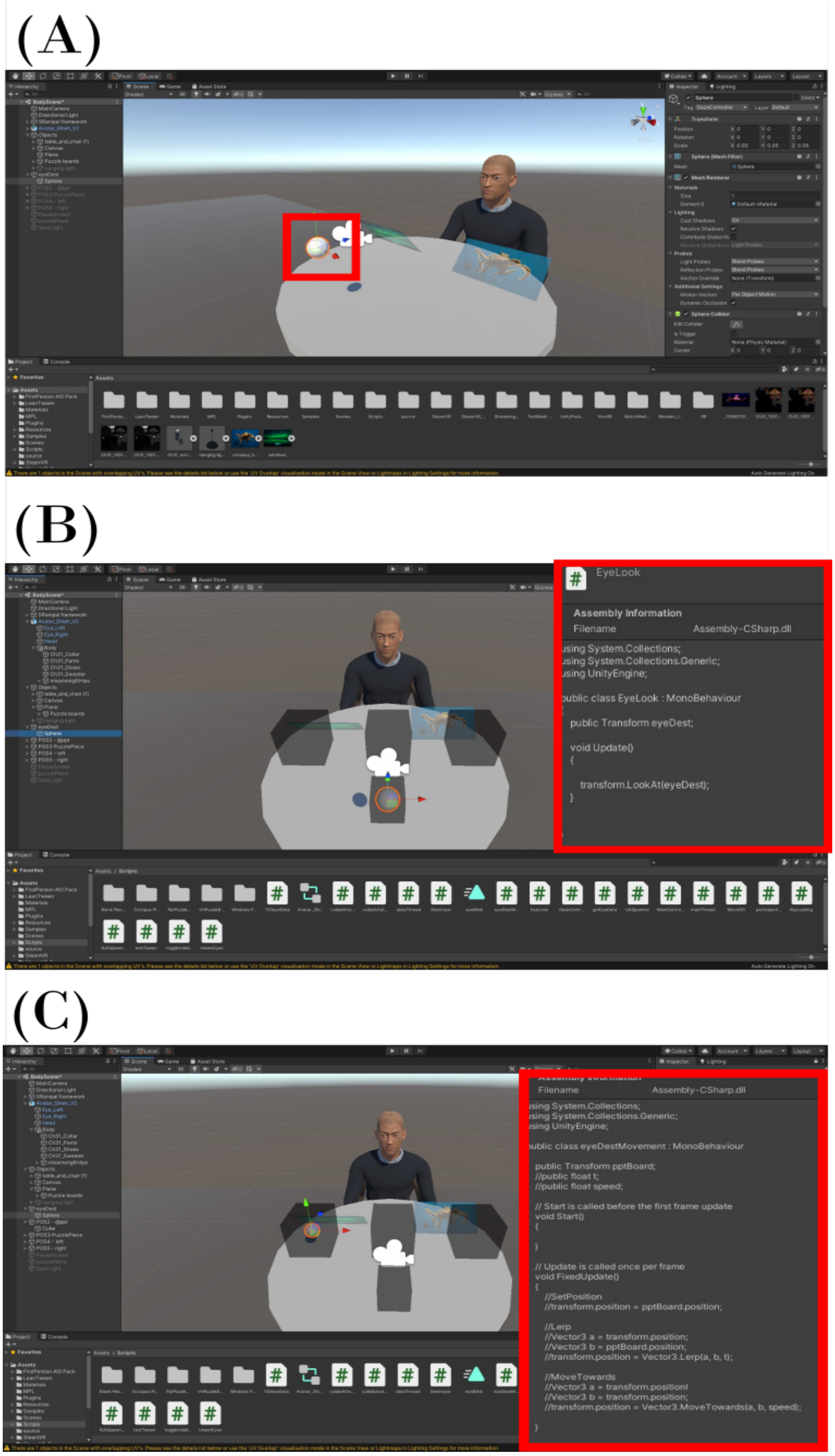

Figure 3.3: Still images of the virtual scene and Unity platform that the gaze sequence was developed on. Highlighted sections of the images focus on the sphere that the virtual human's eye-gaze follows **(a)**, the script that controls the virtual human's eye gaze to follow the sphere **(b)** and the script that controls the sphere's location at either one of the four black cubes **(c)**. The black cubes in scenes **b** and **c** visualise the location points that the virtual human's gaze is directed. The black cubes and the white sphere are invisible during the experimental trials.

3.3.1.1 Experiment 2.1

Experiment 2.1 was the first experiment that used the new sequence structure (Figure 3.1). With this, I continued to use the same method of data collection that was used in Study 1 and briefly reiterated in Section 3.3 .

3.3.1.2 Experiment 2.2

This is a replication of Experiment 2.1 with the inclusion of a lamp in the environment. The upwards gaze was the only gaze direction that did not have an object in its line of sight. The other gaze directions have either (i) either of the puzzleboards, (ii) directly at the participant or (iii) down at the table. This motivated the inclusion of a lamp above the virtual human, so that the upwards gaze would also have an object in it's line of sight. The data was then collected using the same method in Experiment 2.1.

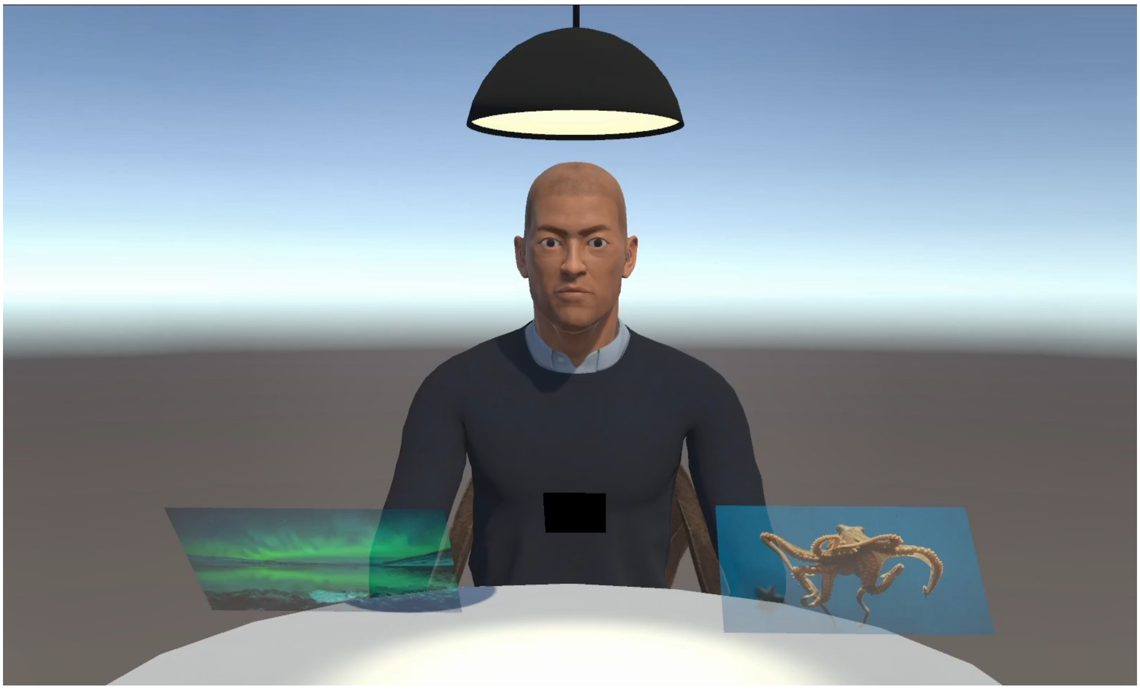

Figure 3.4: A still image of the experiment environment with the addition of a lamp for Experiment 2.2.

3.4 Results

3.4.1 Behavioural Analysis

Maintaining the continuity of the protocol that was used in the first set of experiments, I continued using the same method for the current set of experiments. Similarly, all Non-Collaborative trials were used in the analysis, but only correct trials were included from the Collaborative condition. To recap, the mean and *SD* for each condition were calculated for every participant using Matlab 2019b. However, due to the online experiments producing large outliers that heavily skewed the data, I also calculated the *MAD*, which is a more robust statistic to deal with outliers in comparison to the *SD*. I calculated 3*MADs* at the group level to remain conservative of the data. If a participant had three or more conditions that were above the 3*MADs* calculated for that experiment, the participant was excluded from the data. A more in-depth explanation of the utility of the *MAD* is detailed in Subsection 2.3.3.1. Using the MAD method, twelve participants were removed in Experiment 2.1 and fourteen from Experiment 2.2.

Due to the nature of the Non-Collaborative gaze sequences, the virtual human does not look at either puzzle board. The virtual human either looks up (toward the ceiling) or down (at the puzzle piece). At these central gaze points, the virtual human's gaze is equidistant from both boards. Therefore, in the Non-Collaborative condition, the *correct* answer is not defined and it is meaningless to derive an accuracy measure. However, accuracy was established in Collaborative gaze sequences, where the virtual human looked at one of the puzzle boards. The puzzle board that the virtual human gazed at was assigned as correct. A t-test was then performed to determine the potential effect of Gaze Type (Eye versus NoEye) within the Collaborative condition only. The analysis procedure was identical for experiments 2.1 and 2.2.

3.4.2 Behavioural Results

The response time data for Experiment 2.1 and 2.2 has been visualised as raincloud plots in Figure 3.5.

In Experiment 2.1, I observed a significant main effect of Collaboration, with Collaborative conditions significantly faster than Non-Collaborative trials ($F(1,66)=13.25$, $p<.001$ ($d=.17$). I also found a significant main effect of Gaze Type, where Eye conditions were significantly faster than NoEye

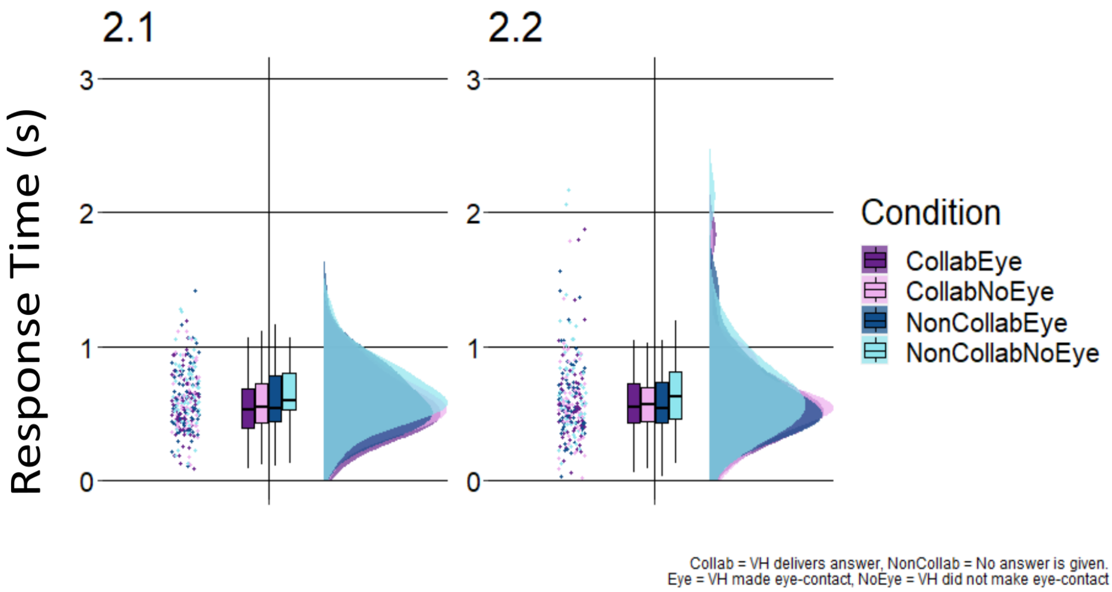

Figure 3.5: The raincloud plots display the response time results from Experiments 2.1 and 2.2 in milliseconds. Collaborative conditions are in purple and Non Collaborative conditions in blue. Darker colours depict Eye conditions and lighter colours NoEye conditions. The individual points (the "rain") represent the raw data for each participant. The box plots the distribution of the data, with upper (75% quantile) and lower hinge (25% quantile). The middle horizontal line represents the median (50% quantile) and the whiskers display a ± 95% confidence interval. Lastly the "cloud" shows the spread of the data. The raincloud plots were created using R Statistics (R Core Team, 2021) and The Raincloud Plot package by (Allen et al., 2021).

conditions ($F(1{,}66)=4.35$, $p<.05$ ($d=.06$). No significant interaction was found between Collaboration-by-Gaze Type.

In contrast, for Experiment 2.2, I found no main effect for Collaboration (F(1,64)=2.66, $p=.11$ ($d=.04$)), Gaze Type ($F(1{,}64)=1.45$, $p=.23$ ($d=.02$)), nor an interaction ($F(1{,}64)=3.50$, $p=.07$ ($d=.05$)).

Accuracy results from Experiment 2.1 found that participant's responses were significantly more accurate in Collaborative, Eye conditions at 88% (M=35.01, $SD=7.51$) compared to Collaborative, NoEye conditions at 84% (M=33.70, $SD=6.98$; $t(79)=3.05$, $p<.001$, Cohens $d=-.34$).

Similarly, in Experiment 2.2, accuracy was significantly higher in Collaborative, Eye conditions at 90% (M=36.18, $SD=5.81$) than Collaborative, NoEye conditions that were 88% (M=35.20, $SD=6.19$; $t(79)=2.36$, $p=.02$, Cohens $d=-.26$).

3.5 Discussion

3.5.0.1 Summary of findings

Experiment 2.1 found that Eye conditions were significantly faster than NoEye. This result was originally hypothesised for Study 1 (1.1-1.5) but was not found in the first two experiments (1.1-1.2). However, after revisiting the gaze sequence with the aim to improve the paradigm, I developed systematic trials that were matched. This was largely achieved by (i) controlling and balancing the gaze shifts within each condition and (ii) having them at a defined and consistent timestamp. A detailed explanation of these changes is written in Section 3.3. As a result of these changes, I observed a significant effect of Gaze Type in Experiment 2.1.

The results of Experiment 2.1 replicated the effect of Collaborative conditions that were found in Study 1 but also highlighted the effect of eye-contact on response times. Notably, I did not see the same effect of Non-Collaborative, Eye conditions that were observed in some of the experiments in Study 1. I speculate that this effect may be regulated by the use of a disembodied head and/or the initial sequence. As intention monitoring was not the focus of our studies, I cannot comment further at this stage but consider its role as a possible future investigation and manipulation within the paradigm.

However, the results from the effect of Gaze Type in Experiment 2.1 may provide evidence for a social effect of gaze. Overall, the gaze sequences that recruited the participant's eye contact was processed differently in comparison to the gaze sequences that did not engage in eye contact. In Particular, we see this difference within Collaborative trials. When the participant was given the location information i.e. the board that the puzzle piece belonged to, the trials that included eye contact (Collaborative, Eye) were still faster in comparison to trials that did not (Collaborative, NoEye). Similarly, when no useful gaze direction is given (Non-Collaborative), participants also responded faster after engaging in eye contact with the virtual human (Non-Collaborative, Eye) compared to trials that did not (Non-Collaborative, NoEye). The overall significant decrease in response times when engaging in eye contact may suggest an influence of eye contact where our attention is possibly heightened and we assign importance to the behaviour (Ristic et al., 2002).

Critical evaluation of the paradigms I have developed has helped identify possible limitations. Firstly, the virtual human's likeability and suitability for participants remain a concern. In feedback from conference poster sessions, experts in the field suggested changes to the virtual human look. Suggestions centred around developing a virtual human that looked more friendly and natural, with concerns that the overall look and feel of the virtual human may be contributing to the uncanny valley effect. An aspect of realism that was not included in Study 2, was the use of the virtual human's eyelids. In Study 1, I made use of sporadic eye blinks, as a result of using recordings derived from human behaviour. However, in Study 2, I programmed the virtual human's eye-gaze and therefore the eyelids would also need to be programmed, in order to blink. This was achieved for Study 2 but unfortunately did not have the desired effect and added an additional element of 'strangeness' to the virtual human. Animating the virtual human's eyelids to reflect a natural movement, required a better understanding of eyelid movement and animation. To capture this accurately, it would have required more time and training to achieve. This wasn't feasible in the given time frame and therefore, a decision was made to keep the eyelids static. Future experiments could address this by working on effective animations for eyelids that consider (i) the lowering of eyelids when gazing in certain directions, particularly when gazing down and (ii) the inclusion of blinks.

In addition, I propose an overall focus on the animation of the virtual human to increase realism. As stated in Chapter 2, I propose that future paradigms should include idle animations of the virtual human, that make use of breathing animations to replace the fixed body used in all of my studies. While the study is not focused on whole-body behaviour, I believe it would be appropriate to include subtle animations that would increase the realism of the virtual environments I display to participants. To do this with the current disembodied head, an animation skeleton would need to be rigged with a chosen or modelled 3D body. This type of work can be done in free and open-source software like Blender. Blender offers computer graphic tool sets for the whole 3D pipeline and includes; animation, 3D modelling and rigging (Blender, 2018). This software could also be used to model new virtual humans and develop the paradigm with different agents. Animation of the virtual human is key to the paradigms' future development. This is beyond the scope of Unity capabilities and requires more time and experience using software like Blender.

3.5.0.2 Implications for the next experiment

Overall, the online experiments outlined in Study 1 and 2 (1.1-1.5 and 2.1-2.2, respectively) have extended the breadth of joint attention literature by demonstrating the effects of dynamic eye-gaze. Study 1 demonstrated a strong effect of Collaborative conditions and presented a potential effect of eye contact that led to refinements for Study 2. The experiments of this chapter (2.1-2.2) have provided me with the final version of the paradigm that will be used in Experiment 3 and displayed in a virtual reality head-mounted display while recording EEG and eye-tracking, concurrently. It is worth noting that the use of online experiments, although initially dictated by the pandemic, are a useful and quick tool to prototype novel paradigms and refine conditions before recruiting in person participants. This medium allows for fast data collection, the narrowing of parameters and the ability to collect data remotely (from within participant's homes) without requiring travel and potentially troublesome logistics.

Chapter 4

4. The role of eye-gaze in joint attention: a EEG-VR study

In the final study, in-person testing was resumed by following COVID-19 standard operating procedures. Original plans for this work would have incorporated both initiating and responding to joint attention. However, due to time constraints, only responding to joint attention was investigated. The refinements to the paradigm that were made in Experiment 2.1 were used and this allowed a direct comparison between a 2D computer screen and an immersive virtual environment using a head-mounted display. In addition to the method of presentation, I also recorded complimentary EEG and eye-tracking data.

4.1 Introduction

This final study concludes the set of experiments for the PhD project. After taking COVID-19 limitations into consideration, my main goal was to replicate the experiment in an immersive virtual environment. The online studies reported in chapters 2 and 3 provided an accessible platform and fast data collection to adapt and develop the paradigm's sequence. After making these changes, I analysed the data and reconsidered each aspect of the sequence, to create a balanced and concise paradigm ready for in-person testing in virtual reality.

Accurately measuring and understanding the role of gaze in joint attention is important in order to gain clarity on non-verbal communication as a whole. Collectively, the online results provide a better insight into the process of joint attention. The take-home points from these experiments have been combined and used to make a more complete paradigm with more modalities. This chapter made use of a head-mounted display to increase participants' presence in the scene and was coupled with recording EEG, which allowed me to pair the neural correlates associated with the behavioural changes elicited by the experiment. The acquisition of electrophysiological data when using an immersive environment, would extend the literature by informing on neural signatures associated with collaborative gaze and eye contact. By identifying the neural correlates of collaborative gaze and eye contact, I could further understand the behaviours that are critical in the process of joint attention. The combination of immersive virtual reality and neuroimaging can be described as 'Neuro-VR', an example is displayed in Figure 4.1. Neuro-VR, in this case EEG-VR, presents a novel way of investigating cognitive research questions whilst collecting multimodal data in an ecologically valid format. Such a pairing allows for the interrogation of many social questions that require dynamic, flexible agents and settings (Gregory, 2021). The implications of this will be discussed further in Chapter 5.

This study also provides a comparison of online and in-person virtual reality experiments. The results from the comparison will be useful for advancing psychology research and to ascertain the usefulness of immersive virtual reality tools in this field. Crucially, the data compares and explores the differences between immersive virtual reality and 2D computer screen presentation. The standard presentation for experiments is on a 2D computer screen and this is true for most neuroimaging studies. The differences found between the two approaches emphasise the need to distinguish between virtual reality equipment used in experiments and the benefits that can be found in each (discussed in Subsection 1.1.4). The findings from this study were directly compared to the results of Experiment 2.1 (a replication of the experiment that was presented on a 2D computer screen). I will discuss the implications of this at the end of the chapter.

The literature has built an understanding of the neural correlates of joint attention largely through static and 2D paradigms. In general, there are reports of mu activity (11-16Hz) in relation to studies that require motor preparation and in particular, hand motor preparation (Pfurtscheller, Neuper, & Krausz, 2000). However, the two main neural frequencies that are largely investigated in relation to joint attention are alpha (8-12Hz) and theta (3-7Hz). Both alpha and theta signatures individually and collectively have been associated with attention and social processing. Changes in alpha, notably, decreases in power have been associated with higher levels of attention (Sauseng et al., 2005). In regards to joint attention, eye gaze has been reported to modulate alpha power. Eye contact both attracts and holds attention (Senju & Hasegawa, 2005; Senju et al., 2005). On a neural level, this has been

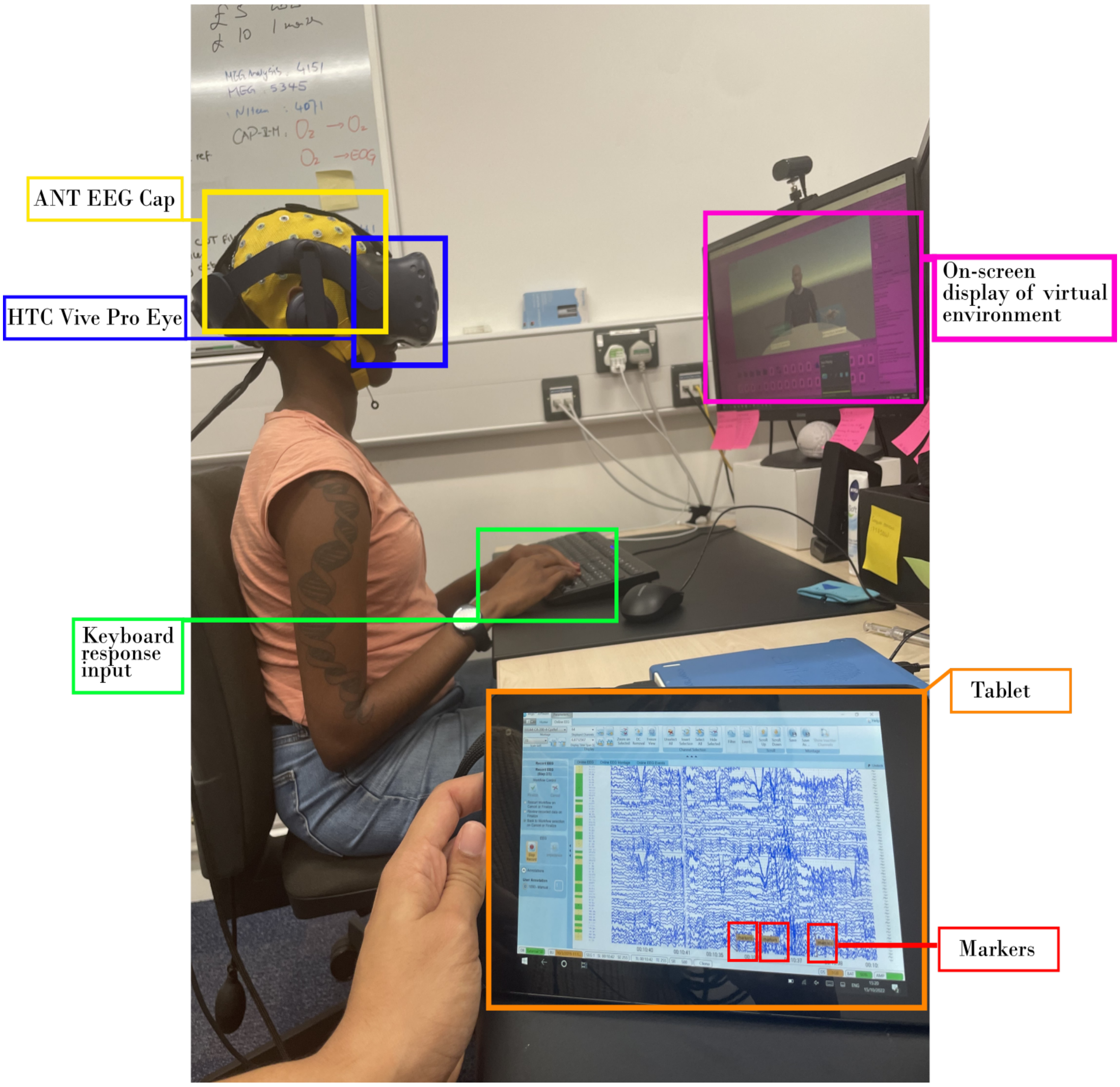

Figure 4.1: An image of the physical experimental set-up. Photographed is a participant seated with the EEG cap on their head and head-mounted display on top. Participants sat at the desk with both fingers on the keyboard for button responses. The blue amplifier is connected to the EEG cap and a small tablet that visualises the live EEG data, as well as triggers (circled in red), that are set up.

observed by pronounced decreases in alpha for direct eye contact compared to averted (Kompatsiari, Bossi, & Wykowska, 2021; Gale, Spratt, Chapman, & Smallbone, 1975). Importantly, Kompatsiari et al (2021) found alpha-band desynchronisation (decreases in alpha power) associated with an agent engaging in eye contact before providing an informative and directional gaze. These decreases were present over left fronto-central and central electrodes. This pattern of alpha decreases is also true when comparing conditions of successful joint attention (both agents looking at the same object) versus no joint attention (agents looking at different objects) (Lachat et al., 2012).

Evidence for theta activity, specifically increases in theta power (synchronisation) are linked with social information integration, over posterior areas (H. Wang, Callaghan, Gooding-Williams, McAllister, &

Kessler, 2016) which are positioned in the literature as an integrative hub for complex social scenarios (Patel, Sestieri, & Corbetta, 2019; Patel et al., 2021). For example, mentalising abilities have been linked with theta synchronisation and have been evidenced by non-invasive theta stimulation facilitating these abilities (Gooding-Williams, Wang, & Kessler, 2016). Evidence of this is conveyed in Gregory et al's (2022) study that investigated both alpha and theta signatures of social (virtual human) and non-social (arrow) cues on working memory. During their encoding interval, they found stronger increases in theta power during the social cue, suggesting theta power may have been modulated by cue type. Their theta effects were in line with research that suggests frontal theta as a mediator for attentional processing and parietal theta for memory encoding (Khader, Jost, Ranganath, & Rösler, 2010). Whereas a lack of theta power and therefore integration has been found to predict impaired social cognition in schizophrenia (Patel et al., 2019). In line with this, studies investigating schizophrenia and bipolar disorder have reported reduced theta power when gaze processing. This was specifically reported between anterior face processing and posterior areas (Grove et al., 2021), specifically, midline anterior and left posterior regions (Lasagna et al., 2023).

Mu rhythms were investigated as participants would be required to prepare their motor response, in this instance, a button press using their index finger. Decreases in mu activity over central electrodes are commonly associated with motor preparation (Pfurtscheller, Brunner, Schlögl, & Lopes da Silva, 2006). However, the specific frequencies are debated when considering whether they overlap with alpha frequencies or not. In this study, I observed mu activity that overlapped with alpha frequencies. Therefore, mu activity was distinguished from alpha based on the topography i.e. bilateral decreases over central electrodes were considered mu rhythms.

The literature on joint attention presents a spectrum of studies that have observed interactions ranging from two human agents (Dravida et al., 2020), to a human agent and a robot agent (Admoni & Scassellati, 2017). While I argue that changing the combination of agents in a study can affect the very behaviours that I am looking to identify, these approaches help us build a full picture of the behaviours associated with all possible agent interactions. A key point mentioned in Chapter 1, Subsection 1.1.2.1, highlights that future research should aim to develop consistent paradigms that apply strict criteria to the manipulations, with a clear aim of the behaviour that is being investigated. This would allow studies to be directly compared across different approaches that may use an alternative agent. As a result, hypotheses were informed based on data (i) from the previous studies in this project (Chapters 2 and 3), (ii) that were formulated as a result of the literature on robots (Kompatsiari et al., 2021) and (iii) using 2D screens in static neuroimaging scanners (Caruana et al., 2015). Joint attention literature that utilises dynamic stimuli, particularly when using a novel tool such as Neuro-VR are relatively rudimentary. While I am aiming to use the literature to guide my expectations of the electrophysiological data, the paradigm is novel therefore, I will also be exploring the main and simple effects within the data. Below are the hypotheses made:

4.2 Hypotheses

4.2.1 Behavioural

H1: Participants will respond faster in Collaborative compared to Non-Collaborative trials.

H2: In Collaborative trials, participants will respond more accurately in Eye vs NoEye conditions.

H3: Across both Collaborative and Non-Collaborative conditions, response times will be faster for Eye vs NoEye conditions.

4.2.2 EEG

H4: In Eye vs. NoEye trials, I predict that Eye trials will have alpha power decreases over left fronto-central and central electrode clusters. This data will be explored.

H4: When comparing Collaborative vs. Non-Collaborative trials, I predict that Collaborative trials will show theta power increases over posterior areas. This data will be explored.

4.3 Methodology and Materials

4.3.1 Physical set up

The HTC Vive Pro Eye was the headset used to present the immersive virtual environment and this was utilised in conjunction with the integrated Tobii eye-tracker (Figure 4.2). The headset used two lighthouses or base stations, which are laser-based positional tracking systems. The lighthouses are placed around the chaperone area (the designated area where the user in the headset can rome) and

Figure 4.2: A combination of photographs taken of the HTC Vive Pro Eye. Photographs A-C show different angles of the headset with A presenting a front view. Picture labeled B shows the head support at the back with the black circle to adjust the headset in length. C is an overhead view with the adjustable strap to help keep the display in position and alleviate some of the weight from the face. Lastly, image D, shows the black dashes around the lenses, which correspond to the eye-tracker.

track the position and orientation of the head-mounted display (see Figure 4.3). Inside of the headset, the screen is a Dual OLED 3.5” (diagonal) with a resolution of 1440 x 1600 pixels per eye (2880 x 1660 pixels combined) at a 90 Hz refresh rate and 110° field of view. The eye-tracking specifications make use of the HTC SRanipal SDK which includes a 120 Hz binocular output frequency, 0.5°-1.1° accuracy and trackable field of view at 110°. The head-mounted display weighed 800g and had a wired connection to a stationary PC that the virtual environment was run on using Unity Technologies (Unity3D, 2019).

