A Critical Examination of Spontaneous Perspective Taking

Abbie C. Millett

A thesis submitted for the degree of Doctor of Philosophy

Department of Psychology

University of Essex

November 2019

Contents

Chapter Two: Perspective Validity, Occluding Barriers, and Avatar Stance

Chapter Three: Eye Tracking

Chapter Four: Spontaneous Perspective Taking in Children and Adults

Chapter Five: Perceived Ownership

Chapter Six: Attribution of Vision and Knowledge in 'Spontaneous Perspective Taking'

Chapter Seven: General Discussion

Acknowledgement and Dedication

First, I would like to thank the Department of Psychology at the University of Essex for the opportunity to examine this topic. It has been an invaluable experience both academically and personally. I would also like to thank all the staff and students at the department who created the ideal atmosphere to tackle all the obstacles associated with a PhD and academia. I would also like to thank my Supervisory Chair Prof Silke Paulmann for her support and advice throughout the PhD.

Next, I would like to thank my supervisors Dr Geoff Cole and Dr Andrew Simpson, who have stood by me throughout a number of different complications, and crises. I would also like to particularly thank Geoff for pushing me to complete early and dedicating a significant proportion of his time to helping me with quick feedback turnarounds. I am also thankful to Dr Jason Braithwaite and Dr Kevin Dent who have taken time to examine my thesis and provide essential feedback.

Additionally, Chapter 5 would not have been possible if not for the generosity of Mr Peter Dewhurst and all the children at Sebert Wood Primary School, thank you. The children truly are an asset to the school and were a pleasure to work with.

I would also like to thank all my friends and family who pushed me to keep going when I needed it most. As well as many a coffee, tea, gin, dinner, yoga and gym date. You have all been a massive part of this process, and I couldn't have done it without all of you, and none more so than my fiancé. Thank you Sam, for coping with my mood swings, and for all that you do for me. I love and appreciate your support, even when I'm stressing about everything and anything.

Finally, I would like to take this opportunity to dedicate this thesis to the memory of my Grandad. He is the origin of my resilience and graced me with unwavering faith even when I didn't believe in myself. We love and miss you always.

Abstract

A number of authors have identified an extension of Theory of Mind (ToM) termed as 'spontaneous perspective taking', in which another's visual perspective is computed both 'rapidly' and 'spontaneously' (Samson, Apperly, Braithwaite, Andrews, & Bodley Scott, 2010). The current work examines this notion. Experiments 1 and 2 employed the ambiguous number paradigm with different manipulations of attention. Evidence was found to suggest the spontaneous assumption of another's visual perspective. However, most importantly, this effect was also identified during conditions when, as the visual perspective taking theory would predict, it should not be apparent. Alternatively, Experiment 3 was unable to identify this effect using a variant of the dot perspective task. The next two experiments increased the measurement sensitivity of spontaneous visual perspective taking using eye movements. Again, similar patterns in the data were identified when the phenomenon should not have been exhibited. Next, Experiments 6 and 7, assessed whether the notion is routed within ToM through experimentation on young children. The developmental findings were unclear, however there were indications that the concept is progressively improved with age. Subsequently, Experiments 8, 9, and 10 adapted the examination of this notion by investigating whether perceived ownership had any effects. Initially, using a novel single response method no significant results were found. However, when using standardised response time measures it was suggested that individuals were exhibiting a spontaneous visual perspective taking response, irrespective of perceived ownership. Lastly, Experiments 11, 12, and 13 introduced an alternative theory suggesting that the agent, as well as any other orientation cue, act as a reference point that anchors and orientates the image. Overall, the present findings challenge the spontaneous visual perspective taking theory; as a number of alternative concepts have also been suggested to contribute towards this phenomenon.

Authors Note

The following should be noted. One, *Experiment 1 and 5* has been published in Kuhn, Vacaityte, D'Souza, Millett, and Cole (2018) thus; some of the content within this thesis replicates our publication. Two, *Chapter 4: Spontaneous Perspective Taking in Adults and Children*, has been drafted in the format of an article, as per the participating school request. We have no intention of publishing this paper. Three, data collection for *Chapter 6: Attribution of Vision and Knowledge in 'Spontaneous Perspective Taking*' was gathered as part of the authors MSc dissertation. However, the work has since been restructured and published in Millett, D'Souza, and Cole (2019). This work was deemed as crucial to the narrative of the thesis under examination and is included within its new format. Consequently, some of the content of *Chapter 6: Attribution of Vision and Knowledge in 'Spontaneous Perspective Taking*' replicates our publication. Four, although the content has been reconstructed, similar points of discussion within *1.5.3 Empirical Consistencies, 1.6.5 Defining an Automatic Process* and *Chapter 7: General Discussion* has been outlined in a published paper (Cole & Millett, 2019) thus; some of the content within this thesis replicates our publication. Lastly, it is worth noting that the current thesis is comprised of 1,097 experimental sessions, with 387 transferred from the authors MSc dissertation, thus 711 new sessions was gathered throughout the development of this thesis.

Chapter One:

Introduction and Literature Review

1.1 Overview

Human interaction and social communication are significantly impacted by effective metalizing abilities, otherwise known as 'Theory of Mind' (ToM; Butterfill & Apperly, 2013; Waytz, Gray, Epley, & Wegner, 2010). This fundamental concept of human nature has been extensively assessed using a variety of procedures. ToM will be the foundation to which the following thesis is based, with the specific focus on the potential shortcomings of the so-called 'Spontaneous Perspective Taking' notion. On-going disputes in visual cognition as well as the application of specific advances within its methodology will be used in the assessment of spontaneous perspective taking. Thus, the primary aim will be to investigate the spontaneous perspective taking theory. Limitations of the theory will also be identified, assessed, and discussed, as well any restrictions the notion has in terms of the wider perspective taking and ToM literature.

The following review will firstly summarise ToM as well as the underlying components of ToM. Perspective taking will then be discussed, before spontaneous perspective taking is outlined and addressed in terms of central methodology, fundamental disputes, and criticisms. Finally, the last section will summarise the key components of this review, as well as highlight the possible trajectories for further investigations within this area.

1.2 Theory of Mind

1.2.1 Overview

ToM can be defined as the ability to infer mental states to others as well as interpreting, explaining and predicting the behaviour of others through those assumptions (Premack & Woodruff, 1978). The key characteristics used when inferring about mental states include beliefs, knowledge, desires, intentions, as well as emotional and motivational states (Apperly, 2011). This theory and process has been researched in a number of ways, from the context of psychiatric and developmental disorders, such as autism and schizophrenia (Baron-Cohen, 1995; Frith, 2004) to differences within human infants and non-human species (Call & Tomasello, 2008; Emery & Clayton, 2009). It is a distinctively unique investigation as not only is it concerned with the mental state of the participant, but also the mental state of an agent. This includes predictions about the behaviours, intentions, desires, emotions and actions that the participant has over the agent. As a result, this process has many advantageous components that are used in successful human interaction. For example, social cooperation and teambuilding, social cohesion, effective communication, and understanding all rely on the individual having a fully functioning ToM ability.

From a young age, children are encouraged to play and interact with one another, forming the foundations of which ToM processes are built. Role-play and group games that use imagination, are essential aspects of childhood development that form the basis of ToM (Leslie, 1987). For instance, role-playing is particularly important as it instigates the process of assuming different mental states, emotions, knowledge, and perspectives that differ from the self (Frith & Frith, 2005). Consequently, once these skills are developed, children can begin to empathise (comprehend the emotion and feelings of others with little to no distinction between the self and others), sympathise (feeling concern for the emotion and feelings of other), and understand the social interactions that surround them (Decety & Michalska, 2010).

One example in which effective ToM processing ensures the efficient functioning of an individual, relates to false belief and deception (Flavell, 1999). Though when examined at a micro level, deception can be perceived to be detrimental, due to false beliefs increasing gaps within social cohesion, it can also increase social stratification and interaction. For example, individuals without this ability will have difficulty understanding mundane social interactions, such as sarcasm and exaggeration (Happé & Firth, 1995). Granting this may not be deemed as an essential aspect of social interaction, however, individuals without this ability will struggle significantly within everyday life, for not all social interaction is taken literally. Moreover, sympathy and empathy are also used in effective social interaction as an extension of ToM (Goldstein & Winner, 2012). For example, if an individual can effectively empathise and sympathise, perception of another's feelings are easier to comprehend. Recall that empathy relates to the comprehension of other individuals' emotional state, with little to no distinction between the self and others, whereas, sympathy relates to feelings of concern for the emotional state of others (Decety & Michalska, 2010). This may in turn, lead to altruistic behaviour that further enhances social cohesion.

As can be seen, ToM has a significant impact on effective social interaction and communication. This is one reason why researchers have extensively investigated the boundaries of this theory, as well as the cognitive processes associated with it. During the following section, on-going criticisms with ToM will be discussed, as these criticisms can extend to the critical evaluation of spontaneous perspective taking.

1.2.2 Mentalizing and Submentalizing

Firstly ToM, otherwise known as mentalizing, can be recategorized into mentalising and submentalizing. Whereas, mentalizing refers to fully functioning ToM abilities dependent upon the processing of alternative mental states, submentalizing refers to general-purpose cognition that simulate the consequences of mentalizing in context (Heyes, 2014). For example, this can include object centred spatial coding. Santiesteban, Shah, White, Bird, and Heyes (2015) investigated this categorisation of submentalizing by attempting to divide the current ToM literature using the director task, into mentalizing and submentalizing research. It was the authors' assumption that previous ToM research, that used the director task, utilises domain specific submentalizing processes such as spatial arrangements rather than the concluded mentalizing ability. As Santiesteban et al. (2015) argued that processing mental states, required for mentalising, is cognitively demanding, whereas submentalizing is not. In other words, instead of assuming the perspective of another, as with mentalizing, the authors suggested that submentalizing uses object-centred spatial coding to process the director task scene (Santiesteban et al. 2015). Using a variation of the director task, the authors instructed participants to move objects in a physical grid located between them and either an actor or camera. To reiterate, the director task involves participants moving objects within a grid, under the instruction of a 'director'. The grid held a mixture of occluded and open panels on both the participant's and alternative viewpoint sides, as well as a number of similar objects that the participant could identify in relation to the director's instruction. For example, two 'balls' could be placed in the grid, one that both the director and the participant could see, and another that only the participant can see. Whether participants considered the alternative (i.e., ‘allocentric’) perspective when choosing a object to move as well as their movements, was used as an indicator of mentalizing and submentalizing abilities. In other words, whether participants moved an object that both the alternative (i.e., ‘allocentric’) perspective and the self (e.g., ‘egocentric’) perspective could see, or an object that only the self (e.g., ‘egocentric’) perspective could see, was used as an indicator of mentalizing and submentalizing.

The authors found that performance in the director task was unchanged irrespective of whether an actor or a camera was present. Thus, an inanimate object provided the same level of effectiveness in regards to orientating the scene as an actor’s gaze direction. Consequently, submentalizing was concluded to hold significance over the director task. Santiesteban et al. (2015) extended their conclusion further to suggest that previous ToM literature may provide evidence for submentalizing and not mentalizing, as mentalizing is significantly more cognitively demanding than submentalizing. However, Santiesteban et al. (2015) have only provided initial evidence to suggest that ToM can be split into the components of mentalizing and submentalizing. Thus, further investigation is needed, especially in relation to alternative methodologies and assessments of ToM processes. This distinction between mentalizing and submentalizing will be revisited in accordance with spontaneous perspective taking within section *1.5.8 Submentalizing*.

1.2.3 Age and Human Development

The development of ToM has also been disputed in terms of age. Previous research has focused on assessing attentional processing using behavioural techniques, such as manipulations to beliefs and false-beliefs (Perner & Ruffmen, 2005). Although developmental psychologists have devised a number of methods in order to assess a child's ToM ability (e.g., Carpendale & Lewis, 2006; Doherty, 2008) the most popular task is through the assessment of false belief. During a false belief task, a child is presented with a story style scenario in which the perspective of a character differs from the perspective of the child. The child's ability to deduct whether the character will search for an object in the place that the character assumes the object is hidden or where the child knows that the object is hidden (the character disappears from the scene when the object is moved to a new hiding place) forms the basis for the judgement of the participants ToM ability (e.g., Wimmer and Perner 1983; Wellman, Hollander, and Schult 1996). Many different variants of this scenario have developed including with the focus on the participants own false belief (e.g., Wimmer and Perner 1983) and the false belief of another (e.g., Siegal and Beattie 1991) yet the standard assumptions have remained relatively the same throughout the literature's development.

The specific age at which ToM develops is debatable. Wimmer and Perner (1983) originally identified that many children judged the false belief scenario incorrectly until around five years of age. However, the age at which this ability forms has been heavily disputed throughout the development of this literature. Consequently, Wellman, Cross and Watson (2001) carried out a meta-analysis of 178 experiments that used false belief tasks to identify the commonly agreed age at which ToM ability develops. Wellman, et al. (2001) concluded that the likelihood of a child giving the correct answer within a false belief task changes most significantly between three and four years of age rather that the originally thought four and five years of age. It is worth noting that Wellman, et al. (2001) also identified that the children's errors were not random, they respond egocentrically, with their own point of view. Wellman et al. (2001) also identified that the little variations upon the fundamental false belief task had little corresponding effect upon the children's response, they performed the same number of errors and within the same bonds of age differences. Thus, it is common belief amongst developmental psychologists that children aged three years and below are unable to differentiate in terms of alternative perspectives and thus have not yet fully developed their ToM abilities.

Alternatively, Senju, Southgate, Snape, Leonard and Csibra (2011) contributed to this discourse, by investigating mental state attribution within a sample of 18-month-old infants. Specifically, the authors attempted to assess the likelihood of false-belief attribution, using a procedure first identified by Heyes (1998). Heyes (1998) proposed a method that could assess an infant's false-belief attribution, which used two visually identical blindfolds. However, one blindfold is transparent whereas the other is opaque. The blindfolds are then transferred to an adult after an infant has worn them previously. Infants who wore the transparent blindfold followed the gaze of the adult, whereas infants who wore the opaque blindfold did not. Meltzoff and Brooks (2008) trialled this procedure and found that infants use previous experience to determine the perception of another through attribution processing.

Senju et al. (2011) adapted Hayes (1998) blindfold procedure to include a search task. Consequently, whilst the adult was wearing the blindfold a toy was hidden. The infants' anticipatory eye movements were recorded as a measure of where the infant predicted the adult would search for the toy. The authors' found that infants with the opaque blindfold expected the adult to search for the toy in the original false-belief spot, whereas with the transparent blindfold, this was not the case. Senju et al. (2011) therefore concluded that 18-month-old infants used experience to infer perceptual judgement of other's false beliefs. Although this research provides evidence towards an anticipatory attention attribution of false beliefs in 18-month-old infants, this is only one concept embedded within ToM, many more still reside that have not been explored. Thus, in terms of the overall development of ToM, the evidence that infants can form perceptual judgments cannot be taken as an indication of overall ToM. Instead it could be argued that Senju et al. (2011) are identifying an early onset of submentalizing ability and not fully formed ToM. The presence of other complex components within ToM such as emotion, perspectives, and desires, which cannot be accounted for within this procedure also counter the assumption of the early onset of ToM abilities. Thus, the age at which ToM abilities develop remains disputed.

1.2.4 Summary

ToM has led to the expansion of many lines of enquiry, from the overall assessment of ToM in terms of human development, to the different components embedded within ToM. One specific avenue of investigation that is particularly significant in terms of the current work is the examination of ‘perspective taking’. This component will now be outlined and discussed in the following section.

1.3 Perspective Taking

1.3.1 Overview

When deconstructing ToM into its relevant components, perspective taking is one particularly crucial aspect used in everyday thought and action (Surtees, Apperly & Samson, 2013). This component not only aids individuals in making judgements regarding those around them in terms of thinking, feeling, and attention, but it can also help with forming predictions about human behaviour and environmental changes. Consequently, perspective taking is one factor that significantly impacts the understanding of human cognition (Tomasello, 2008). Altruistic behaviour and cooperation are two examples that highlight this crucial influence of non-egocentric behaviour that perspective taking influences (Keysar, Lin & Barr, 2003).

Although different lines of enquiry in the perspective taking literature have evolved, for example, visual and spatial perspective taking, the consensus is that perspective taking is the ability to assume another's perspective that differs from our own, in other words 'to put oneself into someone else's shoes'. For this process to occur three components must be present. Firstly, the self or 'egocentric' perspective must have a view that differs from another, known as the 'allocentric' perspective, and lastly a 'target object' that can be perceived differently depending upon the perspective adopted. More specifically, *visual* perspective taking refers to the ability to understand if someone else can see an object, as well as how they can see that object, whereas, *spatial* perspective taking refers to the ability to understand where something is placed in relation to someone else (Surtees et al. 2013).

In conjunction to the identification of different areas of perspective taking, different levels of perspective taking can also be analysed. For example, visual perspective taking can be split into levels, with level 1 relating to whether another individual can see something and Level 2 relating to a deeper understanding that although both the 'self' and 'other' perspective can see an object, the two perspectives may differ (Apperly & Butterfill, 2009). Further disparities between level 1 and 2 visual perspective taking will be discussed as well as the characteristics of spatial perspective taking in the following section. These key discussions will highlight the trajectory of perspective taking research in the formation of the so-called spontaneous perspective taking theory.

1.3.2 Levels of Visual Perspective Taking

As stated previously, visual perspective taking can be deconstructed into two levels, level 1 and level 2 (Salatas & Flavell, 1976; Flavell, Everett, Croft & Flavell, 1981; Yaniv & Shatz, 1990). Previous literature has suggested that the successful assumption of level 1 is obtained earlier in an individual's development, compared with level 2 (Moll & Tomasello, 2006; Surtees & Apperly, 2012). However, in-depth analysis of the differing cognitive shortcuts used in the levels of perspective taking are limited. Michelon and Zacks (2006) assessed the difference in cognitive processes required during visual perspective taking, particularly in relation to the two distinct levels.

Firstly, Michelon and Zacks (2006) focused upon the different processes of level 1 and 2 visual perspective taking. The authors predicted that individuals may use line of sight tracing for allocentric perspective object detection, and perspective transformations, or 'mental rotations', to judge how objects are perceived by another's perspective. In other words, they aimed to see whether level 1 visual perspective taking relies upon line of sight tracing and whether level 2 visual perspectives taking relies on perspective transformations. Participants were firstly asked to identify whether a target object was visible (level 1 visual perspective taking) to an agent before then identifying whether the object was positioned to the left or the right of the agent (level 2 visual perspective taking). Thus, identifying whether the target object is visible assessed level 1 perspective taking and the object positioning for level 2. Results found that response time (RT) increased when the angle between the agent and participant was increased. The authors suggested that this was due to participants performing perspective transformations. In addition, the RT difference found during the angle task was not identified during the visibility task. Thus supporting the authors idea that participants were tracing the line of sight from the agent to the target object which does not impact RT, as it does not require transforming the scene. The authors concluded that perspective transformations are limited to level 2 perspective taking, and supports the idea that mental transformations are only used in level 2 visual perspective taking.

In addition, Michelon and Zacks (2006) also attempted to eradicate the possible use of memory-based strategies within perspective taking tasks, as well as investigate whether increasing the distance in line of sight judgements affect RT. As with the author's previous findings, it was found that level 1 visual perspective taking requires line of sight tracing, and it is influenced progressively with increasing distance between the agent and target object.

One could argue that the perspective taking literature is limited as it has primarily used computerised avatars or agents, in comparison to human agents. This is a limitation as computerised agents lack a ToM, whereas human agents have a fully formed ToM. However, other research has indicated that level 2 visual perspective taking, which was previously concluded to require mental transformations, can occur without the use of an agent (Juurmaa & Lehtinen-Railo, 1994; Presson & Montello, 1994; Amorim, Glasauer, Corpinot, & Berthoz, 1997; Amorim & Stucchi, 1997;

Amorim, Trumbore, & Chogyen, 2000; Wraga et al., 2000; Zacks, Mires, Tversky, & Hazeltine, 2000; Creem et al., 2001; Zacks et al., 2003). Yet, in terms of level 1 visual perspective taking and line of sight tracings, it is not clear whether an agent is essential.

Michelon and Zacks (2006) also assessed this discrepancy in terms of the importance of an agent by replacing the agent with an abstract shape. The task remained the same as previous perspective taking literature, however, an object replaced the agent as the anchor or reference point in which to orientate the scene. Results showed that in terms of level 2 visual perspective taking, the general trend previously found was replicated. However, the effect of line of sight tracings within level 1 visual perspective taking was not identified. Michelon and Zacks, (2006) concluded that line of sight tracings used within level 1 visual perspective taking are dependent upon the presence of a humanised agent in which to trace sight. Subsequently, Michelon and Zacks (2006) argue when an agent is absent; participants rely upon mental transformations to process visual information regardless of the level.

Michelon and Zacks (2006) additionally examined the influence of distance and angle between the target objects in relation to both levels of visual perspective taking. It was found that RT increased with an increased angle between the agent and participant, whereas increasing the overall distance between the agent and the target object did not produce any significant effects. Thus, supporting the overall conclusion that level 2 processing requires mental transformations, for the greater the angle the larger the mental transformation required, hence causing the increase in RT. However, in terms of level 1 processing, the greater angle disparity between the agent and target object did not have any significant effects upon RT. This again supports Michelon and

Zacks' (2006) overall conclusion that level 1 visual perspective taking uses line of sight tracing to process the scene.

In sum Michelon and Zacks (2006) addressed a number of different areas of visual perspective taking, especially in identifying the differences between the levels of processing. This classification of levels is one topic that must be considered in the development of future work. Michelon and Zacks (2006) posed one particularly interesting question: *whether a human presence is essential for perspective taking to occur*. Although other researchers have found that perspective taking is not always dependent upon the presence of another individual, Michelon and Zacks (2006) did find that without a humanised presence, perspective taking judgements are altered. This should be carried forward into future work, for if the so-called perspective taking judgements are made without the presence of an agent; are they computing the perspective of another? Or are there instead alternative extensions and mechanisms of cognition being used?

1.3.3 Visual and Spatial Perspective Taking

When presented with a scene that has an alternative viewpoint, spatial perspective taking states that in order to understand the scene it is processed in reference to someone else (Surtees et al. 2013). In other words, the observer creates frames of reference within that scene. One would assume that the self perspective otherwise known as the 'egocentric' perspective would have priority over the alternative 'allocentric' perspective, which is indeed supported by some empirical research (e.g., Piaget and Inhelder, 1956; Hart and Moore, 1973; Shelton and McNamara; 1997). However, this is not always the case. Some research shows that rats, monkeys, and people on first exposure to the scene form multiple allocentric representations (e.g., Graziano and Gross, 1994; Tversky et al., 1999; Mou et al., 2004). Whereas an egocentric perspective frame of reference describes objects in relation to the body of the self, the allocentric perspective frame of reference would use differing objects in respect to each other, i.e., to the left of the window.

Building upon the previously outlined distinction of level 1 and 2 visual perspective taking, Surtees, Apperly and Samson (2013) attempted to identify whether there are also differences in levels for spatial perspective taking. They particularly focused upon the different levels of perspective taking within both visual and spatial perspective taking domains. Although spatial perspective taking has not been as extensively investigated as visual perspective taking, in terms of the differing levels, the depth of spatial perspective taking has been identified to increase with age. For example, children aged three-four years, can identify whether an object is in front or behind an agent (Harris & Strommen, 1972; Cox, 1981; Bialystok & Codd, 1987) whereas older children can also identify whether the object is to the left or right, as well as in front or behind an agent (Hands, 1972; Harris & Strommen, 1972). For the ease of interpretation, Surtees, et al. (2013) termed this differentiation as 'level 1-type' spatial perspective taking for the former, and 'level 2-type' spatial perspective taking for the latter. To reiterate, 'level 1-type' spatial perspective taking refers to being able to identify whether an object can be perceived by another, for example, is the object in front or behind an agent, whereas 'level 2-type' spatial perspective taking refers to being able to identify where the object is in space in relation to the agent, for example, to the left or right of an agent (Surtees, et al. 2013). Although the levels of spatial perspective taking offer a similar break down in classification to visual perspective taking, the characterisation is very distinct. Consequently, Surtees, et al. (2013) investigated both the distinction of levels of visual perspective taking and the levels of spatial perspective taking within their research.

Initially Surtees, et al. (2013) tested the corresponding effects of level 1, and 2 visual perspective taking and level 1-type, and level 2-type spatial perspective taking when focusing on angle and distance. Participants were required to make judgements about the allocentric perspective depending on stimulus arrangement and experimental questioning. For level 1 visual perspective taking, participants were asked whether an agent could or could not see a target object. For level 1-type spatial perspective taking the same stimuli were presented, however participants were required to make judgements about whether the target object was in front or behind the agent. During trials in which participants identified that the agent could see a target, an extension question was used to further assess level 2 visual perspective taking. For these trials, participants were required to respond with *how* the target was perceived from the orientation of an agent's viewpoint. This accounted for the depth of level 2 visual perspective taking that identifies that although both could view a target, the perception and view of each viewpoint can differ. Lastly, for level 2-type spatial perspective taking participants were asked to which side of the display was the target portrayed, in relation to the agent. In addition, the distance and angle between the agent and target object was also manipulated with near and far distance, and 0°, 60°, 120°, and 180° angles.

Surtees, et al. (2013) found that with the increased angle between the target object and agent, RT also increased. The authors suggested that the perception of the position of a target object in relation to another individual requires a mental rotation to align the two perspectives together. Additionally, there was no significant effect of distance, which the authors suggested could be due to the lack of additional cognitive processing needed, to comprehend the difference in distance, as the line of sight remained the same.

One could argue that Surtees, et al. (2013) initial experiment can be criticised as repeated exposure to the same stimulus could impact the data generated. However, the authors developed an additional experiment in which different stimuli were presented for the different tasks. Yet the same patterns in the data were observed. Thus Surtees, et al. (2013) demonstrated and identified common characteristics and distinctions between visual and spatial perspective taking. Interestingly, the authors provided evidence to suggest that although level 2 visual perspective taking and 'level 2-type' spatial perspective taking are distinctly very different processes they both required a mental rotation. Granting this can account for any increase in RT identified in other research, the authors have also highlighted that RT differences may be a result of the computation of another individual's perspective via mental rotation. For example, the same difference in RT was still present when replacing the agent with an inanimate object, as the mental rotation process remains the same regardless. This could suggest that the process differentiates itself from visual perspective taking in favour of spatial perspective taking, which does not necessarily use ToM processes, for no mental states are present to compute. Yet mental rotation is still used. This controversy in relation to the assumed perspective without a human presence as well as the on-going debate of mental rotation will be discussed in *Section 1.5*, relating specifically to the spontaneous perspective taking theory.