There is a rapid development of alternative (to EEG) mobile brain imaging equipment, with the increased use and development of fNIRS and OPMs. Each of these imaging techniques provide its own benefit, for example, OPMs have the capacity to provide greater spatial resolution and recordings with reduced neck muscle interference (Roberts et al., 2019). However, at the current stage of technological

Figure 4.3: A photograph of the room setup, focusing on lighthouse positions that have been highlighted in red.

development, OPMs have reduced mobility and stronger interference when combined with a head-mounted display. Although the fNIRS hardware allows for mobility, the data acquisition and cleaning capabilities to deal with user movement are still underdeveloped (Abtahi et al., 2020). This would have provided a strong benefit to use fNIRS in movement-based Neuro-VR tasks. However, the technology for OPMs requires further development, and fNIRS compromises temporal resolution and is restricted to limited cortical activity (Wei et al., 2020). As a result, this study used EEG as it still provides excellent temporal and good spatial resolution for oscillatory signal analysis (Lopes da Silva, 2013). Studies have also found that the two modalities (EEG and VR) are compatible for combined usage, where signals <50 Hz have been shown to remain largely unaffected by their simultaneous operation (Hertweck et al., 2019; Weber et al., 2021).

It is now possible to combine the two, to create a novel Neuro-VR approach that adopts an enhanced naturalistic and immersive environment, in order to investigate neural signatures elicited in experimental tasks. Within neuroscience research, virtual reality is being applied effectively to facilitate the treatment of phobias (Botella, Fernández-Álvarez, Guillén, García-Palacios, & Baños, 2017), aiding with rehabilitation after various brain injuries (Laver, George, Thomas, Deutsch, & Crotty, 2012), and consolidating spatial navigation for patients with dementia (Tu, Spiers, Hodges, Piguet, & Hornberger, 2017). These techniques are still new and are continually being developed to better fit the target population. However, the current successes provide a basis to expand the use of Neuro-VR clinically and within psychological research.

This study combined an ANT Neuro 64-electrode EEG cap with the HTC Vive Pro Eye HMD and made use of the in-built eye-tracker. Within the virtual environment, 'colliders' were used to create regions of interest (Figure 4.5). In Unity, colliders are invisible but are physical 3D objects that can be attached to an object in the scene to define the shape. A script can then be assigned to the collider to detect collision from other objects. In this scene, a 'raycast' has been attached to the participant's gaze (see Figure 4.4) and the script attached to the colliders was responsible for recording the time of gaze entry and exit. A separate script was then used to record the participants' eye gaze throughout the experiment. A raycast in this case, is a physics function within Unity that casts a ray from its origin. In this example, the origin is the participant's eyes, which are traced to calculate their line of sight. Ray casting is used more widely within 3D modelling software and the computer graphics industry to determine the objects that are visible within a set trace and allow researchers to measure traces in 3D and dynamic scenes (Roth, 1982).

C# language was then used to code the location of the virtual human's eye-gaze behaviour in Visual Studio. For each participant, a JSON file was utilised to hold their individual and unique experimental sequence. This file included x, y and z coordinates of the positions that the virtual human would gaze toward. These positions were either (i) directly looking at the participant, (ii) gazing laterally and looking at either one of the puzzle boards or (iii) looking up at nothing or looking down at the puzzle piece. To execute this, I coded the virtual human's gaze to follow an invisible object. The coordinates in the JSON file (see Figure 4.6) were then used to direct the invisible object and in turn, guide the virtual human's gaze.

4.3.2 Virtual set-up

The virtual set-up displayed in the headset utilises the iteration of the gaze sequences from Experiment 2.1 detailed in Chapter 3. This included a reorganisation of the gaze sequence that was equally weighted in (i) the number of gaze shifts made, (ii) the time the informative gaze was performed, (iii) counterbalancing of gaze locations in Non-Collaborative conditions, (iv) the length of a trial and (v) computer-controlled gaze.

Figure 4.4: Two still images of the scene set up with the gaze raycast visible. On the left, the camera icon represents the position of the main camera where the scene begins, facing opposite the virtual human. The white line reflects the participant's head positions (measured by the headset) and the fushcia line, that is extended from the camera icon is the gaze raycast which represents the participant's eye-gaze position. This is updated in real time and is the object used to detect when the participants' gaze has entered one of the colliders that have been set up (see Figure 4.5). On the left hand side, the virtual human's puzzleboard collider records the time of entry and time of exit. On the right hand side, is a still image from the participant's point of view. The gaze raycast is focused on the black puzzle square and does not trigger a collider box. In the experiment, the visibility of the gaze raycast is turned off and participant's would not have seen the fushcia line in front of them.

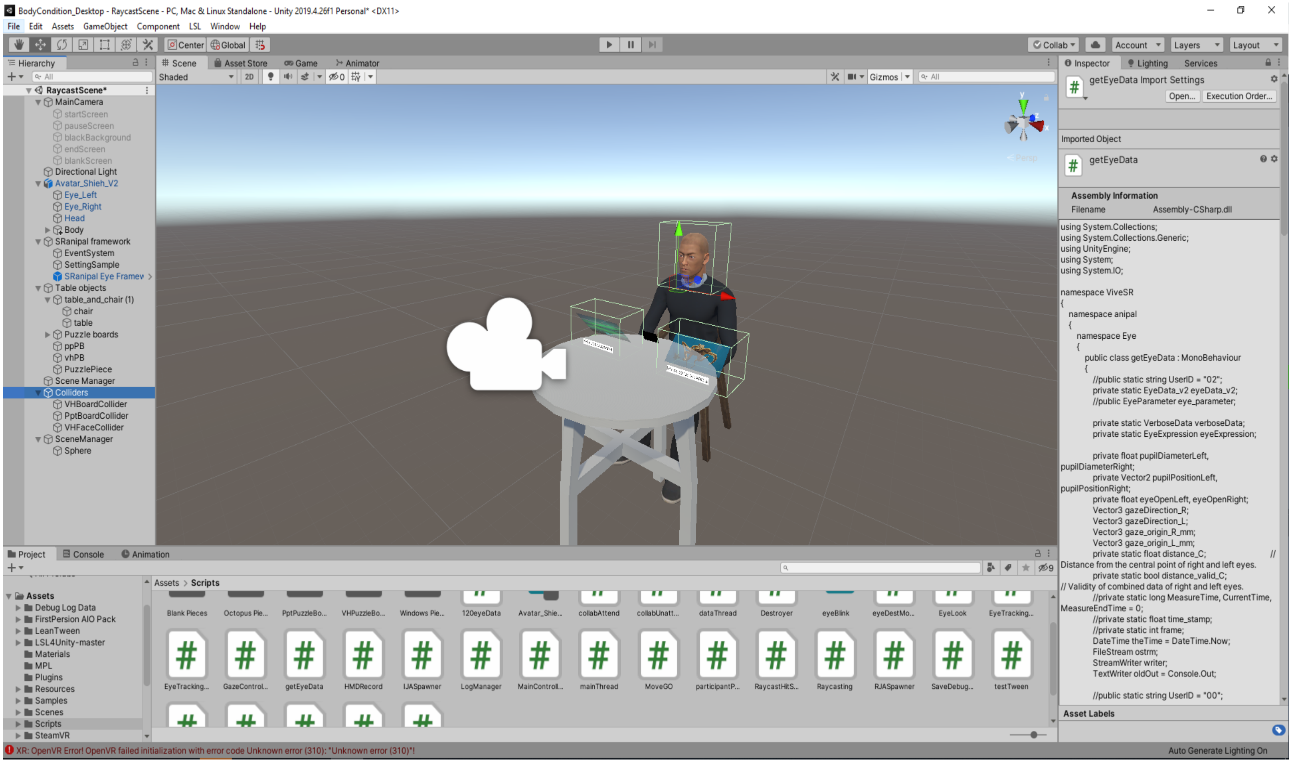

Figure 4.5: A still image of the colliders that were used as regions of interest for the participants' eye-tracking data. They have been made visible in the virtual environment (highlighted in green) for visualising purposes.

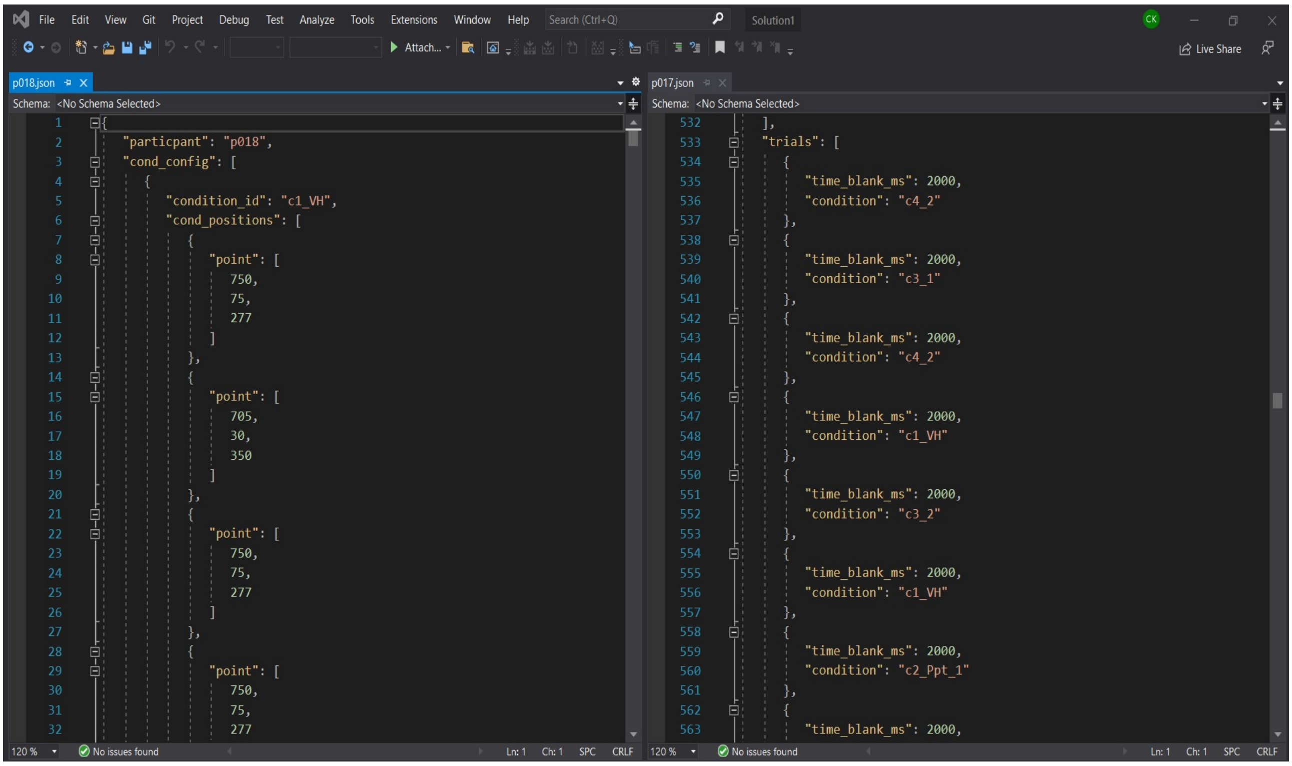

Figure 4.6: Here is an example line of Python code that populated the JSON file for each participant. Two JSON files are displayed in Visual Studio 2019 to demonstrate the two sections of each file. On the left, is the configuration for each condition which details the specific x,y and z coordinates for the gaze locations, And on the right, is the trial list that has been randomised for each participant.

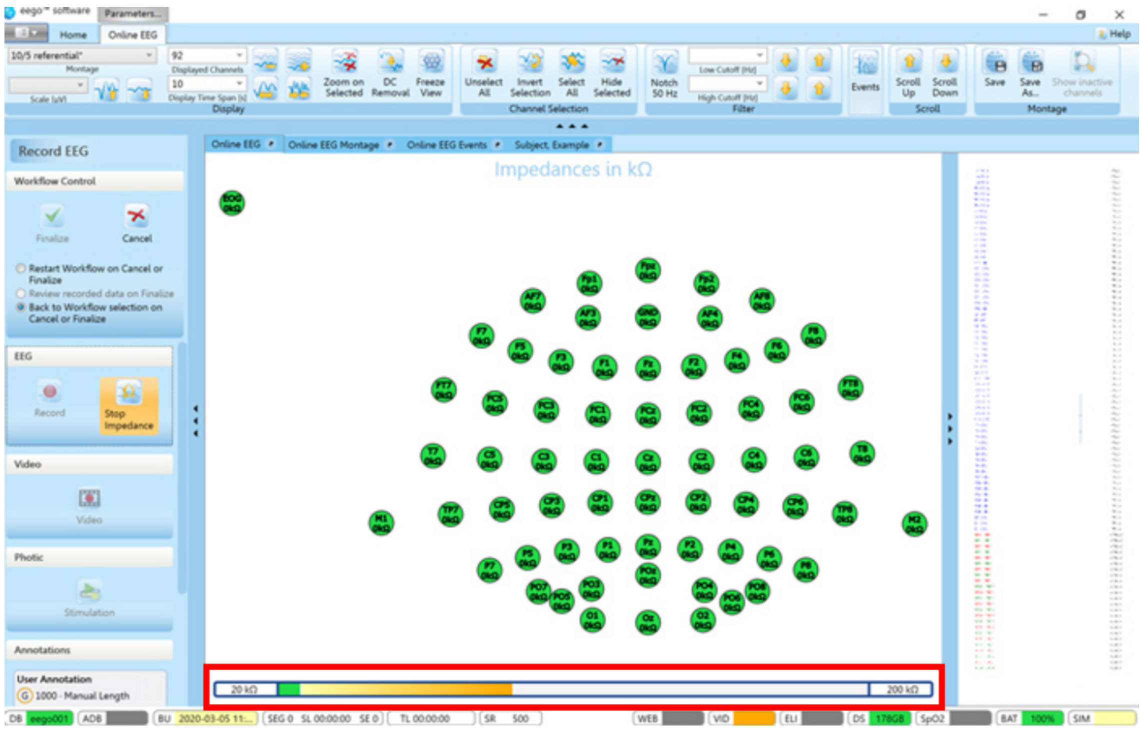

Figure 4.7: A view of the eego software with a virtual representation of each electrode in line with the 10-10 system. Each electrode varies in colour depending on the impedance level. Green, as seen in this image, denotes low impedances below 20Ω. The colour then increases (highlighted in red on the image) as the signal becomes progressively worse until it becomes white which represents no signal at all. **Image adapted from ANT Neuro website (Hengelo, The Netherlands).**

4.3.3 EEG measurement

The electrophysiological data were recorded while participants completed the joint attention task in the head-mounted display. A 64-channel ANT Neuro waveguard mobile EEG system (Hengelo, The Netherlands) recorded the data and digitised it using a sampling rate of 500 Hz. The EEG equipment consisted of the amplifier for data acquisition and a tablet to visualise the recorded data, the experimental markers (pictured in Figure 4.1) and impedance level (Figure 4.7. The individual placement of the electrodes were in line with a modern adaption of the International 10-20 convention (Jasper, 1958), the 10-10 system (Nuwer et al., 1998; Nuwer, 2018). Individual electrode impedances were adjusted below 20Ω prior to recording by injecting 'One Step Clear Gel' into each electrode (Figure 4.7). The data was recorded in an offline set-up, with CPz as the common reference and AFz as the grounded position.

Triggers were embedded within the Unity experimental setup. These triggers were then communicated to the ANT Neuro tablet using the 'Lab Streaming Layer' (LSL) plugin. LSL is an open-source package that can be used in conjunction with a range of software including Unity, PsychoPy and PsychToolbox.

The plugin allows you to sync your coded triggers with EEG recordings and also visualise them on the tablet. Throughout each trial, triggers were placed at set points in the EEG data. The visualisation that LSL permits is highlighted in Figure 4.1. These were (i) at the beginning of the experiment, (ii) when the screen turned grey after the trial sequence, (iii) at the end of the trial and (v) when the participant responded.

4.3.4 Participants

Participants were recruited through an advertised poster (Appendix B) where they were paid £10/ hour to participate in the study. The following inclusion criteria were used: Must be aged between 18 and 40, have normal-to-corrected to normal vision and be able to understand verbal and written information presented in English. Must not be sensitive to motion sickness or have a diagnosis of photosensitive epilepsy, hypertension or traumatic brain injury. In addition to the explanations given for participant exclusion in Chapter 2, a few extra criteria were added to reflect the incorporation of EEG and a head-mounted display. Firstly, I had to ensure that individuals with photosensitive epilepsy were not included as a result of the head-mounted display. Occasionally, disconnections can happen as a result of lost signal between the lighthouses and the headset (e.g. a person walking in front). This can cause the screen to flash black and quickly back to normal (as the connection is restored), potentially triggering an epileptic seizure. Also, a diagnosis of hypertension was included because, a sustained increased blood pressure level can result in significantly different EEG patterns. Lastly, participants may experience motion sickness or nausea from the virtual reality equipment. If they have had previous experience with motion sickness, it is possible that the equipment may induce this feeling again and therefore, are excluded.

In Experiment 3 there were 38 participants analysed, 21 were female and the mean age of all participants was 27 years. Participants age ranged between 18-40 years with a standard deviation of 6 years. All procedures complied with the Declaration of Helsinki and were approved by Aston University ethics committee (Appendix F).

A sample size of 34 was calculated from an a-priori power analysis (G*Power, (Faul, Erdfelder, Lang, & Buchner, 2007)) for the main effect of Collaboration. A small to medium effect size (r=.16) and a high correlation (0.7) were calculated from Experiment 1.1. Although this specific paradigm has not been tested with EEG before, other studies on eye contact have reported effects with less than the proposed number of participants (Kompatsiari et al., 2021). Based on this, I deemed the calculated sample size appropriate.

4.3.5 Data exclusion

For all data, participants were asked to fill out the short Autism Quotient Questionnaire (Baron-Cohen et al., 2001). Individuals who scored above 26 were excluded from the data (Ashwood et al., 2016). The cut off of 26 is one standard deviation above the mean score of neurotypical participants. This exclusion criteria was chosen from the questionnaire in order to only include neurotypical data within this study that represents a homogeneous sample of participants. Forty-three participants were collected and five participants were excluded from before the final analysis for scores higher than 26, leaving a total of thirty-eight participants.

For the behavioural data, trials were excluded if (i) there is no response, (ii) the recorded response time is $>=2SD$ and (iii) if the response is incorrect (this is only applicable on Collaborative trials where a correct answer is possible).

For some participants EEG data may not be usable due to experimental error or environmental factors; however, in these cases, the behavioural data will be used if it is still present after outlier rejection. All attempts were made to be as inclusive as possible for all data collected, though data cleaning procedures mentioned in Subsection 4.4.4.1, may result in further trials being excluded. In this experiment, no participants were removed as a result of noisy EEG data.

4.3.6 Procedure

After expressing interest through email, participants were sent an information sheet (Appendix A). Once happy with the information, a time was organised for the participants to visit the ALIVE labs, located in Aston University main building. On arrival, participants were verbally guided through the process by the experimenter (myself) and given the opportunity to re-read the information sheet, ask any questions and sign the physical consent form. It was emphasised that if they felt uncomfortable or did not want to continue at any point, the experiment would be stopped and this would not be prejudiced against them in any way.

The experiment then began with the AQ questionnaire (Baron-Cohen et al., 2001) (Appendix G) that was presented to them on an Apple iPad. Next, the circumference of their head was measured and the appropriate EEG cap size was placed on the participants' head. Depending on the participants' hair style or religious belief, they were asked to leave their hair naturally and/or remove any head dressings. This information was included in the information sheet sent to them and reiterated at the beginning of the experiment. The cap sizes ranged from small (47 - 51cm), medium (51 - 56cm) and large (56 - 61cm). Once fitted and participants were comfortable, the procedure for adding gel around each electrode was explained and again, they were told that if they felt uncomfortable at any point, to inform the experimenter. After adding gel to the first electrode as a demonstration, participants were asked if they were okay and happy to continue. The experimenter would also intermittently check and ask the participant if the pressure applied to their head from the EEG cap was okay and if they felt comfortable.

After adding the gel to the electrodes, the task was explained and participants were told: *"Thank you for agreeing to take part in the study. You will now be presented with a succession of tasks (trials) and they will go as follows: The virtual human will appear on the screen. Shortly after, the virtual human will be presented with a puzzle piece in front of them (this is only visible to them). You will see the back of the puzzle piece, which will look like a black square. The virtual human will perform an eye-gaze sequence that will suggest whether the puzzle piece belongs to your board, or theirs. They will try to be truthful but might not always get it clearly across. It is your task to decide which puzzle board you think they are suggesting. Place your dominant index finger on the '↑' key and your other index finger on the '↓' key, as was indicated on the print out of the instructions. When the screen goes grey, you will indicate whether you believe the puzzle piece is for yourself (using the '↑' key, pointing to you) or for the virtual human (using the '↓' key, pointing to them). You must do this as quickly and as accurately as possible. You will first go through some practice trials. Once you are comfortable, let the experimenter know and you will begin the main trials. Please press the space bar when you are ready."*

Again, the participant was offered the opportunity to ask any questions and confirm their understanding of the task. The researcher asked participants to remain as still as possible during the task, this was to reduce any electrical activity produced by their muscles and that might be picked up by the EEG electrodes. Examples of these artefacts were demonstrated to participants as they were asked to either move around or clench their jaw while viewing the EEG tablet. Participants then put on the HTC Vive Pro Eye head-mounted display with the help of the experimenter. With the headset fitted and comfortable, participants would first calibrate the software to their eye-gaze. This software is native to the headset and a part of the Vive Pro Eye system. Calibration was done for each participant and involved (i) adjusting the interpupillary distance (IPD), (ii) adjusting the position of the headset on the participant's face and (iii) calibrating the gaze positions. Directions for each of these adjustments were displayed on the screen to the participant (see figure 4.8) and help was provided by the experimenter.

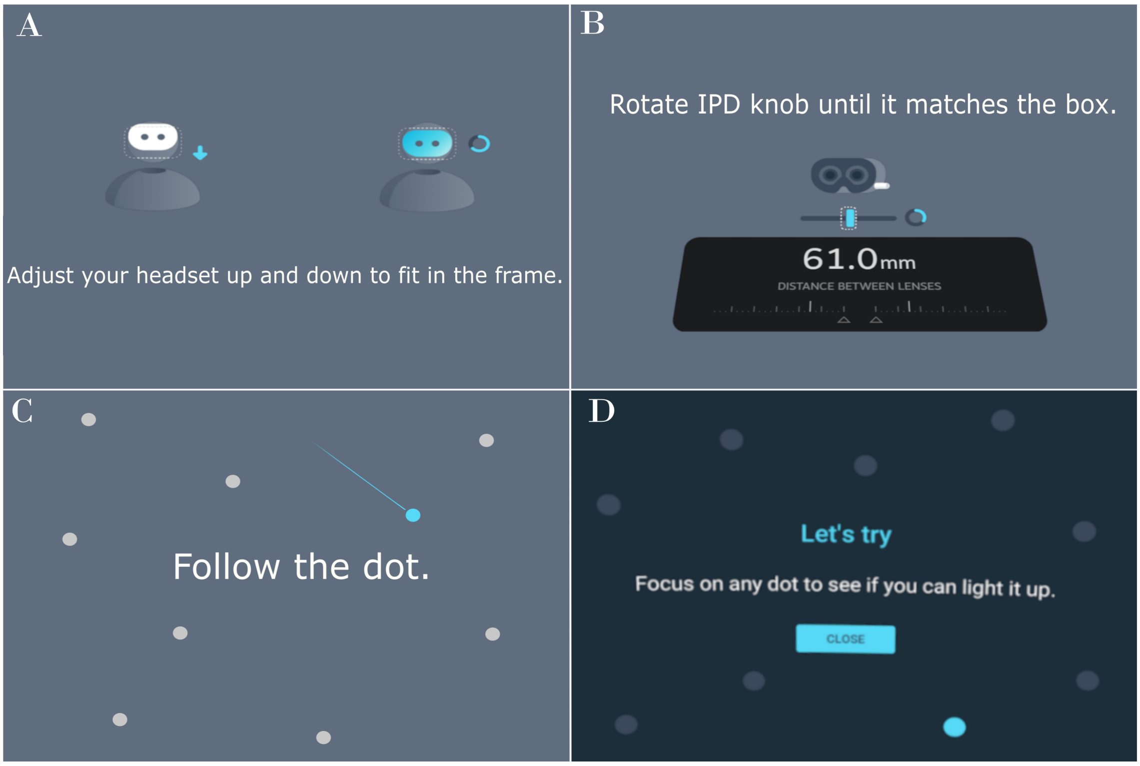

Figure 4.8: A figure of screenshots demonstrating the process of calibration within the head-mounted display. The image shows the sequences of tasks involved in calibration. Image **A**, shows the adjustment of the headset to ensure that it is positioned on the users head in line with their eyes. Image **B**, requires the user to adjust their IPD by turning the knob on the side of the headset until the solid blue line is within the white box. Next, image **C** shows the eye tracking calibration; the users are asked to follow the blue dot around the screen. Lastly, image **D** is displayed once calibration is completed, where users can check to see how well the calibration has worked by focusing on one of the dots to make it light up in blue.

Participants would then have the chance to practice the experiment. Once happy with the task that was required of them, they would begin the full experiment where eye-tracking and EEG would be recorded concurrently (see figure 4.1). The paradigm lasted around 20 minutes in total. Participants wore the headset for around 25 minutes and a maximum of 30 minutes. Upon the conclusion of the experiment, the experimenter would help the participant to remove the head-mounted display and EEG cap. Washing facilities were provided in the laboratory for participants to wash and dry their hair. Participants were thanked, provided with a gift voucher and ushered out of the building.

4.4 Results

4.4.1 Behavioural analysis

An ANOVA was conducted on response times (RT) over the two factors of Collaboration (Collaborative vs.Non-Collaborative) and Gaze Type (Eye-contact vs. No Eye-Contact). For the analysis of response times, I used all Non-Collaborative trials but only correct trials from the Collaborative condition. R Statistics software (R Core Team, 2021) was used to arrange the data into a suitable format. Then, the cleaning process was conducted in Matlab 2019b and involved removing trials that were either (i) incorrect or (ii) outliers, which were defined as responses that were faster than 200ms. The removed trial numbers were then saved and later used in the EEG cleaning process. Next, the descriptive statistics of each condition, for each participant, were calculated using the mean and standard deviation. The outliers for each participant were calculated using the minimum and maximum range by either subtracting the mean from 2 standard deviations or adding 2 standard deviations to the mean, respectively. After removing the outliers, the participant's mean for each condition was calculated again and compiled into a spreadsheet. This spreadsheet was then opened in JASP statistics where a repeated-measures ANOVA was run on the response times.

Due to the nature of the Non-Collaborative gaze sequences, the virtual human does not look at either puzzle board. The virtual human either looks up (toward the ceiling) or down (at the puzzle piece). At these central gaze points, the virtual human's gaze is equidistant from both boards. Therefore, in the Non-Collaborative condition, the *correct* answer is not defined and it is meaningless to derive an accuracy measure. However, accuracy was established in Collaborative gaze sequences, where the virtual human looked at one of the puzzle boards. The puzzle board that the virtual human gazed at was assigned as correct. A two-way t-test was then performed to determine the potential effect of Gaze Type (Eye versus NoEye) within the Collaborative condition.

4.4.2 Behavioural results

A repeated-measures ANOVA revealed a significant main effect for Collaboration, where participants responded significantly faster in Collaborative conditions compared to Non-Collaborative conditions($F(1,37)=27.40$, $p<.001$ ($d=.43$). Additionally, within Gaze Type, participants responded significantly faster during Eye conditions compared to NoEye conditions ($F(1,37)=12.91$, $p<.001$ ($d=.26$).

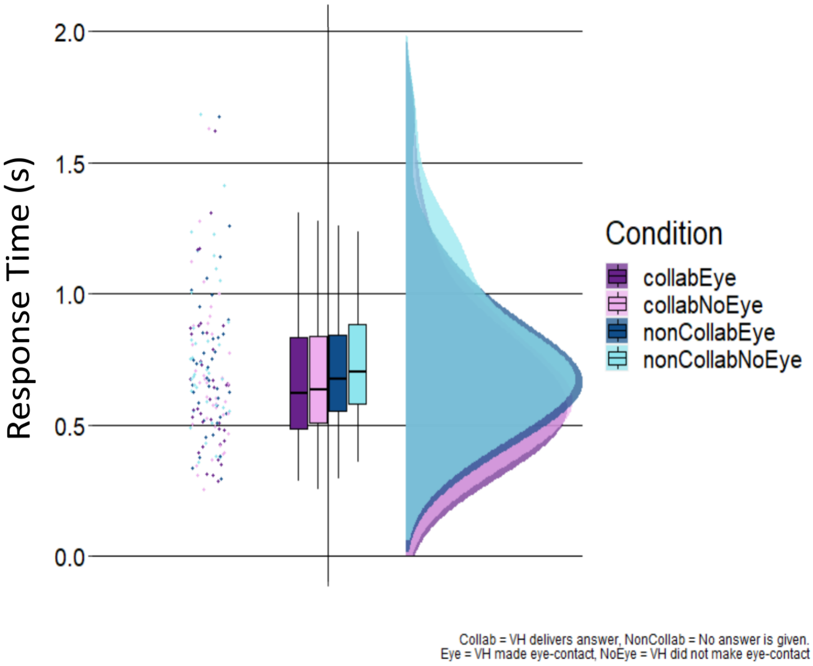

Figure 4.9: Response time data in milliseconds for Experiment 3. The graph shows each 2x2 factor (Collaboration vs. GazeType) with Collaborative conditions in purple and Non-Collaborative conditions in blue. For Gaze Type, Eye conditions can be seen in the darker shades and NoEye conditions in the lighter shades. The graph displays the following means for each factor: Collaborative, Eye (M=0.67, *SD*=0.29); Collaborative, NoEye (M=0.68, *SD*=0.28); Non-Collaborative, Eye (M=0.73, *SD*=0.27); Non-Collaborative, NoEye (M=0.78, *SD*=0.30).