1.3.4 The Development of Perspective Taking

Effective perspective taking has been argued to develop throughout childhood (Wimmer & Perner, 1983; Perner, 1991). It has been documented that children below the age of four have an inability to comprehend that they have different views to others (Wimmer & Perner, 1983; Perner, 1991), are unable to identify ambiguous communication signals (Deutsch & Pechmann, 1982; Sonnenschein & Whitehurst, 1984), and are unable to distinguish between different perspectives (Flavell, 1986). Thus, previous assumptions have suggested that adult and children ToM and perspective taking abilities significantly differ. Consequently, Surtees and Apperly (2012) examined age differences, specifically in relation to egocentrism and automatic perspective taking abilities.

Surtees and Apperly (2012) assess the degree of egocentrism in children and adults by using a cross comparison between ages. To achieve this, parents and carers were required to complete the same task as their child. The task required participants to identify the number of discs embedded in stimuli that either the participant or an avatar could see. Thus, level 1 visual perspective taking was adopted. The RTs for each participant was then used to produce an egocentric bias index, which was compared across the ages. The authors found that all participants processed the self and avatar perspectives. Therefore, when the two perspectives were inconsistent, RTs increased due to egocentric interference (Surtees & Apperly, 2012). Furthermore, Surtees and Apperly (2012) highlighted that participants could use a task strategy dependent upon the avatar to respond, which impacted the egocentric interference effect. Thus, they replicated the same experiment, but replaced the avatar with a non-social orientation cue to see whether this would impact the egocentric interference. Results showed that with the non-social orientation cue RT differences were greater for the younger children, in comparison to the older children and adults, with faster self-judgements and slower allocentric judgements.

Consequently, Surtees and Apperly (2012) propose that egocentrism is not developed throughout childhood, as the youngest children, as young as six, were influenced by another perspective when making judgements about their own perspective. Instead, the authors suggested that it is the counter correction of the interference that is developed throughout childhood. However, one could argue that this ability may have developed prior to the age of participants within this research. Plus, the work only examined one form of perspective taking, other forms, such as level 2 visual perspective taking and spatial perspective taking, have not been assessed. Subsequently, further experimentation to test Surtees and Apperly (2012) findings, as well as other research suggesting that visual perspective taking is developed throughout childhood is needed, before any claims of the developmental foundations of automatic perspective taking are made.

1.3.5 Summary

As this review suggests, perspective taking can be deconstructed into different components, from visual and spatial perspective taking, to the different levels within each component of perspective taking. For the purpose of the current work, it is worth noting that this thesis rejects the central distinction between ‘level 1’ and ‘level 2’ visual perspective taking (e.g., Flavell et al. 1981). Recall that level 1 is defined as knowing that another individual can see an object whereas level 2 refers to knowing how the object looks to the other individual. Although level 1 perspective taking is of course based on a simple truism (i.e., we can know what another person can see), the current thesis argues that only level 2 visual perspective taking can really refer to a visual ‘perspective’. In other words, how something looks to another. To reiterate, level 1 visual perspective taking could just as easily be termed as ‘position’ taking or, as other researchers have termed, a line of sight tracing. Therefore, it is difficult to conceive of how a person can have a visual ‘perspective’ of an object without knowing how it looks. If one does not know how the object will look, it is not a perspective. This is supported by the fact that one only has to know if a straight line (of sight) can be drawn between an agent and object in order to know if they can see it

(Michelon & Zacks, 2006). Of course, it's very useful in everyday parlance to refer to both as 'perspective taking'. Consequently, the current thesis rejects this notion and for the purpose of the empirical work in the following chapters this thesis will primarily focus on a truer reflection of perspective taking using the 'level 2' visual perspective taking distinction. Furthermore, this review of the components of perspective taking led to the a new focus of research, in terms of spontaneity. This area will now be outlined in the following section.

1.4 Spontaneous Perspective Taking

1.4.1 Overview

Perspective taking aids the development of shared knowledge (Clark, 1992), establishing common ground (Clark & Brennan, 1991), and resolves ambiguity within social commutation (Dura, Dale & Kreuz, 2011). Most recently, the trend in the perspective taking literature has shifted from intricate details to unconscious cognition and automaticity, which has been instigated by researchers such as Samson, Teufel, Zhao, and Cole (e.g., Samson, Apperly, Brathwaite, Andrews, & Bodley Scott, 2010; Teufel et al. 2010; Zhao, Cusimano & Malle, 2015; Cole, Smith & Atkinson, 2015). Thus, leading to the development of the spontaneous perspective taking theory. This notion claims that that individuals 'rapidly' and 'spontaneously' assume the visual perspective of another, absent of conscious control. However, despite the common focus of research in this domain, the methodological approaches differ significantly. Each will be discussed, along with identification of underlying criticisms in the assessment of the spontaneous perspective taking theory.

1.4.2 Reflexive Gaze Following

Reflexive gaze is one effect and paradigm used in the assessment of spontaneous perspective taking. In terms of reflexive gaze as an effect, this is the rapid and involuntary attentional shift of a saccadic response (Frischen, Bayliss, & Tipper, 2007). This response has been extensively assessed within the reflexive gaze following/cuing paradigm, usually with the presence of a gaze stimulus (Frischen, Bayliss, & Tipper, 2007). During these experiments, participants are presented with an attentional cue (usually in the form of a gazing agent) and target. The comparison of RT between congruent (when the target and the attentional cue is the same) and incongruent (when the target and the attentional cue differ) trials enables the assessment of the reflexive gaze effect. It is worth noting that some authors term this effect as gaze following (e.g., Samson et al. 2010; Santiesteban, Catmur, Coughlan Hopkins, Bird, & Heyes, 2013) while other term it as gaze cuing (e.g., Cole, Smith & Atkinson, 2015; Furlanetto et al. 2013). However, the general effect remains the same, attention can be rapidly and unintentionally shifted towards the cued at location by the gaze of an agent (Samson et al. 2010; Santiesteban, Catmur, Coughlan Hopkins, Bird, & Heyes, 2013; Furlanetto et al. 2013; Cole, Smith & Atkinson, 2015). Additionally, this phenomenon has been characterised as a sophisticated ToM process by some (Samson et al. 2010; Santiesteban, Catmur, Coughlan Hopkins, Bird, & Heyes, 2013), and an involuntary response by others (Cole, Smith & Atkinson, 2015; Furlanetto et al. 2013). Consequently, Teufel et al. (2010) attempted to classify reflexive gaze in terms of either a complex ToM process or an involuntary response, by employing a variation of the methodology used in the gaze-cuing paradigm, with the addition of a deception task.

Teufel et al. (2010) led participants to believe that they were observing a live video feed of another participant taking part in the experiment. They were in fact viewing a pre-recorded video of an actor. Participants were required to watch the actor and distinguish between two target letters that were presented on either side of the actor's head. Participants were informed that the actor was taking part in an irrelevant auditory task, which required the actor to turn their head to either side. The actor also wore a pair of goggles, which during half of the trials the participant was told inhibited the actor's vision, whereas on the other half of the trials did not obstruct the actor's view. This manipulation therefore impacted the participant's attribution of mental states associated with the actor. Participants were made explicitly aware that the turning of the actor's head was irrelevant to their task. To reiterate, participants were required to rapid distinguish between the two target letters that appeared on either side of the actor's head, by the press of a button, whilst ignoring the actor's movements. RT for target letter detection was recorded and compared in relation to the consistency with the actor's head movements, as well as for the obstructed or unobstructed vision of the actor. Initially, in the first experiment the actor gazed equally as often towards and away from the target letters, whereas in the follow up experiment, that repeated the same task, the actor was twice as likely to gaze away from the target.

During the first experiment Teufel et al. (2010) identified a gaze cuing effect that was modulated by the obstructed vision of the actor. Therefore, Teufel et al. (2010) suggested that the attribution of mental states, in respect to the actor's vision, or in other words the assumption of shared knowledge thorough the embodied perspectives, enhances gaze following. Additionally, during the follow up procedure, when the actor was twice as likely to gaze away from the target, Teufel et al. (2010) found that when participants were led to believe that the actor's vision was obstructed, they were able to divert their attention away from the influence of the actor's gaze. However, during the 'seeing' conditions participants were unable to divert their attention away from the influence of the actor's head movements. Teufel et al. (2010) concluded that during 'seeing' conditions, attributions of mental states and the adopted perspectives significantly impacts the effects of gaze following.

In terms of spontaneous perspective taking, it is clear to state that Teufel et al. (2010) has demonstrated a significant altercentric intrusion effect, or in other words an involuntary assumption of an allocentric perspective. This was a result of participants adopting the actor's perspective during the 'seeing' conditions, as the attribution of the actor's perspective and mental state was not obstructed. However, when the actor's vision was obstructed, this process of assumed perspective was countered and inhibited.

1.4.3 The Dot Perspective Paradigm

The dot perspective paradigm is arguably the most widely employed procedure in the assessment of spontaneous perspective taking. In order to investigate perspective taking, Samson, Apperly, Brathwaite, Andrews, and Bodley Scott (2010) undertook influential research in the exploration of spontaneous perspective taking, using the dot perspective paradigm. The initial assessment of this work was to see whether humans will implicitly assume the perspective of another and to examine the boundaries and corresponding effects of this process. To evaluate this, a variation of the attentional gaze cuing paradigm (Langton & Bruce, 1999) was used. Gaze cuing has been extensively investigated in terms of; body position and neural responses (Perrett, Hietanen, Oram & Benson, 1992), gaze perception (Baron-Cohen, 1995), and joint attention (Allison, Puce & McCarthy, 2000; Moore & Dunham, 1995). Consequently, Samson et al. (2010) adapted this methodology with the different assumptions of egocentric or allocentric perspectives.

Samson et al. (2010) presented participants with a series of blocked trials. At the start of each trial, 'YOU' or 'SHE' appeared on the screen, identifying the perspective to be adopted, followed by a number. Next, a scene depicting an avatar standing in a room with various numbers of discs on the walls was displayed. Participants were instructed to indicate whether the number presented at the start of each trial matched the number of discs either they, or the avatar could see. Within each block, a mixture of 'congruent' and 'incongruent' trials were presented, with congruent trials representing trials in which the number of discs that the participant and avatar could see was the same, whereas incongruent trials presented differing numbers of discs that the avatar and participant could see, through disc presentation occurring outside the avatars direct line of sight.

Initially, Samson et al. (2010) found that participants were faster at responding in terms of the avatar's congruent perspective, compared with their own. However, this initial result focused on the overall phenomenon, by presenting the blocked trials with a mixture of congruency and adopted perspectives. Subsequently, Samson et al. (2010) replicated the original procedure with the differing perspectives blocked accordingly. Again, the same patterns in the data were found in terms of RT and congruency.

Samson et al. (2010) also investigated whether spatial layout was affecting the results, and not the avatar presence. Hence, the original procedure was replicated with the exception that half of the trials replaced the avatar with a large rectangle posing as a distractor stimulus. Still the same pattern in the data in terms of RT and congruency was found, but only when the avatar was present. Consequently, the authors concluded that this difference in RT was due to the participants 'spontaneously' and 'rapidly' assuming the perspective of another, which in this case was the avatar. Subsequently, bypassing the egocentric 'self' perspective. To reiterate, when presented with a congruent stimulus RT decreases, and increases with an incongruent stimulus. Samson et al. (2010) thus argued that this was the result of what they called *spontaneous perspective taking.*

However, Samson et al.'s (2010) interpretation of the data has not gone unchallenged. For example, although gaze cueing has been argued to be due to perspective taking (as above), it has also been used as a critique of perspective taking. Gaze cuing is the concept that observing another individual's attention can and does influence the attention of the observer (Nuku & Bekkering, 2008; Teufel, Alexis, Clayton, & Davis, 2010; Teufel et al., 2009; Teufel, Fletcher, & Davis, 2010). Several experiments have supported this concept and found that attention can be orientated by observing another's eye movements (Friesen, & Kingstone, 1998; Frischen, Bayliss, & Tipper, 2007), body posture (Fischer, Prinz, & Lotz, 2008), and gaze direction (Driver et al. 1999). It is worth noting that gaze cuing can be used during many different everyday social interactions. For instance, an infant is often found to mimic an adult's direction of attention, which supports the development of social bonds with caregivers, (Farroni, Mansfield, Lai, & Johnson, 2003). Subsequently, it could be argued that gaze cuing may be influencing spontaneous perspective taking, for differences in RT may be due to the time required for participants to follow the gaze of the avatar, instead of the time needed to assume the allocentric perspective. Consequently, Santiesteban et al. (2013) assessed whether the spontaneous perspective taking effect could occur when the avatar in the dot perspective task was replaced with an inanimate orientation cue. Instead of using a rectangle void of any orientation, which Samson et al. (2010) used, Santiesteban et al. (2013) used a directional arrow and ran additional conditions where participants were instructed to ignore the central stimulus.

Santiesteban et al. (2013) found that the arrow was just as effective at producing a self-consistency effect, as was an avatar. Additionally, when participants were instructed to ignore the central stimulus, similar patterns were found. Since the arrow condition produced the same pattern of data as the avatar, Santiesteban et al. (2013) argued that the basic effect is not due to mechanisms associated with mental states. On the other hand, the authors found that the effect was not counteracted when participants were specifically instructed to ignore the central stimulus. This supports the notion that the effect observed could be deemed as spontaneous.

Nielsen Slade, Levy, and Holmes (2015) suggested that although previous research has highlighted the influence of non-human orientation stimuli, spontaneous perspective taking has not been investigated in terms of the perceived level of sociality. Consequently, Nielsen et al. (2015) manipulated the degree of sociality and measured its effects on spontaneous perspective taking, whilst simultaneously investigating whether self-reported individual differences affect the degree of perspective taking. Additionally, Nielsen et al. (2015) examined the effects of spontaneous perspective taking upon gaze cuing. It is important to note that as the authors manipulated the level of social influence, they subsequently altered the level of altercentric intrusion upon spontaneous perspective taking, which was used to measure its intrinsic nature. The authors used three variations of directional orientation stimuli; the original avatar deemed as *social*, an arrow deemed as *semi-social* and blocks of colour for *non-social* conditions. All other aspects of the procedure remained a direct replication of the Samson et al. (2010) original dot perspective procedure. Participants were also given the Davis (1983) Interpersonal Reactivity Index self-report questionnaire, focusing primarily on assessing the subscales of *perspective taking* and *empathic concern*.

The authors replicated the standard data pattern irrespective of the inducing stimulus. In addition, Nielsen et al. (2015) found correlations between the subscales of *perspective taking* and *empathic concern* alongside altercentric and egocentric intrusion scores, finding a positive correlation in terms of the *social* but not *semi-social* or *non-social*. Moreover, it has been documented that altercentric intrusion can not only be assessed in relation to RT, but also in terms of gaze duration (Furlanetto, Cavallo, Manera, Tversky & Becchio 2013). Accordingly, using gaze duration, the authors found a positive correlation between self-reported *perspective taking* and *empathic concern* towards altercentric intrusion. Again, this was only localised to the *social* condition, and not *semi-social* or *non-social*. Consequently, Nielsen et al. (2015) suggest that individuals with an increased ability to assume the perspective of another, as well as show higher concern for others, positively correlates with an enhanced ability of spontaneous perspective taking.

Nielsen et al. (2015) concluded that visual perspective taking could be restricted to certain conditions; depending upon perceived sociality of the orientation stimulus. The authors also examined whether the dot perspective paradigm indexes automatic or 'unintentional' processes, as the definition of *automatic* had been questioned by several theories (e.g., Logan, 1985; Bargh, 1992; Bargh & Gollwitzer, 1994). The authors sceptically argued in terms of 'automatic' or 'unintentional' processes as there are many alternative views of what constitutes an automatic process. The surrounding theories of automaticity as well as the differing definitions will be outlined in *Section 1.6.* Overall, Nielsen et al. (2015) concluded in terms of 'unintentional' computations of visual perspective taking and not in terms of spontaneous perspective taking. This is one area that could be further developed in terms of assessing the *intentional* and *unintentional* cognitive systems responsible for ToM and perspective taking.

Again, the criticism that attentional and directional gaze cuing may be affecting the dot perspective paradigm instead of egocentric and altercentric intrusion influenced Baker, Levin and Saylor, (2016). Baker, et al. (2016) investigated this criticism through adaptations of the dot perspective paradigm that specifically examined the effect of line of sight. Instead of utilising a simple irrelevant task, Baker, et al. (2016) devised a complex arrangement of experimental stimuli. Consequently, participants were required to focus their attention towards the perception of their own visual perspective as well as the avatars. To achieve this, Baker, et al. (2016) employed an occluding barrier, which barred a large section of the stimuli for the avatar. Additionally, the barrier was manipulated so that it only blocked the avatars direct line of sight as well as alternatively permitted the avatars line of sight yet blocked all other visual stimuli. Thus, Baker, et al. (2016) were able to assess the significance of line of sight and attentional gaze cuing, and its specific influences on gaze direction and gaze following.

Initially, the authors found an egocentric intrusion effect during inconsistent trials, thus refuting the gaze cuing criticisms of spontaneous perspective taking. However, in terms of the barrier manipulations, support for egocentric intrusion was found for the complete barrier conditions, but not in the 'window' variations. Thus, Baker, et al. (2016) concluded that perspective-taking interference arises when another individual's visual field is inconsistent with our own, regardless of gaze following, therefore providing substantial support for spontaneous perspective taking.

One can argue that the dot perspective paradigm is limited, as participants are directed to adopt either their own perspective or the perspective of another. Consequently, attention has been drawn to the contrasting perspectives. In effect, it could be argued that participants in the basic paradigm are effectively primed to think about perspectives; thus, negating the notion that any effects are spontaneous. To investigate this, one could use the same dot perspective procedure with blocked identification of perspective, or without forced assumption of the contrasting perspectives. Furthermore, Cole et al. (2016) argued that if spontaneous perspective taking occurs without conscious thought or effort, the phenomenon should only be apparent during conditions in which the agent can clearly view the target and only when the attention of participants is not drawn to the differing perspectives. As already noted, in the current dot perspective paradigm, participants are told to assume their own or agent's perspective at the start of each trial. Whereas in the Cole et al. (2016) variant, participants were simply asked about the target, without the forced assumption of perspectives. To put it another way, there was no "YOU" or "SHE" manipulation. This therefore avoids drawing attention to the different perspectives that can be adopted.

To examine agent perspective, Cole et al. (2016) adapted the dot perspective paradigm to include an occluding barrier, similar to Cole et al. (2015). During *seeing* conditions, window-like features were cut out of the barriers, and in *non-seeing* conditions, the barriers remained intact, thus eliminating the ability of the avatar to be able to see any of the discs. Again, a mixture of congruent and incongruent trials was used for comparison. The authors predicted that if so-called spontaneous perspective taking is in fact due to individuals assuming the perspective of another individual, there should be no difference between congruent and incongruent conditions during the barrier manipulation, as the avatar cannot see the critical stimuli.

Cole et al. (2016) found that RT was significantly shorter during congruent *seeing* conditions. However, the authors also found that the barrier structures did not modulate the basic effect. Participants were still significantly faster during congruent conditions, regardless of the closed barriers. Cole et al. (2016) thus suggested that the typical data (i.e., Samson et al., 2010) are not due to the rapid and spontaneous assumption of another individual's perspective, but instead may be dependent upon alternative cognitive shortcuts, such as gaze following. Thus, future research examining the so-called spontaneous perspective taking notion should aim to uncover the alternative cognitive processes contributing towards this phenomenon.

In addition, some authors have manipulated the avatar's head and torso positioning in the dot perspective paradigm. Gardner, Bileviciute, and Edmonds (2018) investigated implicit perspective taking, by incorporating gaze aversion. Gardner et al. also examined whether experimenter reference to perspective primed participants to think about perspective. Consequently, participants were instead informed that they were taking part in a *previously published cognitive task*. In addition, the avatars gaze was manipulated as an assessment of reflexive attentional orienting effects. This was achieved by altering the head positioning of the avatar towards either side of the screen relative to the torso, which faced the observer.

Gardner et al. (2018) found that eliminating the forced assumption of a perspective did not generate the same spontaneous perspective taking-like effects. Plus, manipulating the avatar's head and torso only yielded cue-validity effects at longer stimulus onset asynchronies (SOA), and not for instantaneous conditions. This led the authors to conclude that attentional orientating does have an effect on spontaneous perspective taking. However, the findings associated with longer SOA suggest that this process may in fact be voluntary and not a spontaneous reflex.

As can be seen, the dot perspective paradigm, which is the most widely used procedure, has provided significant insight into the so-called spontaneous perspective taking notion. However, the paradigm is also somewhat limited.

1.4.4 The Ambiguous Number Paradigm

During the ambiguous number paradigm participants are required to respond to a target embedded in a stimulus. However, instead of disc number as with the dot perspective paradigm, numerals are used as the targets. Consequently, during unambiguous conditions numerals with identical identities regardless of orientation are presented, for example '*8*'. In other words, this numeral is always interpreted as

'*8*' regardless of orientation. This therefore allows the egocentric and allocentric perspective to coincide. Whereas, during ambiguous conditions numerals that have differing identities, in respect to orientation, are presented, for example '*9*'. Consequently, this numeral can be interpreted as '*9*', but also '*6*', depending on orientation. Thus, the egocentric and allocentric perspectives differ. To reiterate, refer to Figure 1.1, from the observer's egocentric perspective the ambiguous number is interpreted as '*9*', whereas from the allocentric agent perspective the ambiguous number is interpreted as '*6*'. Subsequently, during the ambiguous number paradigm participant response is used to highlight the perspective adopted. Plus, as with the dot perspective task, differences in RT between conditions can be examined in relation to spontaneous perspective taking-like effects.

Figure 1.1: An example stimulus used in the ambiguous number paradigm. A response of 9 would be interpreted as a response from the observer's egocentric perspective, whereas, from the allocentric agent perspective the ambiguous number would be interpreted as 6.

During an experiment investigating direct and indirect measures of spontaneous perspective taking, as well age differences, Surtees, Butterfill, and Apperly (2012) employed the ambiguous number paradigm. Surtees et al. (2012) argued that the previous difference in performance regarding spontaneous perspective taking could be due to experimental designs, which assess either direct or indirect measures of perspective taking. The authors stated that direct measurements examine an effortful cognitive ability to reason about alternative viewpoints, whereas indirect measures assess the efficient, less cognitively demanding account of ToM abilities. Consequently, the authors adapted the dot perspective task to include the numeral of '*9*' for inconsistent trials and '*8*' for consistent trials, therefore assessing level 2 perspective taking using both direct and indirect measures in a sample of children aged six to eleven, as well as adults. Specifically, participants were presented with auditory stimuli of either '*He sees a Y'* or '*You see a Y'* with the '*Y'* replaced with one of the target numerals. For example, a participant would be presented with the auditory stimulus of '*He sees a 8'* or '*You see a 8'*. After the auditory stimulus participants were then presented with a visual stimulus that depicted an agent standing behind a table facing the participant. The target numeral was either placed on the table in front of the agent, or on an adjacent wall. Participants were required to manually respond using a keypad as to whether the auditory information matched the visual stimulus.

It was found that participants were able to make level 2 judgements using direct measures, but there was no evidence to support automatic indirect measures of level 2 judgements. Additionally, in terms of the age differences, performance was found to progressively improve with increased age. Therefore, Surtees et al. (2012) support the notion of level 2 direct measures of perspective taking, which is developed throughout childhood. However, they do not support the concept of indirect level 2 perspective taking. Thus, depending upon the definition of spontaneous perspective taking, in other words whether the investigation is assessing direct or indirect level 2 perspective taking, the notion can be both supported and refuted by Surtees et al. (2012).

Zhao, Cusimano and Malle (2015) also employed a variant of the ambiguous number paradigm during which participants were presented with a stimulus depicting an agent positioned at a table with the number '*6*' read from the agent's allocentric perspective, and '*9*" from the observer's egocentric perspective. The agent's action and the question posed to participants were manipulated. Either there was no agent, the agent was shown looking away from the number, looking towards the number, or looking towards and reaching for the number. The following three questions were asked, with only one being used per participant; '*What number is on the table?*' '*What number can you see?*' and '*What number can he see?*'

The authors found that when the agent was shown reaching for the number, the frequency of a spontaneous perspective taking response increased, compared with gaze alone. However, when the agent was shown looking towards the number, the frequency of a spontaneous perspective taking response was also increased when comparing with looking away from the number. Zhao et al. (2015) suggested that participants were drawn to the attention of the agent's gaze, however, when an action was introduced, the participant's attention was altered to focus upon the action of the agent, regardless of gaze. This could suggest a hierarchical system that governs spontaneous perspective taking. However, it was found that using the question '*What number can he see?*' forced participants to assume the perspective of the agent and thus the agent's actions did not affect RT, or the response given. Consequently, Zhao et al. (2015) concluded that spontaneous perspective taking varies depending upon the agent's interaction with the ambiguous information.

Zhao et al. (2015) also assessed the effect of emphasised responses that force the assumption of the allocentric perspective in a follow up experiment. Participants were presented with a stimulus showing a series of numbers and a blank space on a table in front of an agent. The numbers presented could be read as '*86*' space '*88*' '*89*' from the agent's allocentric perspective. Participants were asked to fill in the space with the correct number, when the agent was either looking away from the numbers or looking and reaching towards the empty space. The significance within this condition was reduced as task complexity was increased, with the addition of problem solving. Nonetheless, it was observed that RT increased between baseline and both the gaze and goal directed reaching condition, when the perspective of the agent was assumed. Altogether, Zhao et al. (2015) emphasised the importance that agent interaction has upon spontaneous perspective taking, and consequently promote that future work should investigate the influence that social and contextual cues may have over this phenomenon.

Surtees, Samson, and Apperly (2016a) assessed the automaticity claim of spontaneous perspective taking using the ambiguous number paradigm. Participants were asked questions about visual stimuli that they were presented with. They were required to respond in terms of their egocentric perspective, as well as the allocentric perspective of an agent. Interestingly, the authors simultaneously assessed both level 1 (i.e., whether another individual can see an object), and 2 (i.e., judging how another individual can see an object) perspective taking. The questions associated with level 1 perspective taking questioned participants in terms of the number of targets, using a similar procedure to the dot perspective task. Alternatively, the questions associated with level 2 perspective taking questioned participants in terms of an ambiguous numeral. The authors predicted that an increase in RT across egocentric conditions would be identified in respect to consistency and ambiguity. This would suggest egocentrism, as participants are accounting for and ignoring the allocentric perspective. Whereas, the authors predicted that an increase in RT across allocentric conditions, in respect to consistency and ambiguity, would suggest altercentrism, as participants are accounting for and ignoring the egocentric perspective.

Surtees et al. (2016a) found that when participants are directly asked to process in terms of level l and level 2 perspective taking, they can make the correct judgements but exhibit egocentric and allocentric intrusion effects, in respect to the effect on RT. Thus, supporting the perspective taking notion. However, the authors also found that only level 1 perspective taking judgments were absent of conscious control when assessed by indirect measures. Thus, in terms of automaticity, only level 1 perspective taking can be deemed as spontaneous. Consequently, Surtees et al. (2016a) have demonstrated that when measuring perspective taking directly, spontaneous perspective taking was found; yet when indirect measures were employed, only level 1 perspective judgments were found absent of conscious control. Therefore, challenging spontaneous perspective taking in relation to automaticity. This dispute will be further addressed in *Section 1.6*.

1.4.5 Joint Action

Joint action provides another procedure that can be used to investigate spontaneous perspective taking. Surtees, Samson, and Apperly (2016b) asked participants to complete a joint action task that required them to judge whether numbers were smaller or larger than a control number. Interestingly, at no point in the experiment were participants asked to assume the perspective of their partner, yet participant accuracy was significantly greater when the number was consistent for both the participant and partner perspective (e.g., '*8*'). Furthermore, RT was also significantly shorter during consistent variations of the presented number (e.g., '*8*'), rather than inconsistent (e.g., '*6*').

Surtees, et al. (2016b) also assessed the impact of the perceived involvement of the partner in the joint action task, as well as the effect of replacing the joint action task with a distractor task. The authors found that the likelihood of responses imitating the trend of so-called spontaneous perspective taking was progressively more likely with an increased perceived involvement of their partner. Additionally, the authors found that even if the primary focus of the joint task were not on the magnitude of the number (i.e., during the distractor task condition), participants would still spontaneously adopt the perspective of their partner. Consequently, Surtees, et al. (2016b) concluded that spontaneous perspective taking was found during this joint task, with both shared and conflicting perspectives. Thus, accuracy and RT were affected negatively if the perspectives were inconsistent, but also enhanced when the perspectives were consistent.

1.4.6 Summary

As the above review has outlined, several different methodologies have been developed in the investigation of spontaneous perspective taking, from reflexive gaze following to joint action. Yet, with the development of different methodologies, different theoretical issues have also arisen. For example, the issue of whether a task is relevant for the assessment of the phenomenon, to automaticity and identification of the critical components that defines whether a phenomenon is in fact automatic. The current theoretical disputes surrounding spontaneous perspective taking will now be examined in the following sections.

1.5 Theoretical Disputes of Spontaneous Perspective Taking

1.5.1 Overview

Some of the key theoretical disputes of the so-called spontaneous perspective taking phenomenon lay within the following; task relevance, empirical inconsistencies, gaze cueing, mental models, mental imagery and rotation, mental self-rotation, submentalizing, and knowledge attribution. Each will now be outlined, as well as empirical suggestions that could be used to further investigate these debates.

1.5.2 Task Relevance

Relevance is a concept that influences many aspects of psychology that focus on information processing (Schamber, Eisenberg & Nilan, 1990). However, definitions for relevance are limited, as the concept is perceived intuitively (Saracevic, 1996). For the current work relevance involves “an interactive, dynamic establishment of a relation by inference, with intentions toward a context” (Saracevic, 1996, p. 206). In other words, relevance implies a dynamic relationship between an input and output, (Cosijn & Ingwersen, 2000). Consequently, in terms of perspective taking the input, such as the empirical stimuli and the output, such as participant response, can be criticised to be lacking in task relevance. For example, relevance is increased when the experimental task is associated with the alternative perspective and not replaced by an irrelevant distractor task. In terms of perspective taking, Zwickel (2009) assessed this issue by examining the importance of a human body, as did Frischen et al. (2009) who investigated action and action cues.

In regard to spontaneous perspective taking, Zwickel and Muller (2010) examined the impact of task relevance. Participants were required to respond to discs presented on a screen before answering a question about the embedded distraction stimulus. The distraction stimulus was a face with either a fearful or neutral expression, or a rectangle, and the questions posed after the disc identification either increased or decreased the relevance of the experimental task. As with other literature investigating spontaneous perspective taking, the comparison of congruent and incongruent RT was also assessed. However, as the primary focus was on task relevance, Zwickel and Muller (2010) emphasised that the mere presence of a face would produce differences in RTs, regardless of congruency.

The authors found perspective taking effects were apparent when a face was presented with a fearful expression, and not a neutral expression. This would indicate that relevance to the task, such as emotional responses, increases the magnitude of the perspective taking influence. Zwickel and Muller (2010) also identified that merely observing action and the action cues of another, does not necessarily result in spontaneous perspective taking.

Reflecting upon the dot perspective paradigm, participants were first presented with a screen stating which perspective to adopt, ‘YOU’ or ‘HIM/HER’. Consequently, one could argue that the screen is highlighting the relevance of the different perspectives of the avatar and participants, thus contributing to the RT differences. In other words, the screen may have increased the relevance of the task, priming participant response, and therefore reducing the likelihood of an automatic process occurring. Additionally, if the spontaneous perspective taking notion is solely dependent upon a degree of relevance, then finding the RT differences associated with spontaneous perspective taking, when a distractor task is present, would be problematic. As a consequence, relevance may not be the only influencing factor of spontaneous perspective taking.

As spontaneous perspective taking is a relatively new line of investigation, technicalities such as task relevance are yet to be examined extensively. Zwickel and Muller (2010) have highlighted the importance that relevance has upon the phenomenon, but whether a process can be truly 'spontaneous', yet dependent upon definitive factors, is still under examination. Future work in this area should begin to assess whether relevance is bound to perspective taking in terms of vision alone, or whether it can impact other perspective taking abilities that are not primarily based in vision.

1.5.3 Empirical Inconsistencies

Most interestingly some authors have found that the visibility manipulation, otherwise known as the barrier method, applied to the dot perspective paradigm modulates spontaneous perspective taking (Furlanetto, et al. 2016; Baker, Levin, & Saylor, 2016) while others document the opposite (Cole et al. 2015; 2016; Conway et al., 2017). It could be argued that this is due to the fact that it is a common occurrence to generate an effect when a phenomenon is first reported, as it would be unreasonable to expect authors to immediately undertake and report all the work necessary to understand the mechanisms responsible for a phenomenon. Especially when accounting for the current publication trend of null effects. It is also understandable as visual cognition literature often examine an effect's various parameters and 'boundary conditions', initially asking questions such as how long a phenomenon lasts, is it automatic, is it perceptual, attentional, or as a result of a decision process (e.g., Inhibition of return, Posner & Cohen, 1984; attentional blink, Raymond, Shapiro, & Arnell, 1992)? Thus, the replication number of publications increase. However, theories that develop an understanding for results inconsistencies and alternative explanation are also essential. This has been very much lacking within spontaneous perspective taking research. Instead, the field has been dominated by a long list of similar empirical investigations that, aside from their inherent interest, have not generated many explanations.

One possible explanation for these inconsistencies within the dot perspective paradigm reside within reflexive gaze following. Again, recall that during the dot perspective paradigm, participants are required to judge the number of dots from both their own egocentric perspective and, on other trials, from the allocentric agent perspective. Within certain experiments (e.g., Samson et al. 2010; Santiesteban et al. 2014; Nielsen et al. 2015) this occurs within-block such that participants are informed at the start of each trial which perspective they should adopt. Consequently, this procedure could be criticised that participant attention is being drawn to the representation of the agent's perspective even when they are not explicitly instructed to do so. This is as a result of the participants assuming that the adoption of differing perspectives is an important part of the experiment. It is worth noting that the effect of top-down knowledge upon participant attention to features within an experimental set up, and specifically, within the stimuli presented has been well-established since the findings of Folk, Remington, and Johnston (1992). Indeed, the effects of attention work have shown how a stimulus that is nominally task irrelevant can in fact form part of an observer's response cue. Most importantly, this type of attentional influence has been shown to occur in perspective taking paradigms (e.g., Stephenson & Wicklund, 1983). To reiterate, merely instructing participants to consider their own egocentric perspective seems to induce consideration of an alternative allocentric perspective. As a consequence, other authors (e.g., Cole et al. (2015, 2016, 2017; Conway et al., 2017) did not include the manipulation of forced adopted perspectives. Results of these experiments showed perspective taking-like data under this condition (but recall also did so when the agent could not see). However, the extent to which spontaneous perspective taking is depending upon other processes still needs to be further explored. Especially in relation to the inconsistencies within the dot perspective paradigm.

1.5.4 Gaze Cueing

Another influential paradigm that has been argued to influence spontaneous perspective taking is gaze cueing. Recall that gaze cueing is the finding in which the observation of another individual's attention influences the attention for the observer (Nuku & Bekkering, 2008; Teufel, Alexis, Clayton, & Davis, 2010; Teufel et al., 2009; Teufel, Fletcher, & Davis, 2010). The majority of literature investigating this phenomenon presents participants with a face that directs attention to one side of the display. This movement is presented in conjunction with a target that is presented either in the gazed-at direction ('Valid') or on the opposite side of the display ('Invalid'; Frischen & Kingstone, 1998; Langton & Bruce, 1999). Differences in RT consequently lead authors to conclude that seeing gaze movements trigger the attention of the observer to shift accordingly. Consequently, RT for Valid conditions are increased, and decreased for Invalid conditions. Additionally, it has been suggested that gaze direction can be used to imply intentions and goals associated with the object that is being attended to (Calder et al. 2002; Nuku & Bekkering, 2008; Morgan, Freeth, & Smith, 2018). Yet there are also authors that dispute this claim (Driver et al. 1999; Caron, Butler, & Brooks, 2002; Teufel et al. 2010). Cole et al. (2015) combined the use of the gaze cueing procedure with a traditional nonhuman animal attention task, in the form of an occluding barrier. The authors found the same patterns for validity consistent with other gaze cueing research, irrespective of the addition of an occluding barrier. Consequently, Cole et al. (2015) concluded that mental state attribution, in the form of 'seeing' is not (reliably) modulated by the gaze cueing paradigm.

In relation to the so-called spontaneous perspective taking notion, gaze cueing can be argued to be significantly influential. For example, Samson et al. (2010), Teufel et al. (2010), and Gardner et al. (2018) can all be argued to be affected by gaze cueing. This criticism is supported by the work of Cole and colleagues (2015; 2016) who were unable to isolate the spontaneous perspective taking effect to conditions in which the avatar was able to see the target. Instead, the effect was observed in Valid conditions regardless of the visibility manipulations. However, this criticism mainly resides within the reflective gaze following and dot perspective paradigm methodologies. Conversely the ambiguous number paradigm emphasises comprehension, as participants are required to interpret the ambiguous number, thus gaze cuing has not as yet been extensively examined. Therefore, future work would benefit from the addition of occluding barriers in the ambiguous number paradigm, which has previously been explored in the gaze cueing (Cole et al. 2015) and dot perspective method (Cole et al. 2016).

1.5.5 Mental Models

Craik (1943) proposed that humans use small-scale models when processing information in the form of a mental model. Visual stimuli and written descriptions are two examples of the information that can be used in the formation of these small-scale representations. The depth of processing required to form these small-scale representations, is one area of investigation that has been popular in the development of this field. For example, Mani and Johnson-Laird (1982) attempted to investigate the importance of spatial descriptions upon the formation of mental models. They found that a greater depth of processing is required when forming a mental model, which reflects upon the improved recall. Mani and Johnson-Laird (1982) concluded that there are two types of encoding spatial descriptions. Firstly, propositional representations are relatively easy to process but are harder to recall. Secondly mental models, which are harder to process but are easier to recall. Consequently, the work of Mani and Johnson-Laird (1982) would suggest that mental models require a greater depth of processing when being encoded, which increases the recall ability. Craik & Lockhart, (1972) and Johnson-Laird & Bethell-Fox, (1978) support this finding. Once these representations are processed and encoded, they can then be used as a cue or reminder to formulate judgements about a scene (Tversky, 1981). Applying this concept to the spontaneous perspective taking notion, it could be argued that participants may not be assuming the allocentric perspective, as suggested, but instead be developing a mental model of the scene. In other words, the participant is not transforming their sense of self into the position of the avatar or agent, but instead is using the avatar or agent, as well as all other forms of information to create a mental model of the scene. This mental model can therefore be used to form judgements when the participant is asked questions regarding the scene. In this sense the discrepancies in terms of RT, may not be due to the assumption of an allocentric perspective, but instead be due to the processing of the scene, and the mental model transformations required to generate the necessary judgements. However, this is a considerable theoretical debate, which would require examination of brain region activation to support or refute the mental models claim. This debate will now be extended in relation to mental imagery and rotation in the following section.

1.5.6 Mental Imagery and Rotation

Building upon mental models, mental imagery and rotation is another significant issue that can be applied to spontaneous perspective taking. In terms of mental imagery, the form that mental models take has been heavily disputed. Kosslyn (1994) claims that mental models are processed using visual representation, which Dennett (1991) supports. For example, if an individual were asked to think about their car, Kosslyn would claim that the individual would hold a small-scale image of their car in their "mind's eye". However, Pylyshyn (1973) disagrees with this claim, and instead suggests that the individual would use descriptions, prior experiences, and pre-existing knowledge. Thus, Pylyshyn would suggest that when an individual is required to think about their car, they would simple know what model, make and colour it is, due to pre-existing knowledge, and not because of a small-scale image held in their minds eye. Interestingly, advances in neuroimaging have highlighted different activated neural pathways for images and prior knowledge dependency (O'Craven, & Kanwisher, 2000; Kosslyn, & Thompson, 2003), yet the results conflict and the debate of mental imagery remains.

As previously stated, discrepancies in RT during experimentation on spontaneous perspective taking could be a result of mental models and the required transformation of the mental image, and not the assumption of an allocentric perspective. Shepard and Metzler (1971) supports this claim as they found that RT could be progressively influenced with the increased number of mental rotations required for processing. Just and Carpenter (1976) and Hochberg and Gellman (1977) support this claim. Hence, 'spontaneous perspective taking' may actually be a function for the number of mental rotations required to process the mental model, and not due to the computation of the allocentric perspective. However, in order to assess this claim, clarification is needed in terms of the impact of mental transformations. Consequently, future work would benefit from identifying the number of mental rotations required, and combining this information into RT analysis.

1.5.7 Mental Self-Rotation or Object Rotation

An alternative account that may be able to explain the spontaneous perspective taking phenomenon, is object rotation. This is the ability to mentally rotate an object absent of an allocentric perspective in the form of an agent (Shepard & Metzler,1971). Object rotation has been extensive investigated, particularly in relation to spatial perspective taking (Huttenlocher &Presson, 1973; Levine, Jankovic &Palij, 1982; Kessler & Thomson, 2010). Kessler and Colleagues (e.g., Kessler, 2000; Kessler & Thomson, 2010; Kessler, & Rutherford, 2010; Kessler, & Wang, 2012) acknowledged the embodied nature of perspective taking and identified that the deeper the level of processing (e.g., level 2 visual perspective taking and level-2 type spatial perspective taking) the more cognitively demanding the process, and therefore increased effort for the embodied process. However, further classification in terms of the specific aspect of object rotation and the relevantly new strain of literature investigating the spontaneous perspective taking theory is still required.

In contrast to mental rotation of the self, which the majority of perspective taking emphasises in relation to the assumption of the alternative perspective (e.g., Samson et al. 2010; Baker et al. 2016; Gardner et al. 2018), object rotation suggests a different cognitive operation is performed. Instead of a rotation of the self, in reference to either spatial frames of reference for spatial perspective taking (Michelon & Zacks, 2006), or embodied line of sight tracing and mental transformation for visual perspective taking (Surtees et al 2013), object rotation emphasises a centralised rotation of a target object in isolation (Kessler & Thomson, 2010). Consequently, disparities when comparing the differences between these processes have arisen. Kozhevnikov et al. (2006) identified that enhanced perspective taking ability correlates with navigation skills, whereas object rotation ability did not. Kozhevnikov and Hegarty (2001) found that although perspective taking and object rotation abilities are similar, improved performance of one of these skills related to a reduced ability of the other. Additionally, a number of experiments have identified that mental self-rotation, used within perspective taking, is reportedly less cognitively demanding (is fast and accurate) compared with object rotation (Keehner et al. 2006; Wraga, Creem & Proffitt, 1999; Wraga et al. 2005; Zacks & Michelon, 2005). The increased angle required for rotation has also been found to effect mental self-rotation used within perspective taking and object rotation differently. For perspective taking, processing time remains fairly constant (e.g., Graf, 1994; Kozhevnikov & Hegarty, 2001; Keehner et al. 2006; Michelon & Zacks, 2006), whereas for object rotation, a progressive increase in RT correlates with the increased rotated angle (e.g., Shepard & Metzler, 1971; Graf, 1994; Keehner et al. 2006; Michelon & Zacks, 2006). It is these fundamental differences in which spontaneous perspective taking can be applied, as this phenomenon has not yet been extensively investigated in relation to object rotation.

Firstly, as stated above enhanced navigational skills have been correlated with mental self-rotation and perspective taking abilities, but not object rotation (Kozhevnikov et al. 2006). Recall that navigational skills have been used to support spontaneous perspective taking within joint action tasks (e.g., Surtees, et al. 2016b). Thus, this distinction would counter the dispute that object rotation could be applied instead of a self-rotation in the spontaneous perspective taking theory. Alternatively, the distinction that an enhanced ability of one rotation process often leads to a reduced ability of the other (Kozhevnikov & Hegarty, 2001), cannot be applied without specific experimentation of both rotation processes within a spontaneous perspective taking methodological paradigm. Thus, this may be one area to explore within future research. Thirdly, the key distinction that mental self-rotation is cognitively less demanding in that it is fast and more accurate compared with object rotation, is a key characteristic that directly relates to the spontaneous perspective taking phenomenon. Currently, spontaneous perspective taking has been found to be rapid and spontaneous in the assumption of an alternative visual perspective, which correlated to the embodied self-rotation account. However, if future work disputes this claim, object rotation may be one contributing mechanism identified. One way that this distinction could be assessed is through additional conditions in which an ambiguous object replaces the ambiguous number within the ambiguous number paradigm. Lastly, it has been identified that increasing the angle of rotation progressively impacts object rotation RT whereas it does not for mental self-rotation. Hence, this could be one way to disentangle the dispute that object rotation may be influencing the so-called spontaneous perspective taking phenomenon. An experiment could be created, similar to the research carried out by Michelon and Zacks (2006) in which the required rotation of perspective, be that in relation to the embodied perspective or object rotation, is simultaneously manipulated alongside the consistency of perspective for the participant and agent. Thus, if RT is affected by the progressive angle disparity, this would indicate that object rotation may be influencing the so-called spontaneous perspective taking phenomenon and warrant further investigation.

As can be seen, there are a number of cross-comparisons that can be made when investigating perspective taking in terms of mental transformations of the self or target object. Consequently, this is one area of examination that future work critically assessing the spontaneous perspective taking phenomenon may wish to explore.

1.5.8 Submentalizing

As previously outlined, ToM can also be referred to as mentalizing, and deconstructed to include submentalizing. Submentalizing is one component that can be argued to hold significance over perspective taking and spontaneous perspective taking literature. Recall, mentalizing refers to fully functioning ToM abilities, whereas submentalizing refers to general purpose cognition that simulate the consequences of mentalizing in context (Heyes, 2014). Additionally, it is worth noting that mentalizing can also be categorised as explicit and implicit, which is where the dispute of submentalizing has stemmed. The identification of implicit processing highlighted the impact of increased subjectivity and arguably inconclusive assumptions of mentalizing literature. To reiterate, the suggested claims of mentalizing can be argued to be loosely supported though weak evidence and assumptions about participants' behaviour (Heyes, 2014). An example of this that Heyes (2014) uses as support for submentalizing categorisation is the use of young children or infants within mentalizing literature. As the young children or infants are unable to communicate through language, assumptions in relation to the motives behind their behaviours are made in terms of assumed ToM processing, which may have actually been caused through submentalizing (Heyes, 2014). Thus, the challenge of submentalizing has emerged.

In relation to the spontaneous perspective taking literature, the conclusion of adopting another's perspective is a significant implicit claim about the visual processing of information through perspectives and ToM abilities. It could be argued that spontaneous perspective taking authors should also acknowledge other mechanisms of human cognition that may simulate the effects of perspective taking. For example, gaze cuing is one example of an adapted cognitive shortcut that can be used to process a visual scene and simulate spontaneous perspective taking-like effects. Furlanetto, Cavallo, Manera, Tversky, and Becchio (2013) investigated the importance of gaze cuing and spontaneous perspective taking, by examining the perceived level of interaction between the object and the model, and the influence of intention and gaze cuing.

Initially, Furlanetto et al. (2013) presented participants with videos, depicting different levels of interaction between a model and object. Participants were then asked where the target object was in relation to another object. Verbal responses of participants were recorded either in neutral, first, or third person perspective. Consequently, a third person perspective would suggest a spontaneous perspective taking response. The authors found that the increased level of interaction between the model and the target increased the likelihood of a spontaneous perspective taking response.

Although it has been suggested that congruent gaze following paired with action can aid in an individual's inferences of another's intentions (Pierno et al. 2006), incongruence can be argued to increase spontaneous perspective taking at a higher rate, due to the greater need to understand another individual's ambiguous actions. For example, when pairing inconsistent direction of a model's gaze and action, spontaneous perspective taking is essential in terms of understanding the model's intentions. Furlanetto et al. (2013) assessed this claim by presenting participants with videos that depicted either; paired gaze and action of the model and target, ambiguous action, or neutral action. In addition, Furlanetto et al. (2013) blurred the face of the model, which obstructed participants view, but accounted for the manipulation of ambiguous intentions of the model. Furlanetto et al. (2013) found that there was only a marginal increase in participants adopting a third person perspective for the paired gaze and action, compared with the ambiguous intention condition. Thus, spontaneous perspective taking responses increased when gaze and action were incongruent, causing the model's intention to become ambiguous. This shows the complexity of spontaneous perspective taking. The process may not be a simple implicit mentalizing or submentalizing process, as conditional factors, such as the model's intentions in Furlanetto et al. (2013) case, can affect the likelihood of the effect occurring.

Moreover, using a gaze cuing task Bukowski, Hietanen, and Samson (2015) identified the congruency effect associated with spontaneous perspective taking, only when participants had enough time to process the scene. Yet, Bukowski, et al. (2015) also found that if the experimental instruction draws the attention of participants towards the avatar, the time allowing comprehension of the scene was not needed for the congruency affect to be identified. This would suggest that when individuals are given time to comprehend a scene, cognitive load is reduced and automatic cognitive shortcuts, such as gaze cuing and spontaneous perspective taking, are more likely to be implemented. However, if attention is drawn to the cues, automatic cognitive shortcuts are implemented without the need for comprehension. Thus, the following question arises; *are these processes a simulation of ToM, or are automatic ToM abilities not controlled by intentions or general ToM boundaries*?

Overall, as shown by Furlanetto et al. (2013), the so-called spontaneous perspective taking notion is significantly influenced by the model's mundane behaviours, as well as gaze cuing, which supports the criticism of submentalizing that Heyes (2014) suggests. Thus, the spontaneous perspective taking literature, as well as other strains of ToM processes, need to be precise in their assumptions and ensure that a clear identification of whether implicit or explicit processes are driving conclusions. For when concluding in terms of implicit mentalizing, the processes cannot be divided from submentalizing, as they may be simulating the consequences of mentalizing in a specific context. Subsequently, future mentalizing research, including the stain of spontaneous perspective taking, would benefit from precise conclusions in terms of classification of implicit or explicit mentalizing processes.

1.5.9 Knowledge Attribution

Another philosophical issue that can be applied to the spontaneous perspective taking idea is knowledge attribution. Humans are fully functional, cognitive thinkers that have an influential capacity to hypotheses about the intentions and goals of others, through ToM abilities (Calder et al. 2002; Nuku & Bekkering, 2008). As previously stated, these abilities encompass several components from visual perspective, intentions, desires, mental states, and knowledge, to name a few. It is the conjunction of these components that contribute to the overall ability of ToM, and therefore where the issue in question arises. Advocates of the spontaneous perspective taking notion make bold claims that individuals compute the visual experience of another both *rapidly* and *spontaneously*. However, it could be argued that human nature is not as reductionist as claimed by these authors. Instead, comprehension of the intentions and desires, through mental models may also be occurring during spontaneous perspective taking. For example, the Reaching condition of Zhao et al. (2015) received the most allocentric responses, in comparison to the Gaze only condition. If spontaneous perspective taking is in fact solely associated with visual perspectives, then the two conditions should remain similar in terms of allocentric perspective adoption, as both depicted the same gaze direction. However, the intentions of the model can be argued to be simply *known* to the participants in the Reaching condition, which in addition to the gaze direction, could have led to the increased assumption of the allocentric perspective. In other words, the participants used perspective taking as well as knowledge attribution of the model's intentions to process the scene before the allocentric response was given. This can be applied to all other literature that uses action in addition to gaze in the assessment of the so-called spontaneous perspective taking notion. Future work investigating this phenomenon in relation to knowledge attribution should investigate the impact of, intentions, actions, and gaze upon the likelihood of an allocentric response, in isolation and in conjunction with each other. For example, the head of an avatar could be blacked out, leaving the torso performing an action in relation to the ambiguous information.

1.5.10 Summary

In the previous section the on-going debates of task relevance, empirical inconsistencies, gaze cueing, mental models, mental imagery and rotation, mental self-rotation, submentalizing, and knowledge attribution were outlined. Additionally, suggestions were made as to how future research could further these debates in terms of the so-called spontaneous perspective taking concept. However, these are just a few of the issues surrounding spontaneous perspective taking. Future work may benefit from experimenting with other limited debates also. In the following section, the debate concerning automaticity will be outlined.