Accuracy data revealed no significant difference between Collaborative, Eye at 95% correct and Collaborative, NoEye conditions also at 95% correct (Eye (M=37.87, *SD*=2.76), NoEye (M=37.84, *SD*=3.00); $t(38)=.09$, $p=0.93$, Cohens $d=.01$).

4.4.3 Eye-tracking analysis

The data analysed from the eye-tracking data included pre-specified regions of interest (ROI). These were (ROI 1). the head of the virtual human, (ROI 2). the left puzzle board and (ROI 3) the right puzzle board. I initially aimed to analyse the ROIs by (i) the total time looked at (ii) the ROI that was last looked at before responding and (iii) the number of times an ROI was looked at. However, I did not find enough variation in the eye-tracking data to follow through with these plans. Overall, I found that participants largely maintained their gaze on the virtual human's head (ROI 1). Consequently, I did not visualise the data nor run any statistical analysis on these results. The shortcomings of this tool and future suggestions are discussed in detail later in the general discussion (Chapter 5).

4.4.4 Electrophysiological results

4.4.4.1 Pre-processing of EEG

EEG data was read into Matlab2019b®, making use of the Fieldtrip (Oostenveld, Fries, Maris, & Schoffelen, 2011) and Zapline (de Cheveigné, 2020) toolboxes, which were used to preprocess the data. First, I applied Zapline to identify line signals from the head-mounted display at 50 Hz line noise and 90 Hz refresh rate (see Figure 4.10) and then used Fieldtrip to apply a band pass filter between 0.5 - 36.0 Hz (Weber et al., 2021).

The data was epoched for stimulus-locked analysis, the trigger that was used was presented after the virtual human's gaze sequence had finished and the grey screen for responding appeared (see Figure 4.11). Trials were visually inspected for artefacts and those that were large or noisy were removed. Independent component analysis (ICA) was used to identify eye-blink, eye movement and heartbeat components, as well as other known artefacts such as those originating from the muscles, which were then omitted from the data.

4.4.4.2 Time-frequency analysis

Time-frequency analysis was carried out by using Fieldtrip's multi-taper-method convolution using a single Hanning taper. This was applied to frequencies from 2-30 Hz (for every 1 Hz). Power was calculated at each frequency using a sliding time window, which was set at three times the wavelength of each frequency. This was calculated in steps of 50ms. The Hanning taper and sliding time window allowed for time and frequency power estimates within each trial. All of these parameters were set in response to the frequencies of interest for this task, which were 3 Hz and higher. In the event I was interested in frequencies lower than 2 Hz, parameters would need to be adjusted to take the slower wavelength into account.

For each participant, trials were averaged within each condition before comparing conditions statistically. Time-frequency representations were generated for the full time-frequency spectrum, while statistical analysis focused on alpha and theta bands separately. Alpha is defined as a neural oscillation in the frequency range of 8–12 Hz. Theta is defined as a neural oscillation in the frequency range of 3–7 Hz.

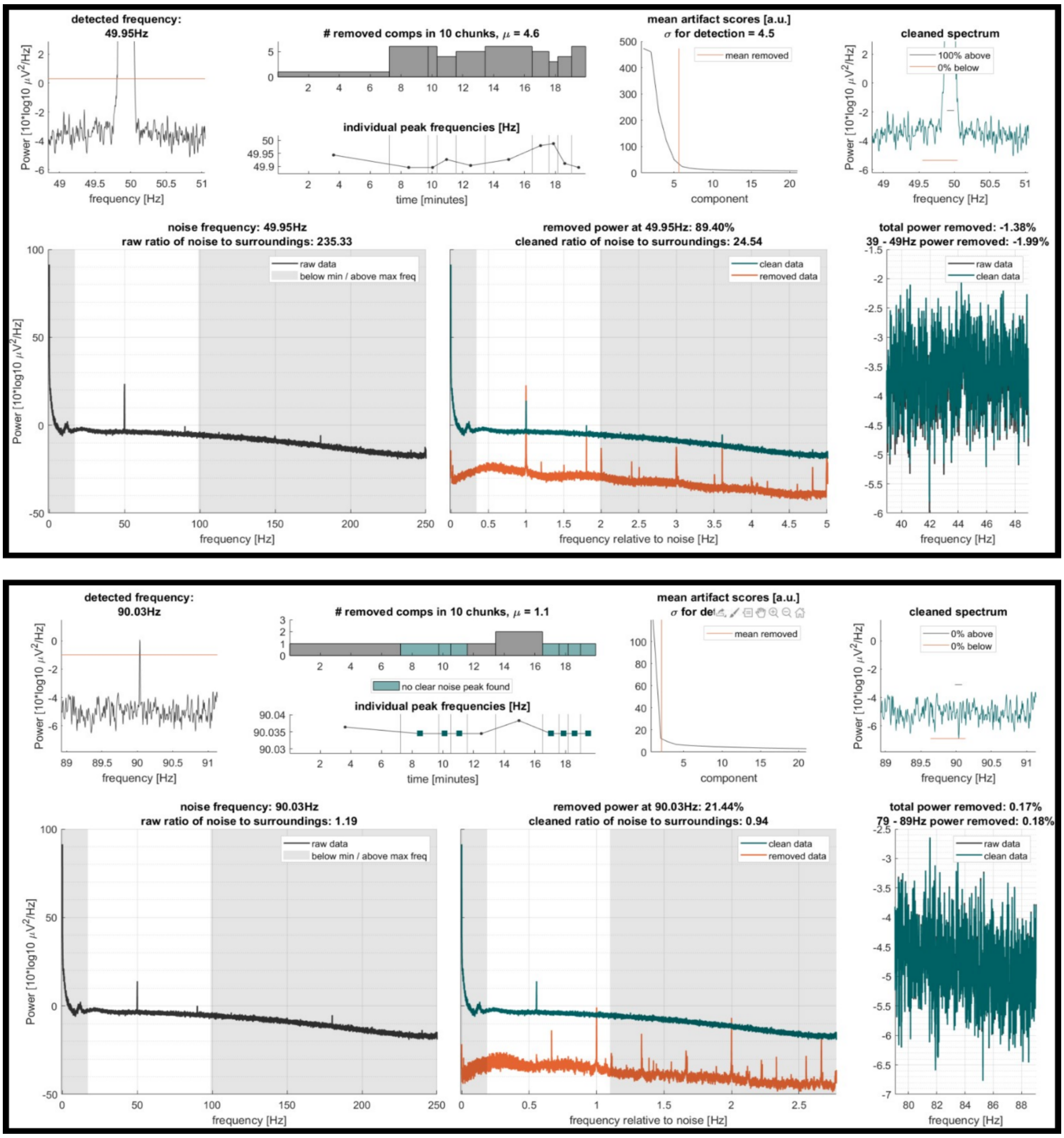

Figure 4.10: Two example outputs of noise removal (50Hz and 90Hz) using the Zapline toolbox.

Additionally, mu rhythms were explored and defined between 11-16 Hz. Spectral analysis was data-driven, which meant that there were no pre-selected time intervals or electrodes. Multiple comparisons were corrected using non-parametric cluster-based permutation tests. These were implemented in the Fieldtrip Toolbox (Oostenveld et al., 2011), with 5,000 permutations calculated (cluster alpha = $p<.05$). This robust statistics approach clusters the data in each of the frequency bands across time points and electrodes, enabling a comparison against our hypotheses.

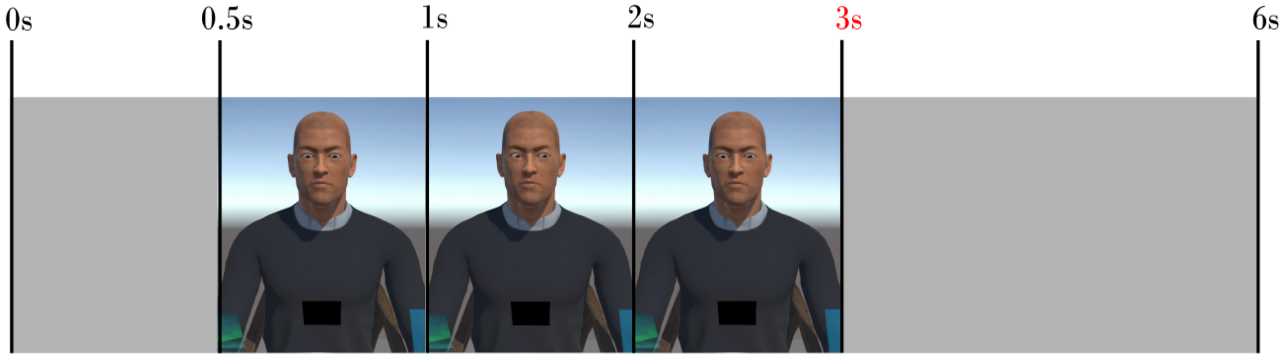

Figure 4.11: A schematic of an epoched segment in a Collaborative Eye condition. Timing is displayed in black, indicating the epoched time from start to finish (0 - 6s) and the chosen trigger (3s) which has been written in red.

4.4.4.2.1 Time-frequency representation

Figure 4.12 shows a time-frequency representation (TFR) over all of the conditions, in comparison to baseline (0-0.5s). The plots convey the overall activity for each of the frequency bands of interest. Both TFRs for anterior (A) and posterior (B) show an overall early increase in theta frequencies in comparison to baseline at ≈0.5-1s (red rectangle). In alpha frequencies, there is an overall decrease from ≈0.5s to 4s (pink rectangle) and in possible beta and mu rhythms (green rectangle) there is an early decrease in comparison to baseline (≈0.5s to 4s) and later increases during the response window (≈4s to 5s). Overall, the TFR plots demonstrate the expected activity from each of the frequency bands of interest. This helped to inform and understand the overall pattern of activity that had been observed in the statistical analysis. However, the data in the following cluster permutations has not been baseline corrected and the calculated comparisons are interpreted in relation to each of the conditions.

4.4.4.3 Cluster permutation results

4.4.4.3.1 Collaborative vs. Non-Collaborative

The second gaze position that the virtual human shifts their gaze to, indicates to the participant whether the virtual human is working collaboratively or non-collaboratively. Differences in stimulus-locked EEG were interpreted in the pre-response period, defined as 0.5-3 seconds. The large window was chosen as I was interested in the development and processing of the gaze that was presented in the condition. This is particularly useful for estimating changes in lower frequencies. In Figure 4.13, I

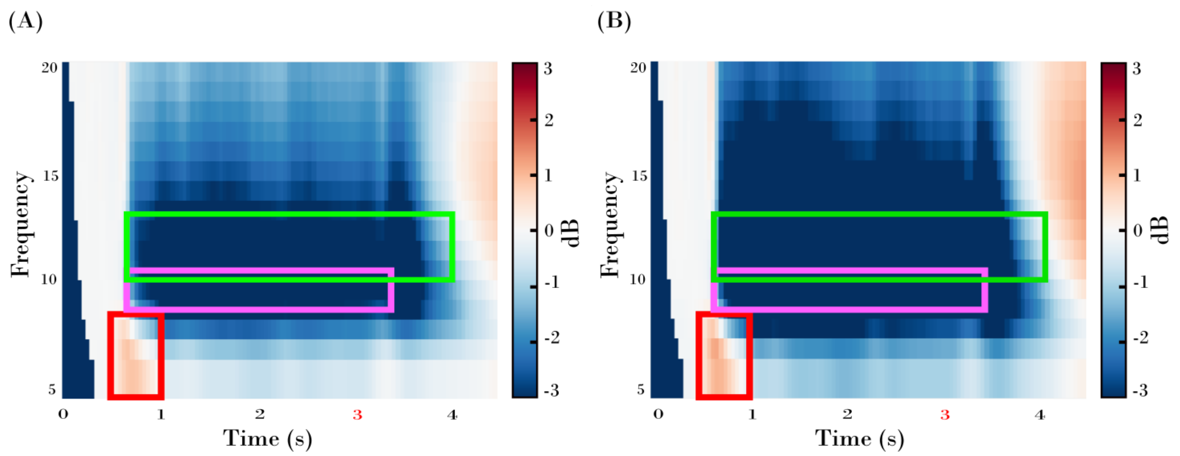

Figure 4.12: Two time-frequency representations (TFRs) are shown of averaged data across all of the conditions, in comparison to baseline. The TFR shows an epoch of electrophysiological activity from 0-4.5s which includes the beginning of the trial (0.5s), the end of the trial and beginning of response window (3s) and the average response time (3.71s). The baseline was set from 0-0.5s. On the left, TFR **A**, shows the averaged data over anterior electrodes (F3, Fz, F4, FC1, FC2, Cz, C4, F1, F2, FC3, FC4, C1, C2) and on the right, TFR **B**, the figure shows the averaged data over posterior electrodes (P7, P3, Pz, P4, P8, POz, O1, O2, P5, P1, P2, P6, PO5, PO3, PO5, PO6, PO7, PO8 Oz). Both TFRs show a red rectangle that highlights the theta increases, a pink rectangle to highlight the alpha decreases and a green rectangle, that highlights the beta/mu early decrease and later increase.

present results from the sensor-level analysis.

Figure 4.13 compares Mu (μ) rhythm signatures (11-16 Hz). The frequency range selected for mu partially overlaps with the range set for the alpha frequency. During Collaborative trials, I observed significant decreases in mu power. Early on in the trial, between 1 and 2 seconds, mu power desynchronisation begins over the left posterior and anterior areas (1.40s). This is soon followed by largely anterior patterns (1.60s) that eventually display bilateral, central power differences (1.80s). Later on in the trial, between 2.28 seconds (when the trial begins to end) and 3.04 seconds (when participants are required to answer), the neural signature changes. At this point, mu begins to display decreases in power in the Non-Collaborative trials over posterior and anterior areas (2.28-2.92s). During the response interval, these effects begin to localise over the left central electrodes.

The time-frequency analysis shows significant alpha decreases (8-12 Hz) in the Non-Collaborative compared to the Collaborative conditions. Alpha power desynchronisation begins over the posterior areas once the virtual human has appeared on the screen. Shortly after, when the Collaborative or Non-Collaborative gaze shift begins at 1 second, this significant effect in alpha power travels towards the anterior electrodes.

In theta frequencies (3-7 Hz), the figure displays a significant increase in power over right posterior

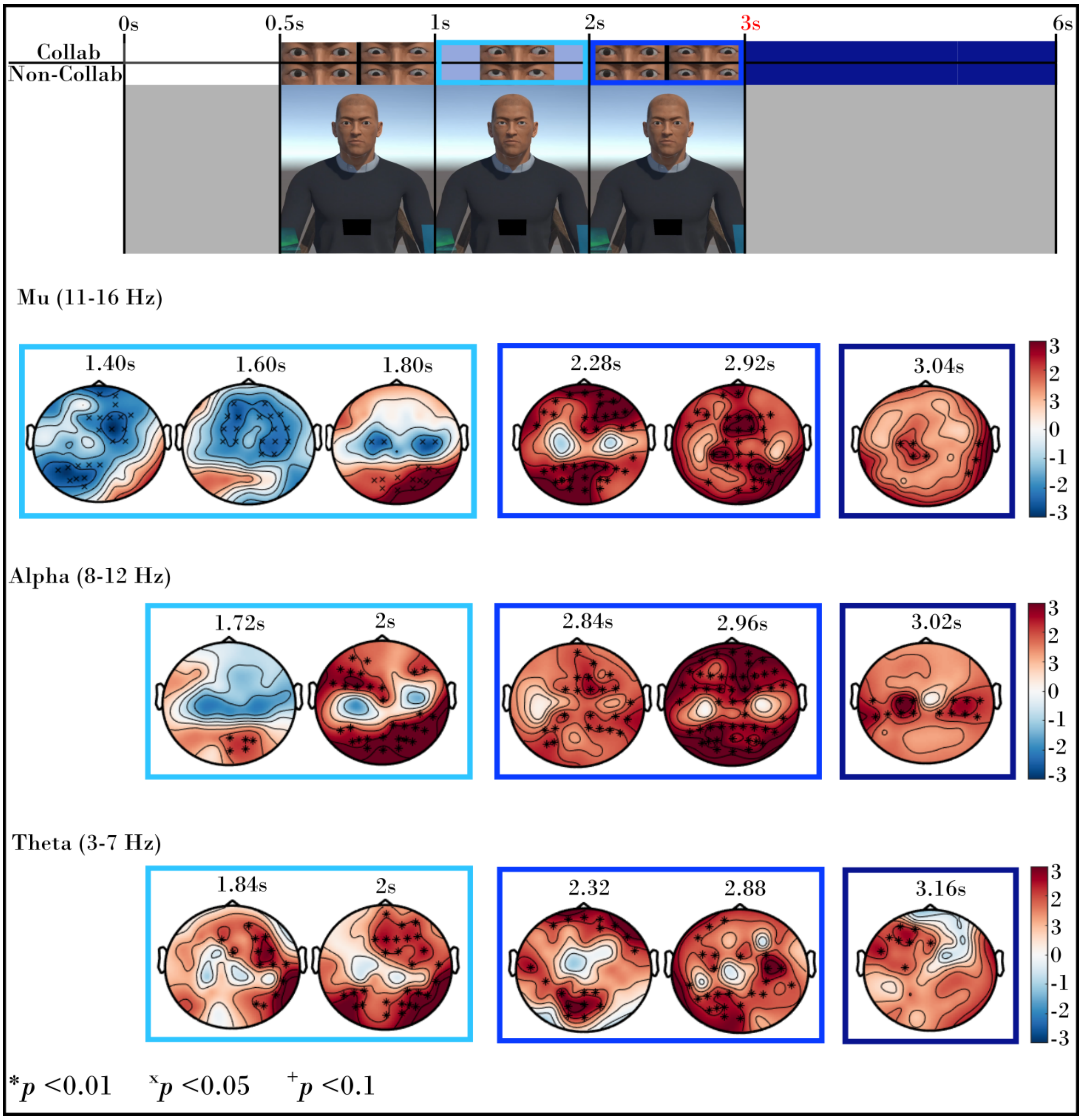

Figure 4.13: At the top of the figure is a trial schematic depicting the conditions being compared (Collaborative vs. Non-Collaborative), with examples of the gaze directions. Below the schematic, the electrophysiological differences between the comparison are displayed. These have been grouped by the frequency bands of interest and by time. All scales represent t-values from the cluster analyses.

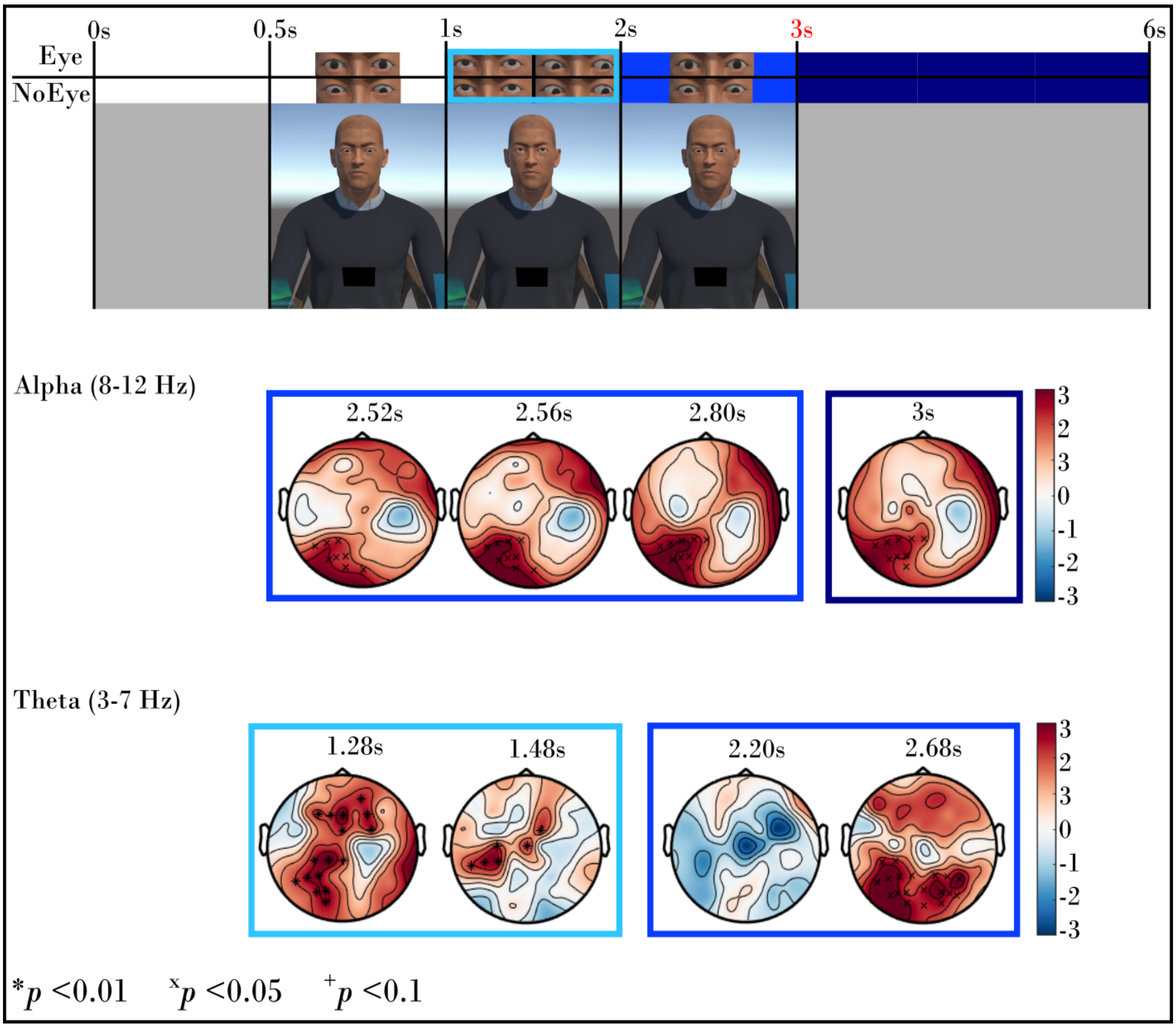

Figure 4.14: At the top of the figure is a trial schematic depicting the conditions being compared, with examples of the gaze directions. Below the schematic are the differences between EEG recorded in Eye and NoEye conditions. Comparisons for all frequency bands of interest are shown. All scales represent t-values from the cluster analyses.

and anterior areas during Collaborative trials. The significant increase in power over anterior areas appears at the end of the Collaborative or Non-Collaborative gaze shift between 1.84 and 2 seconds. Similar theta patterns were also observed in exploratory simple effects (Eye, Collaborative vs. Eye, Non-Collaborative), which are discussed in Subsection 4.4.4.4.3.

4.4.4.3.2 Eye vs. NoEye

A comparison of pre-response activity (0.5-3s) was analysed between Eye and NoEye conditions. The results (presented in Figure 4.14) revealed significant alpha desynchronisation in NoEye conditions. These clusters were present across the left posterior electrodes towards the end of the trial and before the response window between 2.52 and 2.80 seconds. No significant clusters were found in the mu frequency.

Similarly, theta frequencies were analysed in the pre-response period. Within this time, positive clusters were found for comparisons between Eye and NoEye conditions. These were significant theta-power increases that were present over the left posterior electrodes during the shift to the second gaze position (1.28-1.48s). Following the second gaze position (2.20s), theta-power increases in Eye conditions were observed but were not significant. Towards the end of the trial period (2.68s), comparisons revealed significant, theta synchronisation over anterior electrodes.

4.4.4.4 Exploratory analysis

The analysis for each of the main effects did not provide a clear understanding of which factor levels might be driving the observed significant clusters which could be of importance even in the absence of a significant behavioural interaction between the two factors. EEG offers insights into how mental processes unfold precisely in time, rather than just measuring the behavioural outcome at the very end. Given that the respective factor manipulations occurred in distinct trial periods (Gaze Type 0.5-1s and 2-3s vs Collaboration 1-2s) this rationale is particularly warranted here. For this reason, I conducted an exploratory analysis to compare the simple effects (Collaborative, Eye vs. Collaborative NoEye; Non-Collaborative, Eye vs. Non-Collaborative, NoEye; Eye, Collaborative vs. Eye, Non-Collaborative; NoEye, Collaborative vs. NoEye, Non-Collaborative) to gain a better understanding of the factor combinations that may primarily drive the particular neural signatures that have been observed.

Specifically, I did not have any predictions about the role of theta in Eye and NoEye conditions. However, I wanted to review whether these differences were present and if so, what role theta appears to have had in this scenario.

Equally, the only predictions regarding theta were made for Collaborative and Non-Collaborative comparisons. However, I wanted to analyse whether differences in other frequencies, in this case, alpha and mu, were dependent on the collaboration of the virtual human. These results are discussed below.

4.4.4.4.1 Collaborative, Eye vs. Collaborative, NoEye

Theta clusters were observed over anterior and posterior areas. Significant theta increases were present at the end of the trial, where the virtual human either engaged in eye contact or not. Patterns are in

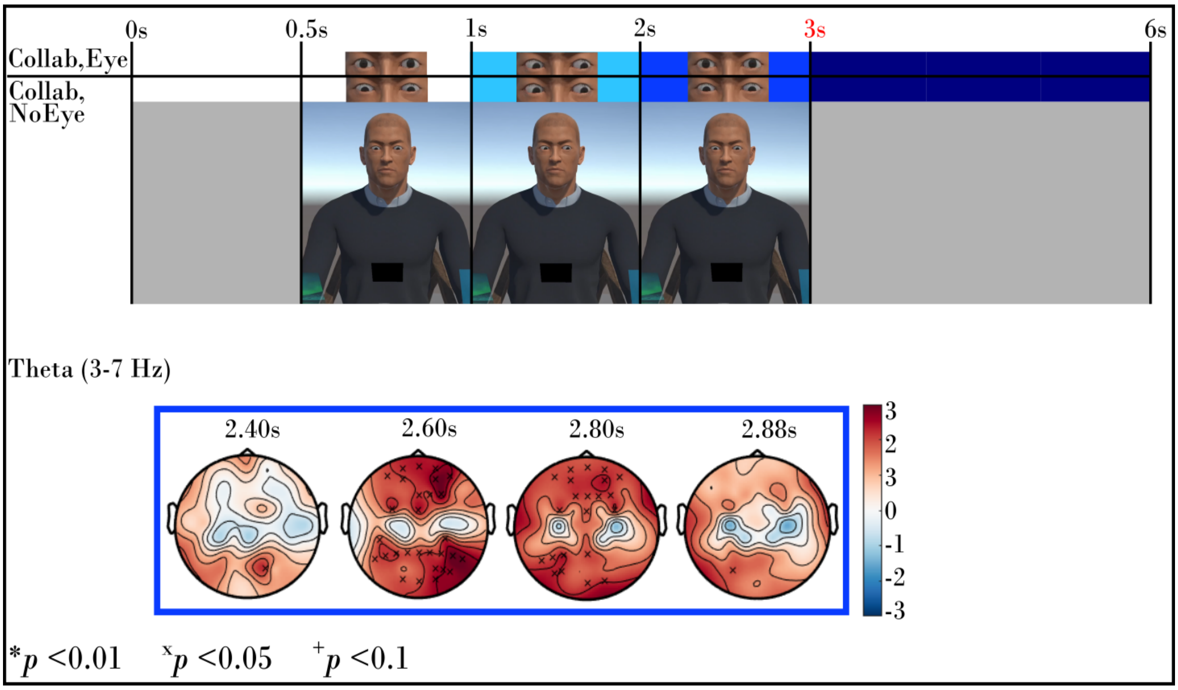

Figure 4.15: The figure displays a trial schematic of a simple effect (Collaborative, Eye and Collaborative, NoEye). Examples of the gaze directions for each condition are shown. Below the schematic are the differences between EEG recorded in both conditions. Differences in theta are shown and no clusters were found in alpha frequencies. All scales represent t-values from the cluster analyses.

line with effects seen in the main effect of Collaboration.

No significant alpha clusters were present in the data.

4.4.4.4.2 Non-Collaborative, Eye vs. Non-Collaborative, NoEye

Analysis of the Non-Collaborative, Eye vs. Non-Collaborative, NoEye simple effect, shown in Figure 4.16 resulted in theta synchronisation over the left posterior and central anterior electrodes at 1.28s. These patterns are similar to what is seen in the main effect of Gaze Type.

No significant alpha clusters were present in the data.

4.4.4.4.3 Eye, Collaborative vs. Eye, Non-Collaborative

In Eye, Non-Collaborative conditions, Figure 4.17 shows significantly stronger decreases in alpha towards the end of the informative gaze shift (1-2s). Here, significant clusters are shown over both anterior and posterior areas. These clusters continue until roughly halfway through the last gaze shift when the number of present clusters in the data is reduced (2.68s).

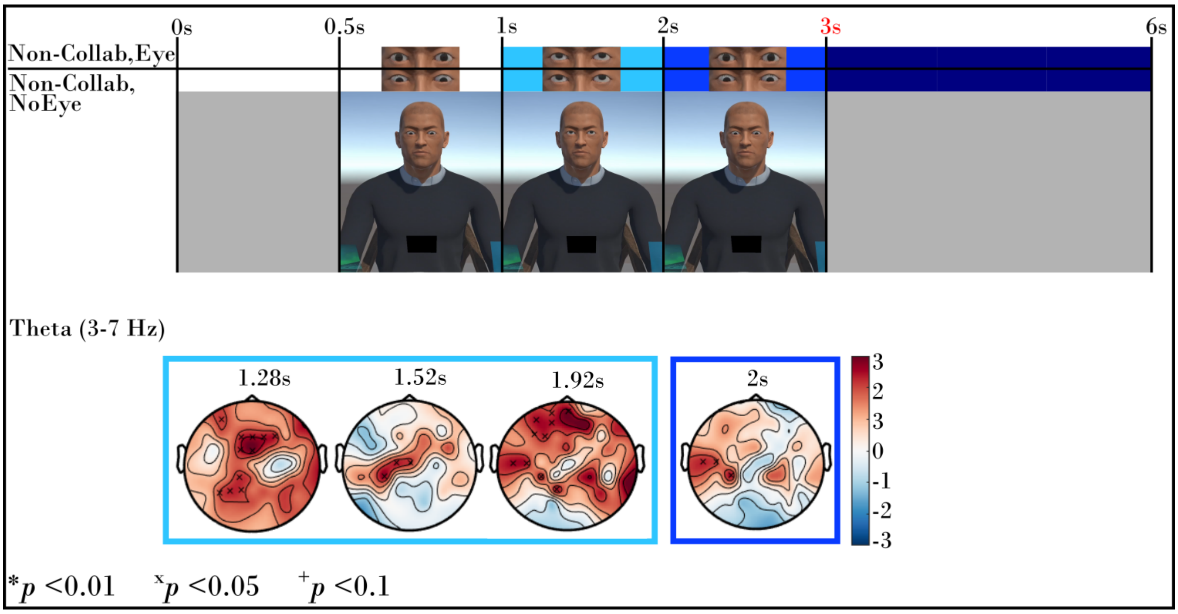

Figure 4.16: At the top of the figure is a trial schematic depicting the conditions being compared, with examples of the gaze directions. This figure displays a simple comparison between Non-Collaborative, Eye and Non-Collaborative, No Eye conditions in theta frequencies. Below the schematic, the electrophysiological differences between the comparison are displayed. All scales represent t-values from the cluster analyses.