1.6 Automaticity

1.6.1 Overview

Automaticity is a controversial theoretical concept with implications for many areas of psychology. It is worth noting that for the current thesis the terms ‘automatic’ and ‘spontaneous’ will be used interchangeably to improve the ease of narrative, with both terms assessing the theoretical concept of automaticity. In terms of the spontaneous perspective taking theory, this debate is specifically concerned with the ‘spontaneous’ and ‘rapid’ claim. However, not only can spontaneous perspective taking be debated in terms what it means to be spontaneous or automatic, but the empirical methodologies can also be challenged by whether ‘automatic’ processes are soundly being investigated. Three theories that can be used to define automaticity will be outlined in the following section. Specifically, the gradual view (Logan, 1985), the conditioned approach (Bargh, 1992), and goal dependency (Bargh & Gollwitzer, 1994). These theories will also be applied to the theoretical background and empirical investigation of the spontaneous perspective taking phenomenon. After these theories are sketched, the central issues that surround defining an automatic or spontaneous process will also be outlined. Lastly, what an ‘automatic’ or ‘spontaneous’ process is, as well as the bounds of this definition, specifically in terms of this thesis will be discussed.

1.6.2 Gradual View

The gradual view of automaticity suggests that automatic processes reside on a continuum, ranging along the breadth of this continuum, which may change depending upon circumstance (Logan, 1985). Consequently, the gradual view suggests that automatic processes are not dichotomous, as they are not bound to a stable positioning upon the continuum. For example, breathing is an automatic process that humans are not consciously aware of when they are not focusing on the behaviour. However, as an individual becomes aware of their breathing, the behaviour becomes less and less automatic and thus the position of breathing on the automatic continuum will change. However, Moor and Houwer (2006) challenged the gradual view of automaticity, arguing that it reduces the overall value of automatic processes, for during a truly autonomous process, awareness of the action should not alter the functioning it any way.

In terms of the dot perspective task, the procedure draws the attention of participants to the allocentric perspectives that can be adopted by highlighting 'YOU' or 'HIM/HER'. Subsequently, the assumed spontaneous perspective taking process can be challenged in terms of whether it is in fact 'spontaneous' or 'automatic', due to the attention of participants being drawn to the different perspectives. Alternatively, Zhao et al. (2016) and Cole et al. (2016) did not draw the attention of participants to the differing perspectives that could be adopted. Thus, the conclusion drawn in terms of spontaneous perspective taking for Zhao et al. (2016), and contradictions in the case of Cole et al. (2016), cannot be challenged by the gradual view of automaticity. Although this is a brief overview of the challenges that the gradual view of automaticity has upon the theoretical framework of spontaneous perspective taking, it highlights the challenges that are faced when claims are made in terms of automaticity. Future research regarding spontaneous perspective taking should ensure that wherever possible participants attention is diverted from differences in perspectives.

1.6.3 Conditioned Approach

It can be argued that the gradual view of automaticity does not draw a definitive line between what is and what is not automatic (Bragh, 1992). In contrast, the conditioned approach, identified by Bragh (1992), suggests that automatic processes must always run to completion when started, yet, are dependent upon preconditions, such as triggers or attentional cues. When applying this theory to the spontaneous perspective taking idea, the need to highlight the features or preconditions that the phenomenon is dependent upon is one area of research to which previous literature lacks. For example, the importance of a human figure, or orientation cue is one question that should be investigated further, as spontaneous perspective taking has been found during cases with a human presence (Samson et al. 2010; Zhao et al. 2015) and without (Santiesteban et al. 2013; Nielsen et al. 2015). Additionally, future work investigating the preconditions that spontaneous perspective taking is dependent upon should also examine whether inanimate objects, as well as other entities with a perceived sociality and therefore perspective, trigger this phenomenon. Subsequently, if the so-called spontaneous perspective taking notion is found in these variations, then is the suggested spontaneous assumption of another's perspective occurring, or are other cognitive mechanisms involved?

1.6.4 Goal Dependency

Bargh (1992) extended the conditioned approach of automaticity stating that these processes can be directly related to the preconscious, post conscious, and goal dependency. Bargh and Gollwitzer (1994) expanded upon the goal dependent aspects of automaticity, stating that conscious goals can often be manipulated by unconscious intentions. Thus, this theory of automaticity can be applied to the spontaneous perspective taking idea in that participants are manipulated to form the conscious goal of performing correctly, which is influenced by unconscious intentions of conforming to emphasized responses. Thus, occluding barriers could be driving unconscious intentions, which in turn alters the conscious goals of performance. In this sense, the unconscious drives that are influenced by the presence of differing stimuli within each paradigm may be influencing the findings, thus leading to conclusions of spontaneous perspective taking. However, this is a very limited criticism, significantly more research into this assessment is required before the criticism can be extended. Future work may benefit from manipulating the procedures to alter the overall goals of the experiment using distraction tasks; therefore, assessing whether this will impact the conclusions drawn.

1.6.5 Defining an Automatic Process

When defining an automatic or spontaneous process there are a number of important preconditions to consider. As Moors and De Houwer (2006) and Reynold and Besner (2006) identified in their extensive reviews, authors have associated automaticity with a large number of characterisations and mechanisms (e.g., stimulus driven, unintentional, fast, unconscious, independent, not affected by practice). Each defining characterisation will now be discussed in relation to the so-called spontaneous perspective taking theory.

Firstly, some authors state that an automatic process is stimulus driven (e.g., Posner & Snyder, 1975; Hasher & Zacks, 1979; Brown, Gore, & Carr, 2002). In terms of spontaneous perspective taking, the current literature would both support and refute this claim as the effect has been found modulated by a stimulus (e.g., Samson et al. 2010; Furlanetto et al 2006; Baker et al. 2016), but also irrespective of a stimulus set up that should abolish the effect (e.g., Cole et al. 2015; 2016; 2017; Conway et al., 2017). Consequently, future work would benefit from differentiating away from the current trend of stimuli variation used in the current literature.

Alternatively, other authors state that an automatic process cannot be intentionally controlled (e.g., Posner & Snyder 1975, Cohen, Servan-Schreiber, &

McClelland,1992; Hasher & Zacks, 1979), and is not affected by practice (e.g., Hasher & Zacks, 1979). Consequently, spontaneous perspective taking concurs with this characterisation. The process has been repeatedly found to be an unintentional process, and identifiable irrespective of practice trials or blocks (e.g., Samson et al. 2010; Santiesteban et al. 2014; Nielsen et al. 2015 etc.).

Additionally, automatic processes have been stated to be independent of other processes (e.g., Posner & Snyder, 1975; Hasher & Zacks, 1979; Logan, 1988; Brown, Gore, & Carr, 2002), and attention (e.g., Shiffrin & Schneider 1977; Laberge & Samuels, 1974; Logan, 1988). In respect to the spontaneous perspective taking literature, there has been little investigation into the phenomenon's association with other processes. Although, reflexive gaze following has been suggested to contribute towards the effect, the extent to which spontaneous perspective taking is dependent upon other processes still needs to be further explored. However, recall that during the dot perspective paradigm participants are required to judge the number of dots from both their own egocentric perspective and, on other trials, from the allocentric agent perspective. Within certain experiments (e.g., Samson et al. 2010; Santiesteban et al. 2014; Nielsen et al. 2015) this occurs within-block such that participants are informed at the start of each trial which perspective they should adopt. Consequently, this procedure could be criticised that the participants attention is drawn to the representation of the agent's perspective even when they are not explicitly instructed to do so. This is as a result of the participants assuming that the adoption of differing perspectives is an important part of the experiment. It is worth noting that this effect of top-down knowledge upon participant attention towards features within an experimental set up, and specifically within the stimuli presented, has been a well-established finding since the conclusion of Folk, Remington, and Johnston (1992).

Indeed, the attention literature has shown how a stimulus that is normally task irrelevant can in fact form part of an observer's response cue. Most importantly, this type of attentional influence has been shown to occur in perspective taking paradigms (e.g., Stephenson & Wicklund, 1983). To reiterate, merely instructing participants to consider their own egocentric perspective can be argued to induce consideration of an alternative allocentric perspective. As a consequence, other authors (e.g., Cole et al. 2015; 2016; Conway et al., 2017) did not include the manipulation of forced adopted perspectives and found perspective taking-like data under this condition (but recall also did so when the agent could not see). Overall, the importance that attention has within the spontaneous perspective taking literature, which may lead critics to dispute the claims of the effect identified being automatic or spontaneous in nature, has been clearly demonstrated. Yet, still further clarification is needed in respect to the alternative methodological paradigms (e.g., ambiguous number paradigm).

Lastly, automatic processes have been characterised as ballistic (e.g., Hasher & Zacks 1979; Brown, Gore, & Carr, 2002), unconscious (e.g., Atkinson & Shiffrin, 1968; Laberge & Samuels, 1974; McCann, Folk, & Johnston, 1992), and fast (e.g., Neely 1977; Logan, 1988). Unfortunately, due to the early stage that spontaneous perspective taking literature is in, little investigation in terms of the characterisation of phenomenon in respect to the unconscious and ballistic nature of the process has been carried out. Therefore, future work may benefit from assessing whether spontaneous perspective taking is ballistic and driven by the unconscious. However, the current literature frequently acknowledges that spontaneous perspective taking is a '*rapid*' process. Thus, the phenomenon does concur with the notion that an automatic or spontaneous effect can be characterised as a fast process.

All in all, there are a number of ways in which to distinguish a spontaneous or automatic process, from a stimulus driven process, to a fast, unintentional process. Thus, when examining the boundaries of an automatic process, including spontaneous perspective taking, future work should be clear within its definition of what characteristics enable the assumption that the process is automatic.

1.6.6 Summary

As reviewed, the subject of automaticity is a substantial theoretical debate holding extensive significance over the spontaneous perspective taking phenomenon. Not only can spontaneous perspective taking be debated in terms what it means to be *spontaneous*, it can also be debated in terms of the characterisation of an automatic process, and whether its methodological paradigms can efficiently assess automaticity. The gradual view of automaticity (Logan, 1985) has highlighted the need to explore whether this phenomenon resides on a continuum, whereas the conditioned approach (Bargh, 1992) has emphasised the importance of preconditions that the notion is dependent upon. Furthermore, the goal dependency assessment of automaticity (Bargh & Gollwitzer, 1994) has exposed the need to alter the methodology of spontaneous perspective taking literature to include distraction tasks. In addition, the need to clarify the definition of spontaneous perspective taking as an automatic process in terms of its preconditions and mechanisms (e.g., stimulus driven, unintentional, fast, unconscious, independent, not affected by practice) has also been highlighted. Thus, along with the theoretical debates outlined in the previous section, the subject of automaticity will be carried forward into the current work, which will be critically examining the spontaneous perspective taking notion.

1.7 Summary and Current Work

As it can be seen, the spontaneous perspective taking notion has emerged rapidly from ToM and perspective taking literature. Although advances have been developed in terms of its basic properties and behavioural characteristics, disputes and criticisms remain. Mental imagery and rotation, the influence of previous knowledge, automaticity, as well as defining and categorising of implicit and explicit processes remain some of the key issues that need to be addressed. Future work on spontaneous perspective taking will benefit from focusing upon these criticisms, along with amalgamating the methodological paradigms as well as introducing different adaptations employed in other contexts. Furthermore, advances within spontaneous perspective taking could also instigate new debates that remain yet to be uncovered, for example, the potential importance of individual differences, as well as the influence of ownership. Nevertheless, spontaneous perspective taking remains a fruitful area of investigation as a result of its impact upon social communication and interaction.

As previously outlined in section *1.3.5 (Perspective Taking) Summary*, the current thesis rejects the central distinction of levels of visual perspective taking on the basis that only level 2 visual perspective taking can refer to a visual ‘perspective’. In addition, the current thesis will be assuming the behaviourist view of visual perspective taking, with the assumption that when an agent is present within a stimulus, participants response is resultantly modulated, (i.e., Salatas & Flavell, 1976; Samson et al. 2010; Cole et al. 2016). Consequently, the following empirical investigation will focus on the exploration of level 2 spontaneous visual perspective taking, using behavioural, developmental and eye movement measures. The only exception to this distinction will be Experiment 3 which will take an alternative approach and assess whether the dot perspective paradigm can be used in the assessment of level 1 spontaneous visual perspective taking. Consequently, the current work will be split into five areas of empirical investigation. Initially, Chapter Two will investigate the use of the dot perspective and ambiguous number paradigms under different manipulations of attention including, the significance of validity, occluding barriers, and avatar positioning. Chapter Three will explore eye movements in the assessment of level 2 spontaneous visual perspective taking. Specifically assessing, gaze cuing, manipulations of the orientation figure, occluding barriers, and the cross-correlation of self-reported measures of empathy. Chapter Four will examine whether level 2 spontaneous visual perspective taking is indeed routed in ToM ability by adapting the dot perspective and ambiguous number paradigms for experimentation on young children. Chapter Five will investigate perceived ownership over a stimulus being jointly viewed to see whether this affects level 2 spontaneous visual perspective taking using both a novel single response and standard RT procedure. Finally, Chapter Six will examine an alternative theory suggesting that the agent can act as a reference point that anchors and orientates an image, using the single response procedure.

Although, each chapter will assess different components of spontaneous perspective taking, the underlying rationale will remain the same. If the so-called spontaneous visual perspective taking notion is in fact the (‘*rapid*’ and ‘*spontaneous*’) assumption of an allocentric visual perspective, the process should only be exhibited with a free viewing human presence, regardless of the current works manipulations. It is also worth noting that for the current thesis the terms ‘automatic’ and ‘spontaneous’ will be used interchangeably to improve the ease of narrative. To reiterate, for the purpose of this thesis spontaneous and automaticity are defined in the same way. Plus, as the central question of spontaneous perspective taking will be challenging the concept in regards to a number of characterisations of automaticity, a clear outline of how the phenomenon is termed as an automatic or spontaneous process should be acknowledged. Consequently, for the purpose of this thesis, spontaneous perspective taking is defined as an automatic process as a result of the following characteristics, it is a fast or '*rapid*' process, it is absent of conscious, intentional control, and is also stimulus driven. These are the bounds to which the current thesis shall examine spontaneous perspective taking specifically relating to the assessment of whether the phenomenon can be soundly concluded in terms of spontaneity.

Chapter Two:

Perspective Validity, Occluding Barriers, and Avatar Stance

2.1 Chapter Overview

The current chapter assesses the spontaneous perspective taking theory by using validity and visibility methods (e.g., occluding physical barriers) together with the combined procedures of the dot perspective task and ambiguous number paradigm. As reviewed in Chapter One, the visibility techniques have been employed with both the dot perspective and gaze cueing paradigms. It has not however been employed with the ambiguous number technique and with what might be called the 'rubbernecking' method of gaze cueing. The experiments will in turn assess whether the suggested spontaneous visual perspective-taking phenomenon is grounded within ToM abilities or driven by alternative cognitive processes.

In Experiment 1, the RT version of the ambiguous number paradigm was employed (i.e., many repeated trials) and the location in which the agent looked was manipulated. Specifically, the agent either looked towards the ambiguous number or away from it. This procedure was repeated in Experiment 2 with the exception that the agent was either able to see the number or her view was occluded by a barrier. In Experiment 3 the dot perspective task was employed and, whilst the agent always looked towards the targets, its body was either facing towards the discs or facing away. Thus, Experiment 3 specifically assessed both this visibility manipulation, but also the influence that the orientation agent, in this case the avatar, has over the directing attention within the spontaneous visual perspective taking theory.

2.2 Experiment 1 – Perspective Validity and the Ambiguous Number Paradigm

2.2.1 Introduction

Two approaches were combined in the assessment of spontaneous perspective taking, as previous research has not examined these paradigms in conjunction with each other. These approaches were the dot perspective and ambiguous number paradigms. The former closely resembles what has often been referred to as the gaze cueing paradigm. It has previously been found (Driver, et al. 1999; Frischen & Kingstone, 1998) that RTs are significantly shorter, for so-called Valid trials (when an agent is gazing towards the critical stimulus) compared with Invalid trials (when the gaze is averted elsewhere). This was employed together with the ambiguous number paradigm. Thus, on half of the trials the number can be interpreted as the same for the participant and agent perspective (i.e., Unambiguous) whilst for the second half the number can be interpreted differently (i.e., Ambiguous). Additionally, half of the trials the agent was looking at the number (i.e., Valid) and the second half she looked elsewhere (i.e., Invalid). Based on previous work, it was predicted that RT would be significantly shorter for Unambiguous trials relative to Ambiguous trials. That is, the ambiguous number effect should be replicated. The same pattern of data should not however be observed when the agent is not viewing the critical stimulus (i.e., during Invalid conditions). If the same pattern of data occurs when the agent is not viewing the number; this will suggest that there is another cognitive mechanism taking place, unlikely to be concerned with the spontaneous visual perspective taking effect.

2.2.2 Method

Participants

A power analysis for a repeated measures within-participants factor ANOVA was conducted in G*Power to determine a sufficient sample size using an alpha level of 0.05, a power of 0.80 and an effect size of 0.6 (Cohen, 1988). Based on these assumptions, the sample size required is 28. Thus, allowing for any unforeseen issues with recruitment, 33 participants were recruited with a mean age of 21.36 (*SD* = 3.23, range = 18-31), with 24 of the sample being female (nine male). Four participants identified themselves as left-handed whilst the remaining 29 identified as right-handed. All participants reported normal, or corrected to normal vision, and were recruited through the University of Essex online volunteer portal known as 'SONA', with participants being reimbursed for their time.

Stimuli and Apparatus

Eight images were generated. Four of these are shown in Figure 2.1. The other four were identical with the sole exception that the ambiguous numbers of '*88*' and '*89*' were employed rather that '*68*' and *'69'*. In terms of the ambiguity manipulation, '*88*' and '*69*' is Unambiguous in that these two numbers can be interrupted the same for the viewer and the agent. Whereas, '*89*' and '*68*' is Ambiguous as the numbers can be interpreted differently depending on the perspective adopted, i.e. the viewers or the agent's perspective. For example, '*68*' can be interrupted as '*68*' from the viewer perspective but '*89*' from the agent perspective. Two pairs of ambiguous numbers were used to ensure that the ambiguity manipulation was counterbalanced across both the left-hand and right-hand responses, whilst simultaneously accounting for the left-to-right ascending order of the numbers. In other words, an ambiguous number was associated with the left-hand response in one pair and the right-hand response in the other, thus controlling for any effect or influence handedness has on RT.

The presented images were 3264 x 2448 pixels in size, with a display ratio of 816 x 612 pixels. For Valid trials the agent was observed as directly looking at the number, whereas for Invalid trials the agent was depicted as looking off to the side of the scene.

The experiment was administered used the SuperLab 5.0 desktop platform, using an Apple iMac with a screen size of 27-inch with the Apple iMac running on version 10.116 with 5120 x 2880 display dimensions. Participant responses and RTs were recorded using a standard keyboard, and the data were analysed using IBM SPSS Statistics (Version 21) computer software.

Figure 2.1: The two top images represent the two critical conditions in the basic ambiguous number paradigm. The left number is Unambiguous in that it is the same for the viewer and the agent (i.e., '69'). In contrast, the number in the right-hand image is Ambiguous; it is '68' for the viewer but '89' for the agent. If this effect is due to a representation of the agent's perspective then no ambiguous number effect should occur when the agent is not looking at the number, as in the two bottom images.

Design and Procedure

A 2 x 2 repeated measures design was used with ambiguity of the number (Ambiguous, Unambiguous), and validity (Valid, Invalid) being manipulated. Ambiguity was manipulated by the number being interpreted as either the same, for the participant and agent perspective (Unambiguous), or differently (Ambiguous). Validity was manipulated by manipulating whether the agent was either gazing towards (Valid), or away from (Invalid) the ambiguous number.

Participants began with their hands positioned on the two response keys at the start of each block. The '*Z*' key was used to indicate a left-hand response corresponding to the figures '*68*' or '*88*', and the 'M' key was used for right-hand response, corresponding to figures '*69*' or '*89*'. During each block the stimulus that depicted the experimental figure remained the same throughout. At the start and end of each trial, the same stimulus was displayed absent of the target number, so that each trial could be easily differentiated. 250 milliseconds after the onset of the neutral image, the experimental figure was presented. The trial ended only after the participant responded using the respective keys on the keyboard. At the end of the trial the neutral image again was shown for 250 milliseconds prior to the onset of the next trial. See Figure 2.2 for a visual representation of the trial sequence used in Experiment 1.

Two blocks used the '*68*' and '*69*' ambiguous number pair, with an Invalid and Valid block variation, which was repeated with the second ambiguous number pair of '*88*' and '*89*', again with the Valid and Invalid variations. Each block presented 96 trials resulting in a total number of 384 trials per participant.

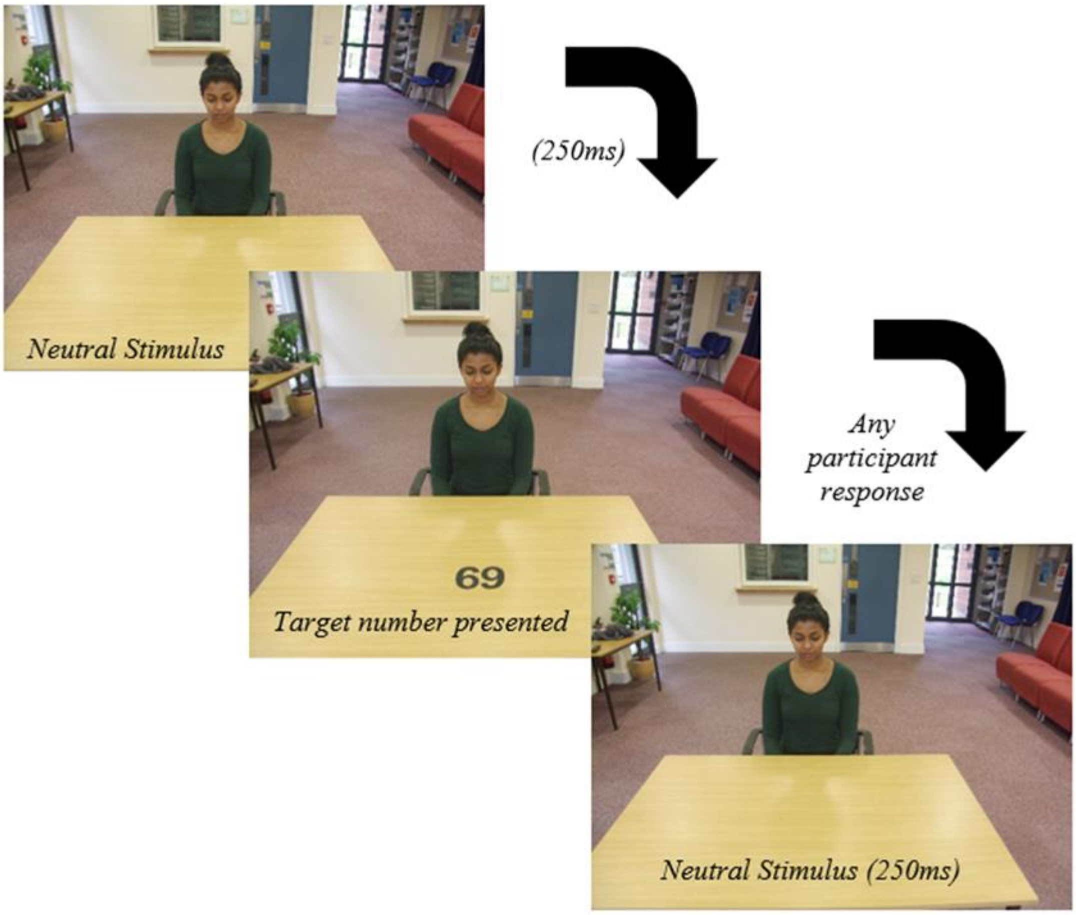

Figure 2.2 The trial sequence used in Experiment 1

Pre-Test Experiment

In accordance with Surtees et al. (2016), a 'pre-test' of the four experimental numbers was administered first, in isolation, absent of any other manipulations. This was undertaken to rule out the possibility that any subsequent effects were due to the numbers we used producing differential in RTs irrespective of any social context. A separate sample of participants (N = 12) was used. This initial assessment was identical to Experiment 1, with the sole exception that only the numbers were presented on a beige background (i.e., no agent). Results found that there was no significant difference in mean RT (unambiguous = 453ms, *SD* = 55; ambiguous = 455ms, *SD* = 54; $t(11) = .38$, $p > .71$).

2.2.3 Results and Discussion

Data Preparation. The data were collated and reorganised using Microsoft Excel and collapsed into a single data file. The data were then transformed to produce mean RTs, excluding any outliners that were 2 standard deviations above or below the mean, for the two levels of each manipulated factor of validity and consistency. Additionally, a percentage for correct self-perspective responses for each condition was generated in order for error rate analysis between factors to be analysed.

Reaction Time. Figure 2.3 depicts the mean RTs for the four conditions. A repeated measures ANOVA with ambiguity and validity as within-participant factors found a significant main effect of ambiguity, $F(1, 32) = 17.4$, $p < .001$, $n_p^2 = .35$, but no significant main effect of validity was found, $F(1, 32) = .54$, $p = .46$, $n_p^2 = .01$. There was also no significant interaction, $F(1, 32) = 1.1$, $p = .29$, $n_p^2 = .03$. As can be seen participants were significantly faster to respond during Unambiguous trials, irrespective of validity.

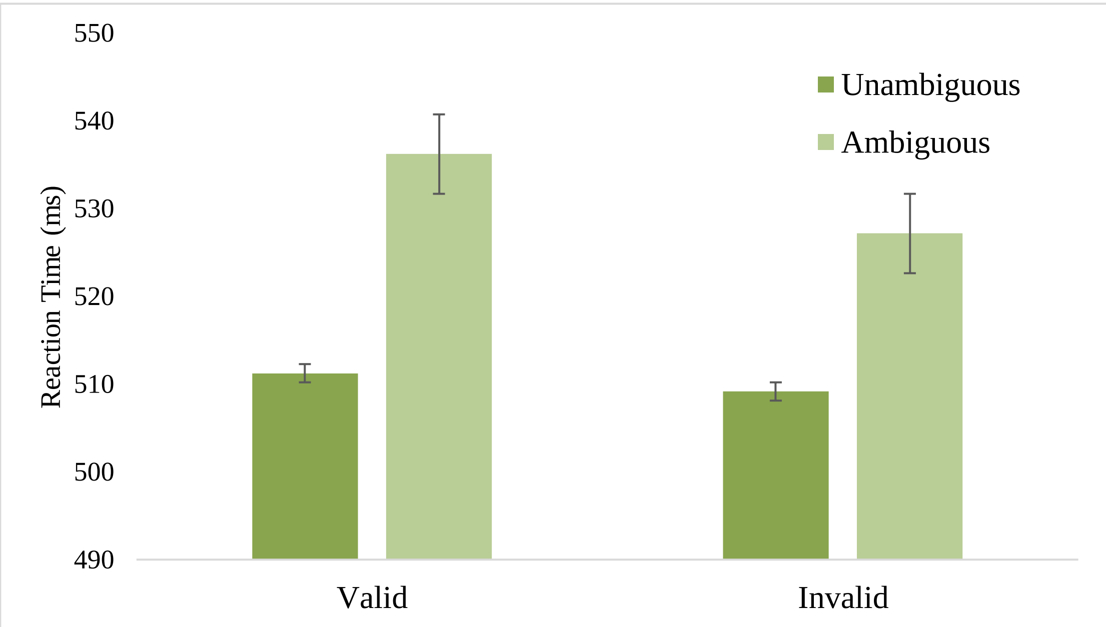

Figure 2.3: Mean RTs for each of the four conditions, with standard error bars included

Error Rate. A repeated measures ANOVA with ambiguity and validity as within-participant factors did not find a significant main effect of ambiguity, $F(1, 32) = 1.2$, $p = .27$, $n_p^2 = .04$, or of validity, $F(1, 32) = .49$, $p = .49$, $n_p^2 = .02$. There was also no significant interaction, $F(1, 32) = 1.19$, $p = .28$, $n_p^2 = .04$. Consequently, there was no significant finding in terms of error rate.