In theta frequencies, significant clusters also begin towards the end of the informative gaze shift (1-2s). The comparison revealed theta increases over anterior and posterior electrodes (1.96-2.12s) and later just anterior electrodes (2.52s).

4.4.4.4.4 NoEye, Collaborative vs. NoEye, Non-Collaborative

Lastly, analysis was performed on the simple comparison of NoEye, Collaboration and NoEye, Non-Collaboration. Here, significant alpha desynchronisation is present from 2 seconds onwards. The clusters begin over posterior electrodes and later incorporate anterior areas towards the end of the trial and just before the participants are required to respond.

No significant theta clusters were present in the data.

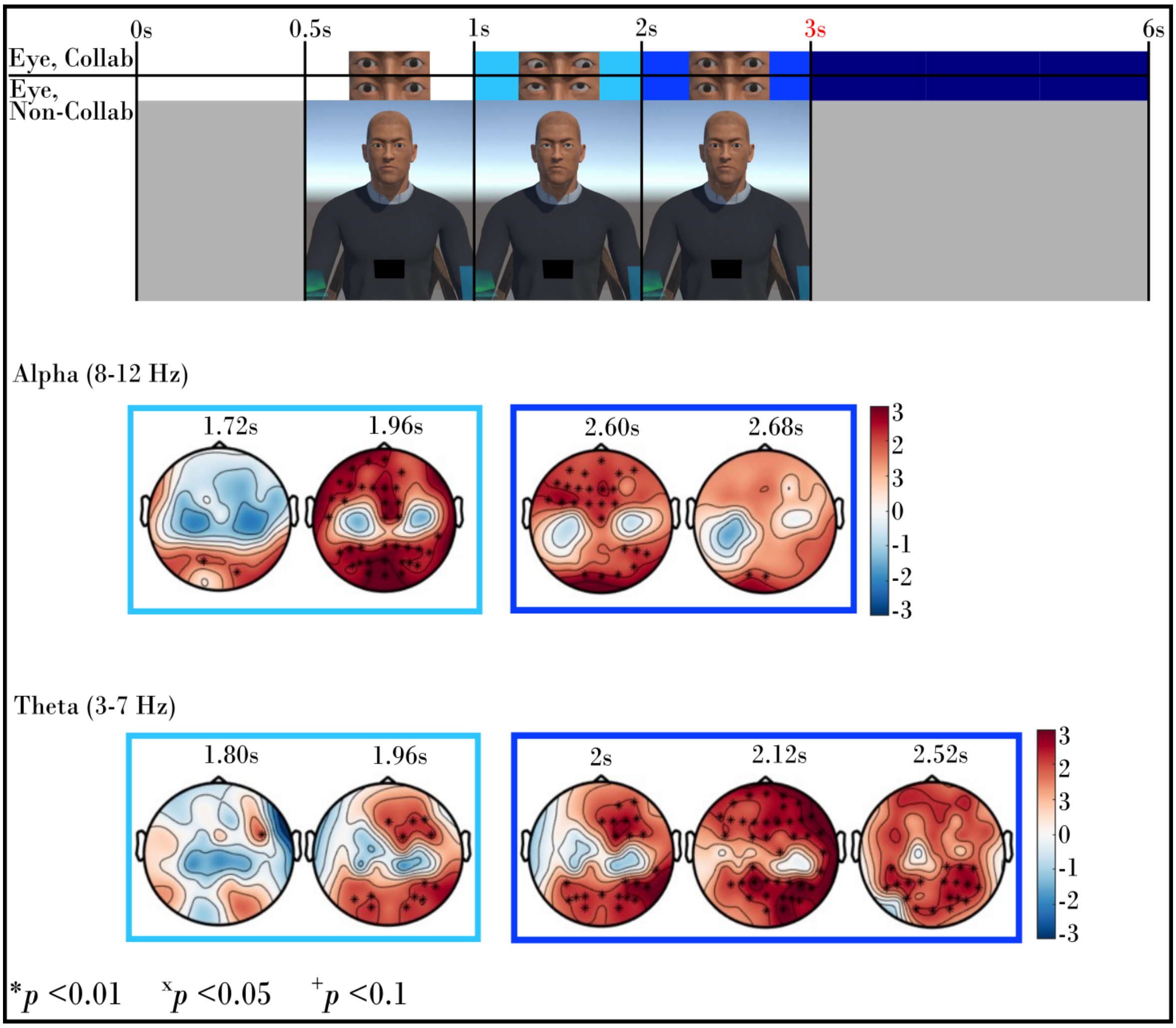

Figure 4.17: At the top of the figure is a trial schematic depicting the conditions being compared (Eye, Collaborative vs. Eye, Non-Collaborative), with examples of the gaze directions. Below the schematic, the electrophysiological differences between the comparison are displayed. These have been grouped by the frequency bands of interest and by time. All scales represent t-values from the cluster analyses.

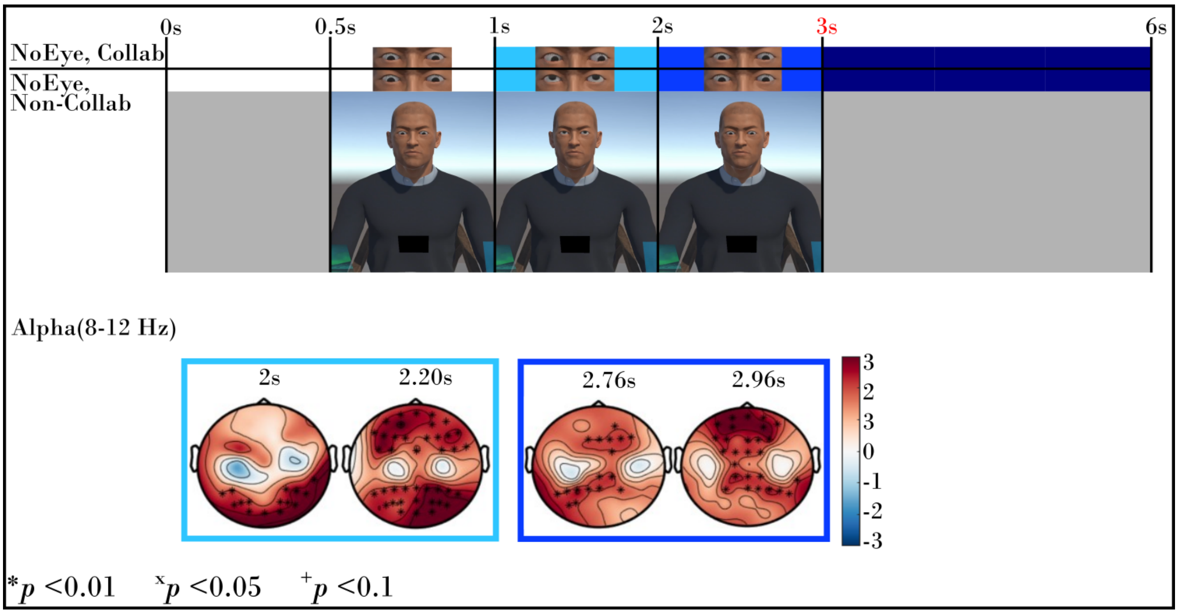

Figure 4.18: The figure displays a trial schematic of a simple effect (NoEye, Collaborative vs. NoEye, Non-Collaborative). Examples of the gaze locations for each condition are shown. Below the schematic are the differences between EEG recorded in both conditions. Differences in alpha are shown and no clusters were found in theta frequencies. All scales represent t-values from the cluster analyses.

4.5 Discussion

4.5.1 Summary of findings

This study investigated the role of eye-gaze on joint attention within an immersive virtual environment. Both behavioural and electrophysiological data were recorded and presented interesting results. The experiment provided a comparison between using a head-mounted display and using a 2D computer screen for stimuli presentation. Overall, the behavioural results revealed an effect of both Collaboration and Gaze Type where Collaborative and Eye conditions were significantly faster, within their respective factors. These results demonstrated the impact of a head-mounted display and therefore a difference between the 2D and 3D presentation of the experiment. In the electrophysiological data, it can be argued that the neurological response implied that Collaboration effects were largely regulated in alpha frequencies, whereas theta was dominant in the differences for Gaze Type. However, these are speculative proposals and would need further investigation to confirm.

4.5.2 Behavioural summary

The behavioural results showed main effects for both Collaboration and Gaze Type. The main effect of Collaboration was found throughout the reported series of experiments (Experiments 1.1-2.2), so it was expected that this effect would be replicated in Experiment 3. However, a significant main effect of Gaze Type was not consistently found throughout but, was found in Experiment 2.1. Similar to the effect of Gaze Type in Experiment 2.1, Experiment 3 found significantly faster response times when the virtual human engaged in eye contact with the participant. Taking both of these results together, I suggest that engaging in eye contact during joint attention may provide a social influence that facilitates the process of joint attention.

When replicating Experiment 2.1 in this study with the addition of a virtual reality head-mounted display, the effect size of both main effects increased in comparison to Experiment 2.1. There was also a decrease in the probability that the effect of Gaze Type was due to random chance. Thus, it is possible that being fully immersed in a virtual reality environment amplifies the effects observed in relation to 2D presentations in online experiments. However, differences between lab-based and home-based experiments may also have an effect.

In addition to the behavioural effects, my argument is also formed on anecdotal evidence from verbal accounts of the participant's experiences. For some, it was their first time using a virtual reality headset and this may have had an effect on their feeling of presence in comparison to a regular user. A post-experiment questionnaire, such as Whitmer and Singers 'Presence Questionnaire' (Witmer & Singer, 1998), should be incorporated into future experiments to objectively measure the participants' presence when using either display. In this study, a measure of presence would have allowed me to explore, interpret and sharpen my understanding of the results. Altogether, I suggest that head-mounted displays, in comparison to 2D computer screens, can be advantageous when observing social behaviours and should be considered in certain cases.

4.5.3 Electrophysiological summary

4.5.3.1 The effect of Collaboration

I had hypothesised that the effect of Collaboration would show theta synchronisation in Collaborative conditions. This was proposed to reflect the increased integration that is a result of the informative gaze shift that is conveyed during these trials. Overall, the results did reveal significantly stronger theta signatures in Collaborative conditions and supported hypothesis H4. However, the topographies observed were similar to the signature patterns and timings that were found in the exploratory analysis of simple effects. As a result of the simple comparisons in Subsection 4.4.4.4.3 I suggest that the theta signatures observed in the main effect of Collaboration may be driven by Eye conditions within the effect, however, this requires further investigation.

Additionally, significant alpha power decreases were observed over posterior areas and then later over anterior areas. This is likely due to alpha power decreases in Non-Collaborative trials which have been related to increased attention (for reviews see (Ward, 2003; Foxe & Snyder, 2011). In Non-Collaborative conditions, participants' attention is heightened once they become aware that the virtual human is not working collaboratively. I propose that this may be an attempt to gain information from the scene that may help them to make a decision. Thus, the alpha power decrease is more pronounced compared to the Collaborative conditions, in which participants are given a clear cue towards the puzzle piece and are possibly preparing to make their response. As a result, I suggest that less attention is administered overall during Collaborative conditions as the gaze direction provides a clear cue for participants to attend to.

Mu rhythms were also investigated as a result of the neural signatures observed in the alpha frequencies. Although this data is exploratory, I will discuss its implications in relation to the comparison of Collaboration to further the overall understanding. During Collaborative conditions, significant differences were revealed in the mu rhythm however, the frequency range is still the subject of debate (Pfurtscheller et al., 2000; Pineda, 2005). In this study, mu and alpha rhythms were found to have overlapping frequencies. As a result, the rhythms have been separated based on their characteristic topographies.

In this experiment, mu clusters exhibit activity differences bilaterally and over central electrodes, which is a distinct pattern typically seen in studies investigating in preparation and execution of voluntary limb movement (Pfurtscheller et al., 2000). Mu power decreases are commonly associated with motor preparation, particularly the movement of the hand (Pfurtscheller et al., 2006). In Study 3, stronger mu power decreases are initially observed in the Collaborative conditions, during the trial period when the virtual human provides the informative gaze cue; thus, suggesting that action preparation can commence in these trials. Later on, at 3.04s (after the trial has finished and participants can respond) positive mu signatures are present over similar electrodes, showing a stronger mu power decrease in

Non-Collaborative conditions when the participant must respond. I suggest that the early mu power mirrors participants' motor preparation in Collaborative trials. These later mu power effects may reflect the delayed motor preparation in Non-Collaborative conditions. This is reflected in the behavioural data that reports significantly faster response times in the Collaborative condition.

A possible alternative, but not mutually exclusive, explanation of the early effects involves beta rhythm (13-18 Hz) and may suggest an effect of attention. The early topographies in Figure 4.13 are loosely consistent with the beta coupling network reported by Gross et al (2004) which could imply an attention-based explanation for the early beta/mu frequency effects at 1.4s and 1.6s rather than a very early motor preparation explanation. These are the earliest effects that have been observed in the EEG data and could indicate that the Collaborative gaze shift (1-2s) draws on the participants' initial attention compared to Non-Collaborative trials. Then, later effects at 1.80s localises over possible sensorimotor areas, indicating motor preparation that is coinciding in time with alpha attention and theta integration effects. This explanation could also be carried across to the later, positive power differences seen between 2.28-3.04s. The earlier positive signatures (2.28-2.92s) also follow a rough topography in line with the beta coupling network pattern and therefore possibly reflect attention. Similarly, later positive effects (3.04s) followed a bilateral, central pattern possibly representing later motor preparation in Non-Collaborative conditions.

4.5.3.2 The effect of Gaze Type

The hypothesis formed around the effect of Gaze Type predicted that Eye conditions would show a stronger alpha power decrease over left fronto-central and central electrodes in line with Kompatsiari et al (2021). This was proposed as a result of the associated pattern of alpha decreases when engaging in joint attention.

However, in the Gaze Type comparison, significant alpha power decreases were observed over posterior electrodes only. Anterior electrodes did not display any difference in alpha power. While this was unexpected, it is a possibility that the Collaboration effects dominated this frequency and overshadowed any Gaze Type effects. Alternatively, these effects could be interpreted as participants finding NoEye conditions more difficult and therefore, paying more attention to decipher secondary information in the scene.

Similar to the theta signatures observed in the main effect of Collaboration, theta patterns in Gaze

Type present significant increases over posterior electrodes. However, in Gaze Type, these patterns begin (1.28s) over the left hemisphere and over the central anterior electrodes. During this time point, participants are being presented with either a Collaborative or a Non-Collaborative trial. However, before this participants would have just engaged in direct eye contact with the virtual human. I believe this early effect may be driven by the initial eye contact made at the beginning of the trial and not the following gaze shift (irrespective of indication). This would be supported by literature that suggests gaze may facilitate integration in theta frequencies (Gregory et al., 2022). In the analyses of simple effects (Non-Collaborative, Eye vs. Non-Collaborative, NoEye) a similar pattern of posterior and anterior theta is observed. The simple comparison showed a potential dominant effect of Eye conditions in theta oscillations. This is in line with reports of eye-gaze (and its default social effect) being processed differently from non-social cues (Kampis & Southgate, 2020). I propose that the theta effects observed in Figure 4.14 may reflect this proposed alternative processing as a result of the rich and more socially complex information associated with eye-gaze.

4.5.3.3 Simple effects

The aim of analysing the simple effects was to explore the data to better understand the main effects. Overall, these comparisons illustrated a potential driving effect of Eye conditions. This is seen throughout the theta signatures in each of the comparisons that include Eye conditions (see theta topographies in Figures 4.13, 4.14, 4.15, 4.16 and 4.17). Additionally, in Non-Collaborative simple effects, theta clusters occur during the trial period where Gaze Type does not differ (1-2s). At this point, the virtual human is looking up or down (Non-Collaborative). This seems to indicate that there are different expectations based on the initial eye contact or lack thereof. The effect implies that eye contact may operate as a potential trigger to participants that information is going to be communicated, thus, expectations to that effect, which on Non-Collaborative trials remain unfulfilled. This is also in line with the lack of significant theta clusters in the NoEye simple comparison (Figure 4.17).

In addition, Non-Collaborative conditions appear to have an effect on alpha power. In these comparisons, alpha power decreases are stronger over both anterior and posterior areas. Similar to alpha signatures in the main effect of Collaboration, these signatures begin at later time points, after the Non-Collaborative gaze shift has completed (2s onwards). This is particularly interesting for NoEye, Non-Collaborative conditions. In these scenarios, participants are not engaged in eye contact with the virtual human and are given no informative gaze cues to direct them to the correct puzzle board. As a result of this combination, I stipulate that the participant's attention is heightened in order to gain information to make a decision (as discussed in the main effect of Collaboration). This is observed through the late and stronger alpha power desynchronisation in NoEye, Non-Collaborative conditions.

4.5.4 Future work

The study used the revised paradigm in Experiment 2.1 to investigate the role of eye-gaze in joint attention within an immersive virtual environment. In itself, this study has demonstrated the feasibility of using an immersive virtual reality paradigm alongside EEG recordings. The virtual reality head-mounted display provided the participants with an increased level of immersion that I had proposed would increase the ecological validity of the experiment. Additionally, it also provided a homogeneous environment where both agents can be present in the virtual world. Using a virtual world allowed the expectations of the virtual human's behaviour to be appropriate to what a member of the general public might assume. For example, I propose that if the virtual human is non-verbal, this behaviour would be accepted and within the participants' expectations. This would be less likely if a human confederate was being used as an agent. Consequently, I believe the limited behaviour of the virtual human may be better received in comparison to that same behaviour in a human being. Having said that, I do believe there are improvements, mainly focused on animation that would enhance the participant's experience when interacting with a virtual human. The effects of animation and other suggested improvements are discussed in further detail in Chapter 5.

Crucially, participants verbally reported that the virtual human was a little bit scary or intense. Some of this was feedback that was presented to me at conferences and during pilot testing. Future work may consider an alternative virtual human that is rated as more 'friendly looking'. This could also be used as a manipulation to investigate virtual humans e.g. age-matched, gender-matched etc. As discussed at the end of Study 2, there are many improvements that can be made to the virtual human pertaining to its aesthetic and animation. In particular, improvements could be made concerning the human-likeness of the animations or lack thereof. In these experiments, the animation was kept very simple and only concerned with the gaze shifts of the virtual human. This kept manipulations and aesthetic changes to a minimum to ensure a high control of validity was achieved. While keeping the animations simplistic allowed these advantages, it is my intuition that the participant's presence in the immersed environment would greatly improve. Specific improvements are explored in Chapter 5.

This study provides further support for the use of Neuro-VR, particularly in social-based experiments e.g. (Gregory et al., 2022). The pairing of EEG and virtual reality could provide relevant information regarding the brain signatures associated with specific behaviours. This can also call attention to the impact that immersive virtual reality may or may not have in these investigations. Long term, these tools and techniques could form part of a toolbox that works to support individuals who have difficulties with social communication. An immersive virtual environment that can be used to provide training through play or gamified therapy is a natural extension for this type of research. This speaks to a number of fields that have already found benefits, for example, stroke therapy (Prof Nick Ward UCL) and dementia (Hugo Spiers, UCL). There are many more areas of research that would benefit from an immersive environment and with training or therapy through play, for example, ASD (Mei et al., 2018) and post-natal depression. Developing a global understanding of the technology's utility would present opportunities to other fields and researchers to further our knowledge of its benefits.

To conclude, I propose that the use of virtual reality when exploring social paradigms may be a critical tool in social neuroscience and the data obtained from this study could be beneficial for future work. With the adoption of this tool in our research, we can better understand these phenomena in atypical populations, particularly those with difficulties in social communication.

Chapter 5

General Discussion

In this chapter, I reflect on decisions made in the development of the novel paradigm and potential design choices that would have been implemented with access to different equipment, budgets and time constraints. I also discuss the results found and what they mean in the contribution to the literature and where future research should be headed.

5.1 Overview

The aims of this PhD project were focused on better understanding joint attention and the behaviours that may drive the ability to successfully engage in this process. Throughout the literature, joint attention is positioned as a core process in the development of social behaviours that are critical to everyday communication but also to our development (Bruner, 1974; Bretherton & Bates, 1979). This makes it an important process to understand on a fundamental level and with increasing complexity. This project focused specifically on elucidating the importance of eye-gaze within the cluster of behaviours that underpin joint attention. To achieve this, virtual humans were presented on computer screens and via a virtual reality headset. The online experiments detailed in this thesis displayed these virtual humans solely on computer screens and were designed to test, refine and build upon the paradigm. From these experiments, I built a strong foundation of the paradigm's validity and reliability based on the manipulations that I had made. The final version of the paradigm eventually came from Experiment 2.1, the redesign of the gaze sequences provided the experiment with balanced conditions and higher experiment control. As a result of these changes, I observed an effect of both Collaboration and Gaze Type. The results positioned the experiment version as an appropriate paradigm to replicate in immersive virtual reality, concurrent EEG acquisition.

Overall, I found that an immersive virtual environment enhances the main effects of Collaboration and Gaze Type. On a neural level, I found significant differences for the factor Collaboration. The results revealed stronger alpha decrease in Non-Collaborative conditions compared to Collaborative conditions. These were present over posterior and anterior clusters towards the end of the trial and suggest that participants were increasingly attentive in Non-Collaborative conditions. A possible explanation of this effect may suggest that alpha-power, and it's association with attention, plays a role in the effect of Collaboration. This is corroborated by the general effects conveyed in the TFRs calculated in relation to baseline in Figure 4.12. As a result, I suggest that the of the lack of information presented in Non-Collaborative trials, increased participants' attention in order to gain information from the scene that would help them to make their decision.

Notably, in Collaboration comparison, I have separated the trial into the "beginning sequence" (1.40-1.80s) and the "end sequence" (2.28-3.04s). Overall, I observed weaker power at the beginning of the trial and stronger power towards the end of the trial in Collaborative conditions. From these observations, two potential explanations are discussed. The first explanation is tentatively interpreted in relation to an overlap of mu and beta frequencies. In both the beginning sequence and end sequence, earlier topographies (1.40-1.60s and 2.28-2.92s) are loosely consistent with a beta coupling network of attention (Gross et al., 2004). These early topographies may indicate heightened attention for (i) Collaborative conditions in the beginning sequence and (ii) Non-Collaborative conditions during the end sequence, indicative of delayed motor preparation for the latter, which is also consistent with the longer response times. Whereas later topographies (1.80s and 3.04s) may reflect motor preparation in mu frequencies for the respective conditions, earlier Collaborative and later Non-Collaborative.

Alternatively, it could be argued that all of the signatures (1.40-3.04s) are representative of mu rhythms and the bilateral, central oscillations are compatible with sensorimotor areas and therefore illustrates potential motor preparation. Mu-power decreases were observed during Collaborative trials, which are in line with the motor preparation to respond in Collaborative conditions. Equally, a similar pattern is observed during the participants' response window, illustrating stronger mu-power decreases in the Non-Collaborative trials. Together, these results corroborate the behavioural data that reports a significant increase (slower response times) for participants during the Non-Collaborative trials.

In addition, theta effects appeared to be dominant in the comparisons that include eye contact (Eye). Early, positive clusters were found in these trials that were not present in NoEye comparisons. The effects of theta are in line with the literature regarding theta integration with social stimuli (H. Wang et al., 2016) and lead to my proposal that theta oscillations might play a significant role in the processing of eye contact and its accompanied social connotations in this experiment (Kampis & Southgate, 2020). To my knowledge, there are only a few studies that have directly linked eye-gaze and theta rhythms (Gregory et al., 2022) but, these findings provide additional motivation to decipher these effects further. Overall, my experimental results, provide further evidence that there is a link between theta rhythms and eye gaze but this would require further investigation. Further exploration of the paradigm and potential manipulations for future studies are discussed in Section 5.3.

In recent years, psychology, cognitive science and many other fields interested in human cognition have begun to explore the use of virtual reality. The utility of virtual reality, in its various different forms, allows researchers to become more creative with their work and to enhance the paradigms and overall research in new ways. This is particularly true of immersive virtual reality which aims to increase the user's presence in the scene. Researchers that are making use of this tool would need to first determine whether a user, or participant in this case, is truly feeling immersed in the scene. If it is agreed that the feeling of presence is high for participants, it is then crucial to investigate a question that makes full use of the virtual reality headset and/or other equipment. Experiments that focus on tasks that include embodiment, navigation and agent interactions, all rely on a level of presence that would potentially benefit from an immersive virtual reality headset. Conversely, tasks that do not require an immersive environment such as the Stroop test and Wisconsin Sorting Card Test, would be better suited to a computer screen or physical questionnaire.

5.2 Critical evaluation

5.2.1 Physical set up

The final experiment of this thesis made use of a Neuro-VR setup that used a HTC Vive Pro Eye head-mounted display. This particular headset weighed $\approx$ 800g with the head strap. Some participants did mention the weight of the headset being on the heavier side and this was taken into consideration in the design of the task (see Chapter 3). One example of such design control is with the total experiment time. I restricted time in the headset to a maximum of 30 minutes. For this paradigm, 30 minutes allowed the experiment to contain enough trials that accounted for the signal-to-noise ratio and also took the participant's potential muscle strain into consideration. Not only would muscle strain be uncomfortable for the participant but, it could also introduce noisy and unwanted electrical signals. This is particularly problematic for EEG. As a result, I aimed for the participant's total time spent in the headset to be 25 minutes. This was made up of 20 minutes for the experiment itself and 5 minutes for eye-tracking calibration and practice trials. Experiments with longer trial times may consider a lighter headset in order to run the experiment longer for neuroimaging data quality. However, this sometimes means a compromise on the screen display and timing quality.

Alternatively, neuroimaging could be used in conjunction with other virtual environments such as the CAVE. Removing the need for a head-mounted display would be beneficial to experimental testing in a variety of ways. The 4-wall physical set-up of a CAVE provides a unique, virtual experience, that captures a participant's entire field-of-view and peripheral vision (see Figure 5.1). Although the CAVE immerses the participant in its virtual world, it still allows for interactions with real-world objects (see Figure 5.2), analogous to augmented reality. An example of this can be demonstrated from a short project I worked on during my PhD. The project focused on investigating step gait in older adults and required the creation and development of two scenes (i) a pleasant environment to walk through (park) and (ii) a gloomy environment to walk through (see Figure 5.2). I developed both of these scenes from first principles and made them compatible with the use of physical steps, which were included for participants to walk on. While the environment provided one manipulation of the task, another manipulation was the type of pattern that was projected onto the physical step. My Neuro-VR paradigm could be adapted to incorporate such features.

Another strength of CAVE systems are the glasses used. They are lightweight and are a minimal piece of equipment compared to head-mounted displays (see Figure 5.1). The removal of weighted equipment on the participant's head and screens that are further away would allow researchers to run longer experiments. Some may also argue that CAVE systems have a higher sense of presence as there is no potential muscle strain associated that could remind participants of there whereabouts. A Neuro-VR combination that used a CAVE, particularly EEG, would benefit from the minimal disturbance of virtual reality equipment on the electrodes. Additionally, some CAVEs offer a generous space that the participant could move around in. The use of fNIRS or OPMs at a later date, would be more preferable for a movement-based experiment, as, unlike EEG, they are not as sensitive to movement. However,

Figure 5.1: A picture demonstrating a use traversing a CAVE environment whilst wearing stereoscopic 3D glasses

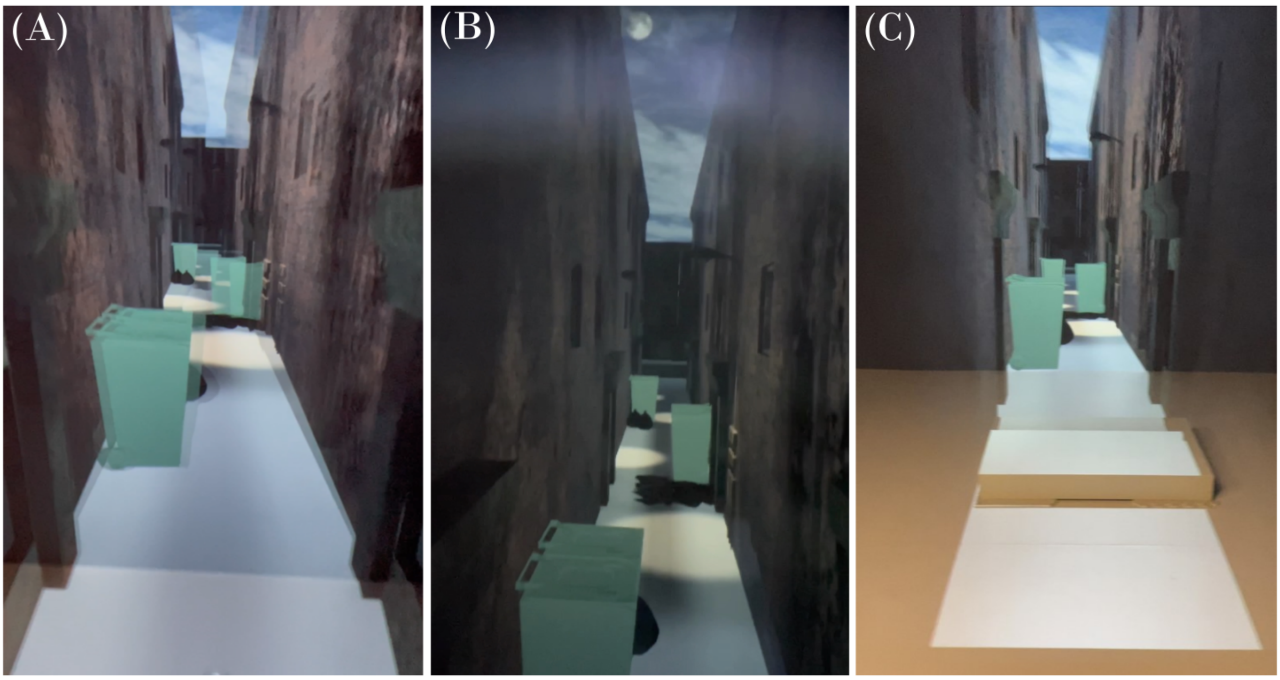

Figure 5.2: Three still photographs of the CAVE scene. Image **A**, displays the CAVE scene without stereoscopic glasses, image **B** displays the CAVE with stereoscopic glasses and image **C** displays the the CAVE scene in addition to a physical step.

due to the cost, maintenance and space required to own a CAVE, there are only a few facilities that possess and utilise this equipment.