Overall, taken in isolation the significant ambiguity finding suggests that participants were assuming the visual perspective of the agent in the photograph as the current literature has previously suggested. In other words, during Ambiguous trials, RT was significantly longer when compared with Unambiguous trials, in which the visual perspective of the agent and participants are identical. However, the same pattern in RT was found irrespective of validity. If the assumption of the agent's visual perspective is driving the result of this experiment, then no such effect should have occurred in the Invalid conditions, as the agent is not directly looking at the ambiguous number. This coincides with the findings of Cole et al. (2015) whom also found that perspective taking type data trends occurred in conditions where a perspective could not be assumed. Therefore, this experiment does not support the spontaneous perspective taking concept, that an individual *rapidly* and *spontaneously assumes* the visual perspective of another.

Further Analysis

Although results from the 'pre-test' control experiment (see *2.2.2 Method*) suggest that the ambiguous numbers used do not in isolation generate differential RTs, additional analysis was undertaken to assess whether the present results occur in both of the ambiguous number pairs.

'68'-'69' Ambiguous Number Pair

Reaction Time. Figure 2.4 depicts the mean RT for the four conditions using the '*68'*-'*69'* ambiguous number pair. A repeated measures ANOVA with ambiguity and validity as within-participant factors found a significant main effect of ambiguity, $F(1, 32) = 5.9$, $p = .02$, $n_p^2 = .16$, but no significant main effect of validity was found, $F(1, 32) = 2.02$, $p =. 17$, $n_p^2 = .06$. There was also no significant interaction, $F(1, 32) = .52$, $p = .48$, $n_p^2 = .02$. In other words, participants were significantly faster to respond during Unambiguous trials using the number of '*69*', when compared with Ambiguous trials using the number '*68*', irrespective of validity.

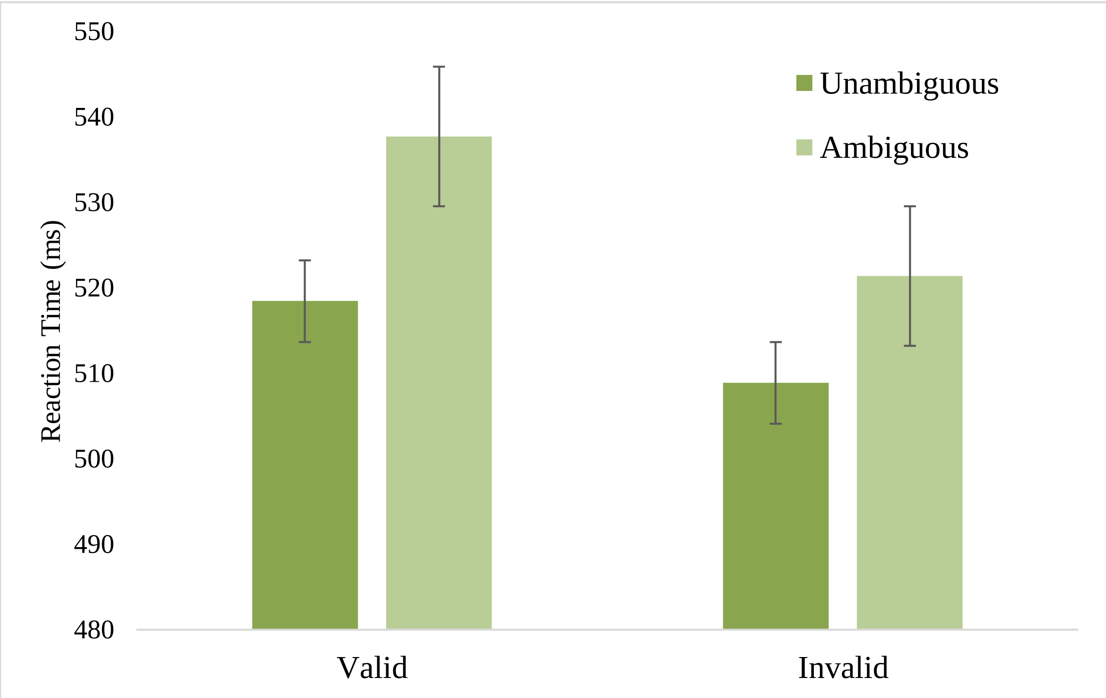

Figure 2.4: Mean RTs for each of the four conditions for the '68' and '69' ambiguous number pair, with standard error bars included

Error Rate. A repeated measures ANOVA with ambiguity and validity as within-participants factors did not find a significant main effect of ambiguity, $F(1, 32) = .24$, $p = .63$, $n_p^2 = .008$, or of validity, $F(1, 32) = 1.85$, $p = .18$, $n_p^2 = .06$. There was also no significant interaction, $F(1, 32) = .15$, $p = .70$, $n_p^2 = .005$. As can be seen, there was no significant difference in error rates for the '*68*' and '*69*' ambiguous number pair.

'88'-'89' Ambiguous Number Pair

Reaction Time. Figure 2.5 depicts the mean RT for the four conditions using the *'88'-'89'* ambiguous number pair. A repeated measures ANOVA with ambiguity and validity as within-participant factors found a significant main effect of ambiguity, $F(1, 32) = 12.97$, $p < .001$, $n_p^2 = .29$, but no significant main effect of validity was found, $F(1, 32) = .04$, $p = .84$, $n_p^2 = .001$. There was also no significant interaction, $F(1, 32) = .72$, $p = .40$, $n_p^2 = .02$. In other words, participants were significantly faster to respond during Unambiguous trials with the ambiguous number of '*88*', when compared with Ambiguous trials using '*89*', irrespective of validity.

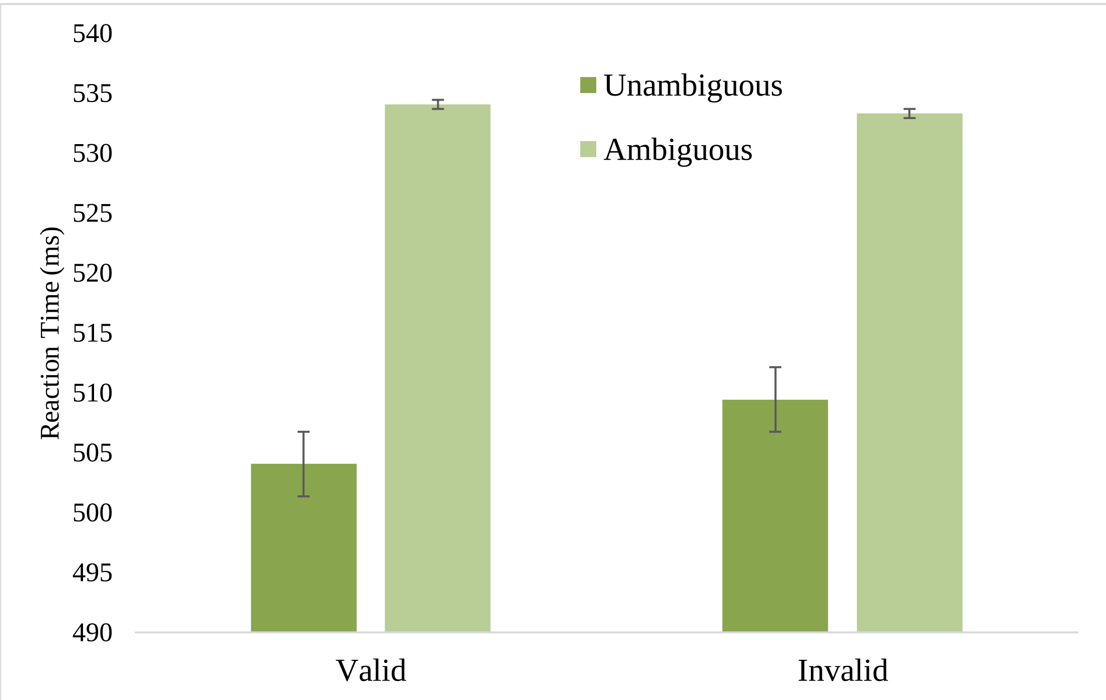

Figure 2.5: Mean RTs for each of the four conditions for the '88' and '89' ambiguous number pair, with standard error bars included

Error Rate. A repeated measures ANOVA with ambiguity and validity as within-participant factors found a significant main effect of ambiguity, $F(1, 32) = 4.63$, $p = .04$, $n_p^2 = .13$, but no significant main effect of validity was found, $F(1, 32) = 3.57$, $p = .07$, $n_p^2 = .10$. There was also no significant interaction, $F(1, 32) = 1.49$, $p = .23$, $n_p^2 = .05$. Therefore, participants were significantly more error prone during trials using the ambiguous number of '*89*', compared to trials using the ambiguous number of '*88*', irrespective of validity.

The results of this additional analysis have found that similar patterns in the data were exhibited within both of the ambiguous number pairs. Consequently, it can be assumed that neither of the ambiguous number pairs, or the spatial mapping of response hand was driving the results obtained in Experiment 1.

Overall, in isolation the ambiguity effect could suggest that participants were assuming the visual perspective of the agent. However, the same pattern was found irrespective of validity. If the assumption of the agent's visual perspective is driving the result of this experiment, no such effect should have occurred when the agent was not directly looking at the ambiguous number. This coincides with the findings of Cole et al. (2015) whom also found perspective taking-like data during conditions where a perspective could not be assumed. Therefore, this experiment does not support the previously concluded spontaneous visual perspective taking theory. Instead, Experiment 1 suggests that alternative cognitive mechanisms not associated with assumed perspectives are driving the effects identified within this manipulation.

Experiment 2 will repeat the procedure employed in the current experiment with the main difference being that visibility will be manipulated instead of validity. Rather than the agent looking away, a physical barrier will be located between the target and the agent. Importantly, the agent's direction of attention will remain the same i.e., directed to the target, in both conditions. If spontaneous perspective taking is indeed the spontaneous and rapid assumption of another's visual perspective, the barrier should abolish this phenomenon. However, if alternative cognitive processes, such as knowledge attribution, are driving these results then the addition of the barrier should not hinder the findings, and the same pattern in RTs should be observed.

<u>2.3 Experiment 2 – An Occluding Barrier and the Ambiguous Number Paradigm</u>

2.3.1 Introduction

The same experimental procedure used in Experiment 1 was repeated in Experiment 2, however an alternative method for manipulated visual perspectives was applied. Instead of manipulating the gaze of the agent, during so-called Invalid trials, the visibility method was used (i.e., Cole et al. 2016). Specifically, the agent held a newspaper, resting on the table, and looked down towards it. This therefore controlled for where the agent was attending. As many authors have pointed out (e.g., Cole et al, 2016), the dot perspective task and gaze cueing paradigm confound the agent's view with the location of visual attention. Consequently, in the present experiment, the agent always looks to the same position, i.e., down towards the table/newspaper. To reiterate, Experiment 2 specifically assessed the possible confounds of diverting the agent's attention to see whether this manipulation is driving the discrepancies in spontaneous visual perspective taking literature.

All other aspects of the procedure, in terms of blocking of trials and ambiguous numbers, were kept as a direct replication of Experiment 1. It is predicted that if the suggested spontaneous perspective taking theory is indeed due to the representation of another's visual perspective, obscuring the view of the ambiguous number for the agent will eradicate any spontaneous perspective taking-like effects. However, if similar patterns in the data are found, irrespective of whether the agent can see the number or not, the results would suggest that an alternative cognitive process could be driving spontaneous perspective taking effects in this experiment and arguably previous research also.

2.3.2 Method

All aspects of the Method were as described for Experiment 1, with the following three exceptions. One, a new sample of participants was gathered. Two, visibility was manipulated instead of validity using an occluding barrier rather than the gaze direction of the agent and three, participants were informed that the agent could not see the target number for half of the trials (they were not told how this was achieved).

In terms of the sample, 38 participants were recruited with a mean age of 20.00 (*SD* = 1.86, range = 18-27), with 27 of the sample being female (11 male). Six participants identified themselves as left-handed whilst the remaining 32 identified themselves as right-handed. All 38 participants reported normal, or corrected to normal vision, and were recruited through the University of Essex online volunteer portal known as ‘SONA’, with participants being reimbursed for their time.

Additionally, instead of manipulating validity Experiment 2 manipulated visibility, in which the agent was depicted as holding and gazing towards a newspaper resting on the table directly in front of the ambiguous number. Consequently, the agent’s ability to perceive the number was directly obstructed, hereafter known as Non-visible. See Figure 2.6 for an example of the different stimuli that was used for each condition.

Figure 2.6: The two top images represent the two critical conditions in the basic ambiguous number paradigm. If the ambiguous number effect is due to a representation of the agent's visual perspective, then no such effect should occur when the agents view is obstructed by the occluding barrier in the form of a newspaper in the two bottom images.

2.3.3 Results and Discussion

Data Preparation. The data were collated and reorganised using Microsoft Excel and collapsed into a single data file. The data were then transformed to produce mean RTs, excluding any outliners that were 2 standard deviations above or below the mean, for the two levels of each manipulated factor of visibility and ambiguity. Additionally, a percentage for correct self-perspective responses for each condition was generated in order for analysis of error rate between factors.

Reaction Time. Figure 2.7 depicts the mean RT for the four conditions. A repeated measures ANOVA with ambiguity and visibility as within-participant factors found a significant main effect of ambiguity, $F(1, 37) = 25.73$, $p < .0001$, $n_p^2 = .41$, but no significant main effect of visibility was found, $F(1, 37) = .8$, $p = .38$, $n_p^2 = .02$. A significant interaction was also found, $F(1, 37) = 4.13$, $p < .05$, $n_p^2 = .1$.

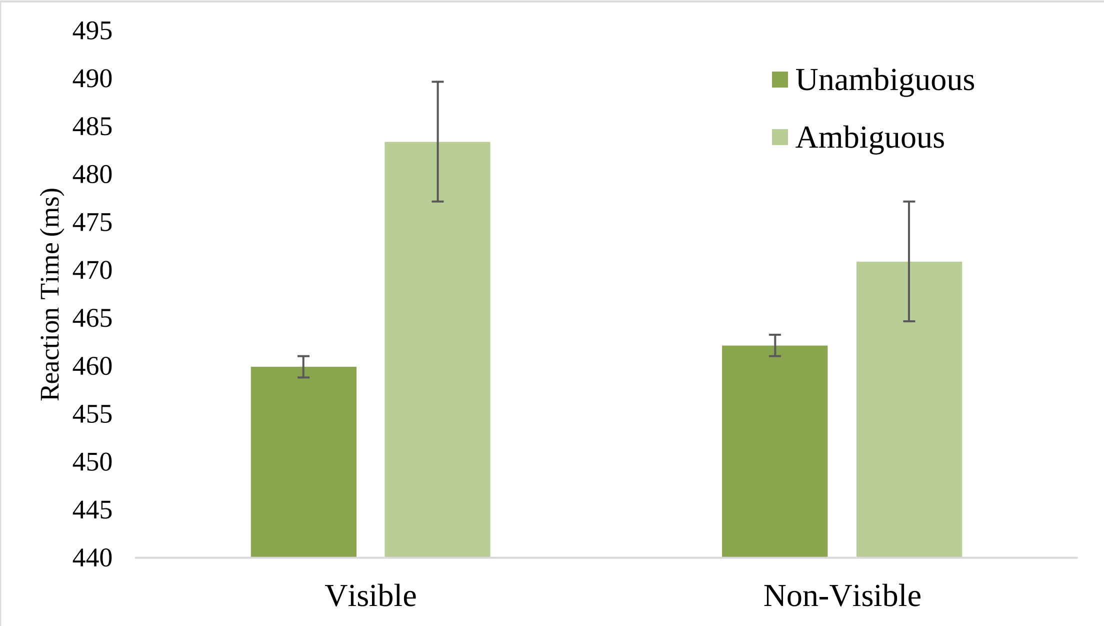

Figure 2.7: Mean RTs for each the four conditions, with standard error bars included

To examine the significant interaction three planned comparison *t*-tests were conducted. Firstly, there was a significant difference in RT for Visible, Unambiguous (*M*=459.93, *SD*=55.57) and Visible, Ambiguous (*M*=483.43, *SD*=60.36) conditions; $t(37) = -3.89$, $p > .001$, $d = 0.63$, $BF_{10} = 68.31$. Thus, showing that RT for Visible Unambiguous conditions was significantly shorter than Visible Ambiguous conditions. This is further supported by a Bayesian *t*-test which found the data to be 68 times more likely under the alternative hypothesis. Additionally, there was a significant difference between Non-Visible, Unambiguous (*M*=462.16, *SD*=65.77) and Non-Visible, Ambiguous (*M*=470.93, *SD*=64.49) conditions; $t(37) = -2.79$, $p >$

.008, *d* = 0.45, BF_{10} = 4.979. In other words, these results show that within the Non-Visible condition, there was a significant difference between ambiguity conditions with Unambiguous being significantly shorter than Ambiguous conditions. This is further supported by a Bayesian *t*-test which found that the data to be 5 times more likely under the alternative hypothesis. Alternatively, there was no significant difference in RT for Visible, Ambiguous (*M*=483.43, *SD*=60.36) and Non-Visible, Ambiguous (*M*=470.93, *SD*=64.49) conditions; *t*(37) = 1.85, *p* = .07, *d* = 0.3, BF_{10} = 0.81, and is further supported by a Bayesian *t*-test which found the data to be 1.2 times more likely under the null than the alternative hypothesis.

As can be seen participants were significantly faster to respond during Unambiguous conditions, irrespective of the occluding barrier. Plus, the significant interaction demonstrates that participants are significantly faster to respond during Unambiguous trials in both Visible and Non-Visible conditions. However, most importantly there was no significant difference in RT between Ambiguous Visible and Ambiguous Non-Visible conditions. These results therefore dispute the spontaneous visual perspective taking claim. The claim is clear, if this phenomenon is in fact the spontaneous assumption of an alternative visual perspective, the effect should not have been identifiable during Non-Visible conditions at all, where the agent was unable to view the ambiguous number.

Error Rate. A repeated measures ANOVA with ambiguity and visibility as within-participant factors did not find a significant main effect of ambiguity, $F(1, 37) = .24$, $p = .63$, $n_p^2 = .006$, or visibility, $F(1, 37) = 1.83$, $p = .18$, $n_p^2 = .05$. There was also no significant interaction found, $F(1, 37) = 3.23$, $p = .08$, $n_p^2 = .08$. Consequently, there was no significant finding to report in terms of error rate.

Overall, in isolation the significant ambiguity finding suggests that participants were assuming the visual perspective of the agent as previous literature has suggested. In other words, if we take the non-barrier conditions only, the data suggests that participants were spontaneously assuming the agent's visual perspective during Ambiguous conditions, leading to the longer RTs in comparison to Unambiguous conditions. However, importantly, this ambiguity effect was found irrespective of whether the agent could see the ambiguous number of not. To reiterate the rationale for the present experiment, if the assumption of the agent's visual perspective is indeed driving the results of this experiment, then no such effect should have occurred when the alternative perspective cannot view the ambiguous number. Therefore, Experiment 2 supports the findings of Experiment 1, in that alternative cognitive processes must be driving the spontaneous visual perspective taking-like effects observed. This again coincides with the findings of Cole, et al. (2015) whom also found perspective taking-like effects during conditions in which the alternative perspective was unable to view the target.

Further Analysis

As with Experiment 1, additional analyses were undertaken to ensure that the results found overall were exhibited in both ambiguous number pairs.

'68'-'69' Ambiguous Number Pair

Reaction Time. Figure 2.8 depicts the mean RT for the four conditions using the '*68'-'69'* ambiguous number pair. A repeated measures ANOVA with ambiguity and visibility as within-participant factors found a significant main effect of ambiguity, $F(1, 37) = 28.27$, $p < .0001$, $n_p^2 = .43$, but no significant main effect of visibility, $F(1, 37) = .22$, $p = .65$, $n_p^2 = .006$. There was also no significant interaction, $F(1, 37) = 1.14$, $p = .29$, $n_p^2 = .03$. In other words, participants were significantly faster to respond during trials with the '*69*' ambiguous number, in comparison with '*68*'. In addition, the presence of an occluding barrier did not affect this data trend.

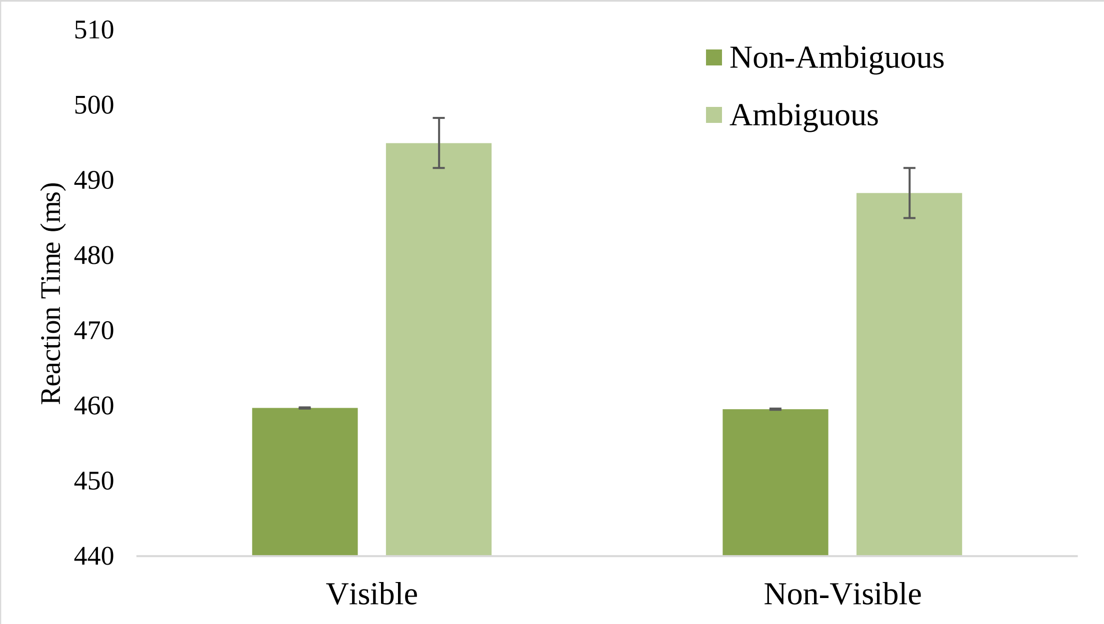

Figure 2.8: Mean RTs for each of the four conditions for the '68' and '69' ambiguous number pair, with standard error bars included

Error Rate. A repeated measures ANOVA with ambiguity and visibility as within-participant factors did not find a significant main effect of ambiguity, $F(1, 37) = .71$, $p = .4$, $n_p^2 = .02$, or of visibility, $F(1, 37)= 2.82$, $p = .1$, $n_p^2 = .07$. There was also no significant interaction, $F(1, 37)= .14$, $p = .71$, $n_p^2 = .004$. As can be seen there was no significant difference in error rate for trials using the '*68*' and '*69*' ambiguous number pair.

'88'-'89' Ambiguous Number Pair

Reaction Time. Figure 2.9 depicts the mean RT for the four conditions using the '*88*'-'*89*' ambiguous number pair. A repeated measures ANOVA with ambiguity and visibility as within-participant factors did not find a significant main effect of ambiguity, $F(1, 37) = .008$, $p = .93$, $n_p^2 = .0001$, or of visibility, $F(1, 37) = 1.26$, $p = .27$, $n_p^2 = .03$. There was also no significant interaction, $F(1, 37) = 3.05$, $p = .09$, $n_p^2 =$

.08. Thus, there was no significant finding in terms of RT when isolating the '*88*' and '*89*' ambiguous number pair.

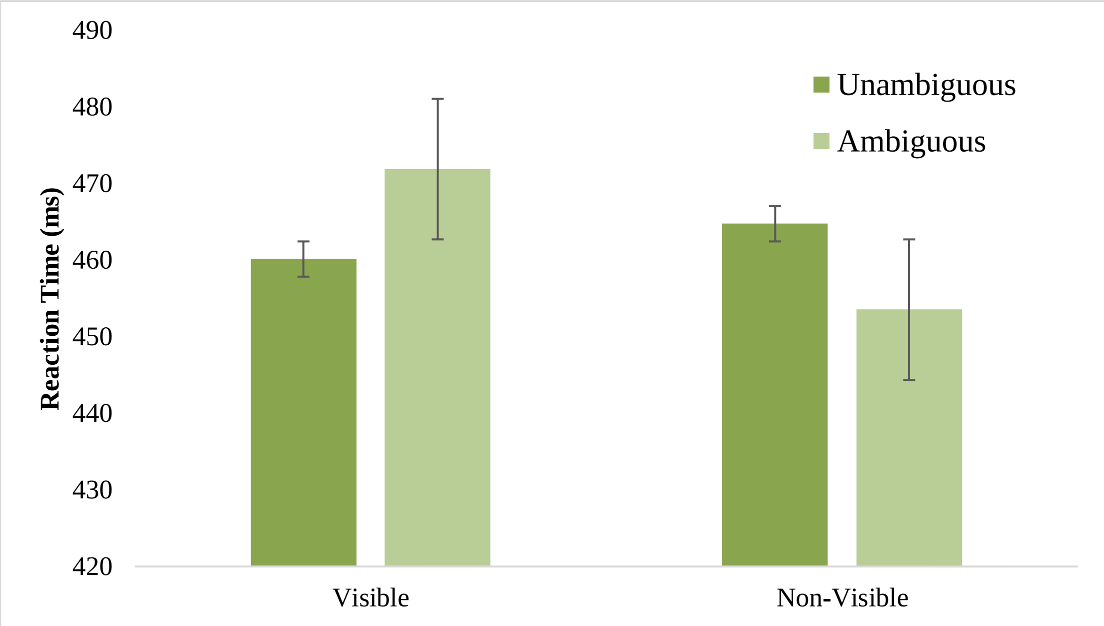

Figure 2.9: Mean RTs for each of the four conditions for the '88' and '89' ambiguous number pair, with standard error bars included

Error Rate. A repeated measures ANOVA with ambiguity and visibility as within-participant factors did not find a significant main effect of ambiguity, $F(1, 37) = 2.69$, $p = .11$, $n_p^2 = .07$, or of visibility, $F(1, 37) = .88$, $p = .35$, $n_p^2 = .02$. There was also no significant interaction, $F(1, 37) = 5.37$, $p = .03$, $n_p^2 = .13$. Thus, there were no significant finding reported in terms of error rate when isolating the '88' and '89' ambiguous number pair.