Using a range of tools required careful consideration of the exclusion criteria to ensure participant safety. For example, it is not uncommon for virtual reality users to experience motion sickness. In this experiment, the participant's real-world motion was matched with movement occurring in the virtual reality space, which helped mitigate against such a response. As this is the most common cause of motion sickness, the vast majority of participants would be unaffected with such a change. However, for some, this can still persist and participants were informed on the information sheet (Appendix A) and again in person to make the experimenter aware such that the experiment could be stopped.

EEG considerations focused on neurological differences and equipment capabilities. Neurological differences that may present different EEG recordings (e.g. traumatic brain injury) or present a potential confound to the study question (e.g. ASD) were excluded. I explored the importance of the neurological exclusion criteria of Study 3 in detail in Subsection 4.3.4.

With regards to equipment capabilities, EEG as a technique is the gold standard for investigating social questions but it is not without its challenges. The inclusiveness of EEG as a research modality is becoming an increasingly discussed concern. The EEG cap used in Study 3 was designed and developed by ANT Neuro for mobile EEG recording. ANT Neuro provides three adult sizes (small, medium and large) that were designed to fit heads within a bracket of sizes. For example, the medium cap fits a circumference of 51 to 56 centimetres. Electrodes are placed around the stretchable cap in line with the 10-20 system (Jasper, 1958). In theory, the caps allow researchers to place the electrodes systematically on each participant however, this is not the case when taking all hair types into consideration. This EEG cap and most EEG equipment were designed for, and work best when being placed on European straight hair. Designs (of most imaging equipment) are typically centred around European phenotypes which makes it increasingly difficult to use these pieces of equipment on Afro or more curly and coarse hair textures. This is incredibly exclusionary and reduces the diversity of the population sample used as these hair textures are typical in people of African descent.

As a direct result of this lack of inclusion and other systemic racist behaviours that are entrenched within our society, racial disparities continue to prevail in research and clinical settings. This is particularly apparent with EEG, as it is used as the first step for the diagnosis of many disorders. Additionally, it presents a non-invasive and inexpensive way to measure neural signatures and so it is commonly used in both clinical and research settings (Etienne et al., 2020; Petit, Gagnon, Fantini, Ferini-Strambi, & Montplaisir, 2004; S. J. M. Smith, 2005). In research studies, these design flaws are manifested in homogeneous populations that are not representative across race or culture. And it is these research findings that are then used to support, justify and guide the implementation of industrial outputs such as pharmaceuticals (Brown et al., 2023), artificial intelligence (Buolamwini & Gebru, 2018) and healthcare practices (Penner et al., 2023). The lack of diversity as a result of inclusive equipment typically results in participants being turned away on the day, which is a waste of time, money and can be embarrassing for the participant. In extreme but not uncommon cases, it can also lead to misdiagnosis. As a result, it is important to highlight the effects of exclusionary neuroimaging tools and bring awareness of the need to adapt existing techniques whilst developing new methods and practises for researchers.

In order to reduce the bias of the population sample, it is critical for the experimenter to research afro hair textures and plan ahead. This includes becoming familiar with afro hair textures and understanding the accommodations that may be needed. Part of this is to provide participants with appropriate information beforehand (in the participant information sheet, an example is included in Appendix A) and to be prepared to adapt to the hair texture on the day. There may be a tendency to choose a larger cap to account for thicker and coarser hair texture however, it is sometimes appropriate to plait or canerow the hair in order to minimise the volume and to allow the cap to fit more comfortably and closer to the participant's scalp (see Figure 5.3). More recommendations can be found in a rolling developed EEG handbook created to address this limitation in neuroimaging research and clinical trials (Richardson, 2021).

Additionally, novel electrodes have been developed at Carnegie Mellon University, coined 'Sevo Electrodes' (Etienne et al., 2020). Engineers developed a novel electrode that provided a reduced impedance over the EEG gold standard, gold-cup electrodes. Researchers have combined Sevo electrodes with an electrode-bearing hair clip that fits neatly within parted hair, depicted on the right hand side of Figure 5.3. This combination of canerowed hair and using a Sevo electrode resulted in a significant decrease (>15x lower) in impedance measures compared to uncanerowed and canerowed hair that used gold-cup electrodes. Broadly speaking, such technological developments are desperately needed in the neuroimaging field to reduce these disparities and promote conscious diverse data collection.

The need for technological development also transfers over to fNIRS. This is particularly important in terms of the hardware. For fNIRS, the probes possess two critical issues that are required to

Figure 5.3: Image adapted from (Etienne et al., 2020). Shows the Sevo electrode and electrode-bearing hair clip placed between two canerows on Afro hair.

be addressed. Firstly, fNIRS utilises infrared light technology to estimate hemodynamic activity in response to a task (Scarapicchia et al., 2017). Broadly speaking, the technology uses the amount of light reflected or absorbed on the scalp to obtain physiological signals (Ba, Wang, Karinca, Bozkurt, & Kadambi, 2021). However, in the development of fNIRS, differences in melanin and pigmentation were not taken into consideration. As a result, recordings on darker skin are systematically higher in noise than on lighter skin due to the greater absorption of light for darker skin (Webb, Etter, & Kwasa, 2022). In addition to the technique used, the physical build of fNIRS is not inclusive of different hair textures. To record blood oxygenation, probes are used as part of the sensor. However, the length of the probe was designed with straight, European hair in mind. Longer probes in fNIRS are needed to penetrate through Afro and coarser hair textures for closer scalp contact (T. C. Parker & Ricard, 2022). This lack of diversity and consideration in the development of medical tools leads to poorer health outcomes for ethnic minorities (Carter, Lau, Johnson, & Kirkinis, 2017).

Currently, OPMs are the latest mobile, non-invasive neuroimaging tool that are being developed. With current interest focused on OPMs, it is crucial that diversity is considered and implemented into the updated designs of the equipment. Progression in neuroimaging is indicative of inclusion. In order to gain representative samples and accurate measurements, tools need to be developed with those who do not fit the white European standard in mind.

5.2.2 Virtual set up

Even with the perfect physical set-up, what is displayed to the participant is just as important for the successful investigation of my research question. There are therefore many considerations and adaptations that would ensure the creation of an ecologically valid virtual environment. Animation is one such modification and a key aspect of the development of virtual environments that contain dynamic scenes and objects. Throughout this project, there has been a particular focus on the virtual human's appearance and mobility. While I consciously focused on the animation of the virtual human's eyes, there is a need to attend to the animation of the scene as a whole. This would include the development of a virtual human with complete rigging and blending shapes. Rigging a virtual human is essentially the attribution of a skeleton. It connects the virtual human's "limbs" together so that animations can incorporate the appropriate areas that are near. In turn, this would help the animations appear more human-like. Whereas blending shapes provides frame-by-frame positioning, improving the fluidity of the animation. Had I been afforded the time to learn and develop these skills, they would have been applied to animating the (i) eyelids, (ii) breathing and (iii) idle behaviour (slight posture swaying). As a result, the virtual human would have been more advanced which would have added an additional level of realism to the paradigm. These modifications would be welcome additions to my Neuro-VR paradigm and would be great starting points for follow up investigations of this research question.

Within the literature, there are studies that have demonstrated the positive utility of the HTC Vive Pro Eye's embedded eye-tracking (Sipatchin, Wahl, & Rifai, 2021). Researchers investigated the temporal precision of the eye-tracker and reported a 58.1ms delay, deeming it a suitable ready-to-go tool. However, to make use of the eye-trackers full potential, careful thought into the experimental design and an expertise in back-end development are required. In Study 3, I recorded gaze movements and retrieved an unexpected result. Collectively, the gaze of all participants was predominantly focused on the virtual human's head. My interpretation of this finding has led me to conclude that the task was simple enough that participants did not need to alternate between the virtual human and the puzzle boards. While this result isn't necessarily rich in data, it does shine a light on how participants behaved in the head-mounted display. Given more time, I would have appropriately explored, analysed and refined the way that the data was collected within the headset and virtual environment. This could have taken the form of a short pilot experiment, where eye-tracking data was collected, which would have allowed the time to develop this further. Now that I know that the participant's focused on the virtual human's head, it would have been interesting to adjust the method of collection and measure which sections of the face, participants mainly centred their gaze. This type of data could later serve as a comparison between those who score higher on the AQ versus those who score lower. Research that has compared differences in visual attention to facial features have reported delayed neural responses in the N170 event-related potential in autistic people (T. Parker et al., 2021). Such investigations could help us better understand the differences between neurotypical and atypical gaze behaviour and help develop social and diagnostic tools to help those with ASD.

While critically evaluating my research decisions, attention must also be paid to the overall experimental paradigm. When developing the experiments and reviewing the data afterwards, a range of alternative routes to investigate joint attention were discussed. Decisions around the design of the experiment explored its potential gamification. The long-term goal of the project was to create a toolkit for individuals with difficulties in social communication. The toolkit would have offered a virtual environment to practise and gain experience in interactions if desired. As a result, the level at which the paradigm mimicked a game or interactive space could be adjusted. After reviewing the literature, it was decided to first understand the mechanisms of a key social behaviour, joint attention. Beginning with the foundations allowed us to pinpoint a particular behaviour (eye-gaze) and build a novel paradigm that understands its effects within the wider cluster of joint attention-related behaviours. The methodical analysis that was applied to the conception of the experiment was also applied to the design of the paradigm.

Fundamental experiment factors were chosen for the paradigm's design. The factor of Collaboration was built up of Collaborative and Non-Collaborative gaze shifts. It was clear that Collaborative gaze shifts would follow the prototypical behaviour that joint attention is described as in the literature. However, the way that Non-Collaborative conditions would be designed and measured was ambiguous. This is best demonstrated in the discussion of measuring the ground truth in Non-Collaborative conditions. In the experiment, the virtual human gaze was equidistant between both puzzle boards and therefore provided no ground truth (see Figure 3.2). This, in itself, could have been used as a manipulation of the experiment. Instead of having the virtual human's eyes gaze directly in between both puzzles, the angle of the eye-gaze could have been adjusted slightly towards or away from a given puzzle board. This would have resulted in a technical ground truth, depending on which puzzle board the virtual human's gaze was slightly more positioned toward. Having gaze behaviour that does not yield a ground truth in Non-Collaborative conditions removed the potential of an additional manipulation. Using an adjusted gaze location as a manipulation, the degree to which the gaze is directed towards a given puzzle board could have been increased or decreased. Such a manipulation would give light to whether there is a threshold angle at which gaze becomes informative enough for a participant to accurately guess the gazed-at location.

Equally, the speed at which the virtual human gazes at an object, be that the participant or target, may affect the processing of the gaze. When exploring the effect of the initial sequence in Experiment 1.2, the assumed meaning of faster or slower gaze was considered. For example, exploring whether there was a speed at which eye-gaze becomes redundant or alternatively, is interpreted as significant information. Such a manipulation would inform many investigations of eye-gaze and is important as the field begins to develop increasingly dynamic studies of gaze.

In real-world environments, joint attention and gaze behaviour rarely takes place within a vacuum. We are regularly interrupted by personal cues such as mobile notifications, other interactions such as louder conversations and sudden onsets of noise, like car alarms. As we aim to communicate with other individuals, there are a range of stimuli that can divert our attention. Such exogenous cues could be incorporated into my paradigm as distractions and are manipulations that were originally planned for the study. They could be visualised as the inclusion of background objects that had the potential to be factors of (i) movement (static vs. dynamic), (ii) consistency (always on screen vs. disappearing) and (iii) size (large vs. small). These distractions would have resulted in exogenous-driven, Non-Collaborative gaze shifts and could serve as a comparison to Collaborative trials or a direct comparison within and between their respective factors. This was not explored in the current project as the main goal was to initially understand the fundamental effect of gaze. However, these manipulations serve as an increase in complexity where levels of difficulty could be explored and would support the long-term goal of training through gamification.

The background of real-life interaction is typically busy and levels of difficulty in a gamified version of the paradigm could offer a gradual introduction to these everyday distractions. A staggered approach could prove useful to atypical populations and could also be applied to other factors within the paradigm. For example, the level of collaboration from a virtual human. In real life, varied levels of collaborative behaviour are displayed as either cooperative or uninterested and distracted behaviour. Such behaviour from a virtual human may invoke the participant into questioning whether their agent is paying attention to the task and providing clear and informative cues (gaze shifts) or if are they looking elsewhere and not engaging with them. In line with this, allows the ability to control the number, location and onset of distractions in the environment; sudden appearances may invoke overt shifts of attention. These manipulations provide scenarios that can then be explored using EEG and eye-tracking, helping us further elucidate the neural correlates of joint attention and the effects of disruption in real-time. A Neuro-VR approach, in this case, provides the appropriate tools to answer these questions.

5.3 Future outlook

Overall, there are a range of changes and considerations that could have been implemented in the development of my paradigm. However, the decisions that I made, have been detailed in this thesis to provide a foundation for the further investigation of eye gaze within dynamic virtual environments. If this work was to be directly expanded upon in the near future, there are a range of changes that could enhance these studies, both in their immersiveness and data output. Therefore, for researchers that choose to build on this work, I propose 5 key modifications:

- *A multimodal imaging approach to better understand the neural correlates of joint attention.*

  Multimodal imaging presents the opportunity to explore the neural correlates of behaviour with high spatial and temporal precision. This is achievable with the pairing of fNIRS and EEG. As continued tests are conducted to improve the ecological validity of virtual environments, it is important to gain a good understanding of how this is expressed in oscillatory signatures in relation to behavioural changes recorded. Measuring cortical changes in response to the manipulation of factors within a paradigm does not only benefit the research question but also our understanding of how we, as humans, interact with virtual humans.

- *Enhanced animation, specifically applied to the eyelids and an idle animation.*

  In these foundational experiments, I had chosen to only allow one dynamic feature, the animation from the virtual human's eyes. This allowed me to confidently associate behavioural and neurological changes to the manipulations made. However, now that I have seen a consistent pattern of results throughout the project, the next step would be to develop new, virtual human animations. It is my belief that small, yet powerful animations would increase a participant's presence in the environment. Specifically, I propose that efforts should be made to incorporate the virtual human's eyelids. On a fundamental level, I believe the natural movement of an eyelid when looking in different directions should be incorporated into future work. By extension, the inclusion of eye blinks should also be a consideration. However, variations in speed and length of a blink could also attribute certain characteristics to the virtual human and should be carefully thought out. Such modifications would increase the agent's display of human-like behaviour but must be balanced against the uncanny valley effect, which may reduce a participant's connection to a scene.

- *Including a questionnaire to measure the participant's presence.*

  In line with these suggested changes to the virtual human's animation, the inclusion of a questionnaire to measure the presence of a participant is necessary. Presence is an agent's response to a certain level of immersion (Slater, 2003). For example, understanding how aware an agent is of the real-world events that are occurring around them while viewing a virtual environment (Witmer & Singer, 1998). This would help to capture a complete understanding of the differences between 2D computer screens and immersive virtual environments when interacting with a virtual human. An experiment could be set up where participants are exposed to either method of presentation followed by a questionnaire that measures their feeling of presence. A questionnaire can also act as a prompt for participant feedback, encouraging dialogue around behaviours/animations they deemed strange or noticed were missing in the experiment. Such feedback is valuable for future developments and manipulations to create a paradigm that will promote natural interaction in relation to the given environment.

- *Development of one of the proposed manipulations.*

  Recommended manipulations have been detailed above in the critical evaluation of my work (see section 5.2). In order to push the paradigm to the next step, I believe the complexity of the paradigm should be addressed. Two examples of how this can be done are (i) by adjusting the angle at which Non-Collaborative conditions are presented and therefore installing a ground truth, or (ii) by including exogenous distractions in the environment, with the aim to investigate the effects of disrupted joint attention and deciphering between intentional and averted gaze. Both manipulations offer the ability to increase and decrease the complexity of the task and are complementary to a later gamified version.

- *Enhancement of the eye-tracking acquisition.*

  Lastly, I advise that future work takes time to incorporate eye-tracking. Recording eye-tracking produces data that can be used to benefit a range of manipulations. This could be exploratory when using a new paradigm or when comparing stimulus presentation methods (e.g. 2D vs. immersive virtual reality). Likewise, it can be used to corroborate findings that are based on previous literature (e.g. understanding gaze preferences between two population samples (Clin, Maes, Stercq, & Kissine, 2020)). However, there are many benefits that relate directly to the aforementioned suggested manipulations. When using distractions, it provides information on when and where attention is being diverted. This would feedback into the development of the task to understand which distractions can hinder or potentially aid successful joint attention and in what kind of scenarios. Equally, in an experiment that utilises a ground truth of Non-Collaborative conditions, eye-tracking can help determine the angle at which the difficulty level is rendered useless or useful. Additionally, it can also provide insight into whether there are troubleshooting sequences, the length of time spent looking at an ROI and the last ROI looked at before making a decision. Generally, eye-tracking would provide greater insight into the typical gaze behaviours of joint attention. Importantly, this data could also be used to compare atypical gaze behaviour and explore any differences that may be present.

Altogether, there are many future avenues that can be expanded upon from this paradigm. This is a result of the strong foundational experimentation throughout the project that has offered an elementary design that can be easily built upon. This should appeal to the field's researchers who are interested in extending our understanding of joint attention. This work demonstrates a clear benefit to investigating cognitive questions using a Neuro-VR toolkit and immersive virtual environments. There are clear pathways to continue this development that leads to the creation of increasingly, ecologically valid studies.

5.4 Conclusion

To conclude, this project revealed that eye-gaze does have an effect on joint attention. Through a series of experiments, the results revealed an effect of Collaboration, where participants responded significantly faster in Collaborative conditions. In regards to Gaze Type, I initially did not see an effect however, as I began to modify and refine the paradigm the effect became clear. This was reinforced by the use of an immersive virtual environment which may have increased the participant's presence in the scene and working with the virtual human. In particular, engaging in eye contact with the virtual human appeared to facilitate the participant's ability to respond during trials. This is also true within the Collaborative condition. When participants were given the correct answer, they performed faster when the sequence included eye contact compared to when the sequence did not include eye contact. From this, I suggest that eye-gaze and specifically eye contact are influential in successfully engaging in joint attention and the efficiency of such processes relies on them. From this, future research on joint attention should carefully consider the latency and frequency that an agent engages in eye contact. Overall, the literature requires more attention to be paid to the effects of eye contact as from this research I have found it to be a more fundamental and possibly social behaviour of joint attention.

Crucially, this project has demonstrated the success of Neuro-VR, using the combination of a virtual reality head-mounted display and EEG. This particular pairing has been fruitful in understanding the process of joint attention and offered objective data regarding its neural underpinnings. Neuro-VR as an experimental tool could be transposed to other research questions, particularly those that may require additional measurements when verbal investigation of the topic is not possible. In these scenarios, virtual reality offers a tool to circumvent these difficulties by communicating through immersive and dynamic environments. Developing paradigms with non-verbal behaviours, allows the research to be inclusive of those with reduced communicative abilities. Therefore, the paradigm not only serves to comprehend the process of joint attention in typical populations but also in atypical populations as well.

Ultimately, my research provides a basis for creating more immersive paradigms that investigate cognitive, social behaviour. It also positions Neuro-VR as a strong experimental tool that sits between static and naturalistic experiments, providing the right amount of experimental control. I plan to continue to innovate in this space, creating immersive, inclusive environments that glean rich data outputs that can be applied to benefiting atypical populations and improving our overall understanding of cognitive neuroscience.

References

Abtahi, M., Bahram Borgheai, S., Jafari, R., Constant, N., Diouf, R., Shahriari, Y., & Mankodiya, K. (2020). Merging fNIRS-EEG brain monitoring and body motion capture to distinguish parkinsons disease. , *28*(6), 1246–1253. Retrieved 2020-06-18, from https://ieeexplore.ieee.org/document/9067065/ doi: 10.1109/TNSRE.2020.2987888

Adams, R. B., & Kleck, R. E. (2005). Effects of direct and averted gaze on the perception of facially communicated emotion. , *5*(1), 3–11. Retrieved 2021-04-07, from http://doi.apa.org/getdoi.cfm?doi=10.1037/1528-3542.5.1.3 doi: 10.1037/1528-3542.5.1.3

Admoni, H., & Scassellati, B. (2017). Social eye gaze in human-robot interaction: A review. , *6*(1), 25. Retrieved 2021-09-30, from http://dl.acm.org/citation.cfm?id=3109975 doi: 10.5898/JHRI.6.1.Admoni

Adolphs, R. (2002). Neural systems for recognizing emotion.
doi: 10.1016/s0959-4388(02)00301-x

Allen, M., Poggiali, D., & Whitaker, K. e. a. (2021). *Raincloud plots: a multi-platform tool for robust data visualization.* Wellcome Open Res 2021, 4:63. Retrieved from 10.12688/wellcomeopenres.15191.2

Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. *Groups, leadership, and men*, 177–190.

Ashwood, K. L., Gillan, N., Horder, J., Hayward, H., Woodhouse, E., McEwen, F. S., ... Murphy, D. G. (2016). Predicting the diagnosis of autism in adults using the autism-spectrum quotient (AQ) questionnaire. , *46*(12), 2595–2604. Retrieved 2023-04-22, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4988267/ doi: 10.1017/S0033291716001082

Ba, Y., Wang, Z., Karinca, K. D., Bozkurt, O. D., & Kadambi, A. (2021). *Overcoming difficulty in* *obtaining dark-skinned subjects for remote-PPG by synthetic augmentation.* arXiv. Retrieved 2023-04-26, from http://arxiv.org/abs/2106.06007 (Number: arXiv:2106.06007)

Bacos, C. A. (2020). Enhancing social attention using eye- movement modeling and simulated dyadic social interactions. , 6.

Bailenson, J. N., Blascovich, J., Beall, A. C., & Loomis, J. M. (2003). Interpersonal distance in immersive virtual environments. , *29*(7), 819–833. Retrieved 2020-06-18, from http://journals.sagepub.com/doi/10.1177/0146167203029007002 doi: 10.1177/0146167203029007002

Bakeman, R., & Adamson, L. B. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interaction. , *55*(4), 1278. Retrieved 2020-04-30, from https://www.jstor.org/stable/1129997?origin=crossref doi: 10.2307/1129997

Baron-Cohen, S. (1995). *Mindblindness: an essay on autism and theory of mind.* MIT Press.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? , *21*(1), 37–46. doi: 10.1016/0010-0277(85)90022-8

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, malesand females, scientists and mathematicians. , *31*(1), 5–17.

Barr, D. J. (2008). Pragmatic expectations and linguistic evidence: Listeners anticipate but do not integrate common ground. , *109*(1), 18–40. Retrieved 2020-06-26, from https://linkinghub.elsevier.com/retrieve/pii/S0010027708001698 doi: 10.1016/j.cognition.2008.07.005

Barry, R. A., Graf Estes, K., & Rivera, S. M. (2015). Domain general learning: Infants use social and non-social cues when learning object statistics. , *6*. Retrieved 2023-04-23, from https://www.frontiersin.org/articles/10.3389/fpsyg.2015.00551

Bartsch, K., & Wellman, H. M. (1995). *Children talk about the mind.* Oxford University Press. (Google-Books-ID: GZkmKBY664kC)

Bayliss, A. P., Murphy, E., Naughtin, C. K., Kritikos, A., Schilbach, L., & Becker, S. I. (2013). "gaze leading": Initiating simulated joint attention influences eye movements and choice behavior. , *142*(1), 76–92. Retrieved 2022-05-05, from http://doi.apa.org/getdoi.cfm?doi=10.1037/a0029286 doi: 10.1037/a0029286

Becchio, C., Bertone, C., & Castiello, U. (2008). How the gaze of others influences object processing. *Trends in cognitive sciences*, *12*(7), 254–258.

Bickmore, T., Kimani, E., Shamekhi, A., Murali, P., Parmar, D., & Trinh, H. (2021). Virtual agents as supporting media for scientific presentations. , *15*(2), 131–146. Retrieved 2023-02-24, from https://doi.org/10.1007/s12193-020-00350-y doi: 10.1007/s12193-020-00350-y

Billeci, L., Narzisi, A., Tonacci, A., Sbriscia-Fioretti, B., Serasini, L., Fulceri, F., ... Muratori, F. (2017a). An integrated eeg and eye-tracking approach for the study of responding and initiating joint attention in autism spectrum disorders. *Scientific Reports*, *7*(1), 13560.

Billeci, L., Narzisi, A., Tonacci, A., Sbriscia-Fioretti, B., Serasini, L., Fulceri, F., ... Muratori, F. (2017b). An integrated EEG and eye-tracking approach for the study of responding and initiating joint attention in autism spectrum disorders. , *7*(1), 13560. Retrieved 2020-01-29, from http://www.nature.com/articles/s41598-017-13053-4 doi: 10.1038/s41598-017-13053-4

Biswas, S. S. (2023). Role of chat gpt in public health. *Annals of Biomedical Engineering*, 1–2.

Blender. (2018). *Blender - a 3d modelling and rendering package.* Stichting Blender Foundation. Retrieved from http://www.blender.org

Boland, J. E., Fonseca, P., Mermelstein, I., & Williamson, M. (2022). Zoom disrupts the rhythm of conversation. , *151*, 1272–1282. (Place: US Publisher: American Psychological Association) doi: 10.1037/xge0001150

Bombari, D., Schmid Mast, M., Canadas, E., & Bachmann, M. (2015). Studying social interactions through immersive virtual environment technology: virtues, pitfalls, and future challenges. , *6*, 869. doi: 10.3389/fpsyg.2015.00869

Botella, C., Fernández-Álvarez, J., Guillén, V., García-Palacios, A., & Baños, R. (2017). Recent progress in virtual reality exposure therapy for phobias: A systematic review. , *19*(7), 42. Retrieved 2020-06-27, from http://link.springer.com/10.1007/s11920-017-0788-4 doi: 10.1007/s11920-017-0788-4

Botvinick, M., P Jha, A., Bylsma, L., A Fabian, S., E Solomon, P., & M Prkachin, K. (2005). Viewing facial expressions of pain engages cortical areas involved in the direct experience of pain. doi: 10.1016/j.neuroimage.2004.11.043

Bouchard, S., Dumoulin, S., Robillard, G., Guitard, T., Klinger, E., Forget, H., ... Roucaut, F. X. (2017). Virtual reality compared with in vivo exposure in the treatment of social anxiety disorder: A three-arm randomised controlled trial. , *210*(4), 276–283. doi: 10.1192/bjp.bp.116.184234

Bowman, D. A., & McMahan, R. P. (2007). Virtual reality: How much immersion is enough? , *40*(7), 36–43. (Conference Name: Computer) doi: 10.1109/MC.2007.257

Brandone, A. C., & Stout, W. (2022). The origins of theory of mind in infant social cognition: Investigating longitudinal pathways from intention understanding and joint attention to preschool theory of mind. , *0*(0), 1–22. Retrieved 2022-12-05, from https://doi.org/10.1080/15248372.2022.2146117 (Publisher: Routledge _eprint: https://doi.org/10.1080/15248372.2022.2146117) doi: 10.1080/15248372.2022.2146117

Bretherton, I., & Bates, E. (1979). The emergence of intentional communication. , *1979*(4), 81–100. Retrieved 2020-04-30, from http://doi.wiley.com/10.1002/cd.23219790407 doi: 10.1002/cd.23219790407

Brookes, M. J., Leggett, J., Rea, M., Hill, R. M., Holmes, N., Boto, E., & Bowtell, R. (2022). Magnetoencephalography with optically pumped magnetometers (OPM-MEG): the next generation of functional neuroimaging. , *45*(8), 621–634. Retrieved 2023-04-15, from https://www.cell.com/trends/neurosciences/abstract/S0166-2236(22)00102-3 (Publisher: Elsevier) doi: 10.1016/j.tins.2022.05.008

Brown, K., Etrouth, S., Jayachandran, P., Moore, J., Oni-Orisan, A., Vasist, L., ... Dresser, M. (2023). Diversity in clinical pharmacology: A call to action. , *113*(3), 483–485. Retrieved 2023-04-29, from https://onlinelibrary.wiley.com/doi/abs/10.1002/cpt.2852 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cpt.2852) doi: 10.1002/cpt.2852

Bruinsma, Y., Koegel, R. L., & Koegel, L. K. (2004). Joint attention and children with autism: A review of the literature. , *10*(3), 169–175. Retrieved 2022-01-13, from https://onlinelibrary.wiley.com/doi/abs/10.1002/mrdd.20036 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrdd.20036) doi: 10.1002/mrdd.20036

Bruner, J. (1974). From communication to language—a psychological perspective. , *3*(3), 255–287. Retrieved 2021-09-29, from https://linkinghub.elsevier.com/retrieve/pii/0010027774900122 doi: 10.1016/0010-0277(74)90012-2

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Proceedings of the 1st conference on fairness, accountability and transparency* (pp. 77–91). PMLR. Retrieved 2023-04-29, from https://proceedings.mlr.press/v81/buolamwini18a.html (ISSN: 2640-3498)

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. , *63*(4), i–174. Retrieved 2021-12-20, from https://www.jstor.org/stable/1166214 (Publisher: [Society for Research in Child Development, Wiley]) doi: 10.2307/1166214

Carter, R. T., Lau, M. Y., Johnson, V., & Kirkinis, K. (2017). Racial discrimination and health outcomes among racial/ethnic minorities: A meta-analytic review. , *45*(4), 232–259. Retrieved 2023-04-26, from `https://onlinelibrary.wiley.com/doi/abs/10.1002/jmcd.12076` (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jmcd.12076) doi: 10.1002/jmcd.12076

Caruana, N., Alhasan, A., Wagner, K., Kaplan, D. M., Woolgar, A., & McArthur, G. (2020). The effect of non-communicative eye movements on joint attention. , *73*(12), 2389–2402. Retrieved 2021-09-25, from `https://doi.org/10.1177/1747021820945604` (Publisher: SAGE Publications) doi: 10.1177/1747021820945604

Caruana, N., Brock, J., & Woolgar, A. (2015). A frontotemporoparietal network common to initiating and responding to joint attention bids. , *108*, 34–46. Retrieved 2020-01-16, from `https://linkinghub.elsevier.com/retrieve/pii/S1053811914010337` doi: 10.1016/j.neuroimage.2014.12.041

Caruana, N., & McArthur, G. (2019). The mind minds minds: The effect of intentional stance on the neural encoding of joint attention. Retrieved 2019-10-25, from `http://link.springer.com/10.3758/s13415-019-00734-y` doi: 10.3758/s13415-019-00734-y

Caruana, N., McArthur, G., Woolgar, A., & Brock, J. (2017a). Detecting communicative intent in a computerised test of joint attention. , *5*, e2899. Retrieved 2020-03-11, from `https://peerj.com/articles/2899` doi: 10.7717/peerj.2899

Caruana, N., McArthur, G., Woolgar, A., & Brock, J. (2017b). Simulating social interactions for the experimental investigation of joint attention. , *74*, 115–125. Retrieved 2021-04-07, from `https://www.sciencedirect.com/science/article/pii/S0149763416304778` doi: 10.1016/j.neubiorev.2016.12.022

Charman, T. (1997). The relationship between joint attention and pretend play in autism. , *9*(1), 1–16. doi: 10.1017/S095457949700103X

Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Cox, A., & Drew, A. (2000a). Testing joint attention, imitation, and play as infancy precursors to language and theory of mind. *Cognitive development*, *15*(4), 481–498.

Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Cox, A., & Drew, A. (2000b). Testing joint attention, imitation, and play as infancy precursors to language and theory of mind. , *15*(4), 481–498. Retrieved 2022-05-13, from `https://www.sciencedirect.com/science/article/pii/S0885201401000375` doi: 10.1016/S0885-2014(01)00037-5

Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Drew, A., & Cox, A. (2003). Predicting language outcome in infants with autism and pervasive developmental disorder. *International Journal of Language & Communication Disorders*, *38*(3), 265–285.

Christodoulou, C. (2001). Functional magnetic resonance imaging of working memory impairment after traumatic brain injury. , *71*(2), 161–168. Retrieved 2023-04-15, from `https://jnnp.bmj.com/lookup/doi/10.1136/jnnp.71.2.161` doi: 10.1136/jnnp.71.2.161

Cilia, F., Touchet, C., Vandromme, L., & Driant, B. L. (2020). Initiation and response of joint attention bids in autism spectrum disorder children depend on the visibility of the target. , *5*, 2396941520950979. Retrieved 2021-10-25, from `https://doi.org/10.1177/2396941520950979` (Publisher: SAGE Publications Ltd) doi: 10.1177/2396941520950979

Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge.

Clin, E., Maes, P., Stercq, F., & Kissine, M. (2020). No preference for direct versus averted gaze in autistic adults: a reinforced preferential looking paradigm. , *11*(1), 91. Retrieved 2023-04-22, from `https://doi.org/10.1186/s13229-020-00398-3` doi: 10.1186/s13229-020-00398-3

Converse, B. A., Lin, S., Keysar, B., & Epley, N. (2008). In the mood to get over yourself: Mood affects theory-of-mind use. , *8*(5), 725–730. Retrieved 2020-06-26, from `http://doi.apa.org/getdoi.cfm?doi=10.1037/a0013283` doi: 10.1037/a0013283

Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., & Liaw, J. (2004). Early social attention impairments in autism: Social orienting, joint attention, and attention to distress. , *40*(2), 271–283. Retrieved 2020-04-30, from `http://doi.apa.org/getdoi.cfm?doi=10.1037/0012-1649.40.2.271` doi: 10.1037/0012-1649.40.2.271

de Cheveigné, A. (2020). ZapLine: A simple and effective method to remove power line artifacts. , *207*, 116356. Retrieved 2020-05-07, from `https://linkinghub.elsevier.com/retrieve/pii/S1053811919309474` doi: 10.1016/j.neuroimage.2019.116356

De Paolis, L. T., & Bourdot, P. (Eds.). (2018). *Augmented reality, virtual reality, and computer graphics: 5th international conference, AVR 2018, otranto, italy, june 24–27, 2018, proceedings, part II* (Vol. 10851). Springer International Publishing. Retrieved 2020-06-29, from `http://link.springer.com/10.1007/978-3-319-95282-6` doi: 10.1007/978-3-319-95282-6

Deák, G. O., Flom, R. A., & Pick, A. D. (2000). Effects of gesture and target on 12- and 18-month-olds' joint visual attention to objects in front of or behind them. , *36*(4), 511–523. Retrieved

2021-10-25, from http://doi.apa.org/getdoi.cfm?doi=10.1037/0012-1649.36.4.511 doi: 10.1037/0012-1649.36.4.511

di Lanzo, J. A., Valentine, A., Sohel, F., Yapp, A. Y. T., Muparadzi, K. C., & Abdelmalek, M. (2020). A review of the uses of virtual reality in engineering education. , *28*(3), 748–763. Retrieved 2023-04-29, from https://onlinelibrary.wiley.com/doi/abs/10.1002/cae.22243 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cae.22243) doi: 10.1002/cae.22243

Dravida, S., Noah, J. A., Zhang, X., & Hirsch, J. (2020). Joint attention during live person-to-person contact activates rTPJ, including a sub-component associated with spontaneous eye-to-eye contact. , *14*. Retrieved 2021-03-22, from https://www.frontiersin.org/articles/10.3389/fnhum.2020.00201/full (Publisher: Frontiers) doi: 10.3389/fnhum.2020.00201

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. , *6*(5), 509–540. Retrieved 2021-09-20, from http://www.tandfonline.com/doi/abs/10.1080/135062899394920 doi: 10.1080/135062899394920

Edwards, S. G., Seibert, N., & Bayliss, A. P. (2020). Joint attention facilitates observed gaze direction discrimination. , *73*(1), 80–90. Retrieved 2020-01-17, from http://journals.sagepub.com/doi/10.1177/1747021819867901 doi: 10.1177/1747021819867901

Etienne, A., Laroia, T., Weigle, H., Afelin, A., Kelly, S. K., Krishnan, A., & Grover, P. (2020). Novel electrodes for reliable EEG recordings on coarse and curly hair. , *2020*, 6151–6154. doi: 10.1109/EMBC44109.2020.9176067

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. , *39*(2), 175–191. doi: 10.3758/bf03193146

Foot, P. (1967). The problem of abortion and the doctrine of the double effect. , *5*, 6.

Fortenbaugh, F. C., DeGutis, J., Germine, L., Wilmer, J., Grosso, M., Russo, K., & Esterman, M. (2015). Sustained attention across the lifespan in a sample of 10,000: Dissociating ability and strategy. , *26*(9), 1497–1510. Retrieved 2023-04-15, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4567490/ doi: 10.1177/0956797615594896

Foulsham, T., Walker, E., & Kingstone, A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. , *51*(17), 1920–1931. Retrieved 2023-02-23, from https://www.sciencedirect.com/science/article/pii/S0042698911002392 doi: 10.1016/j.visres.2011.07.002

Foxe, J., & Snyder, A. C. (2011). *Frontiers | the role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention.* Retrieved 2023-04-28, from https://www.frontiersin.org/articles/10.3389/fpsyg.2011.00154/full

Friesen, C. K., & Kingstone, A. (1998). The eyes have it! reflexive orienting is triggered by nonpredictive gaze. , *5*(3), 490–495. Retrieved 2021-06-22, from http://link.springer.com/10.3758/BF03208827 doi: 10.3758/BF03208827

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. , *133*(4), 694–724. Retrieved 2021-09-25, from http://doi.apa.org/getdoi.cfm?doi=10.1037/0033-2909.133.4.694 doi: 10.1037/0033-2909.133.4.694

Frith, C. D., & Frith, U. (2007). Social cognition in humans. , *17*(16), R724–R732. Retrieved 2020-04-16, from https://linkinghub.elsevier.com/retrieve/pii/S0960982207014923 doi: 10.1016/j.cub.2007.05.068

Gale, A., Spratt, G., Chapman, A. J., & Smallbone, A. (1975). EEG correlates of eye contact and interpersonal distance. , *3*(4), 237–245. Retrieved 2023-04-23, from https://www.sciencedirect.com/science/article/pii/030105117590023X doi: 10.1016/0301-0511(75)90023-X

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. doi: 10.1016/j.tics.2004.07.002

Gallup, A. C., Hale, J. J., Sumpter, D. J. T., Garnier, S., Kacelnik, A., Krebs, J. R., & Couzin, I. D. (2012). Visual attention and the acquisition of information in human crowds. , *109*(19), 7245–7250. Retrieved 2023-02-23, from https://www.pnas.org/doi/full/10.1073/pnas.1116141109 (Publisher: Proceedings of the National Academy of Sciences) doi: 10.1073/pnas.1116141109

Garau, M., Slater, M., Vinayagamoorthy, V., Brogni, A., Steed, A., & Sasse, M. A. (2003). The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 529–536). Association for Computing Machinery. Retrieved 2022-12-06, from https://doi.org/10.1145/642611.642703 doi: 10.1145/642611.642703

García, A. S., Fernández-Sotos, P., González, P., Navarro, E., Rodriguez-Jimenez, R., & Fernández-

Caballero, A. (2022). Behavioral intention of mental health practitioners toward the adoption of virtual humans in affect recognition training. , *13*. Retrieved 2023-02-24, from https://www.frontiersin.org/articles/10.3389/fpsyg.2022.934880

Geary, R. C. (1935). The ratio of the mean deviation to the standard deviation as a test of normality. *Biometrika*, *27*(3/4), 310–332.

Gobel, M., Kim, H., & Richardson, D. (2015). The dual function of social gaze | elsevier enhanced reader.
doi: 10.1016/j.cognition.2014.11.040

Gooding-Williams, G., Wang, H., & Kessler, K. (2016). THETA-rhythm makes the world go round: Dissociative effects of TMS theta versus alpha entrainment of right pTPJ on embodied perspective transformations. , *30*(5), 561–564. Retrieved 2020-06-17, from http://link.springer.com/10.1007/s10548-017-0557-z doi: 10.1007/s10548-017-0557-z

Grafman, J., Jonas, B., & Salazar, A. (1990). Wisconsin card sorting test performance based on location and size of neuroanatomical lesion in vietnam veterans with penetrating head injury. , *71*, 1120–1122. Retrieved 2023-04-15, from https://journals.sagepub.com/doi/epdf/10.2466/pms.1990.71.3f.1120 doi: 10.2466/pms.1990.71.3f.1120

Gredebäck, G., Fikke, L., & Melinder, A. (2010). The development of joint visual attention: a longitudinal study of gaze following during interactions with mothers and strangers: The development of joint visual attention. , *13*(6), 839–848. Retrieved 2021-10-18, from https://onlinelibrary.wiley.com/doi/10.1111/j.1467-7687.2009.00945.x doi: 10.1111/j.1467-7687.2009.00945.x

Green, R., Shaw, D. J., & Kessler, K. (2023). Dissociating visual perspective taking and belief reasoning using a novel integrated paradigm: A preregistered online study. , *235*, 105397. Retrieved 2023-04-22, from https://www.sciencedirect.com/science/article/pii/S0010027723000318 doi: 10.1016/j.cognition.2023.105397

Gregory, S. (2021). Investigating facilitatory versus inhibitory effects of dynamic social and non-social cues on attention in a realistic space. Retrieved 2021-08-27, from https://link.springer.com/10.1007/s00426-021-01574-7 doi: 10.1007/s00426-021-01574-7

Gregory, S., & Jackson, M. C. (2021). Increased perceptual distraction and task demand enhances gaze and non-biological cuing effects. , *74*(2), 221–240. Retrieved

2023-04-23, from http://journals.sagepub.com/doi/10.1177/1747021820959633 doi: 10.1177/1747021820959633

Gregory, S., Kelly, C. L., & Kessler, K. (2021). Look into my “virtual” eyes: What dynamic virtual agents add to the realistic study of joint attention. , *2*, 156. Retrieved 2021-12-16, from https://www.frontiersin.org/article/10.3389/frvir.2021.798899 doi: 10.3389/frvir.2021.798899

Gregory, S., Wang, H., & Kessler, K. (2022). EEG alpha and theta signatures of socially and non-socially cued working memory in virtual reality. , *17*(6), 531–540. Retrieved 2022-11-18, from https://doi.org/10.1093/scan/nsab123 doi: 10.1093/scan/nsab123

Griffin, J. W., & Scherf, K. S. (2020). Does decreased visual attention to faces underlie difficulties interpreting eye gaze cues in autism? , *11*(1), 60. Retrieved 2021-03-22, from https://doi.org/10.1186/s13229-020-00361-2 doi: 10.1186/s13229-020-00361-2

Gross, J., Schmitz, F., Schnitzler, I., Kessler, K., Shapiro, K., Hommel, B., & Schnitzler, A. (2004). Modulation of long-range neural synchrony reflects temporal limitations of visual attention in humans. , *101*(35), 13050–13055. Retrieved 2023-04-27, from https://www.pnas.org/doi/10.1073/pnas.0404944101 (Publisher: Proceedings of the National Academy of Sciences) doi: 10.1073/pnas.0404944101

Grove, T. B., Lasagna, C. A., Martínez-Cancino, R., Pamidighantam, P., Deldin, P. J., & Tso, I. F. (2021). Neural oscillatory abnormalities during gaze processing in schizophrenia: Evidence of reduced theta phase consistency and inter-areal theta-gamma coupling. , *6*(3), 370–379. Retrieved 2023-04-22, from https://www.sciencedirect.com/science/article/pii/S2451902220302482 doi: 10.1016/j.bpsc.2020.08.013

Hare, B. (2017). Survival of the friendliest: Homo sapiens evolved via selection for prosociality. , *68*(1), 155–186. Retrieved 2023-03-10, from https://doi.org/10.1146/annurev-psych-010416-044201 (_eprint: https://doi.org/10.1146/annurev-psych-010416-044201) doi: 10.1146/annurev-psych-010416-044201

Hedden, T., & Gabrieli, J. D. E. (2004). Insights into the ageing mind: a view from cognitive neuroscience. , *5*(2), 87–96. Retrieved 2023-04-15, from http://www.nature.com/articles/nrn1323 doi: 10.1038/nrn1323

Hertweck, S., Weber, D., Alwanni, H., Unruh, F., Fischbach, M., Latoschik, M. E., & Ball, T. (2019).

Brain activity in virtual reality: Assessing signal quality of high-resolution EEG while using head-mounted displays. In *2019 IEEE conference on virtual reality and 3d user interfaces (VR)* (pp. 970–971). IEEE. Retrieved 2019-11-18, from https://ieeexplore.ieee.org/document/8798369/ doi: 10.1109/VR.2019.8798369

Hills, P. D., Clavin, M. V. Q., Tufft, M. R. A., Gobel, M. S., & Richardson, D. C. (2022). Video meeting signals: Experimental evidence for a technique to improve the experience of video conferencing. , *17*(8), e0270399. Retrieved 2023-02-21, from https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0270399 (Publisher: Public Library of Science) doi: 10.1371/journal.pone.0270399

Hofsten, C., Dahlström, E., & Fredriksson, Y. (2005). 12-month-old infants' perception of attention direction in static video images. , *8*(3), 217–231. Retrieved 2021-10-18, from http://doi.wiley.com/10.1207/s15327078in0803$_2$ doi: 10.1207/s15327078in0803$_2$

Holper, L., Muehlemann, T., Scholkmann, F., Eng, K., Kiper, D., & Wolf, M. (2010). Testing the potential of a virtual reality neurorehabilitation system during performance of observation, imagery and imitation of motor actions recorded by wireless functional near-infrared spectroscopy (fNIRS). , *7*(1), 57. Retrieved 2023-04-26, from https://doi.org/10.1186/1743-0003-7-57 doi: 10.1186/1743-0003-7-57

Hughes, C., & Dunn, J. (1998). Understanding mind and emotion: Longitudinal associations with mental-state talk between young friends. , *34*, 1026–1037. (Place: US Publisher: American Psychological Association) doi: 10.1037/0012-1649.34.5.1026

Itier, R. J., Alain, C., Kovacevic, N., & McIntosh, A. R. (2007). Explicit versus implicit gaze processing assessed by ERPs. , *1177*, 79–89. Retrieved 2023-01-10, from https://www.sciencedirect.com/science/article/pii/S0006899307016460 doi: 10.1016/j.brainres.2007.07.094

Jasper, H. (1958). The twen twenty electrode system of the international federation. , *10*(2), 371–375. doi: https://doi.org/10.1016/0013-4694(58)90053-1

Jones, E. A., & Carr, E. G. (2004). Joint attention in children with autism: Theory and intervention. , *19*(1), 13–26. Retrieved 2021-03-19, from https://doi.org/10.1177/10883576040190010301 (Publisher: SAGE Publications Inc) doi: 10.1177/10883576040190010301

Jyoti, V., & Lahiri, U. (2020). Virtual reality based joint attention task platform for children with autism. , *13*(1), 198–210. Retrieved 2020-06-28, from https://ieeexplore.ieee.org/document/8697116/ doi: 10.1109/TLT.2019.2912371

Kampis, D., & Southgate, V. (2020). Altercentric cognition: How others influence our cognitive processing. , *24* (11), 945–959. Retrieved 2023-04-23, from https://www.sciencedirect.com/science/article/pii/S1364661320302175 doi: 10.1016/j.tics.2020.09.003

Kasari, C., Gulsrud, A., Freeman, S., Paparella, T., & Hellemann, G. (2012). Longitudinal follow-up of children with autism receiving targeted interventions on joint attention and play. , *51* (5), 487–495. Retrieved 2020-06-18, from https://linkinghub.elsevier.com/retrieve/pii/S0890856712001451 doi: 10.1016/j.jaac.2012.02.019

Khader, P. H., Jost, K., Ranganath, C., & Rösler, F. (2010). Theta and alpha oscillations during working-memory maintenance predict successful long-term memory encoding. , *468* (3), 339–343. Retrieved 2023-04-27, from https://www.sciencedirect.com/science/article/pii/S0304394009014943 doi: 10.1016/j.neulet.2009.11.028

Khowaja, M., & Robins, D. (2013). Checklist for autism in toddlers (CHAT). In F. R. Volkmar (Ed.), *Encyclopedia of autism spectrum disorders* (pp. 571–575). Springer New York. Retrieved 2021-09-25, from http://link.springer.com/10.1007/978-1-4419-1698-$3_2$63 doi: 10.1007/978-1-4419-1698-$3_2$63

Kiili, K. (2007). Foundation for problem-based gaming. , *38* (3), 394–404. Retrieved 2023-04-29, from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8535.2007.00704.x (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-8535.2007.00704.x) doi: 10.1111/j.1467-8535.2007.00704.x

Kim, G., Buntain, N., Hirshfield, L., Costa, M. R., & Chock, T. M. (2019). Processing racial stereotypes in virtual reality: An exploratory study using functional near-infrared spectroscopy (fNIRS). In D. D. Schmorrow & C. M. Fidopiastis (Eds.), *Augmented cognition* (pp. 407–417). Springer International Publishing. doi: 10.1007/978-3-030-22419-$6_2$9

Kim, S., & Kim, E. (2020). The use of virtual reality in psychiatry: A review. , *31* (1), 26–32. Retrieved 2023-04-29, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7324842/ doi: 10.5765/jkacap.190037

Kompatsiari, K., Bossi, F., & Wykowska, A. (2021). Eye contact during joint attention with a humanoid robot modulates oscillatory brain activity. , *16* (4), 383–392. Retrieved 2022-04-13, from https://doi.org/10.1093/scan/nsab001 doi: 10.1093/scan/nsab001

Krauss, R. M., & Glucksberg, S. (1977). Social and nonsocial speech. , *236*(2), 100–105. Retrieved 2022-10-07, from https://www.jstor.org/stable/24953896 (Publisher: Scientific American, a division of Nature America, Inc.)

Kristen, S., Sodian, B., Thoermer, C., & Perst, H. (2011). Infants' joint attention skills predict toddlers' emerging mental state language. , *47*(5), 1207–1219. Retrieved 2022-05-13, from https://www.proquest.com/docview/889451684/abstract/2E3C4821F9104E8EPQ/1 (Num Pages: 1207-1219 Place: Washington, US Publisher: American Psychological Association (US)) doi: https://doi.org/10.1037/a0024808

Kuhlen, A. K., & Brennan, S. E. (2013). Language in dialogue: when confederates might be hazardous to your data. , *20*(1), 54–72. Retrieved 2022-09-12, from https://doi.org/10.3758/s13423-012-0341-8 doi: 10.3758/s13423-012-0341-8

Lachat, F., Hugeville, L., Lemarechal, J.-D., Conty, L., & George, N. (2012). Oscillatory brain correlates of live joint attention: A dual-EEG study. , *6*, 156. Retrieved 2021-10-13, from https://www.frontiersin.org/article/10.3389/fnhum.2012.00156 doi: 10.3389/fnhum.2012.00156

Laidlaw, K. E. W., Foulsham, T., Kuhn, G., & Kingstone, A. (2011). Potential social interactions are important to social attention. , *108*(14), 5548–5553. Retrieved 2021-04-07, from http://www.pnas.org/cgi/doi/10.1073/pnas.1017022108 doi: 10.1073/pnas.1017022108

Lasagna, C. A., Grove, T. B., Semple, E., Suzuki, T., Menkes, M. W., Pamidighantam, P., ... Tso, I. F. (2023). Reductions in regional theta power and fronto-parietal theta-gamma phase-amplitude coupling during gaze processing in bipolar disorder. , *331*, 111629. Retrieved 2023-04-23, from https://www.sciencedirect.com/science/article/pii/S0925492723000392 doi: 10.1016/j.pscychresns.2023.111629

Laver, K., George, S., Thomas, S., Deutsch, J. E., & Crotty, M. (2012). Virtual reality for stroke rehabilitation. , *43*(2), e20–e21. Retrieved 2023-04-17, from https://www.ahajournals.org/doi/full/10.1161/STROKEAHA.111.642439 (Publisher: American Heart Association) doi: 10.1161/STROKEAHA.111.642439

Leavitt, K., Qiu, F., & Shapiro, D. L. (2021). Using electronic confederates for experimental research in organizational science. *Organizational Research Methods*, *24*(1), 3–25.

Lee, Y.-G., & Lyons, K. W. (2004). Smoothing haptic interaction using molecular force calculations. , *36*(1), 75–90. Retrieved 2020-06-29, from https://linkinghub.elsevier.com/retrieve/pii/S0010448503000800 doi: 10.1016/S0010-

4485(03)00080-0

Leong, V., Byrne, E., Clackson, K., Georgieva, S., Lam, S., & Wass, S. (2017). Speaker gaze increases information coupling between infant and adult brains. , *114* (50), 13290–13295. Retrieved 2023-02-03, from https://www.pnas.org/doi/full/10.1073/pnas.1702493114 (Publisher: Proceedings of the National Academy of Sciences) doi: 10.1073/pnas.1702493114

Lewy, A. L., & Dawson, G. (1992). Social stimulation and joint attention in young autistic children. , *20* (6), 555–566. Retrieved 2021-09-25, from http://link.springer.com/10.1007/BF00911240 doi: 10.1007/BF00911240

Limniou, M., Roberts, D., & Papadopoulos, N. (2008). Full immersive virtual environment CAVETM in chemistry education. , *51* (2), 584–593. Retrieved 2020-06-29, from https://linkinghub.elsevier.com/retrieve/pii/S0360131507000747 doi: 10.1016/j.compedu.2007.06.014

Lindner, P., Miloff, A., Fagernäs, S., Andersen, J., Sigeman, M., Andersson, G., ... Carlbring, P. (2019). Therapist-led and self-led one-session virtual reality exposure therapy for public speaking anxiety with consumer hardware and software: A randomized controlled trial. , *61*, 45–54. Retrieved 2020-06-19, from https://linkinghub.elsevier.com/retrieve/pii/S0887618517306321 doi: 10.1016/j.janxdis.2018.07.003

Lopes da Silva, F. (2013). EEG and MEG: Relevance to neuroscience. , *80* (5), 1112–1128. Retrieved 2020-06-18, from https://linkinghub.elsevier.com/retrieve/pii/S0896627313009203 doi: 10.1016/j.neuron.2013.10.017

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., Leventhal, B. L., DiLavore, P. C., ... Rutter, M. (2000). The autism diagnostic observation schedule generic: A standard measure of social and communication de[U+FB01]cits associated with the spectrum of autism. , *30* (3), 205–223. Retrieved 2020-04-30, from http://link.springer.com/10.1023/A:1005592401947 doi: 10.1023/A:1005592401947

Marsh, K. L., Richardson, M. J., & Schmidt, R. C. (2009). Social connection through joint action and interpersonal coordination. , *1* (2), 320–339. Retrieved 2023-02-03, from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1756-8765.2009.01022.x (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1756-8765.2009.01022.x) doi: 10.1111/j.1756-8765.2009.01022.x

McDonald, K., Newby-Clark, I. R., Walker, J., & Henselwood, K. (2018). It is written all over your face: Socially rejected people display microexpressions that are detectable after training in the micro expression training tool (METT). , *48*(2), O175–O181. Retrieved 2023-04-27, from https://onlinelibrary.wiley.com/doi/abs/10.1002/ejsp.2301 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ejsp.2301) doi: 10.1002/ejsp.2301

McPartland, J., Cheung, C. H. M., Perszyk, D., & Mayes, L. C. (2010). Face-related ERPs are modulated by point of gaze. , *48*(12), 3657–3660. Retrieved 2023-01-10, from https://www.sciencedirect.com/science/article/pii/S0028393210003155 doi: 10.1016/j.neuropsychologia.2010.07.020

Mei, C., Zahed, B. T., Mason, L., & Quarles, J. (2018). Towards joint attention training for children with ASD - a VR game approach and eye gaze exploration. In *2018 IEEE conference on virtual reality and 3d user interfaces (VR)* (pp. 289–296). IEEE. Retrieved 2020-03-25, from https://ieeexplore.ieee.org/document/8446242/ doi: 10.1109/VR.2018.8446242

Meindl, J. N., & Cannella-Malone, H. I. (2011). Initiating and responding to joint attention bids in children with autism: A review of the literature. , *32*(5), 1441–1454. Retrieved 2021-06-22, from https://linkinghub.elsevier.com/retrieve/pii/S0891422211000886 doi: 10.1016/j.ridd.2011.02.013

Meltzoff, A. N., & Brooks, R. (2008). Self-experience as a mechanism for learning about others: A training study in social cognition. , *44*(5), 1257–1265. Retrieved 2021-09-29, from http://doi.apa.org/getdoi.cfm?doi=10.1037/a0012888 doi: 10.1037/a0012888

Moore, C., Dunham, P. J., & Dunham, P. (2014). *Joint attention: Its origins and role in development.* Psychology Press.

Morales, M., Mundy, P., Crowson, M. M., Neal, A. R., & Delgado, C. E. F. (2005). Individual differences in infant attention skills, joint attention, and emotion regulation behaviour. , *29*(3), 259–263. Retrieved 2020-06-18, from http://journals.sagepub.com/doi/10.1177/01650250444000432 doi: 10.1177/01650250444000432

Morales, M., Mundy, P., Delgado, C. E. F., Yale, M., Messinger, D., Neal, R., & Schwartz, H. K. (2000). Responding to joint attention across the 6- through 24-month age period and early language acquisition. , *21*(3), 283–298. Retrieved 2021-10-25, from https://www.sciencedirect.com/science/article/pii/S0193397399000404 doi: 10.1016/S0193-3973(99)00040-4

Morgan, E. J., Smith, D. T., & Freeth, M. (2021). Gaze cueing, mental states, and the effect of autistic traits. *Attention, Perception, & Psychophysics*, 1–9.

Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley [from the field]. , *19*(2), 98–100.

(Conference Name: IEEE Robotics Automation Magazine) doi: 10.1109/MRA.2012.2192811

Mundy, P., & Jarrold, W. (2010). Infant joint attention, neural networks and social cognition. , *23*(8), 985–997. Retrieved 2020-01-09, from `https://linkinghub.elsevier.com/retrieve/pii/S0893608010001607` doi: 10.1016/j.neunet.2010.08.009

Mundy, P., Kasari, C., Sigman, M., & Ruskin, E. (1995). Nonverbal communication and early language acquisition in children with down syndrome and in normally developing children. *Journal of Speech, Language, and Hearing Research*, *38*(1), 157–167.

Mundy, P., & Newell, L. (2007). Attention, joint attention, and social cognition. , *16*(5), 269–274. Retrieved 2020-03-20, from `http://journals.sagepub.com/doi/10.1111/j.1467-8721.2007.00518.x` doi: 10.1111/j.1467-8721.2007.00518.x

Murman, D. L. (2015). The impact of age on cognition. , *36*(3), 111–121. Retrieved 2023-04-15, from `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4906299/` doi: 10.1055/s-0035-1555115

Murray, D. S., Creaghead, N. A., Manning-Courtney, P., Shear, P. K., Bean, J., & Prendeville, J.-A. (2008). The relationship between joint attention and language in children with autism spectrum disorders. , *23*(1), 5–14. Retrieved 2021-10-18, from `http://journals.sagepub.com/doi/10.1177/1088357607311443` doi: 10.1177/1088357607311443

Newson, J., & Newson, E. (1975). Intersubjectivity and the transmission of culture: on the social origins of symbolic functioning. , *28*, 437–446.

Nichols, A. L., & Maner, J. K. (2008). The good-subject effect: Investigating participant demand characteristics. *The Journal of general psychology*, *135*(2), 151–166.

Niforatos, E., Palma, A., Gluszny, R., Vourvopoulos, A., & Liarokapis, F. (2020). Would you do it?: Enacting moral dilemmas in virtual reality for understanding ethical decision-making. , 12.