Overall, these data again reveal an ambiguous number effect; RTs were longer when the participant and agent can interpret the number differently. This on its own is consistent with past reports of visual perspective taking. Furthermore, the significant interaction identified that this effect is reduced when the model cannot see the number, which again is suggestive of an allocentric visual perspective computation. However, the visual perspective taking notion is clear in what it predicts when a alternative perspective cannot see the ambiguous number; there should be no effect at all. Yet, in Experiment 2 a visual perspective taking-like effect was found when the agent was unable to see the ambiguous number; i.e., when the newspaper obstructed the view. Therefore, this effect cannot be concluded to be the rapid and spontaneous computation of the agent's allocentric visual perspective. Other cognitive processes must be driving the results of this experiment. However, contrary to Experiment 1, a significant interaction between ambiguity and visibility was also found in Experiment 2. Yet, further analysis identified that the ambiguity effect between Unambiguous and Ambiguous conditions was found in both of the visibility conditions, and no significant difference between Ambiguous Visible and Ambiguous Non-Visible conditions were found. Thus, although the effect was reduced in the Non-Visible condition, again the visual perspective taking theory is clear that the effect should not have been identifiable in this condition. Consequently, it can be concluded that Experiment 2 again does not support the spontaneous visual perspective taking theory, as the same effect was identified during condition in which the allocentric visual perspective could not see the ambiguous number.

In Experiment 3, the importance of the occluding barrier will be revisited. Specifically, the occluding barrier manipulation and variations to the avatar's stance will be applied within the dot perspective task. This will assess whether alternative cognitive processes, such as gaze following, or mental rotation can be used to simulate visual perspective taking-like effects.

2.4 Experiment 3 – Avatar Stance, Occluding Barriers, and the Dot Perspective Task

2.4.1 Introduction

Following on from Experiments 1 and 2, Experiment 3 also assessed the spontaneous visual perspective taking notion using the physical barrier method, i.e., Cole et al. (2016). However, Experiment 3 applied this method to the dot perspective task, with two additional manipulations of consistency and the avatar's bodily stance. Recall from Chapter One, consistency has been applied to a number of experiments investigating spontaneous perspective taking using variants of the dot perspective paradigm, and refers to whether the avatar and participants can see the same number of discs (i.e., Consistent) or not (i.e., Inconsistent). Thus, in terms of the spontaneous visual perspective taking notion it is expected that RT will be shorter during Consistent conditions in comparison to Inconsistent conditions. Next, as with Experiment 2, Visibility refers to whether the view of the targets is obstructed (i.e., Non-Visible) or able to be freely viewed by the avatar (i.e., Visible) through the use of occluding barriers. However, contrary to Experiment 2, Experiment 3 used additional walls within the stimuli to form the occluding barriers in this variation of the dot perspective task. Lastly, avatar stance was also manipulated. A number of experiments have highlighted that head orientation of an avatar or agent directs attention (Langton & Bruce, 1999) more so than the body orientation of the agent or avatar (Cooney, Brady & Ryan, 2017). Most interestingly however, this directional cueing has been found to be strongest during conditions in which the head is not aligned with the body of the avatar (Hietanen, 2002; Ponianowska, Germeys, Verfaillie & Newell, 2012). To reiterate, conditions in which the avatar's body was facing forward with the head averted towards the side (i.e., "Rubbernecking") was found to be a stronger attentional cue than when the avatars body and head position faced the target (i.e., Facing). Gardner et al. (2018) support this avatar body manipulation within the investigation of spontaneous visual perspective taking, as they found that manipulating the avatar's head and torso only yielded cue-validity effects at longer SOAs, and not for instantaneous conditions. Consequently, Experiment 3 manipulated the avatars body in the dot perspective task also, to include both Facing and Rubbernecking conditions.

Subsequently, three predictions were made in regards to Experiment 3. One, if spontaneous visual perspective taking is found in regards to the presented avatar, a consistency effect should be identified, with RTs being significantly shorter during Consistent conditions in comparison to Inconsistent conditions. Second, if the spontaneous visual perspective taking notion is due to the computation of another individual's visual perspective, the consistency effect should be exhibited in both the facing and rubbernecking conditions. However, to comply with previous findings the Rubbernecking condition should produce a larger cuing effect in comparison to Facing conditions, as the averted head increases the cuing effect of the avatar's visuals perspective. Lastly, as with Experiment 2, any spontaneous visual perspective taking effect should be completely eliminated when the agent cannot see the target due to the occluding barriers. If, however, the same pattern of data occurs regardless of the occluding barriers; this will again suggest that there is another cognitive mechanism taking place, unlikely to be concerned with ToM and the assumption of an alternative visual perspective.

2.4.2 Method

Participants

38 participants were recruited with a mean age of 20.00 (*SD* = 1.86, range = 18-27), with 27 of the sample being female (11 male). Six participants identified themselves as left-handed whilst the remaining 32 identified themselves as right-handed. All 38 participants reported normal, or corrected to normal vision, and were recruited through the University of Essex online volunteer portal known as 'SONA', with participants being reimbursed for their time.

Stimuli and Apparatus

40 images were created using Daz 3D software with the extension pack named '*The Filing Room'* built within DAZ Studio desktop platform. Six examples of these images are shown in Figure 2.10. Eight out of the 40 images were filler conditions in which the stimulus was depicted void of any targets in the form of red discs.

Consistency, the first of the three factors, was manipulated by presenting differing numbers of red discs on the outer walls of the stimulus. For Consistent trials the discs were only presented in front of the avatar, and for Inconsistent trials the discs were presented in front and behind the avatar. In other words, either one or two discs were presented in front of the avatar for Consistent trials or were presented behind the avatar for Inconsistent trials. In terms of visibility, an additional two walls were presented in the stimulus, acting as occluding barriers obstructing the view of the target discs for the avatar. During Visible trials the extra walls were not present, allowing the avatar to freely view the discs, whereas during Non-Visible trials the extra walls barred the view of the discs for the avatar. Lastly, the avatars stance was also manipulated. All of the stimuli depicted a male avatar standing in the middle of a room. However, during Facing conditions the avatar was depicted with both his body and head oriented to the wall, perpendicular to the viewer. Whereas, during Rubbernecking conditions the avatar's body was depicted as facing the viewer, with the avatars head and gaze averted to the sidewalls, using a full neck extension. Consequently, 40 stimuli images were created with different variations of the manipulated factors of consistency, visibility, avatar stance, as well as orientating the avatar to either the left-hand or right-hand side of the screen and with either one or two target discs present within the stimuli.

The experiment was run on the SuperLab 5.0 desktop platform, using an Apple iMac with a screen size of 27-inch with the Apple iMac running on version 10.116 with 5120 x 2880 display dimensions. Participant responses and RTs were recorded using a standard keyboard, and the data were analysed using IBM SPSS Statistics (Version 21) computer software.

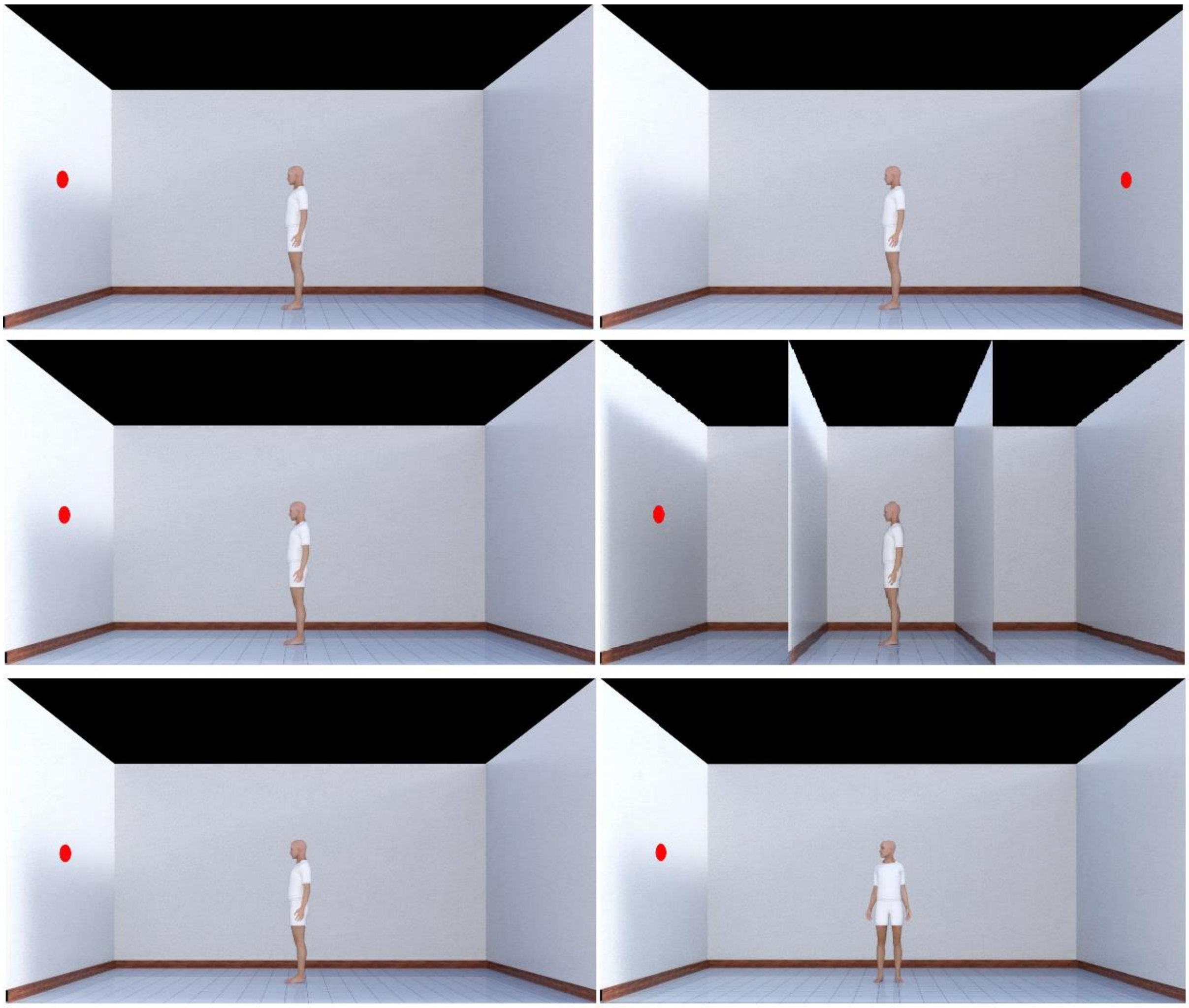

Figure 2.10: The top two images represent the two critical conditions in the basic dot perspective task. The left image is Consistent as the viewer and the avatar can see the same number of discs. Whereas, the right is Inconsistent as the viewer and the avatar can see different numbers of discs, i.e., the participants can see one disc behind the avatar, and the avatar cannot see any. The middle two images represent visibility conditions. The left image is Visible as the avatar can freely view the discs presented on the walls, whereas the right image is Non-visible as the avatar is obstructed from viewing the discs by a barrier in the form of additional walls. Lastly, the two bottom images represented the avatar stance conditions. The left image is Facing, as the avatars body and head orientation is facing the discs on the wall, whereas the right image is Rubbernecking as the avatar's body is facing the viewer with the head and neck is averted to the side.

Design and Procedure

A 2 x 2 x 2 design with consistency (Consistent, Inconsistent), visibility (Visible, Non-Visible) and avatar stance (Facing, Rubbernecking) as within participant factors was used. See *Stimuli and Apparatus* for the factor manipulation particulars.

Trials began with a blank screen for 500ms followed by a digit (*1* or *2*) for a further 750ms. After which the stimulus was presented, depicting the different variations of consistency, visibility and avatar stance with the addition of the target information in the form of red discs. The stimulus was presented for a maximum of 5,000ms, or until a response from the participant was detected. Participants were required to respond with the *'Z'* key on a standard keyboard for '*YES*' responses and the *'M'* key for '*NO*' responses to whether the digit (*1* or *2*) presented at the start of the trial correlated to the number of red discs in the stimulus. Once the predefined time had elapsed or a participant response was detected a blank screen was presented for a further 50ms before the onset of the next trial. Participants were informed to respond as quickly as possible, without losing accuracy. Additionally, in contrast to Samson et al. (2010) there was no mention of an alternative perspective or the forced assumption of differing perspectives.

All experimental conditions were blocked in one and were presented only after full instructions, and a practice block of 10 trials, were administered. Additionally, a break was included after the practice block in order for participants to ask the experimenter any questions, before the onset of the experimental block. Finally, after the experimental block, participants were presented with exit instructions and a full debrief.

A total of 288 trials were administered, including 32 filler trials in which no red discs were presented. Trials were randomly ordered and comprised of 32 trials for each combination of the three factors of consistency, visibility and avatar stance, as well as an even number of '*YES*' and '*NO*' responses. In addition, the orientation of the avatar was counterbalanced, with half of the trials oriented to the left-hand side of the screen and the other half to the right.

2.4.3 Results and Discussion

Data Preparation. The data were collated and reorganised using Microsoft Excel and collapsed into a single data file. The data were then transformed to produce mean RTs, excluding any outliners that were 2 standard deviations above or below the mean, for the different levels of each manipulated factor of consistency, visibility and avatar stance. Once scores for each level of each factor were obtained, the data was filtered to ensure that only participants with scores for each variant were used in the overall analysis. Additionally, a percentage score for correct egocentric perspective responses was also calculated for error rate analysis. Lastly, Filler trials were also examined to determine whether the same RT and error data trends were present without the target information.

Reaction Time. Figure 2.11 depicts mean RTs for each condition. A repeated measures ANOVA with visibility, consistency and avatar stance as within-participant factors did not find a significant main effect of consistency, $F(1, 37) = .005$, $p = .94$, $n_p^2 = .001$, avatar stance, $F(1, 37) = .01$, $p = .92$, $n_p^2 = .001$, or visibility, $F(1, 37) = .13$, $p = .72$, $n_p^2 = .004$. In terms of the two way interactions, consistency and visibility was not significant, $F(1, 37) =.12$, $p = .73$, $n_p^2 = .003$, neither was consistency and avatar stance, $F(1, 37) =3.21$, $p = .08$, $n_p^2 = .08$, or visibility and avatar stance, $F(1, 37) =2.19$, $p = .15$, $n_p^2 = .06$. Lastly, there was no significant three way interaction, $F(1, 37) = .08$ $p = .78$, $n_p^2 = .002$. Overall, there was no significant finding for RT between consistency, avatar stance, or visibility manipulations.

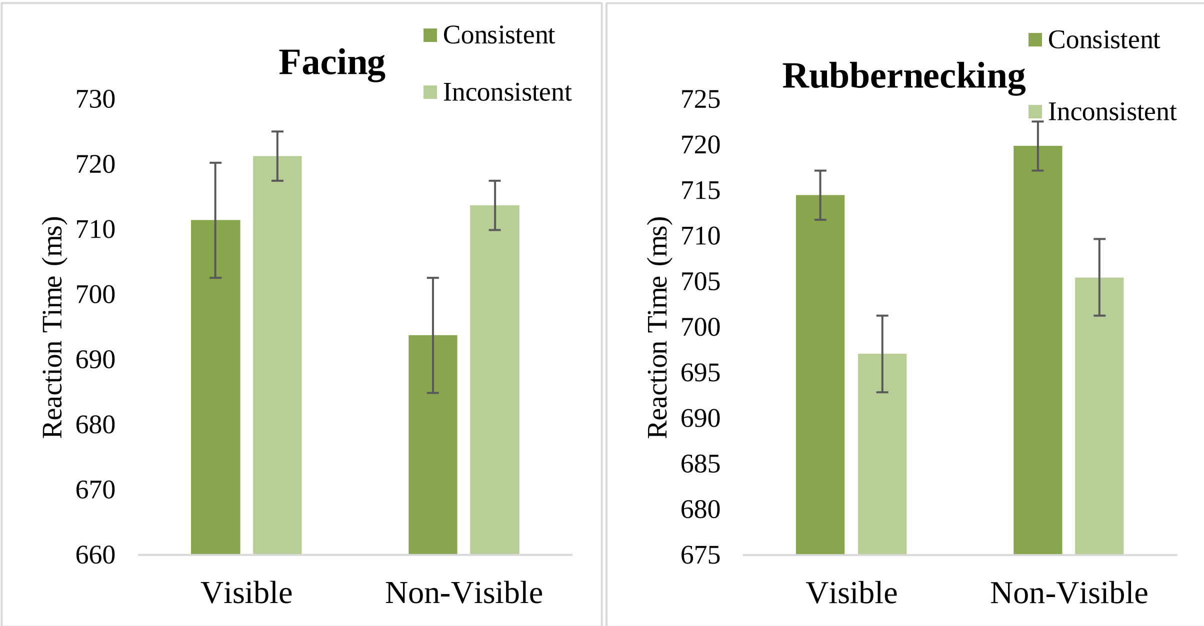

Figure 2.11: Mean RTs for Facing on the left and Rubbernecking on the right, with standard error bars included

Error Rate. A repeated measures ANOVA with visibility, consistency and avatar stance as within-participant factors did not find a significant main effect of consistency, $F(1, 37) = .03$, $p = .86$, $n_p^2 = .001$, avatar stance, $F(1, 37) = .02$, $p = .88$, $n_p^2 = .001$, or visibility, $F(1, 37) = .009$, $p = .93$, $n_p^2 = .001$. In terms of the two way interactions, consistency and visibility was not significant, $F(1, 37) =.13$, $p = .73$, $n_p^2 = .003$, neither was consistency and avatar stance, $F(1, 37) =1.09$, $p = .3$, $n_p^2 = .03$, or visibility and avatar stance, $F(1, 37) =.07$, $p = .79$, $n_p^2 = .002$. Lastly, there was no significant three way interaction, $F(1, 37) = 1.64$, $p = .21$, $n_p^2 = .04$. Consequently, there was no significant finding for error rate between consistency, avatar stance, or visibility manipulations.

Filler Trials. Figure 2.12 depicts the mean RTs for filler trials. A repeated measures ANOVA on the RT of filler trials with visibility and avatar stance as within-participant factors did not find a significant main effect of avatar stance, $F(1, 37)$ =

.04, p = .84, $n_p{}^2$ = .001, or visibility, F(1, 37) = 2.81, p = .1, $n_p{}^2$ = .07. There was also no significant interaction, F(1, 37) = .001, p = .99, $n_p{}^2$ = .001. Additionally, in relation to error rates a repeated measures ANOVA on filler trials again with visibility and avatar stance as within-participant factors did not find a significant main effect of avatar stance, F(1, 37) = .04, p = .84, $n_p{}^2$ = .001, or visibility, F(1, 37) = 2.81, p = .1, $n_p{}^2$ = .07, or an interaction, F(1, 37) = .001, p = .99, $n_p{}^2$ = .0001. Taken together, the results of the filler trials identified that there was no significant difference in terms of the error rate or RT between manipulations of avatar stance and visibility when no target information was presented.

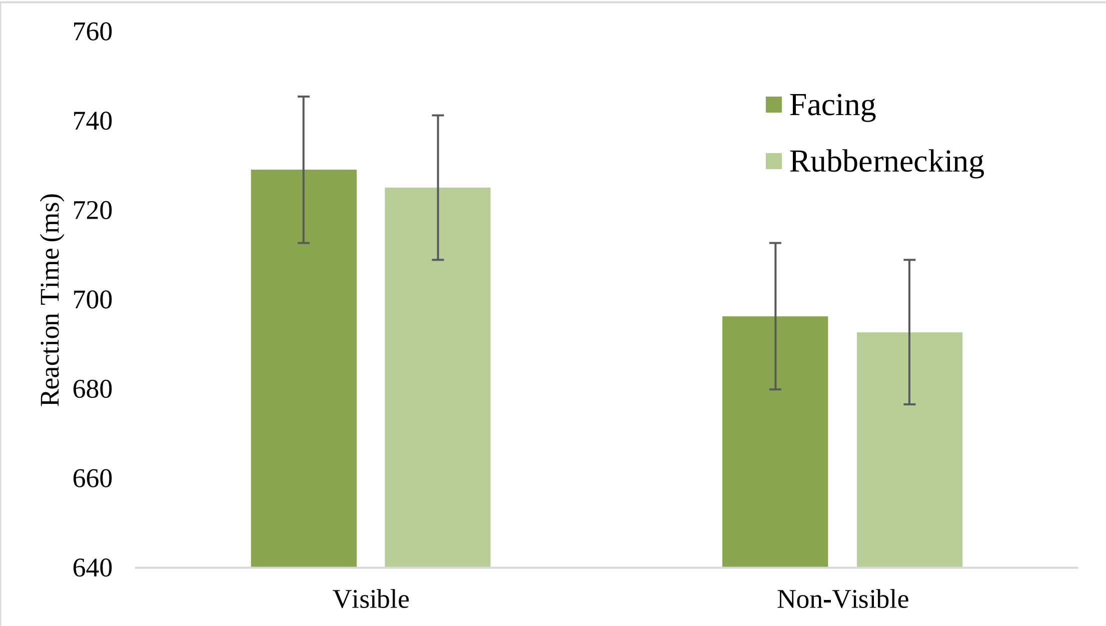

Figure 2.12: Mean RTs for filler trials, with standard error bars included

Overall, the results Experiment 3 did not find significance in terms of RT or error rate regarding the factors of consistency, visibility and avatar stance. However, these findings may not be due to the overall phenomenon in question but instead may be associated with design flaws embedded within this experiment.

One possible design limitation that could be impacting the results of this experiment is the blocking procedure. Previous literature such as Samson et al. (2010), Surtees et al. (2016) and Gardner et al. (2018) used separate blocks for different factors, whereas Experiment 3 collapsed all factors into one block. This was due to the experiment's specific assessment of the *spontaneous* element within the phenomenon. To reiterate, if this phenomenon is spontaneous, the addition of influencing factors within one experimental block should not hinder the observation of the phenomenon's effect. Nevertheless, when using a single block, participants are required to focus their attention for an extended period, which could have led to fatigue. Additionally, in previous work participants are indirectly allowed a break during the transition between blocks. In future work, this issue could be addressed by collapsing all factors into a single block, but also embedding frequent breaks within the block to account for fatigue. In addition, during the breaks, participants could be asked to perform an irrelevant pseudo task in which the participant's attention is diverted elsewhere. Subsequently, Experiment 3 neither supports nor refutes the spontaneous perspective taking claim.

Future work should continue to address the key questions driving Experiment 3. Specifically, whether spontaneous visual perspective taking, is in fact routed within perspective taking and ToM abilities or is a by-product of alternative cognitive mechanisms such as gaze cuing and mental rotation. In addition, the definition of *spontaneous* should also be extrapolated, in terms of what it explicitly means. *Is this phenomenon driven absence of objective goals? Is it outside of conscious control?* Thus, limitations for how this phenomenon is spontaneous should also be addressed.

2.5 General Discussion

Overall, Experiment 1 challenges the spontaneous visual perspective taking notion as participants were slower to react when the target was ambiguous, in that it could be interpreted differently depending on the perspective adopted. However, this effect was identified irrespective of the averted attention of the agent, which should have inhibited the effect from taking place. Experiment 2 also challenges the spontaneous visual perspective taking notion. It was again found that participants were slower to react during ambiguous conditions. However, Experiment 2 found this effect regardless of the occluding visual barrier, which obstructed the agent's view of the target. Consequently, Experiments 1 and 2 both identified the spontaneous visual perspective taking-like effect during conditions in which the agent could not see the ambiguous number. Thus, Experiments 1 and 2 cannot conclude in terms of the spontaneous assumption of another individual's visual perspective. As how can the spontaneous assumption of another's visual perspective be driving this effect when the alternative visual perspective is not directly interacting with the target information? Lastly, the results of Experiment 3 did not find significance in terms of consistency, visibility, and avatar stance. However, previous research supports the effects of consistency (Samson et al. 2010), visibility (Cole et al. 2016), and avatar stance (Gardner et al. 2018) within spontaneous visual perspective taking. Consequently, one could argue that design flaws may have influenced the results of Experiment 3 in which significance was not found, as was discussed.

Recall that the central claim of spontaneous visual perspective taking is that individuals *rapidly* and *spontaneously* assume the visual perspective of another (Samson et al. 2010). Thus, the phenomenon should only be identified during conditions in which the allocentric perspective can see or is directly interacting with the target. Thus, when this effect, or simulations of this effect are found, under conditions in which the phenomenon should not be present, one begins to question whether the effect has been found, or alternative theories are responsible. One possible suggestion for an alternative theory responsible that has been discussed in *Section 1.5* was gaze cuing. As reviewed in Chapter 1, gaze cuing is the concept that observing another individual's attention can and does influence the attention of the observer (Nuku & Bekkering, 2008; Teufel, Alexis, Clayton, & Davis, 2010; Teufel et al., 2009; Teufel, Fletcher, & Davis, 2010). Thus, the averted gaze in Experiment 1 could be questioned in terms of gaze following, but this concept cannot account for the discrepancies in terms of the spontaneous visual perspective taking-like effects identified in Experiment 2. To examine the spontaneous perspective taking and gaze cuing crossover further, future work could extend the manipulated gaze conditions with barriers further, including misdirection, and orientation manipulations.

Overall, this chapter assessed the spontaneous visual perspective taking theory in respect to two methodological paradigms, as well as identified issues within the methodologies and phenomenon. Experiment 1 and 2 found that spontaneous visual perspective taking-like effects during conditions in which it should have been abolished and Experiment 3 did not find any significance in terms of consistency, avatar stance, or visibility, using the dot perspective task. Yet, issues regarding the methodologies and central assumptions remain. As behavioural measures can be limited by noise within the sample, the following chapter will attempt to increase the sensitivity of spontaneous visual perspective taking measures by using eye movements to measure RT.