Nuwer, M. R. (2018). 10-10 electrode system for EEG recording. , *129*(5), 1103. Retrieved 2023-04-22, from `https://www.sciencedirect.com/science/article/pii/S1388245718300907` doi: 10.1016/j.clinph.2018.01.065

Nuwer, M. R., Comi, G., Emerson, R., Fuglsang-Frederiksen, A., Guérit, J.-M., Hinrichs, H., . . . Rappelsburger, P. (1998). IFCN standards for digital recording of clinical EEG. , *106*(3), 259–261. Retrieved 2023-04-22, from `https://www.sciencedirect.com/science/article/pii/S0013469497001065` doi:

10.1016/S0013-4694(97)00106-5

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. , *2011*, 1–9. Retrieved 2022-12-15, from http://www.hindawi.com/journals/cin/2011/156869/ doi: 10.1155/2011/156869

Parker, T., Crowley, M., Naples, A., Rolison, M., Wu, J., Trapani, J., & McPartland, J. (2021). The n170 event[U+2010]related potential reflects delayed neural response to faces when visual attention is directed to the eyes in youths with ASD. , *14*. doi: 10.1002/aur.2505

Parker, T. C., & Ricard, J. A. (2022). Structural racism in neuroimaging: perspectives and solutions. , *9*(5), e22. (Publisher: Elsevier) doi: 10.1016/S2215-0366(22)00079-7

Patel, G. H., Arkin, S. C., Ruiz-Betancourt, D. R., Plaza, F. I., Mirza, S. A., Vieira, D. J., ... Javitt, D. C. (2021). Failure to engage the temporoparietal junction/posterior superior temporal sulcus predicts impaired naturalistic social cognition in schizophrenia. , *144*(6), 1898–1910. Retrieved 2023-04-23, from https://academic.oup.com/brain/article/144/6/1898/6168992 doi: 10.1093/brain/awab081

Patel, G. H., Sestieri, C., & Corbetta, M. (2019). The evolution of the temporoparietal junction and posterior superior temporal sulcus. , *118*, 38–50. Retrieved 2023-04-23, from https://www.sciencedirect.com/science/article/pii/S0010945219300462 doi: 10.1016/j.cortex.2019.01.026

Penner, F., Wall, K. M., Guan, K. W., Huang, H. J., Richardson, L., Dunbar, A. S., ... Rutherford, H. J. V. (2023). Racial disparities in EEG research and their implications for our understanding of the maternal brain. , *23*(1), 1–16. Retrieved 2023-04-20, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9684773/ doi: 10.3758/s13415-022-01040-w

Petit, D., Gagnon, J.-F., Fantini, M. L., Ferini-Strambi, L., & Montplaisir, J. (2004). Sleep and quantitative EEG in neurodegenerative disorders. , *56*(5), 487–496. Retrieved 2023-04-20, from https://www.sciencedirect.com/science/article/pii/S0022399904000182 doi: 10.1016/j.jpsychores.2004.02.001

Pfurtscheller, G., Brunner, C., Schlögl, A., & Lopes da Silva, F. H. (2006). Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks. , *31*(1), 153–159. Retrieved 2023-04-23, from https://www.sciencedirect.com/science/article/pii/S1053811905025140 doi:

10.1016/j.neuroimage.2005.12.003

Pfurtscheller, G., Neuper, C., & Krausz, G. (2000). Functional dissociation of lower and upper frequency mu rhythms in relation to voluntary limb movement. , *111* (10), 1873–1879. Retrieved 2023-04-26, from https://www.sciencedirect.com/science/article/pii/S1388245700004284 doi: 10.1016/S1388-2457(00)00428-4

Pineda, J. A. (2005). The functional significance of mu rhythms: Translating "seeing" and "hearing" into "doing". , *50* (1), 57–68. Retrieved 2023-04-26, from https://www.sciencedirect.com/science/article/pii/S0165017305000573 doi: 10.1016/j.brainresrev.2005.04.005

Povinelli, D. J., & Vonk, J. (2003). Chimpanzee minds: suspiciously human? , *7* (4), 157–160. Retrieved 2023-02-09, from https://www.sciencedirect.com/science/article/pii/S1364661303000536 doi: 10.1016/S1364-6613(03)00053-6

R Core Team. (2021). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. Retrieved from URL https://www.R-project.org/.

Rey, B., & Alcañiz, M. (2010). Research in neuroscience and virtual reality.

Richardson. (2021). *EEG hair project - hello brain lab: the UCF BRaIN lab.* Retrieved 2023-04-20, from https://hellobrainlab.com/research/eeg-hair-project/

Richesin, M., Baldwin, D., & Wicks, L. (2022). The role of contextual information in a virtual trolly problem: A psychophysiological investigation. , *17* (3). (ISSN: 1747-0919) doi: https://doi.org/10.1080/17470919.2022.2076733

Risko, E. F., Laidlaw, K. E. W., Freeth, M., Foulsham, T., & Kingstone, A. (2012). Social attention with real versus reel stimuli: toward an empirical approach to concerns about ecological validity. , *6*, 143. Retrieved 2021-09-30, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3360477/ doi: 10.3389/fnhum.2012.00143

Risko, E. F., Richardson, D. C., & Kingstone, A. (2016). Breaking the fourth wall of cognitive science: Real-world social attention and the dual function of gaze. , *25* (1), 70–74. Retrieved 2021-09-30, from https://doi.org/10.1177/0963721415617806 (Publisher: SAGE Publications Inc) doi: 10.1177/0963721415617806

Ristic, J., Friesen, C. K., & Kingstone, A. (2002). Are eyes special? it depends on how you look at it. , *9* (3), 507–513. Retrieved 2023-04-23, from https://doi.org/10.3758/BF03196306 doi: 10.3758/BF03196306

Roberts, G., Holmes, N., Alexander, N., Boto, E., Leggett, J., Hill, R. M., ... Brookes, M. J. (2019). Towards OPM-MEG in a virtual reality environment. , *199*, 408–417. Retrieved 2020-06-18, from https://linkinghub.elsevier.com/retrieve/pii/S1053811919304951 doi: 10.1016/j.neuroimage.2019.06.010

Rose, V., Trembath, D., Keen, D., & Paynter, J. (2016). The proportion of minimally verbal children with autism spectrum disorder in a community-based early intervention programme. , *60*(5), 464–477. doi: 10.1111/jir.12284

Roth, S. D. (1982). *Ray casting as a method for solid modeling.* General Motors Corporation, Research Laboratories Division.

Rubio-Fernández, P. (2017). The director task: A test of theory-of-mind use or selective attention? , *24*(4), 1121–1128. Retrieved 2020-06-26, from http://link.springer.com/10.3758/s13423-016-1190-7 doi: 10.3758/s13423-016-1190-7

Ruhland, K., Peters, C. E., Andrist, S., Badler, J. B., Badler, N. I., Gleicher, M., ... McDonnell, R. (2015). A review of eye gaze in virtual agents, social robotics and HCI: Behaviour generation, user interaction and perception. , *34*(6), 299–326. Retrieved 2021-04-07, from https://onlinelibrary.wiley.com/doi/10.1111/cgf.12603 (Publisher: John Wiley and Sons, Ltd) doi: 10.1111/cgf.12603

Salazar, M., Shaw, D., Gajdoš, M., Mareček, R., Czekóová, K., Mikl, M., & Brázdil, M. (2021). You took the words right out of my mouth: Dual-fMRI reveals intra- and inter-personal neural processes supporting verbal interaction. , *228*(117697). Retrieved 2023-02-23, from http://www.scopus.com/inward/record.url?scp=85098970750partnerID=8YFLogxK doi: 10.1016/j.neuroimage.2020.117697

Sanko, J. S., Shekhter, I., Kyle Jr, R. R., Di Benedetto, S., & Birnbach, D. J. (2013). Establishing a convention for acting in healthcare simulation: merging art and science. *Simulation in Healthcare*, *8*(4), 215–220.

Santos, L. R., & Hauser, M. D. (1999). How monkeys see the eyes: cotton-top tamarins' reaction to changes in visual attention and action. , *2*(3), 131–139. Retrieved 2023-02-09, from http://link.springer.com/10.1007/s100710050033 doi: 10.1007/s100710050033

Sauseng, P., Klimesch, W., Stadler, W., Schabus, M., Doppelmayr, M., Hanslmayr, S., ... Birbaumer, N. (2005). A shift of visual spatial attention is selectively associated with human EEG alpha activity. , *22*(11), 2917–2926. Retrieved 2023-04-22, from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1460-9568.2005.04482.x (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1460-9568.2005.04482.x) doi: 10.1111/j.1460-9568.2005.04482.x

Scaife, M., & Bruner, J. (1975). The capacity for joint visual attention in the infant. , *253*(5489), 265–266. Retrieved 2020-04-30, from http://www.nature.com/articles/253265a0 doi: 10.1038/253265a0

Scarapicchia, V., Brown, C., Mayo, C., & Gawryluk, J. R. (2017). Functional magnetic resonance imaging and functional near-infrared spectroscopy: Insights from combined recording studies. , *11*, 419. doi: 10.3389/fnhum.2017.00419

Schilbach, L. (2015). Eye to eye, face to face and brain to brain: novel approaches to study the behavioral dynamics and neural mechanisms of social interactions. , *3*, 130–135. Retrieved 2021-04-07, from https://www.sciencedirect.com/science/article/pii/S2352154615000492 doi: 10.1016/j.cobeha.2015.03.006

Senju, A., & Hasegawa, T. (2005). Direct gaze captures visuospatial attention. , *12*(1), 127–144. Retrieved 2022-04-13, from http://www.tandfonline.com/doi/abs/10.1080/13506280444000157 doi: 10.1080/13506280444000157

Senju, A., Hasegawa, T., & Tojo, Y. (2005). Does perceived direct gaze boost detection in adults and children with and without autism? the stare-in-the-crowd effect revisited. , *12*(8), 1474–1496. Retrieved 2022-04-13, from http://www.tandfonline.com/doi/abs/10.1080/13506280444000797 doi: 10.1080/13506280444000797

Seuren, L. M., Wherton, J., Greenhalgh, T., & Shaw, S. E. (2021). Whose turn is it anyway? latency and the organization of turn-taking in video-mediated interaction. , *172*, 63–78. Retrieved 2023-02-03, from https://www.sciencedirect.com/science/article/pii/S0378216620302782 doi: 10.1016/j.pragma.2020.11.005

Shaw, J. A., Bryant, L. K., Malle, B. F., Povinelli, D. J., & Pruett, J. R. (2017). The relationship between joint attention and theory of mind in neurotypical adults. , *51*, 268–278. Retrieved 2022-05-13, from https://linkinghub.elsevier.com/retrieve/pii/S1053810016303920 doi: 10.1016/j.concog.2017.02.012

Shepherd. (2010). Following gaze: gaze-following behavior as a window into social cognition. Retrieved 2021-06-22, from http://journal.frontiersin.org/article/10.3389/fnint.2010.00005/abstract doi:

10.3389/fnint.2010.00005

Shteynberg, G. (2010). A silent emergence of culture: The social tuning effect. , *99*, 683–689. (Place: US Publisher: American Psychological Association) doi: 10.1037/a0019573

Singh, L., & Chee, M. (2016). Rise and fall: Effects of tone and intonation on spoken word recognition in early childhood. , *55*, 109–118. Retrieved 2023-03-10, from https://www.sciencedirect.com/science/article/pii/S0095447015001096 doi: 10.1016/j.wocn.2015.12.005

Sipatchin, A., Wahl, S., & Rifai, K. (2021). Eye-tracking for clinical ophthalmology with virtual reality (VR): A case study of the HTC vive pro eye's usability. , *9*(2), 180. Retrieved 2023-04-15, from https://www.mdpi.com/2227-9032/9/2/180 doi: 10.3390/healthcare9020180

Siposova, B., & Carpenter, M. (2019). A new look at joint attention and common knowledge. , *189*, 260–274. Retrieved 2021-07-08, from https://linkinghub.elsevier.com/retrieve/pii/S0010027719300745 doi: 10.1016/j.cognition.2019.03.019

Slater, M. (2003). A note on presence terminology.

Smith, S. J. M. (2005). EEG in the diagnosis, classification, and management of patients with epilepsy. , *76*, ii2–ii7. Retrieved 2023-04-20, from https://jnnp.bmj.com/content/76/suppl$_2$/ii2 (Publisher: BMJ Publishing Group Ltd) doi: 10.1136/jnnp.2005.069245

Smith, T., & Mital, P. K. (2013). Attentional synchrony and the influence of viewing task on gaze behavior in static and dynamic scenes. , *13*(8), 16. Retrieved 2023-02-23, from https://doi.org/10.1167/13.8.16 doi: 10.1167/13.8.16

Spezio, M. L., Adolphs, R., Hurley, R. S., & Piven, J. (2007). Abnormal use of facial information in high-functioning autism. , *37*(5), 929–939. Retrieved 2020-06-19, from http://link.springer.com/10.1007/s10803-006-0232-9 doi: 10.1007/s10803-006-0232-9

Stephenson, L. J., Edwards, S. G., & Bayliss, A. P. (2021). From gaze perception to social cognition: The shared-attention system. , *16*(3), 553–576. Retrieved 2023-02-06, from https://doi.org/10.1177/1745691620953773 (Publisher: SAGE Publications Inc) doi: 10.1177/1745691620953773

Sterling, L., Dawson, G., Webb, S., Murias, M., Munson, J., Panagiotides, H., & Aylward, E. (2008). The role of face familiarity in eye tracking of faces by individuals with autism spectrum disorders. , *38*(9), 1666–1675. doi: 10.1007/s10803-008-0550-1

Stevanovic, M., & Peräkylä, A. (2015). Experience sharing, emotional reciprocity, and turn-taking. , *6*. Retrieved 2023-02-03, from https://www.frontiersin.org/articles/10.3389/fpsyg.2015.00450

Sun, Y., Greaves, D. A., Orgs, G., de C. Hamilton, A. F., Day, S., & Ward, J. A. (2023). Using wearable sensors to measure interpersonal synchrony in actors and audience members during a live theatre performance. , *7*(1), 27:1–27:29. Retrieved 2023-04-24, from https://dl.acm.org/doi/10.1145/3580781 doi: 10.1145/3580781

Swanson, M. R., & Siller, M. (2013). Patterns of gaze behavior during an eye-tracking measure of joint attention in typically developing children and children with autism spectrum disorder. , *7*(9), 1087–1096. Retrieved 2021-03-22, from https://www.sciencedirect.com/science/article/pii/S1750946713001098 doi: 10.1016/j.rasd.2013.05.007

Tager-Flusberg, H. (2007). Evaluating the theory-of-mind hypothesis of autism. , *16*, 311–315. (Place: United Kingdom Publisher: Blackwell Publishing) doi: 10.1111/j.1467-8721.2007.00527.x

Theuring, C., Gredebäck, G., & Hauf, P. (2007). Object processing during a joint gaze following task. , *4*(1), 65–79. Retrieved 2021-10-25, from http://www.tandfonline.com/doi/abs/10.1080/17405620601051246 doi: 10.1080/17405620601051246

Tierney, T. M., Holmes, N., Mellor, S., López, J. D., Roberts, G., Hill, R. M., ... Barnes, G. R. (2019). Optically pumped magnetometers: From quantum origins to multi-channel magnetoencephalography. , *199*, 598–608. Retrieved 2023-04-29, from https://www.sciencedirect.com/science/article/pii/S1053811919304550 doi: 10.1016/j.neuroimage.2019.05.063

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. , *28*(5), 675–691. doi: 10.1017/S0140525X05000129

Tu, S., Spiers, H. J., Hodges, J. R., Piguet, O., & Hornberger, M. (2017). Egocentric versus allocentric spatial memory in behavioral variant frontotemporal dementia and alzheimer's disease. , *59*(3), 883–892. Retrieved 2023-04-17, from https://content.iospress.com/articles/journal-of-alzheimers-disease/jad160592 (Publisher: IOS Press) doi: 10.3233/JAD-160592

Unity3D. (2019). *Unity technologies.*

VandenBos, G. R. (2007). APA dictionary of psychology.

Vaughan Van Hecke, A., Mundy, P., Block, J., Delgado, C. E. F., Parlade, M., Pomarea, Y. B., & Hobsson, J. A. (2012). Infant responding to joint attention, executive processes, and self-regulation in preschool children | elsevier enhanced reader. , *35*(2), 303–311. doi: 10.1016/j.infbeh.2011.12.001

Volkmar, F. R., Lord, C., Bailey, A., Schultz, R. T., & Klin, A. (2004). Autism and pervasive developmental disorders. , *45*(1), 135–170. Retrieved 2023-02-23, from `https://onlinelibrary.wiley.com/doi/abs/10.1046/j.0021-9630.2003.00317.x` (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1046/j.0021-9630.2003.00317.x) doi: 10.1046/j.0021-9630.2003.00317.x

Vuilleumier, P., & Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: Evidence from functional neuroimaging. , *45*(1), 174–194. Retrieved 2023-02-21, from `https://www.sciencedirect.com/science/article/pii/S0028393206002296` doi: 10.1016/j.neuropsychologia.2006.06.003

Võ, M. L.-H., Smith, T. J., Mital, P. K., & Henderson, J. M. (2012). Do the eyes really have it? dynamic allocation of attention when viewing moving faces. , *12*(13), 3. Retrieved 2023-02-23, from `https://doi.org/10.1167/12.13.3` doi: 10.1167/12.13.3

Wahl, S., Michel, C., Pauen, S., & Hoehl, S. (2013). Head and eye movements affect object processing in 4-month-old infants more than an artificial orientation cue. , *31*(2), 212–230. Retrieved 2023-04-23, from `https://onlinelibrary.wiley.com/doi/abs/10.1111/bjdp.12001` (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/bjdp.12001) doi: 10.1111/bjdp.12001

Wang, H., Callaghan, E., Gooding-Williams, G., McAllister, C., & Kessler, K. (2016). Rhythm makes the world go round: An MEG-TMS study on the role of right TPJ theta oscillations in embodied perspective taking. , *75*, 68–81. Retrieved 2023-04-23, from `https://www.sciencedirect.com/science/article/pii/S0010945215003895` doi: 10.1016/j.cortex.2015.11.011

Wang, X., Xing, W., & Laffey, J. M. (2018). Autistic youth in 3d game-based collaborative virtual learning: Associating avatar interaction patterns with embodied social presence: Autistic youth in 3d game-based learning. , *49*(4), 742–760. Retrieved 2020-06-19, from `http://doi.wiley.com/10.1111/bjet.12646` doi: 10.1111/bjet.12646

Ward, L. M. (2003). Synchronous neural oscillations and cognitive processes. , *7*(12), 553–559. doi: 10.1016/j.tics.2003.10.012

Webb, E. K., Etter, J. A., & Kwasa, J. A. (2022). Addressing racial and phenotypic bias in human neuroscience methods. , *25*(4), 410–414. Retrieved 2023-04-26, from `https://www.nature.com/articles/s41593-022-01046-0` (Number: 4 Publisher: Nature

Publishing Group) doi: 10.1038/s41593-022-01046-0

Weber, D., Hertweck, S., Alwanni, H., Fiederer, L. D. J., Wang, X., Unruh, F., ... Ball, T. (2021). A structured approach to test the signal quality of electroencephalography measurements during use of head-mounted displays for virtual reality applications. , *15*. Retrieved 2023-02-17, from https://www.frontiersin.org/articles/10.3389/fnins.2021.733673

Wei, Y., Chen, Q., Curtin, A., Tu, L., Tang, X., Tang, Y., ... Wang, J. (2020). Functional near-infrared spectroscopy (fNIRS) as a tool to assist the diagnosis of major psychiatric disorders in a chinese population. Retrieved 2020-06-18, from http://link.springer.com/10.1007/s00406-020-01125-y doi: 10.1007/s00406-020-01125-y

Whalen, C., Schreibman, L., & Ingersoll, B. (2006). The collateral effects of joint attention training on social initiations, positive affect, imitation, and spontaneous speech for young children with autism. , *36*(5), 655–664. Retrieved 2021-10-18, from http://link.springer.com/10.1007/s10803-006-0108-z doi: 10.1007/s10803-006-0108-z

Wilms, M., Schilbach, L., Pfeiffer, U., Bente, G., Fink, G. R., & Vogeley, K. (2010). It's in your eyes—using gaze-contingent stimuli to create truly interactive paradigms for social cognitive and affective neuroscience. , *5*(1), 98–107. Retrieved 2020-03-30, from https://academic.oup.com/scan/article/5/1/98/1732514 doi: 10.1093/scan/nsq024

Wilson, C. J., & Soranzo, A. (2015). The use of virtual reality in psychology: A case study in visual perception. , *2015*, e151702. Retrieved 2023-04-29, from https://www.hindawi.com/journals/cmmm/2015/151702/ (Publisher: Hindawi) doi: 10.1155/2015/151702

Witmer, B., & Singer, M. (1998). Measuring presence in virtual environments. , *7*(3). doi: https://doi.org/10.1162/105474698565686

Zeitrâg, C., Jensen, T. R., & Osvath, M. (2022). Gaze following: A socio-cognitive skill rooted in deep time. , *13*. Retrieved 2023-03-10, from https://www.frontiersin.org/articles/10.3389/fpsyg.2022.950935

Appendix A

Supporting documents

The following appendices are a collection of documents that were used to aid the work undertaken in this thesis. Documents include a participant information sheet, a poster to recruit volunteers, consent forms, standard operating procedures, a protocol, debrief sheets and ethical committee confirmation.

Puzzle Joint Attention Task in Virtual Reality

Participant Information Sheet

Invitation

We 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 a member of the research team, whose 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?

Joint attention is a process/action that we perform on a daily basis. It is common for individuals with disorders such as Autism to demonstrate difficulties in performing such actions. Understanding how this is characterised in everyday scenarios with typical individuals, will provide us with a basis to further understand how this is expressed in atypical individuals. Using virtual reality to display scenarios provides a modern and novel approach to tackling this problem in the lab.

Why have I been chosen?

You are being invited to participate because you have expressed an interest in the study from recruitment advertisement. We require a group of healthy volunteers from the public, who should meet the following requirements:

- You are between 18-40 years of age.
- No history of epileptic seizures.
- Have normal, or normal to corrected vision. (If you wear glasses, please let us know in advance – you can either wear contacts or we will need to check your glasses frame size inside of the HMD).
- Do not have a diagnosis of photosensitive epilepsy, traumatic brain injury or hypertension.
- Are not sensitive to motion sickness and do not experience ‘cyber sickness’.
- Must be able to understand verbal and written information presented in English

REC/IRAS ID: [#1637], [V3], [24-01-2022]

What will happen to me if I take part?

Adhering to the government guidelines, all participants and experimenters are required to wear a mask throughout the task

We will arrange for you to come to the Aston Laboratory for Immersive Virtual Environments (ALIVE), situated in the Main Building, Aston University, Birmingham. The study requires approximately one hour of testing (including setup). You will be given a choice of times between the working hours of 09:00 – 17:00 on a given week day. Together, you and the researcher will find a time that suits you both. On the day, the task set up will go as follows:

1. Virtual reality and task practice (10 mins)
2. Electroencephalography set up (30-40 mins)
3. Main Puzzle Joint attention task practice and ASQ questionnaire (1 hour)

Practice Tasks

We want you to be comfortable in VR, especially if you have not used it before. Throughout the practice task and the main task, you will be seated. You will have time to try out the equipment and practice the main task to ensure you feel comfortable and know what you have to do. We will be assisting you when putting the headset on and will answer any questions you may have.

Virtual Reality Puzzle Joint Attention Main Task with EEG and VR.

Before the main experiment, you will be asked to complete the Autism Spectrum Quotient Questionnaire. This is **not** a diagnostic test. The results of the questionnaire will not be shared with the participant. If you have any concerns, please contact your GP.

After familiarising yourself with the virtual reality equipment and completing the questionnaire, we will begin preparing you to wear the electroencephalography (EEG) cap that you will wear throughout the entire task. For the EEG, we will measure your head circumference to fit the EEG cap. The cap will then be placed on your head and adjusted into position by the researcher. Once the EEG cap is on, gel is injected into the electrodes on the cap using blunt syringes (you will be able to wash this out of your hair at the end of the study). This creates conductive gel channels to your scalp so that we can detect electrical activity on your scalp. This will take approximately 30-40 minutes to set up.

Once this is done, you will wear a HMD (head mounted display) and will be presented with main part of the experiment.

Once you are comfortable, you will remain seated and you can begin the main task where you will be presented with a series of scenarios for you to solve. The virtual environment and virtual human will remain consistent, however, their behaviour will vary These behaviours will be explained to you at the beginning of the experiment. Either you or the virtual human will be presented with a piece of a puzzle. When you are presented with the puzzle piece, your task is to communicate whose puzzle board the piece belongs to. Likewise, when the virtual human is presented with the puzzle piece, they will communicate where it belongs and you will be tasked to figure out whether that piece is for your puzzle board or for them to keep. Half way through the experiment, you will be offered and encouraged to take a break. If at any point during the experiment, you would like to stop, please do so by letting the researcher know.

REC/IRAS ID: [#1637], [V3], [24-01-2022]

VR Equipment

HTC Vive Pro Eye

The HTC Vive Pro Eye is a Head Mounted Display, commonly used in the gaming industry. It allows us to present an image to each eye separately to create a compelling 3D visual scene. When using the HTC Vive Pro Eye, you will not be able to see your own body, but you will be able to see virtual hands.

What is EEG?

EEG, short for electroencephalography, literally means "the descriptive science/recording ("-graphy") of electricity ("electro-") relating to the brain ("encephalo-"). It is a common, non-invasive technique used for recording brain activity.

Participants wear a tight cap with many sensors which need to touch the scalp. Direct contact is not always possible so we apply a small amount of conductive gel to each sensor. This is done using a blunt syringe with which we part the hair to the scalp and squeeze the gel. Due to the cap we currently use, it may not be possible to collect data from participants with particular hairstyles, such as recently plaited braids, recently twisted dreadlocks any head dresses or canerows/cornrows as the sensors may not be able to touch the scalp. If braids or dreadlocks are older and have some regrowth we can potentially collect data if a connection can be established between the sensors and the scalp. Furthermore, any participants who cover their hair e.g. headdresses for fashion or for religious beliefs will be required to remove their scarves, turbans or wraps if they wish to take part, however, this is done with only the experimenters present.

You will need to wash the gel out after wearing the cap. There are hair washing facilities in the ALIVE, including a large sink fitted with a showerhead, shampoo, hairdryer and towels. If you wish to bring any specific hair products, please do.

Particular hair styles and products in hair, such as wax, gel and hairspray interfere with the sensors, so we ask participants to avoid applying these on the day of the study. Also, when wearing the cap, it is important to avoid moving its position on the scalp.

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 would still be free to withdraw at any time during the study and up to two weeks after without giving a reason. To indicate you would like to withdraw from the study, you can verbally shout out to the researcher.

Will my taking part in this study be kept confidential?

Yes. A code will be attached to all the data you provide to maintain confidentiality.

Your personal data (name and contact details) will only be used if the researchers need to contact you to arrange study visits or collect data by phone. Analysis of your data will be undertaken using coded data.

The data we collect will be stored in a secure document store (paper records) or

REC/IRAS ID: [#1637], [V3], [24-01-2022]

electronically on a secure encrypted mobile device, password protected computer server or secure cloud storage device.

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.

What are the possible benefits of taking part?

There are no direct benefits to taking part, but we hope that you will enjoy taking part.

What are the possible risks and burdens of taking part?

- There are no significant risks to taking part in this study.
- In VR there is a risk of slipping or colliding with objects. You will be sitting for this task, we will remove all obstacles and will be observing your movements to avoid this.
- Some people experience nausea when in VR. This is minimised in our experiment because all movement is caused by you. During the training stages we will make sure you are well and comfortable. If not, we will end the study.
- The EEG cap requires gel to be applied to the scalp using a blunt syringe. The gel is not harmful, but it will feel cold to the touch.
- Some people may find the tasks in this study challenging, particularly where tasks are unfamiliar. This is normal and not a bad reflection on you. All data will be analysed as a group with no individual results being analysed or reported.

What will happen to the results of the study?

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.

A lay summary of the results of the study will be available for participants when the study has been completed and the researchers will ask if you would like to receive a copy.

Expenses and payments

If required, 24 SONA credits will be credited after completion of the study.

A payment of £15 will be made to those individuals who do not require SONA credits. This will be paid after completion of the study.

Who is funding the research?

REC/IRAS ID: [#1637], [V3], [24-01-2022]

The study is being funded by: Aston University and Engineering and Physical Science Research Council.

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. You can find out more about how we use your information in Appendix A.

Who has reviewed the study?

This study has received a favourable opinion from the Aston University 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 research team and they will do their best to answer your questions. Contact details can be found at the end of this information sheet.

If the research team are 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

Cliona Kelly	kellyc2@aston.ac.uk	[redacted]
Prof. Tim Meese	t.s.meese@aston.ac.uk	
Dr. Ulysses Bernardet	u.bernardet@aston.ac.uk	
Razan Elobeid	[redacted]@aston.ac.uk	
Rebekah Harnett	[redacted]@aston.ac.uk	
ALIVE laboratory	N/A	0121 204 5045

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.

REC/IRAS ID: [#1637], [V3], [24-01-2022]

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.

Appendix B

Poster for the Puzzle Joint Attention Task in Virtual Reality and EEG requesting volunteers

VOLUNTEERS NEEDED FOR NEUROSCIENCE EXPERIMENT

ALIVE

Puzzle Joint Attention in Virtual Reality (VR).

We are investigating how eye-gaze is used in joint attention using virtual reality and using electroencephalography (EEG).

Inclusion criteria:

- 18-40 years of age.
- No history of epileptic seizures
- Normal to corrected vision.
- Do not have a diagnosis of photosensitive epilepsy, traumatic brain injury or hypertension.
- Are not sensitive to motion sickness and do not experience 'cyber sickness'.
- Must be able to understand verbal and written information presented in English.

This study has received a favourable opinion from the Aston University Research Ethics Committee.

Study time:

Approx. 2 hrs and **£20 voucher.**

Where:

ALIVE labs, Aston University.

To sign up please email Clíona Kelly at: kellyc2@aston.ac.uk.

Clíona Kelly kellyc2@aston.ac.uk
Clíona Kelly kellyc2@aston.ac.uk
Clíona Kelly kellyc2@aston.ac.uk
Clíona Kelly kellyc2@aston.ac.uk
Clíona Kelly kellyc2@aston.ac.uk
Clíona Kelly kellyc2@aston.ac.uk
Clíona Kelly kellyc2@aston.ac.uk
Clíona Kelly kellyc2@aston.ac.uk
Clíona Kelly kellyc2@aston.ac.uk
Clíona Kelly kellyc2@aston.ac.uk

REC/IRAS ID: [#1637], [V2], [02-12-2021]

Appendix C

Consent Form for the Puzzle Joint Attention Task in Virtual Reality and EEG

Aston University
BIRMINGHAM UK

Puzzle Joint Attention Task in Virtual Reality and EEG

Consent Form

Name of Chief Investigator: Cliona Kelly

Please initial boxes

1.	I confirm that I have read and understand the Participant Information Sheet (Version 3, 24/01/2022) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	
2.	I understand that my participation is voluntary and that I am free to withdraw at any time during the study and up to two weeks after the study, without giving any reason and without my legal rights being affected.	
3.	I agree to my personal data and data relating to me collected during the study being processed as described in the Participant Information Sheet.	
4.	I agree to take part in this study.	