Chapter Three:

Eye Tracking

3.1 Chapter Overview

A central principle of experimental psychology is that different dependent measures have different sensitivities with respect to indexing the phenomenon in question. The fundamental aim of the present chapter was to increase sensitivity to any potential spontaneous visual perspective taking effect by using eye tracking.

The analysis of eye movements, for example, first fixation upon a target, is one way to increase the sensitivity of RT. The eye-mind assumption states that delays in processing are not apparent during the instance of a first fixation upon a target (Just & Carpenter, 1980), i.e., as soon as a fixation is made the mind is processing what is being seen. Thus, eye movement measures are able to highlight processes that otherwise are not found. Whereas, with manual RT, the time taken to process the stimuli and respond is measured, not just the first instance of processing. Consequently, when looking at the behavioural manipulations of spontaneous visual perspective taking literature, the differences in RTs may be due to a multitude of factors associated with cognitive processing of visual stimuli as well as motor response.

Subsequently, two experiments examining eye movements were devised with three common factors of, stimulus-onset asynchrony (SOA), validity, and self-reported measures of empathy. Firstly, as previous literature has claimed the visual perspective taking effect is *spontaneous*, variations in SOA will enable the assessment of whether the spontaneous visual perspective taking effect can be observed as a rapid process as well as after a delay in stimulus onset. This has been supported by a number of authors (e.g., Teufel et al. 2010; Samson et al. 2010; Santiesteban et al. 2013; Nieslen et al. 2015) who argue that spontaneous visual perspective taking occurs *spontaneously* alongside the processing of other components in the stimuli, including the target. Whereas, in the gaze cueing paradigm for example, the effect of attentional cueing is found when the gaze cue is presented between 50 and 800ms prior to the target (Frischen, Bayliss, & Tipper, 2007) and not when the gaze cue is presented simultaneously with the target (Xu, Tanaka, & Mineault, 2012). However, as stated in Chapter 1, Bukowski et al. (2015) investigated spontaneous visual perspective taking and SOA using the dot perspective task and found the effect to be apparent during longer SOAs. Nevertheless, their research used behavioural measures that may have been unable to index the phenomenon at shorter SOAs. Thus, it is predicted that if this effect is indeed *spontaneous*, it should be observed during the shortest SOA conditions. A corollary to this is that if the effect is *non-spontaneous*, the effect should not be observed at the earliest SOA but maybe at the later SOA when top-down processes begin to occur.

Additionally, and crucially, spontaneous visual perspective taking has been hypothesised to be an extension of ToM processes (i.e., Teufel et al. 2010; Samson et al. 2010; Santiesteban et al. 2013; Nieslen et al. 2015). Thus, an enhanced awareness of others should in turn positively correlate towards a greater susceptibility to assume the visual perspective of another. As a consequence, self-reported measures of empathy will be correlated against susceptibility to assume the allocentric perspective, using the Davis (1980) Interpersonal Reactivity Index. This measure can be deconstructed into the components of, *Perspective Taking, Fantasy, Empathic Concern* and *Personal Distress*. Interestingly, Nieslen et al. (2015) correlated the subscales of *Perspective Taking,* and, *Empathic Concern* only, finding a positive correlation to susceptibility of assuming the allocentric visual perspective. In addition, Mattan, Rotshtein, and Quinn (2016) examined to what extent visual perspective taking is modulated by trait-level empathy, and found that, higher trait-level empathy improved visual perspective taking performance. Consequently, it is predicted that a higher susceptibility to assume the allocentric perspective will positively correlate to all of the self-reported Davis (1980) Interpersonal Reactivity Index subscales.

Finally, as with the current spontaneous visual perspective taking literature, the comparison of validity differences will be examined alongside eye movement measures to assess whether the effect is still present. It is predicted that the same effect of spontaneous visual perspective taking will be found that previous work documents, with Valid conditions producing a significantly shorter RT, compared to Invalid conditions.

Consequently, Experiment 4 will examine spontaneous visual perspective taking using eye movement, SOA, validity, human and non-human directional cues, covert and overt tasks and self-reported measures of empathy. Whereas Experiment 5 will assess spontaneous visual perspective taking using eye movement, SOA, validity, occluding barriers and self-reported measures of empathy.

3.2 Experiment 4 –Task Differences, Non-Human Orientation Cues, Perspective Validity, and Empathy

3.2.1 Introduction

In addition to eye movements, SOA, validity, and self-reported measures of empathy, Experiment 4 also manipulated the task and orientation stimulus in the investigation of spontaneous visual perspective taking. In order to assess the efficiency of eye movements as an assessment of the phenomenon in question, an Overt and Covert task was administered. Both tasks used the comparison of RT during different validity manipulations as an indication of the spontaneous visual perspective taking notion, as previous literature documents (i.e., Teufel et al. 2010; Samson et al. 2010; Santiesteban et al. 2013; Nieslen et al. 2015). However, for the Overt task RT recorded the time taken for first fixation upon the target, absent of manual response. Whereas, for the Covert task, participants were required to maintain their focus on the central point of the screen throughout the duration of the condition, and manually respond, when the target was detected. Thus, RT during the Covert task recorded the time taken for the target to be detected, as well as time taken for a response to be given. Consequently, manual RT (Covert task) and fixation RT (Overt task) for target detection can be compared, which in turn can be used to assess the claim that eye movement measures have a greater sensitivity at indexing this phenomenon. Hence validating the investigation of eye movement in the spontaneous visual perspective taking notion.

Additionally, Experiment 4 manipulated the orientation stimulus. Previous authors such as Santiesteban et al. (2013) and Nieslen et al. (2015) both explored the effect that altering the orientation stimulus can have upon the notion of spontaneous perspective taking. As previously mentioned in *Chapter One*, Santiesteban et al.

(2013) examined the difference that arrows and an avatar have upon the phenomenon, finding that the arrow was just as effective at producing the spontaneous visual perspective taking notion. Moreover, Nieslen et al. (2015) manipulated the level of sociality by assessing the differences that an arrow, rectangle, and avatar have upon spontaneous visual perspective taking. Nieslen et al. (2015) replicated previous findings of spontaneous visual perspective taking, in all conditions, regardless of the inability to assume the perspective of non-social orientation stimuli. Taken together, the notion has been identified during conditions absent of a human agent, where it should not be apparent. Thus, Experiment 4 examined this discrepancy by comparing the use of an Arrow and human Face as an orientation cue. However, to reiterate, as previous work has found the effect to be apparent during non-human orientation conditions, the use of eye movement analysis in Experiment 4 will increase the sensitivity of the experiment and therefore increase the chances of highlighting any differences between conditions. Hence, if the spontaneous visual perspective taking phenomenon is indeed the *rapid* and *spontaneous* assumption of another's visual perspective, routed in ToM, the effect should only be observed during conditions in which an alternative human perspective is present. In other words, the Face orientation cue should provide conditions in which the notion can be observed, as there is an alternative human visual perspective, but the Arrow condition should not. On the other hand, if Experiment 4 supports Santiesteban et al. (2013) and Nieslen et al. (2015) and identifies the effect during the Arrow condition this would suggest that alternative cognitive processes are driving the spontaneous visual perspective taking notion.

Therefore, if eye movement measures do indeed have an increased sensitivity at indexing this phenomenon, RT should be significantly shorter within the Overt task, in comparison to the Covert task. Additionally, if spontaneous visual perspective taking is the rapid and spontaneous assumption of another individual's visual perspective, routed in ToM, the effect should be dependent on the human agent and therefore only be identified during conditions with a human presence, and abolished during conditions with a non-human orientation cue. Furthermore, a positive correlation between self-reported empathy and susceptibility of the effect should also be found. In addition, if this effect is indeed *non-spontaneous* it should only be observed during the longest SOA manipulations. Finally, the same effect of validity that previous literature documents should be found, with Valid conditions producing significantly shorter RT, compared to Invalid conditions.

3.2.2 Method

Participants

Thirty-two participants were recruited with a mean age of 20.88 (SD = 2.42, range = 18-27). Twenty-three participants were female (nine male), with three identifying as left-handed (twenty-nine as right-handed). Furthermore, in terms of ethnicity, 22 identified themselves as white, four as Asian, three as Black and three who identified as a non-listed ethnic group. All participants reported normal or corrected to normal vision and were recruited through the University of Essex online volunteer portal known as 'SONA', and were reimbursed for their time.

Stimuli and Apparatus

Sixteen variations of stimuli were generated, using different configurations of the manipulated factors. All trials began with a fixation point in the middle of a black screen, which participants were required to fixate upon before the onset of the stimuli. Next, either a line or a male face was presented, both of which can be seen in Figure 3.1. After either an 80 or 300ms delay, the eyes on the face moved to direct to the side of the display, or arrowheads appeared, again directing to one side of the display. A target simultaneously appeared in the form of a letter 'C' on one side of the screen. A flow chart of this process can be viewed in Figure 3.2.

Photoshop and Experiment Builder (SR-Research) were used to create the visual stimuli. Eye movements were recorded with a head-mounted, video-based eye tracker (EyeLink 1000; SR Research Ltd., Osgoode, Ontario, Canada), using a sampling rate of 1000 Hz. They were recorded monocularly and analysed using Eyelink Data Viewer (SR-Research). All stimuli were displayed on a 19-inch monitor with a 1024 x 768-pixel screen resolution, with the stimuli being presented at 25 frames per second, using Experiment Builder presentation software (SR-Research), with a viewing distance of approximately 57 cm.

Figure 3.1: The neutral stimuli prior to the onset of the directional cues used in Experiment 4, with the face on the left and line that arrowheads were added to on the right

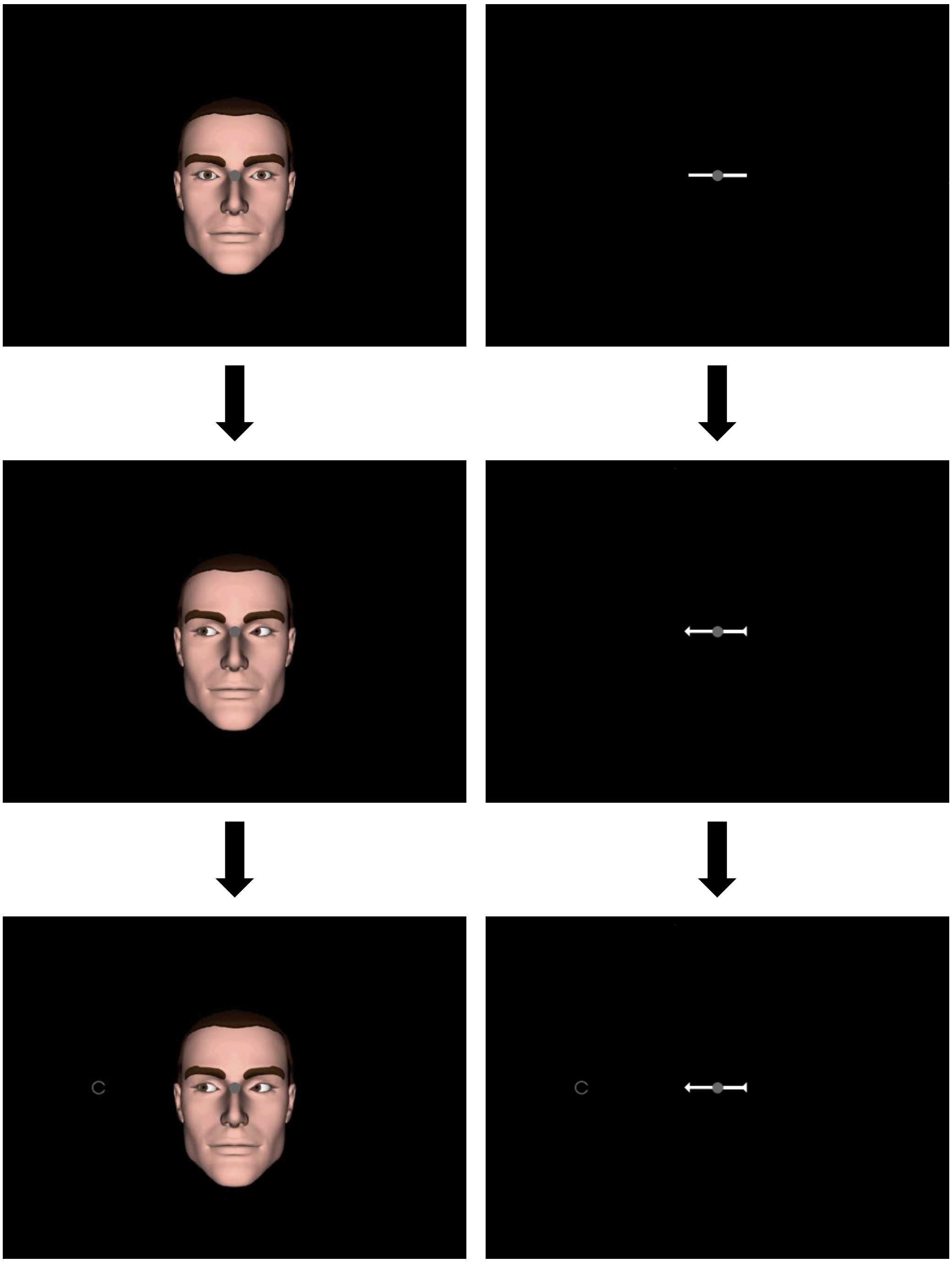

Figure 3.2: The format of a Valid trial for the Face (left) and Arrow condition (right)

Design and Procedure

A 4 x 2 repeated measures design was used, with task (Covert, Overt), figure (Face, Arrows), validity (Valid, Invalid) and SOA (80ms, 300ms) being manipulated. Consequently, 16 variations of experimental trial with the four manipulated factor variations were used.

Participants were positioned in a chinrest approximately 60cm away from the computer monitor. Prior to the onset of the practice trials, participant eye movements were calibrated and validated using a nine-point calibration procedure. In addition, each trial began with a drift correction and started only once the participants gaze was fixated on the central point of the screen, identified by a black disk.

During the experiment, participants took part in four blocks, consisting of two tasks, a Covert and Overt task. Additionally, within each task a condition using a distractor of a human Face and Arrow were employed.

Covert Task. Participants were required to focus solely on the middle of the screen and press the space bar on a standard keyboard, as soon as they saw the target 'C' appears out of the corner of their eyes. Participants were informed prior to the start of each block which side the target would appear on and were explicitly informed that it would only appear on that side, until informed otherwise. During the trial, in the centre of the screen a human face or line was presented. The eyes on the face moved, and arrowheads appeared on the line pointing to one side of the screen, both of which acted as a directional cue. Participants were informed of this and told to ignore the face and arrows. Directly in the centre of the screen a black disk was presented to help participants focus their attention. Each trial ended when the space bar was pressed. If no response was given, the trial ended automatically after a delay of 1,000ms. This enabled the presentation of filler trials, absent of a target; ensuring attention was maintained throughout the duration of the experiment.

Overt Task. Similar to the Covert Task, participants were required to press the space bar as quickly as possible after the presentation of the target. However, participants were also required to move their gaze from the centre of the screen towards the target before pressing the space bar and ending the trial.

A total of 296 counterbalanced trials were used, split into four blocks of 68 (four of which were filler trials), with six practice trials preceding each block. The trials were blocked depending on the position of the target, towards the left-hand side of the screen for two blocks and the right for two, as well as two blocks (i.e., one with the letter ‘C’ on the left-hand side of the screen and one on the right) using the Face, and two with the Arrows. Participant responses could therefore be cross-correlated depending upon the presence of the Arrows/ Face, as well as accounting for left or right-hand biases. After each block was completed, a break was given. Prior to the start of each block, participant eye movements were re-calibrated and validated using the nine-point calibration procedure. Participants were lastly asked to complete the Davis (1980) self-report Interpersonal Reactivity Index, before being debriefed.

Participant eye movements were analysed using data viewer 2.4.0.198 (SR Research) and cross-analysed with the Davis (1980) Interpersonal Reactivity Index results using IBM SPSS Statistical (Version 21) software.

3.2.3 Results and Discussion

Data Preparation. The data were collated and analysed using DataViewer (SR-Research), Microsoft Excel and IBM SPSS Statistics (Version 21) computer software. Firstly, using DataViewer (SR-Research) the central fixation point, as well as each target area was defined as a Region of Interest. Using the Reaction Time Manager and Reaction Time Definitions, a RT measurement between target onset and first fixation upon the target was created for the Overt task. Additionally, for the Covert task Regions of Interest were used to identify any trials in which fixations left the central point, and thus were excluded from further analysis. As each trial ended once the target was detected, dwell time was not processed, as it was not meaningful. Both the fixation RT for Overt conditions and manual RT for covert conditions were extracted from DataViewer (SR-Research). These RT were then collated and used to create mean RT for each manipulation of the four factors of task, figure, validity, and SOA, using Microsoft Excel pivot tables. Any outliners that were 2 standard deviations above or below the mean were excluded from the overall analysis. Four participants were excluded from analysis as either they were not successful in fixating upon the correct target for over 50% of the total trials, or the equipment did not effectively track saccades. IBM SPSS Statistical (Version 21) software was then used for cross analysis of conditions and correlation to the Davis (1980) self-report Interpersonal Reactivity Index.

Reaction Time. Figure 3.3 shows the mean RTs for manual RT for the Covert task and first fixation for the Overt task, for the eight different conditions. An ANOVA with task, figure, validity and SOA as within-participant factors found a significant main effect of task, $F(1, 35) = 23.45$, $p < .001$, $n_p^2 = .40$, a significant main effect of validity, $F(1, 35) = 27.54$, $p < .001$, $n_p^2 = .44$, and a significant main effect of SOA, $F(1, 35) = 149.84$, $p < .001$, $n_p^2 = .81$, but no significant main effect of figure, $F(1, 35) = 2.64$, $p = .11$, $n_p^2 = .07$. In terms of the interactions, only task by validity, $F(1, 35) = 9.78$, $p < .004$, $n_p^2 = .22$ and task by SOA, $F(1, 35) = 4.71$, $p < .04$, $n_p^2 = .12$, were significant. All other interactions were non-significant (all Fs < 1.4 and all ps > .23). Perhaps most importantly, with respect to the central rationale there was no validity x figure interaction. In other words, the Face did not induce greater cueing than Arrows. This is true for both Covert and Overt orienting.

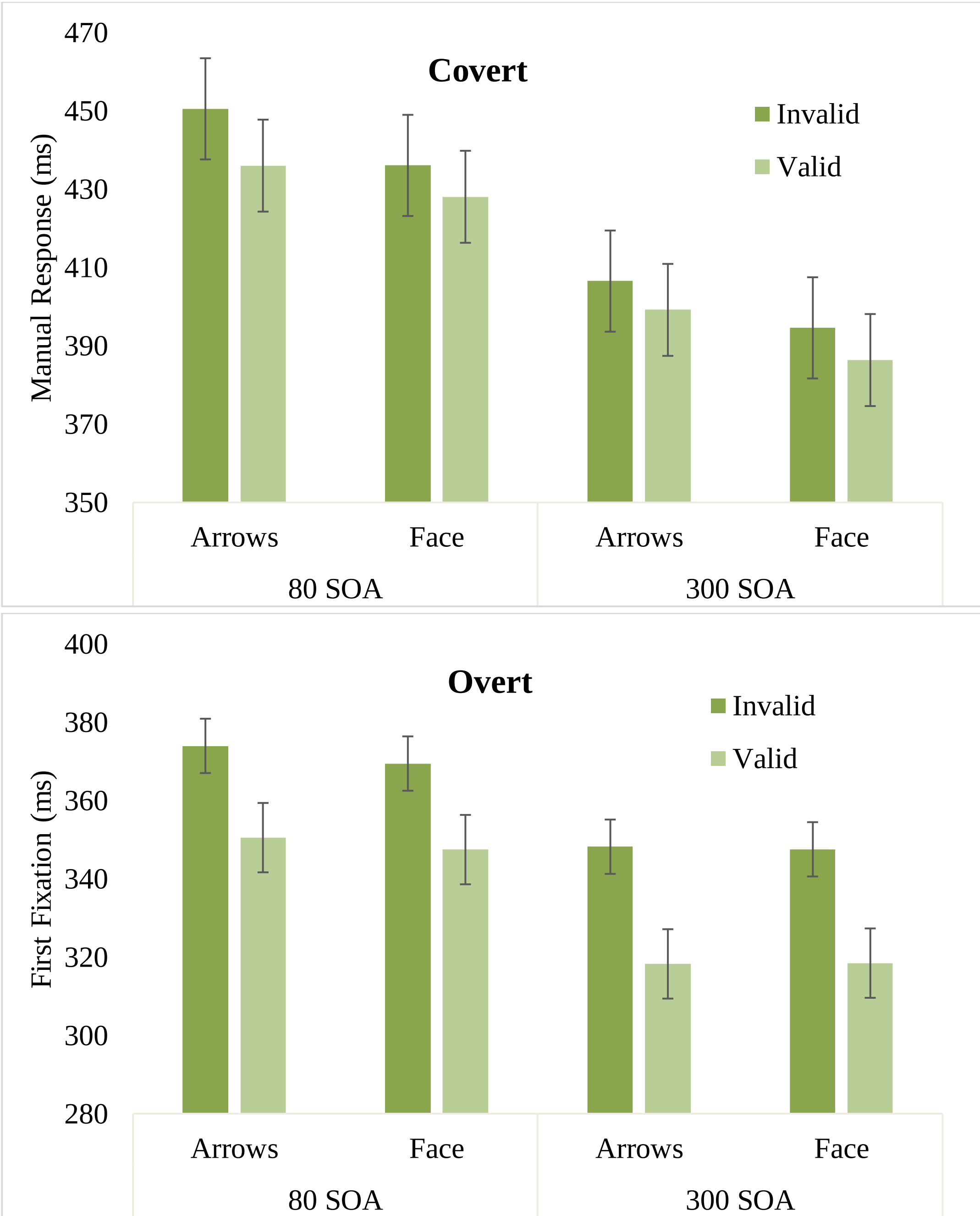

Figure 3.3: Mean RTs for Covert manual response time (top) and first fixation for Overt task (bottom) for each of the eight conditions, with standard error bars included

To reiterate, the significant main effect of task shows that participants were faster to fixate upon the target in the Overt task compared to manual RT in Covert task. Additionally, the significant main effect of validity demonstrates that participants were significantly faster to respond with a key press or fixation during

Valid conditions, where the distractor direction and target matched, in comparison to Invalid conditions, where the distractor direction and target mismatched. Furthermore, the significant main effect of SOA highlights that, participants were significantly faster to respond with a key press or fixation during 300ms conditions in comparison to 80ms. The significant interactions of task by validity and task by SOA are also to be expected due to the medium effect size of the main effects of task and validity and large effect of SOA.

Davis (1980) Interpersonal Reactivity Index. Additionally, Pearson's correlation coefficients were used to assess the relationship between participant's susceptibility to gaze cue, which recall, is supposedly due to visual perspective taking (i.e., ToM), using the saccade data and self-reported measures of empathy. To calculate a participant's susceptibility, the difference between validity conditions were used to create an index. Additionally, the Davis (1980) Interpersonal Reactivity Index was deconstructed, creating scores for each of the four components of, *Perspective Taking*, *Fantasy*, *Empathic Concern* and *Personal Distress*. No significant correlation was found between the computed validity index and *Perspective Taking* (r = -.19, p = .27), *Fantasy* (r = -.003, p = .99), *Empathic Concern* (r = -.19, p = .28), or *Personal Distress* (r = .18, p = .31). Consequently, in terms of the self-reported measures of empathy, and the participant's visual perspective taking susceptibility, no relationship was found.

Overall, Experiment 4 found that RT was shortest for fixating upon the target in the Overt task compared to manual RT in the Covert task. Thus, validating the assessment of eye movements at indexing the spontaneous visual perspective taking phenomenon. In addition, RT was significantly shorter during Valid conditions, where the distractor direction and target matched, in comparison to Invalid conditions, where the distractor direction and target mismatched. Furthermore, RT was significantly shorter during 300ms SOA conditions in comparison to 80ms. On the contrary, there was no reported correlation between self-reported measures of empathy and susceptibility to spontaneously assume an allocentric visual perspective, and there was no difference in terms of a human and non-human directional cue. In consequence, the effect identified in Experiment 4 is unlikely to be due to the spontaneous assumption of an alternative visual perspective, as the same effect was observed absent of an alternative human presence.

A non-spontaneous visual perspective taking explanation for the effects observed in Experiment 4 is that the cue orientates attention without any ToM mechanisms being involved. That is, the eyes and arrowheads redirect the viewer towards or away from the target, thus influencing RT. Alternatively; gaze cuing can be argued to be influencing the results within this experiment also. It has been commonly concluded that individuals are drawn to the direction of another individual's attention, which is modulated through gaze (Driver et al. 1999; Friesen & Kingstone, 1998; Hietanen, 1999, 2002; Langton & Bruce, 1999). Thus, the eyes draw the attention of the viewer to either side of the screen, impacting RT. However, this explanation does not explain why the arrowheads produced the same effects.

It is worth noting that Experiment 4 can be criticised in terms of the stimulus arrangement. It can be argued that it is difficult to conclusive state what the avatar can and cannot see. As the avatar only moves its eyes to each side of the display it can therefore be criticised that the avatar cannot see the target. Instead the avatar may only be able to see an image of the participant in front of them. Consequently, the importance of peripheral vision within this stimulus set up is questioned. However, this is a criticism that could be used within all behavioural experiments that use computerised representations of stimuli. For example, any research that uses the traditional gaze cuing procedure (Frischen & Kingstone, 1998; Langton & Bruce, 1999) often superimposes the target on top of the basic stimulus. Thus, the validity of the overall stimulus can be questioned. However, this is a question that could be adapted and applied to all research within the field of visual perspective taking and reflexive gaze following. Additionally, this criticism could often lead to questioning the effectiveness of stimuli with computerised avatars also. To reiterate, can a computerised avatar have a mental state in order to associate ToM processes to? If the avatar does not have a mental state how can a participant 'assume' the avatar's visual perspective? Again, this is a common issue within all visual perspective taking literature, including spontaneous visual perspective taking and the current work. Further discussion on what constitutes a '*perspective*' will be discussed in *Chapter 7 General Discussion*.