____________________	_______________	__________________
Name of participant	Date	Signature
____________________	_______________	__________________
Name of Person receiving consent.	Date	Signature

REC/IRAS ID: [#1637], [V3], [24-01-2022]

Appendix D

The Standard Operating Procedure for EEG Data Collection

STANDARD OPERATING PROCEDURE

Cognition and Neuroscience Research Group

SOP Title: EEG Data Collection Under Relaxed Social Distancing

1. INTRODUCTION

While the legally mandated social distancing guidelines are no longer in place we have a responsibility for the safety of staff and participants.

2. SCOPE

The SOP refers to all participants of studies within ALIVE labs where HMD, EEG, EMG equipment and other wearable devices are used and where direct contact with participants is unavoidable. This includes studies with adults (from 18 years of age) as well as children from 5 years of age.

3. RESPONSIBILITIES

The SOP refers to experiment sessions that will be carried out by researchers who are either the study lead, co-investigators, research students, or research assistants.

- Staff will read and adhere to this SOP.
- Staff should not attend work if they display any signs and/ or symptoms of Covid-19.

4. PRECAUTIONS

Prior to attending an appointment, participants/parent will be contacted by email or by phone and asked if they or anyone they have personally been in contact with:

- Has confirmed or suspected coronavirus
- Is self-isolating/shielding
- Has a high temperature and/or a new, continuous cough
- Has experienced a sudden loss of smell and/or taste
- Has any other COVID19 symptom listed by the NHS

Should any participants/parent answer 'yes' then they will not be invited to attend at this time and current guidance from Public Health England will be consulted to calculate the time period that would need to pass prior to contacting them again and re-assessing their situation.

Participants who should be shielding will not be encouraged to come onto campus until government guidelines allow.

Participants should not be encouraged to attend against the advice of Public Health England

Version 3.2 18th May 2021 Page 1 of 5

The name and contact details of participants should be retained (securely) and debrief procedures should include the name and contact details of the Experimenter to allow for contact tracing should that become necessary.

Participants should be encouraged to inform us if the experience covid symptoms. This will help us to monitor risk. Cases should be reported to covid-19Reporting@aston.ac.uk.

5. SPECIFIC PROCEDURE

All surfaces and equipment will be cleaned with disinfecting wipes prior to study participant arrival and repeated between, and after, appointments. Participants will be offered an optional face covering and will be asked to wash hands or use disinfectant gel on arrival, and when necessary, during and after the appointment.

Sessions will be as brief as possible within the constraints of the study's design.

Any equipment to be worn by the participants (e.g., headphones, wearable devices) must be fitted and removed by the participant if possible. If the experimenter must fit an equipment to the participant, the experimenter is expected to wear a mask.

Enter the lab: experimenter to disinfect the door handles after self or the participant enters the lab; participant to wash hands under proper instructions (6 step hand hygiene as shown in the washing facility room)

The experimenter is expected to wash their hands and wear a face mask

Brief and consent: experimenter to check if participant is fit for the study; participant will be debriefed and to read and sign the consent form (print or digitally signed on site).

Handling and preparation of the Material (where needed):

EEG Device
To prepare the EEG cap the experimenter is expected to wear a face mask and dispose of any used wipe into the clinical waste bin (the one with a yellow lid, next to the reception desk) .

HMD Device
To prepare the HMD the experimenter must use disinfectant wipes to clean the HMD and use fresh protective cover.

EMG

To prepare device, cables etc. the experimenter is expected to wear a face mask and dispose of any used wipe into the clinical waste bin (the one with a yellow lid, next to the reception desk).

Force platform
To prepare the device the experimenter is expected to wear a face mask clean all surfaces of the device and dispose of any used wipe into the clinical waste bin (the one with a yellow lid, next to the reception desk).

Wearable devices (e.g. vibrating motors)
To prepare the devices the experimenter is expected to wear a face mask, and dispose any used wipe to clinical waste bin (the one with a yellow lid, next to the reception desk).

Vicon markers
To prepare the devices the experimenter must disinfect their hands and all the material (e.g. scissors) before markers handling and preparation, and dispose any used wipe to the clinical waste bin (the one with a yellow lid, next to the reception desk).

Physiotherapy bed

To prepare the devices the experimenter must assure that the new paper cover is on, wear PPEs as described above, and dispose any used wipe to the clinical waste bin (the one with a yellow lid, next to the reception desk).

Drinks (tea/coffee/water): participants can be offered drinks as normal. Ensure cups are cleaned thoroughly with water and soap. Participants can also bring their own water bottle or mug. Touch points should be wiped down using disinfectant wipes provided.

After participant left:

1. Experimenter to disinfect the door handles after the participant exits ALIVE;
2. **Switch off** the eego Amplifier, and wipe the amplifier and its connectors with Clinell wipes;
3. Used syringes and blunt needles must be washed immediately after the experiment using soap and hot water (60 – 90 °c);

4. EEG caps must be washed and disinfected immediately after the experiment following the standard procedure. Please see Appendix for details;
5. Assure all disposable accessories (e.g. EMG electrodes, double side adhesive tape, etc.) are disposed to the clinical waste bin (the one with a yellow lid, placed next to the reception desk).
6. Measuring tape used for anatomical reference location must be disinfected (spray or wipes) immediately after the experiment.
7. Wearable devices and relevant accessory used must be disinfected (spray or wipes) immediately after the experiment.
8. Any used towels and HMD covers must be put in the designated laundry bag (the green one) and must be taken to the washing facility (wash at 60 or 90 degrees, then dry in the tumble dryer using the 'cotton' option).
9. If used, and mouse and keyboard must be wiped using Clinell disinfectant wipes;
10. Other devices including: tablet (must be switched off), telephone in the office area, Kat mini, Cyberith Virtualizer, and the driving simulator set must be disinfected accordingly (if used or touched during the experiment);

In case of emergency: please ring 222 using the phone in the office area.

6. CHANGE HISTORY

This SOP is version 3.21 and dated 15th September 2021.

APPENDIX EEG CAPS CLEANING AND DISINFECTION STANDARD ROCEDURE

1. Immediately after your recording process, rinse off any gel in each electrode using lukewarm water, then soak the cap in lukewarm soapy water for 5 minutes. Afterwards, give the cap some gentle rub and rinse the soap off from the cap again with lukewarm water.
2. Leave the cap inside-out on the cap rest till dry;
3. Prepare the disinfectant solution in the stainless container using the Sekusept™ Aktiv powder provided. The solution should contain 2% of the disinfectant power, i.e., add 24ml of powder to 1.0 litre of water. Water should be about 20-25°C. There are two stainless containers in ALIVE.
4. Soak the dried cap in the above prepared disinfectant solution for 15 minutes, rinse the cap thoroughly with clean water, then leave the cap inside-out on a cap rest to dry. Wipe the cap rest with a Clinell wipe before you put the cap on it.

Disinfectants will shorten the life of the cap so please do not leave it in the disinfectant solution for too long.

5. Please put the disinfected cap in a cap box to store. Caps remain on the cap rest will be regarded as to-be-disinfected and cannot be used for testing.
6. Please note that if you discover that ALIVE runs out of disinfectants, only the following should be used if you decide to use your own: Cidex OPA (Advanced Sterilization Products), Metricide (Metrex), Cavicide (Metrex), Pursept AF (by Schülke & Mayr), Sekusept activ (by Ecolab) or Korsolex Extra (BODE Chemie). Please contact Dr Hongfang Wang if you have any questions.
7. Rinse the stainless steel container, dry, and leave to the place where you find it.

Appendix E

Example of the debrief sheet given to participants.

Debrief Sheet

Investigating joint attention in an online puzzle task.

Thank you for your participation!

Your experimental data will provide us with a better understanding of how individuals communicate using joint attention. This information will be used with other research to better understand how typical populations perform the process of joint attention.

Experimental task

The study was designed to observe the reaction time and accuracy in responding to and initiating joint attention with others.

The nature of this study is exploratory yet supported by previous research which indicated there were significant differences in gaze allocation between children in control and experimental groups.

The purpose of this experiment is to explore non-verbal social behaviours, through collection of EEG, eye-tracking data and behavioural data (response times and accuracy). Developing our understanding of joint attention in social communication could help provide insight into conditions that have difficulties with communication for example, Autism Spectrum Disorder.

We would like to remind you that the experimental task was designed to be difficult and finding it difficult isn't cause for alarm. If you are worried about your cognitive performance, please contact your GP or Aston University Counselling Services.

Questionnaires

You completed the Autism Spectrum Quotient Questionnaire (AQ: ; Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J, & Clubley, E., 2001) which is designed designed in order to assess if there are any autistic traits in the general population that may have affected ability to follow eye gaze. According to the national autistic society, autism is a lifelong developmental disability that affects how a person communicates with, and relates to, other people. It also affects how they make sense of the world around them. **We are not testing for, or diagnosing autism, we are simply using the questionnaire to test for any traits that may affect our outcomes.** The data from these questionnaires is averaged across participants, this means that nothing in the data will be linked back to one individual.
Analysis will involve a repeated measures ANOVA on task accuracy and response time scores. The independent variable will be collaboration and eye contact (whether

REC/IRAS ID: [#1637], [V2], [02-12-2021]

the virtual human was collaborating with you or not and whether the virtual human looked at you or not).

The EEG data will be grouped and averaged across everyone who has participated in this study. We will look at everyone's data together to see if there is a common pattern of activity that happens when you are answering where the puzzle piece belongs. And we will see if this differs depending on the condition you were presented.

If you are concerned by your answers, please contact your GP.

If you are an Aston staff or student, you also have the option to contact Aston University Counselling Services. They can be contracted via email: counselling@aston.ac.uk

Questions: If you have questions about how or why this research was conducted, what happened in the study, or would like more information about this research in general, please contact Cliona Kelly (kellyc2@aston.ac.uk).

Right to withdraw

You still have the right to withdraw your data from this study at a later date (up to 2 weeks after today). If you wish to do so, please contact Cliona Kelly stating your unique participant ID number.

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 research team and they will do their best to answer your questions. Contact details can be found at the end of this information sheet.

If the research team are 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.**Researcher contact information**

Cliona Kelly	kellyc2@aston.ac.uk	[redacted]
Dr.Ulysses Bernardet	u.bernardet@aston.ac.uk	0121 204 3893
Razan Elobeid	[redacted]@aston.ac.uk	
Rebekah Harnett	[redacted]@aston.ac.uk	
ALIVE laboratory	N/A	0121 204 5045

If you have any further questions please ask one of the research team, thank you.

REC/IRAS ID: [#1637], [V2], [02-12-2021]

Appendix F

Ethics approval.

Aston University
BIRMINGHAM UK

Aston University
Birmingham B4 7ET
United Kingdom

+44 (0)121 204 3000
www.aston.ac.uk

1 February 2022

Professor Tim Meese
Student: Cliona Kelly
College of Health and Life Sciences

Dear Cliona,

Study title:	Puzzle Joint Attention Task in Virtual Reality
REC REF:	#1637

Confirmation of Favourable Ethical Opinion

On behalf of the Committee, I am pleased to confirm a favourable opinion for the above research on the basis of the application described in the application form, protocol and supporting documentation listed below.

Approved documents

The final list of documents reviewed and approved by the Committee is as follows.

Document	*Version*	*Date*
EEG Participant information sheet	3	24/01/2022
EEG Consent Form	3	24/01/2022
EEG poster	2	02/12/2021
EEG Protocol	2	02/12/2021
HMD Participant information sheet	3	24/01/2022
HMD Consent form	3	24/01/2022
HMD Protocol	2	06/12/2021
HMD poster	2	02/12/2021
Risk assessment form	1	12/10/2021
AQ questionnaire	2	02/12/2021

After starting your research please notify the University Research Ethics Committee of any of the following:

- Amendments. Any amendment should be sent as a Word document, with the amendment highlighted or showing tracked changes. The amendment request must be accompanied by a covering letter along with all amended documents, e.g. protocols, participant information sheets, consent forms etc. Please include a version number and amended date to the file name of any amended documentation (e.g. "Ethics Application #100 Protocol v2 amended 17/02/19.doc").

  Amendment requests should be outlined in a "Notice of Amendment Form" available by emailing ethics@aston.ac.uk.

- Unforeseen or adverse events e.g. disclosure of personal data, harm to participants.

- New Investigators

- End of the study

Please email all notifications or queries to ethics@aston.ac.uk and quote your UREC reference number with all correspondence.

Wishing you every success with your research.

Yours sincerely

[redacted]

R E Case (Feb 1, 2022 12:59 GMT)

Becky Case
Acting Chair, University Research Ethics Committee

Appendix G

An example of the Autism Quotient questionnaire.

The Adult Autism Spectrum Quotient (AQ)

Participant ID:.. Sex:..

Date of birth:.................................. Today's Date................................

How to fill out the questionnaire

Below are a list of statements. Please read each statement very carefully and rate how strongly you agree or disagree with it by circling your answer.

DO NOT MISS ANY STATEMENT OUT.

Examples

E1. I am willing to take risks.	definitely agree	slightly agree	**slightly disagree**	definitely disagree
E2. I like playing board games.	definitely agree	**slightly agree**	slightly disagree	definitely disagree
E3. I find learning to play musical instruments easy.	definitely agree	slightly agree	slightly disagree	**definitely disagree**
E4. I am fascinated by other cultures.	**definitely agree**	slightly agree	slightly disagree	definitely disagree

REC/IRAS ID: [#1637], [V2], [02-12-2021]

1. I prefer to do things with others rather than on my own.	definitely agree	slightly agree	slightly disagree	definitely disagree
2. I prefer to do things the same way over and over again.	definitely agree	slightly agree	slightly disagree	definitely disagree
3. If I try to imagine something, I find it very easy to create a picture in my mind.	definitely agree	slightly agree	slightly disagree	definitely disagree
4. I frequently get so strongly absorbed in one thing that I lose sight of other things.	definitely agree	slightly agree	slightly disagree	definitely disagree
5. I often notice small sounds when others do not.	definitely agree	slightly agree	slightly disagree	definitely disagree
6. I usually notice car number plates or similar strings of information.	definitely agree	slightly agree	slightly disagree	definitely disagree
7. Other people frequently tell me that what I've said is impolite, even though I think it is polite.	definitely agree	slightly agree	slightly disagree	definitely disagree
8. When I'm reading a story, I can easily imagine what the characters might look like.	definitely agree	slightly agree	slightly disagree	definitely disagree
9. I am fascinated by dates.	definitely agree	slightly agree	slightly disagree	definitely disagree
10. In a social group, I can easily keep track of several different people's conversations.	definitely agree	slightly agree	slightly disagree	definitely disagree
11. I find social situations easy.	definitely agree	slightly agree	slightly disagree	definitely disagree
12. I tend to notice details that others do not.	definitely agree	slightly agree	slightly disagree	definitely disagree
13. I would rather go to a library than a party.	definitely agree	slightly agree	slightly disagree	definitely disagree
14. I find making up stories easy.	definitely agree	slightly agree	slightly disagree	definitely disagree
15. I find myself drawn more strongly to people than to things.	definitely agree	slightly agree	slightly disagree	definitely disagree
16. I tend to have very strong interests which I get upset about if I can't pursue.	definitely agree	slightly agree	slightly disagree	definitely disagree
17. I enjoy social chit-chat.	definitely agree	slightly agree	slightly disagree	definitely disagree
18. When I talk, it isn't always easy for others to get a word in edgeways.	definitely agree	slightly agree	slightly disagree	definitely disagree
19. I am fascinated by numbers.	definitely agree	slightly agree	slightly disagree	definitely disagree

REC/IRAS ID: [#1637], [V2], [02-12-2021]

20. When I'm reading a story, I find it difficult to work out the characters' intentions.	definitely agree	slightly agree	slightly disagree	definitely disagree
21. I don't particularly enjoy reading fiction.	definitely agree	slightly agree	slightly disagree	definitely disagree
22. I find it hard to make new friends.	definitely agree	slightly agree	slightly disagree	definitely disagree
23. I notice patterns in things all the time.	definitely agree	slightly agree	slightly disagree	definitely disagree
24. I would rather go to the theatre than a museum.	definitely agree	slightly agree	slightly disagree	definitely disagree
25. It does not upset me if my daily routine is disturbed.	definitely agree	slightly agree	slightly disagree	definitely disagree
26. I frequently find that I don't know how to keep a conversation going.	definitely agree	slightly agree	slightly disagree	definitely disagree
27. I find it easy to "read between the lines" when someone is talking to me.	definitely agree	slightly agree	slightly disagree	definitely disagree
28. I usually concentrate more on the whole picture, rather than the small details.	definitely agree	slightly agree	slightly disagree	definitely disagree
29. I am not very good at remembering phone numbers.	definitely agree	slightly agree	slightly disagree	definitely disagree
30. I don't usually notice small changes in a situation, or a person's appearance.	definitely agree	slightly agree	slightly disagree	definitely disagree
31. I know how to tell if someone listening to me is getting bored.	definitely agree	slightly agree	slightly disagree	definitely disagree
32. I find it easy to do more than one thing at once.	definitely agree	slightly agree	slightly disagree	definitely disagree
33. When I talk on the phone, I'm not sure when it's my turn to speak.	definitely agree	slightly agree	slightly disagree	definitely disagree
34. I enjoy doing things spontaneously.	definitely agree	slightly agree	slightly disagree	definitely disagree
35. I am often the last to understand the point of a joke.	definitely agree	slightly agree	slightly disagree	definitely disagree
36. I find it easy to work out what someone is thinking or feeling just by looking at their face.	definitely agree	slightly agree	slightly disagree	definitely disagree
37. If there is an interruption, I can switch back to what I was doing very quickly.	definitely agree	slightly agree	slightly disagree	definitely disagree
38. I am good at social chit-chat.	definitely agree	slightly agree	slightly disagree	definitely disagree

REC/IRAS ID: [#1637], [V2], [02-12-2021]

39. People often tell me that I keep going on and on about the same thing.	definitely agree	slightly agree	slightly disagree	definitely disagree
40. When I was young, I used to enjoy playing games involving pretending with other children.	definitely agree	slightly agree	slightly disagree	definitely disagree
41. I like to collect information about categories of things (e.g. types of car, types of bird, types of train, types of plant, etc.).	definitely agree	slightly agree	slightly disagree	definitely disagree
42. I find it difficult to imagine what it would be like to be someone else.	definitely agree	slightly agree	slightly disagree	definitely disagree
43. I like to plan any activities I participate in carefully.	definitely agree	slightly agree	slightly disagree	definitely disagree
44. I enjoy social occasions.	definitely agree	slightly agree	slightly disagree	definitely disagree
45. I find it difficult to work out people's intentions.	definitely agree	slightly agree	slightly disagree	definitely disagree
46. New situations make me anxious.	definitely agree	slightly agree	slightly disagree	definitely disagree
47. I enjoy meeting new people.	definitely agree	slightly agree	slightly disagree	definitely disagree
48. I am a good diplomat.	definitely agree	slightly agree	slightly disagree	definitely disagree
49. I am not very good at remembering people's date of birth.	definitely agree	slightly agree	slightly disagree	definitely disagree
50. I find it very easy to play games with children that involve pretending.	definitely agree	slightly agree	slightly disagree	definitely disagree

Developed by:
The Autism Research Centre
University of Cambridge

REC/IRAS ID: [#1637], [V2], [02-12-2021]

Appendix H

The document outlining the project protocol.

Protocol Puzzle Joint Attention Task In Virtual Reality

Experiment Protocol

REC/IRAS ID: [#1637], [V2], [01-12-2021]

Puzzle Joint Attention Task In Virtual Reality

1. **Purpose.**

Many conditions and disorders have highlighted reduced communicative abilities as part of the diagnosis, for example, in Autism Spectrum Disorder (ASD), Schizophrenia and ADHD (Linder et al., 2019). The outcome of this reduction has provided many individuals with the feeling of isolation and finding difficulty in social interactions. For this reason, it is important that we improve our understanding of how this reduction is manifested behaviourally, but also to investigate whether there are neural underpinnings responsible for these processes initially, in the typical brain.

Eye-gaze is an important non-verbal form of social communication (Friesen & Kingstone, 1998). Eye-gaze can provide social information to guide individuals in a conversation and provide an additional form of information, for example, the individuals' feelings or to direct attention (Bakeman & Adamson, 1984).

Understanding this, there is the potential to develop an experiment that can specifically investigate non-verbal social behaviours, such as eye-gaze, to investigate the form of social communication, joint attention. The use of EEG analysis also has the potential to provide an additional data source for healthcare professionals during interventions.

2. **Protocol Outline.**

Before arrival, the participant will be provided with the Volunteer Information Sheet. They will be asked to provide their informed consent and then we will begin the study. The participant will become familiar with being in virtual reality in a training session. They will then practice the main Virtual Reality Joint Attention Task. We will then begin the main task. There will be two tasks within the puzzle paradigm: the initiating joint attention task and responding joint attention task. A debrief will then be provided and the study will end.

3. **Equipment Protocol.**

3.1. HTC Vive Pro Eye

The HTC Pro Eye is a head mounted display which rapidly presents images to each eye individually in order to present a compelling 3D visual scene. When worn, it completely blocks out any visual input from the real world. The device itself is lightweight and comfortable and is held closely against the face with adjustable straps about the head.

3.2. CAVE

The participant will have read the Volunteer Information Sheet which describes the CAVE and the necessary procedures for its use, but these will be emphasised again. The participant will be asked to either remove their footwear or wear the provided overshoe slippers while walking on the projector screen floor. They will also be reminded that they must not get to close to the walls. This is to avoid injury and damaging the coating on the walls. A warning screen will appear if they do become too close to the walls, but this is to be avoided as it will interfere with the tasks. In addition, the investigator will be observing the whole time and will advise the participant if these protocols are not being adhered to and possible end the study. While in the CAVE they will need to wear a pair of glasses which enable each eye to see a different image on the projector screen. They also have a 'rigid body' setup of infra-red reflective markers which enable reliable six degrees of freedom motion capture (6DOF = X, Y, Z location, pitch, roll, yaw).

REC/IRAS ID: [#1637], [V2], [01-12-2021]

3.3. General VR Protocol

When leaving the VR, people can experience 're-entry problems'. Re-entry problems result from readjustment to the physical and social properties of the real world from those of the virtual world (Behr, Nosper, Klimmt, & Hartmann, 2005). This can manifest as a loss of sense of direction, change in body perception, poor hand-eye coordination, mismatched emotional state and confusion. All or none of these effects may be present and the participant may or may not notice them at the time. Therefore, the investigator will be ready to assist them with balance as they exit the VR equipment and direct them back to the ALIVE office. They will be invited to sit down and rest while drinks are offered. We will then give them their footwear or ask them to remove the slippers.

3.4. EEG

The participant will have read the Participant Information Sheet which explains how we will be collecting the EEG data. This information will be repeated each time the EEG cap is worn.

To setup the cap, the participant will be asked to sit down in the ALIVE office. A clean (40°C wash and machine dried) towel will then be placed over their shoulders to prevent gel getting on their clothes. If they have long hair, or a style which covers the eyes or ears, they will be asked to rearrange it to not obstruct the sensors. We can help with this if required. The circumference of their head will be measured and the correct sized cap selected. This cap will then be placed loosely onto the participant's head and then more tightly positioned while fitting the chinstrap. The placement of the sensors and the fit of the cap will be checked, changing cap size if necessary.

Conductive gel will be applied to each electrode to minimise the resistance between each electrode and the scalp. To ensure the gel is working as intended, the cap will be connected to the amplifier and the amplifier to a computer which displays the resistance of each electrode. While monitoring this display, the gel will be applied using a blunt syringe. The syringe will be cleaned before use and the tip sterilised each time more gel needs to be drawn. The syringe is used to part the participant's hair to the scalp and create a gel tunnel back to the electrode. This involves gentle back and forth action which, if care is not taken, may scratch the scalp. Every effort will be made to avoid this and the participant will be encouraged to report any discomfort or pain.

The participant will be told the importance of minimising blinking, swallowing, jaw clenching and talking during EEG recording. To illustrate this, and for the participant's own interest, we will ask them to do each of these actions in turn while viewing a live feed of the activity of each sensor. They should see bursts of noise each time. Similarly, they will be asked not to move the cap until we take it off and to avoid touching the cables. They will wear a backpack containing the amplifier and computer (laptop or tablet) to avoid any trailing cables.

When the cap is ready to be removed, we will invite them to sit down in the ALIVE office again and place the towel on their shoulders. The chinstrap will be undone and the cap carefully removed and placed onto its holder. The participant will be shown the washing facilities (large sink fitted with a showerhead, shampoo, hairdryer and towels) and assisted if needed. They will be asked to use their own comb which they should have brought with them, as per instruction in the Volunteer Information Sheet. They can bring and use any additional cleaning and hair products/accessories that we do not provide.

After use, the cap will be washed thoroughly in warm soapy water and left to air dry before being stored. Syringes will be washed after each use as well.

REC/IRAS ID: [#1637], [V2], [01-12-2021]

4. Participant Information and Informed Consent.

Participants will be sent the Volunteer Information Sheet at least one week before attending the stud. Contact details provided on this will enable them to ask any questions they may have before travelling.

On arrival, the investigator will explain the purpose of the study and ensure participants are familiar with the content of the Volunteer Information Sheet. Once any concerns or questions have been addressed, participants will be provided with the appropriate consent form and a cop signed by the investigator. The participant will also be assigned a pre-determined participant number. This will be a number which is not based on any personal or identifiable feature of the participant.

The collection of EEG data will follow the same procedures in order to anonymise the data. A 3D scan of participants is taken at the beginning of the experiment. This data is anonymised following the steps that have been suggested by Fieldtrip. The consent form and volunteer information sheet will make it clear that we will only be analysing the data at group level and individual performances are not analysed.

5. Virtual Reality Familiarity Exercise.

Some people who use virtual reality experience motion sickness or 'cybersickness'. This is caused when there is a mismatch between information from the user's perception of self-motion and the information being presented to them in virtual reality. For example, latency between input and movement, or acceleration not caused by their own movement. We have taken steps to prevent this: no motion in our VR tasks involve acceleration or rotation not caused by the participants' own motion; the equipment used has extremely low latency between motion capture and presentation; participants who are highly sensitive to motion sickness are excluded before recruitment.

In addition, most people have little or no experience of virtual reality. It can take some time to become adjusted to VR and we want participants to familiar with VR, both for their own comfort and to ensure high quality data collection. They need to be able to move around confidently and know how to interact with the world proficiently.

To address these two points – motion sickness and familiarity – there will be a VR familiarity exercise at the beginning of the study. In this there will be demonstrations of diverse virtual environments which can be navigated and interacted with. The participant will be explicitly asked whether they are experiencing any motion sickness during this exercise and if they are then the study will not continue. This task will be at the beginning of the study to avoid wasting the participant's time should they experience motion sickness or not feel comfortable in VR. This task should take approximately 10 minutes, but may take more or less time, depending on the participant's reports of comfort.

6. Virtual Reality Joint Attention Task Practice.

The Joint Attention Task is designed to promote collaborative game play with a virtual human. We are comparing scenarios where the virtual human is collaboration level varies, whether the virtual human engages in eye contact and whether you are responding to- or initiating joint attention. For the task practice, participants will go through 5 trials. After finishing they will have the opportunity to ask questions. This practice should take approximately 10 minutes.

REC/IRAS ID: [#1637], [V2], [01-12-2021]

7. Virtual Reality Joint Attention Task Full.

Participants will first take part in the practice before the setup of EEG.

7.1. EEG set up

The EEG cap setup and baseline measurements will be completed in accordance with the equipment protocols detailed previously.

7.2. Joint Attention Task

This task will use VR equipment and so the relevant equipment protocols will be used to set this up. Once this is done, the task will begin with the first trial. There will be a total of 160 trials of varying length which will take approximately one hour. Between each trial there will be a break of approximately 15 seconds. This timing may not be exact because each trial has a variable loading time.

Each trial will be presented as a scenario within a simple and dynamic 3D visual scene. For example, the participant will be placed sat down facing the virtual human at a table. Once this scene has loaded there will be a short five second pause before participants are given the instructions for the trials.

Responding to joint attention: "*The virtual human will be presented with a puzzle piece that belongs to either their own or your puzzle board. They will then communicate to you where the puzzle piece belongs. You must indicate where you believe the puzzle piece belongs*". Based on the animations and their instruction, the participant will decide where the puzzle piece belongs by using the Vive controller to click the button corresponding to their answer ("mine", "VH").

Initiating joint attention: "*You will now be presented with a puzzle piece that belongs to either yours or the virtual human's puzzle board. Your task is to communicate to the virtual human where the puzzle piece belongs. The virtual human will then indicate where they believe the piece belongs*".

The animation will always follow the same structure: the virtual human will perform a checking sequence that involves looking at the puzzle piece and then at both puzzle boards. The virtual human will then perform one of the four trial sequences (*collaboration* = collaborative vs. non-collaborative; *gaze type* = attended vs. unattended). Depending on the condition, either the virtual human or participant will respond.

After the initiating joint attention trials only, participants will be asked "*Do you think the virtual human understood your signals?*". After each scenario has finished, there will be another five second break and then the next trial will be loaded. At any point the experiment can be paused by the investigator to give the participant a break or if something goes wrong with the equipment. Otherwise the experiment will run for the first 80 trials (halfway, ~30 mins) at which point there will be an enforced break. During this break the participant will be invited to sit. After the enforced break, the trials will run for the final 80 scenarios, ending the experiment.

REC/IRAS ID: [#1637], [V2], [01-12-2021]

8. References

Bakeman, R., & Adamson, L. B. (1984). Coordinating Attention to People and Objects in Mother-Infant and Peer-Infant Interaction. *Child Development*, *55*(4), 1278. https://doi.org/10.2307/1129997

Behr, K.-M., Nosper, A., Klimmt, C., & Hartmann, T. (2005). Some Practical Considerations of Ethical Issues in VR Research. Presence: Teleoperators and Virtual Environments, 14(6), 668–676. https://doi.org/10.1162/105474605775196535.

Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review*, *5*(3), 490–495. https://doi.org/10.3758/BF03208827

Lindner, Philip, Alexander Miloff, Simon Fagernäs, Joel Andersen, Martin Sigeman, Gerhard Andersson, Tomas Furmark, and Per Carlbring. 2019. 'Therapist-Led and Self-Led One-Session Virtual Reality Exposure Therapy for Public Speaking Anxiety with Consumer Hardware and Software: A Randomized Controlled Trial'. *Journal of Anxiety Disorders* 61 (January): 45–54. https://doi.org/10.1016/j.janxdis.2018.07.003.

REC/IRAS ID: [#1637], [V2], [01-12-2021]