In the following experiment, eye movement measures will again be used in the investigation of spontaneous visual perspective taking and correlated alongside self-reported measures of empathy. However, in addition and more specifically, the validity effects observed in Experiment 4 will be examined in relation to occluding visual barriers. This is a result of the idea that if spontaneous visual perspective taking is the *spontaneous* and *rapid* assumption of an allocentric visual perspective, the presence of an occluding visual barrier should abolish this phenomenon. However, if alternative cognitive processes such as gaze cuing or reorientation points are driving these effects, then the addition of the barrier will not hinder the effects of the phenomenon, and the same pattern in RTs should be observed regardless.

<u>3.3 Experiment 5 – Occluding Barriers, Perspective Validity, and Empathy</u>

3.3.1 Introduction

The barrier method (i.e., Cole et al. 2016) has produced a number of interesting observations in the assessment of spontaneous visual perspective taking (i.e., Experiment 2 and 3). Consequently, Experiment 5 applied occluding barriers together with eye movement measures. Specifically, an agent was located between two potential targets, with an occluding barrier positioned in between the agent and one of the targets. Previous applications of the barrier method used what might be called 'passive' barriers outside the control of the agent, whereas the agent in Experiment 5 is depicted as actively holding the barrier. In other words, she is instrumental in preventing herself from seeing the target. Additionally, participants were explicitly informed that the barrier obstructs the view of a target. If ToM processes are underpinning the visual perspective taking effect, enhancing the interaction between the agent and barrier could in turn increase the likelihood of the effect being observed. This notion is supported by Zhao et al. (2015) findings that increasing the interaction between the agent and ambiguous target progressively increased the chances of a spontaneous visual perspective taking response occurring. In addition, as with Experiment 4, SOA self-reported measures of empathy, and validity will also be examined.

As was predicted for Experiment 4, if spontaneous visual perspective taking is indeed *rapid* and *spontaneous* and routed in ToM, a positive correlation between self-reported empathy and susceptibility of the effect will be found. A corollary to this is that if the effect is *non-spontaneous* it should be only observed during the longer SOA manipulation. Furthermore, the same effect of validity that previous literature documents, with Valid conditions producing a significantly shorter RTs compared to

Invalid conditions should also be found. Finally, in conjunction with other replications of the occluding barrier method, conditions in which the barrier inhibits the view of the target any spontaneous visual perspective taking-like effects should be abolished.

3.3.2 Method

Participants

Thirty-six participants were recruited with a mean age of 21.47 (SD =2.76, range = 18-30). Twenty participants were female (sixteen male), with three identified as left-handed (thirty-three right-handed). In addition, the variation of ethnicity in the sample were; seven Asian, four Black, one Hispanic, twenty-one White and three who identified themselves as a non-listed ethnic group. All participants reported normal or corrected to normal vision and were recruited through the University of Essex online volunteer portal known as 'SONA', and were reimbursed for their time.

Stimuli and Apparatus

Eight variations of stimuli were generated, using different variation of the three manipulated factors. See Figure 3.4 for an example of the basic framework used in the stimuli. All the stimuli began depicting an agent sat behind a table facing and looking directly at the viewer. To each side of the agent, a red balloon was presented, two in total. Additionally, the agent was depicted actively holding an occluding visual barrier, which obstructed the agent from seeing one side of the display and consequently one of the red balloons. In half of the trials the barrier was presented on the left-hand side of the screen and the other half on the right. Directly in the middle of the stimuli, superimposed over the central point of the screen was a black disk, which depicted the central fixation point used at the start of each trial.

Photoshop was used to manipulate the luminance of the two balloons, which acted as the targets. Eye movements were recorded with a head-mounted, video-based eye tracker (EyeLink 1000; SR Research Ltd., Osgoode, Ontario, Canada), using a sampling rate of 1000 Hz. They were recorded monocularly and analysed using Eyelink Data Viewer (SR-Research). All of the stimuli were displayed on a 19-inch monitor with a 1024 x 768-pixel screen resolution, with the stimuli being presented at 25 frames per second, using Experiment Builder presentation software (SR-Research), with a viewing distance of approximately 57 cm.

Figure 3.4: The basic arrangement used in Experiment 5. The agent is shown holding an occluding barrier which participants are made explicitly aware inhibits the view of the balloon behind it, for the agent. During the experiment the agent will move her gaze to either the left- or right-hand side of the image and the balloons luminance will change, one becoming brighter, the other darker, the latter acting as the target.

Design and Procedure

A 2 x 2 x 2 repeated measures design was used, with validity (Valid, Invalid), visibility (Visible, Occluded), and SOA (80ms, 300ms) being manipulated.

Participants were positioned in a chinrest, approximately 60cm from the display screen. Prior to the onset of the practice trials, participant eye movements were calibrated and validated using a nine-point calibration procedure. In addition, each trial began with a drift correction and started only once the participant had fixated on the centre of the screen, identified by a black disk.

Each trial began with the agent facing the participant for 1500ms. During each trial the agent's head moved towards the left or right-hand side of the screen, thus changing her directional attention. In addition, one of the red balloons illuminated whilst the other became darker. Participants were required to focus on the black disk in the middle of the screen, ignoring all other visual stimuli until the balloons changed luminance. Once this occurred participants were required to fixate upon the darker balloon as quickly as possible and press the space bar on a standard keyboard in front of them, ending the trial. Participants were explicitly informed that the agent's head movements were random, and the occluding barrier obstructed the agent's view of one of the balloons. Figure 3.5 represents the sequence of a single trial used in Experiment 5.

The experiment comprised of six practice trials, followed by 128 experimental trials with 64 trials with the barrier on the left and 64 with the barrier on the right. The order for which side the barrier started on was counterbalanced. After the first block was administered, a break was given. Prior to the start of the second block, participant eye movements were again re-calibrated and validated using the nine-point calibration procedure. Lastly, participant eye movements were analysed using DataViewer 2.4.0.198 (SR-Research) and cross-analysed with the Davis (1980) Interpersonal Reactivity Index results using IBM SPSS Statistical (Version 21) software.

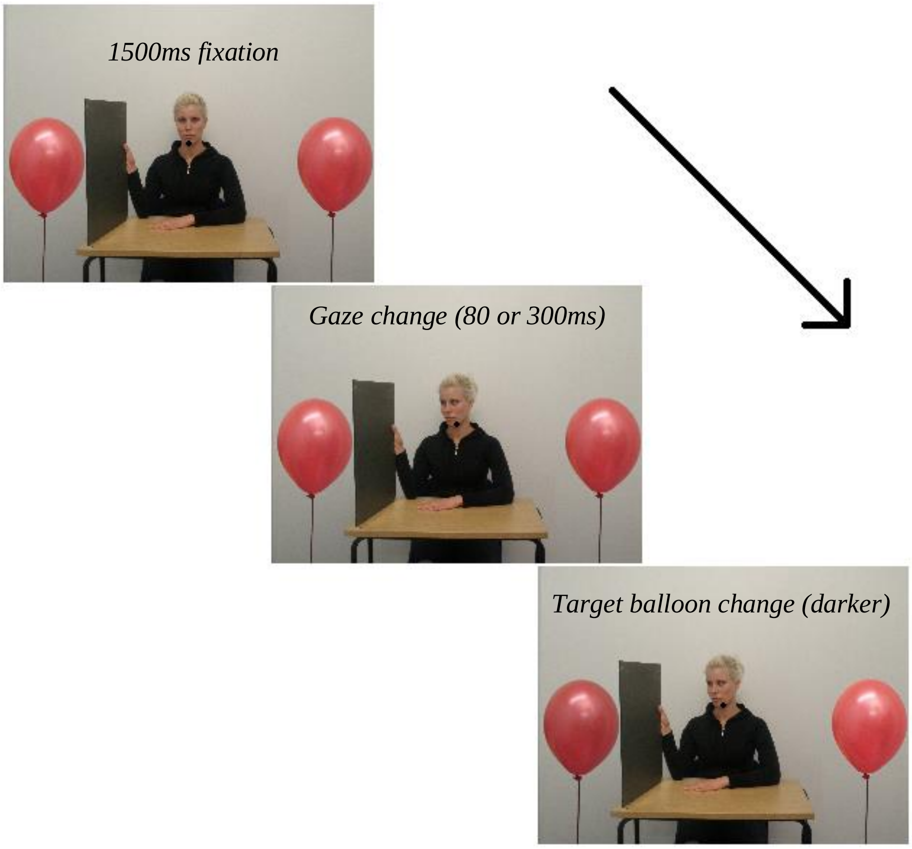

Figure 3.5: The sequence of an Occluded, Valid trial

3.3.3 Results and Discussion

Data Preparation. The data were collated and analysed using DataViewer (SR-Research), Microsoft Excel and IBM SPSS Statistical (Version 21) software. Firstly, using DataViewer (SR-Research) each balloon was defined as a Region of Interest and using the Reaction Time Manager and Reaction Time Definitions, a RT measurement between target onset and participant's first fixation upon the target was created. As each trial ended once the target was detected, dwell time was not processed, as it was not meaningful. Both the fixation RT and the manual RT, which ended each trial was extracted from DataViewer (SR-Research). These RTs were then collated and used to create mean RT for each manipulation of the three factors of validity, visibility and SOA, using Microsoft Excel pivot tables. Any outliners that were 2 standard deviations above or below the mean were excluded from the overall analysis. Four participants were excluded from analysis as either they were not successful in fixating upon the correct target for over 50% of the total trials, or the equipment did not effectively track saccades. IBM SPSS Statistical (Version 21) software was then used to assess the effects of the factors upon time to fixate upon the target and manual RT as well as correlations against the self-reported Davis (1980) Interpersonal Reactivity Index.

Target Fixation. Figure 3.6 shows mean RT to fixation on the correct target. An ANOVA with validity, visibility and SOA as within-participant factors found a significant main effect of validity, $F(1, 31) = 8.76$, $p < .006$, $n_p^2 = .22$, thus showing the standard gaze cuing effect. There was also a significant main effect of SOA, $F(1, 31) = 18.03$, $p < .001$, $n_p^2 = .37$, but there was no significant main effect of visibility, $F(1, 31) < .001$, $p = 1.00$, $n_p^2 = .001$. There was no SOA by validity interaction, $F(1, 31) = 2.74$, $p = .11$, $n_p^2 = .09$, or SOA by visibility, $F(1, 31) = .21$, $p = .65$, $n_p^2 = .007$. Most importantly there was no validity by visibility interaction, $F(1, 31) = .23$, $p = .64$, $n_p^2 = .007$, thus showing that the effect of validity was not effected by whether the target was visible or occluded for the agent. Finally, there was no SOA by validity by visibility interaction, $F(1, 31) = .06$, $p = .81$, $n_p^2 = .002$. Overall, these data show that participants were significantly faster to fixate upon the target when the agent gazed towards the target as well as during conditions in which the target onset was set to 300ms. Thus, demonstrating the standard gaze cuing effect. However, the presence of the barrier, which occluded the agent from viewing the target, did not affect RT to fixate upon the target.

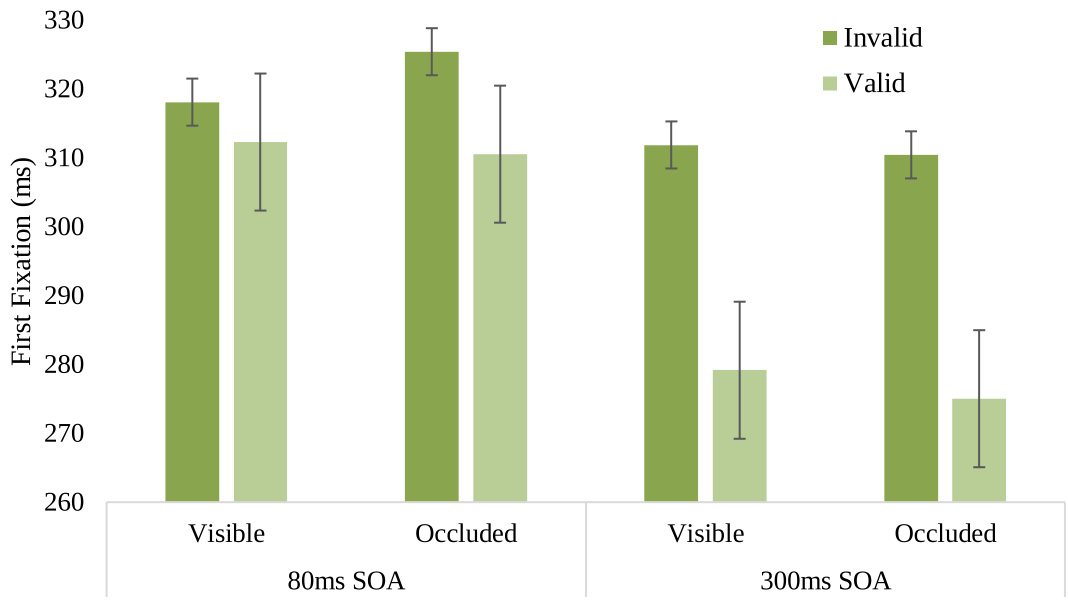

Figure 3.6: Mean fixation RTs for each of the eight conditions, with standard error bars included

Davis (1980) Interpersonal Reactivity Index. Pearson's correlation coefficients assessed the relationship between susceptibility to spontaneously assume an allocentric visual perspective using the fixation data and self-reported measures of empathy using the Davis (1980) Interpersonal Reactivity Index. To calculate a participant's susceptibility to spontaneously assume an allocentric visual perspective, the difference between Valid and Invalid conditions were used to create an index. Additionally, the Davis (1980) Interpersonal Reactivity Index was deconstructed into its components of, *Perspective Taking*, *Fantasy*, *Empathic Concern* and *Personal Distress*. No significant correlation was found between susceptibility index, and *Perspective Taking* (r = .05, p = .79), *fantasy* (r = .02, p = .92), *Empathic Concern* (r = .16, p = .34), or *Personal Distress* (r = .11, p =. 56). Consequently, in terms of the self-reported measures of empathy and the susceptibility of spontaneous visual perspective taking, no relationship was found.

Manual Reaction Time. Figure 3.7 shows mean manual RT. An ANOVA with validity, visibility and SOA as within-participant factors found a significant main effect of validity, $F(1, 31) = 9.31$, $p < .005$, $n_p^2 = .23$, thus showing the standard gaze cuing effect. There was also a significant main effect of SOA, $F(1, 31) = 18.23$, $p < .001$, $n_p^2 = .37$, but there was no significant main effect of visibility, $F(1, 31) = .04$, $p = .84$, $n_p^2 < .001$. There was no SOA by validity interaction, $F(1, 31) = 1.76$, $p = .19$, $n_p^2 = .05$, or SOA by visibility interaction, $F(1, 31) < .002$, $p = .96$, $n_p^2 = .001$. Most importantly there was no validity by visibility interaction, $F(1, 31) = .41$, $p = .53$, $n_p^2 = .01$, thus showing that the effect of validity was not effected by whether the target was visible or occluded for the agent. Finally, there was no SOA by validity by visibility interaction, $F(1, 31) = .01$, $p = .92$, $n_p^2 = .001$. In sum, as with the fixation results, Experiment 5 found that participants were significantly faster to response manually when the agent gazed towards the target, as well as during conditions in which the target onset was set to 300ms. Thus, demonstrating the standardized gaze cuing effects. However, most importantly the presence of the occluding barrier, which inhibited the agent from viewing the target, did not affect RT.

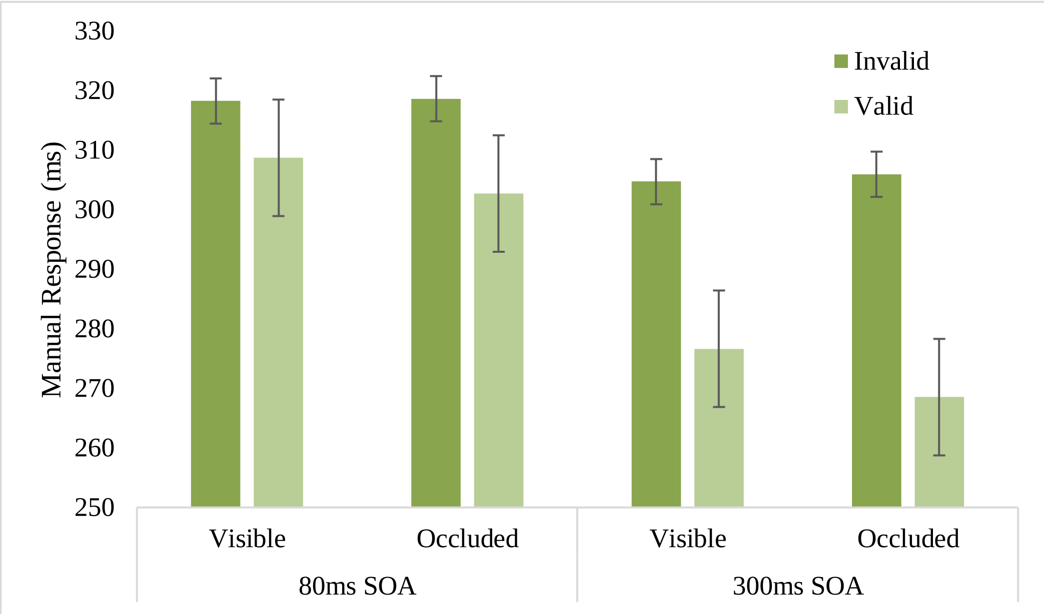

Figure 3.7: Mean manual RTs, for each of the eight conditions, with standard error bars included

Davis (1980) Interpersonal Reactivity Index. Pearson's correlation coefficient assessed the relationship between susceptibility of spontaneous visual perspective taking using the manual RT and self-reported measures of empathy from the Davis (1980) Interpersonal Reactivity Index. Again, as with the fixation data, a participant's susceptibility to spontaneously assume another individual's visual perspective was calculated as the difference between Valid and Invalid conditions. No significant correlation was found between the computed susceptibility index and *Perspective Taking* (r = .06, p = .77), *Fantasy* (r = .12, p = .94), *Empathic Concern* (r = .16, p = .38), or *Personal Distress* (r = .11, p =. 56). Consequently, in terms of the self-reported measures of empathy, and spontaneous visual perspective taking susceptibility, no relationship was found.

Overall, these data show that the agent's gaze shifted the attention of the observer in that same direction. Most importantly, this effect was not hindered with the presence of the occluding barrier, which barred the view of the target object for the agent. Subsequently, conclusions in terms of spontaneous visual perspective taking cannot be concluded; an individual cannot represent the visual perspective of another when that individual cannot see the target. It could be argued that gaze cuing is one phenomenon that may be driving the results within this experiment as well as others investigating spontaneous visual perspective taking. It has been commonly concluded that individuals are drawn to the direction of another individual's attention, which is modulated through gaze (Driver et al. 1999; Friesen & Kingsonte, 1998; Hietanen, 1999, 2002; Langton & Bruce, 1999). For example, in this experiment, as the line of sight from the agent to the target is identical, irrespective of the presence of the occluding barrier, the effect remains the same. This notion is further supported by the results of Experiment 5, which found no significance in terms of empathy and susceptibility to spontaneously assume the visual perspective of another. In other words, if the effect that has been identified within this research is not grounded in ToM processes but instead associated with attention, then empathy, which is rooted in ToM, should not be found to be associated. However, to increase the reliability of this conclusion, future work would benefit from examining whether spontaneous visual perspective taking-like effects are found even in participants without a fully formed ToM, for example, young children and infants.

3.4 General Discussion

The current chapter applied eye movement measures to the investigations of so-called spontaneous visual perspective taking. It can be argued that such measures increase sensitivity to human attention in comparison to behavioural measurements (Just & Carpenter, 1980). Experiment 4 primarily examined the significance of validity, human and non-human directional cues, and SOA in both Covert and Overt tasks, whereas Experiment 5 assessed the importance of validity, occluding barriers and SOA. In addition, alongside the use of eye movement, empathy was also correlated with susceptibility to spontaneously assume another individual's visual perspective using the Davis (1980) self-report Interpersonal Reactivity Index in both Experiment 4 and 5.

As was expected, Experiment 4 found that fixation RT in the Overt task was significantly shorter when compared with manual RT in the Covert task. Thus, validating the use of eye movement measures in the investigation of spontaneous visual perspective taking. In addition, RT was significantly shorter during Valid conditions, where the distractor direction and target matched, in comparison to Invalid conditions, where the distractor direction and target mismatched. However, there was no correlation between empathy and susceptibility to spontaneously assume an allocentric visual perspective for both Experiments 4 and 5, and there was also no difference in response between human and non-human directional cues in Experiment 4. Furthermore, Experiment 5 found that the validity of the agent's gaze shifted the attention of the observer in the same direction. Most importantly, this effect was not hindered with the presence of the occluding barrier, which barred the view of the target for the agent.

Taken together the current chapter highlighted that the effects observed using manual RT in the investigation of spontaneous visual perspective taking, can be observed using eye movement measures also. However, both experiments have challenged the central claim of spontaneous visual perspective taking. If spontaneous visual perspective taking is indeed associated with ToM, then Experiment 4 should have only observed the effects associated with spontaneous visual perspective taking during the Face condition, and not also with the Arrow cue. Additionally, in Experiment 5, the occluding barrier should have abolished the effect of spontaneous visual perspective taking, as an individual cannot spontaneously assume the visual perspective of another when that other person cannot see the target. Moreover, both experiments should have also observed a correlation between measures of empathy and susceptibility to visual perspective taking, as ToM improves with enhanced empathetic ability.

As was previously discussed in Experiment 4, instead of individuals assuming the visual perspective of another, which is governed by ToM, the results of the present chapter may simply be due to attentional orienting. In other words, the directional cue emphasised by gaze direction or the arrowhead direction may have overruled all other visual stimuli in the processing of the task, thus enhancing RT. However, to assess this claim further, future work could aim to address the question as to whether ToM is essential for the so-called spontaneous visual perspective taking phenomenon to be observed. In the following chapter, the claim that spontaneous visual perspective taking is routed within ToM will be explored using developmental procedures.

Chapter Four:

Spontaneous Perspective Taking in Children and Adults

4.1 Chapter Overview

Currently the present thesis has examined the so-called spontaneous visual perspective taking notion in accordance to its theoretical assumptions, as well as examine its central empirical paradigms. In the following chapter, another central claim to spontaneous visual perspective taking will be assessed. Specifically, if this phenomenon resides within ToM, then it should also be developed throughout childhood when other aspects of ToM are forming. In the following sections the developmental trajectories of ToM, Perspective Taking and Spontaneous Perspective Taking will be outlined.

Theory of Mind

Recall from Chapter 1 that ToM impacts social cohesion, processing of alternate viewpoints and perspectives as well as levels of sympathy and empathy. From a very early age, children are encouraged to play in social groups; which is one key stage of growth that nurtures ToM development. For example, role-play as well as games in which children use their imagination promote the development of ToM (Leslie, 1987). This is due to the child assuming different mental states, emotions, levels of knowledge, and perspectives that may differ from others and the self (Frith & Frith, 2006). Thus, using these abilities to distinguish differences between the self and others, the child can begin to empathise and understand adult social interactions. For example, false belief and deception have been associated with effective ToM (Flavell, 1999). These individuals are able to understand mundane social interactions such as sarcasm and exaggeration (Happé & Frith, 1995), and thus increase their social cohesion and stratification. In addition, Goldstein and Winner (2012) also emphasised the importance of sympathy and empathy as an extension of ToM, as individuals are able to identify and understand how others feel in different situations, which leads to the promotion of altruistic behaviours, that in turn strengthens social bonds.

Although developmental psychologists have devised a number of methods in order to assess a child's ToM ability (e.g., Carpendale & Lewis, 2006; Doherty, 2008) the most popular task is through the assessment of false belief. During a false belief task, a child is presented with a story style scenario in which the perspective of a character differs from the perspective of the child. The child's ability to deduct whether the character will search for an object in the place that the character assumes the object is hidden or where the child knows that the object is hidden (the character disappears from the scene when the object is moved to a new hiding space) forms the basis for the judgement of the participants ToM ability (e.g., Wimmer and Perner 1983; Wellman, Hollander, and Schult 1996). Many different variants of this scenario have developed including with the focus on the participants own false belief (e.g., Wimmer & Perner 1983) and the false belief of another (e.g., Siegal and Beattie 1991) yet the standard assumptions have remained relatively the same throughout the literature's development.

The specific age at which ToM develops is debatable. Wimmer and Perner (1983) originally identified that many children judged the false belief scenario incorrectly until around five years of age. However, the age at which this ability forms has been heavily disputed throughout the development of this literature. Consequently, Wellman, Cross and Watson (2001) carried out a meta-analysis of 178 experiments that used false belief tasks to identify the commonly agreed age at which ToM ability develops. Wellman, Cross and Watson (2001) concluded that the likelihood of a child giving the correct answer within a false belief task changes most significantly between three and four years of age rather that the originally thought four and five years of age. It is worth noting that Wellman, Cross and Watson (2001) also identified that the children's errors were not random. It was identified that they respond egocentrically, with their own point of view. Wellman et al. (2001) also identified that the little variations upon the fundamental false belief task had little corresponding effect on the research findings. Children still preform the same number of errors and within the same bonds of age differences. Thus, it is common belief amongst developmental psychologists that children aged three years and below are unable to differentiate in terms of alternative perspectives and thus have not yet fully developed their ToM abilities.

However, more recently, the focus of ToM development has shifted towards infants. Senju, Southgate, Snape, Leonard, and Csibra (2011) investigated ToM and age with the use of traditional mental state attribution tasks within 18-month-old infants. The procedure that Senju et al. (2011) adopted was an extension of a proposed procedure by Heyes (1998). This method used two visually identical blindfolds, with one being opaque and the other transparent. Once the infants wore one of the blindfolds it is then transferred to an adult confederate. Heyes (1998) suggested that infants who wore the transparent blindfold would follow the gaze of the confederate, whereas the infants that wore the opaque blindfold would not. Consequently, this would suggest that infants are able to use past experiences to predict the perception of another through an attribution process. Senju et al. (2011) adopted this procedure with the addition of a search task. Once the adult confederate was wearing the blindfold, a toy was hidden. Using eye tracking, the infants' anticipatory eye movements were recorded as an objective measure of assessing where the infant would predict the confederate would search for the toy. Senju et al. (2011) found that infants who wore the opaque blindfold gazed towards the original false-belief hiding spot, whereas the infant who wore the transparent blindfold looked predominantly towards the new hiding spot. Hence the authors concluded that 18-month-old infants were able to use past experiences, in this case wearing the blindfold, to form perceptual judgments of another individual's false beliefs. Despite this, in terms of the overall development of ToM, the evidence that infants can form perceptual judgments cannot be taken as an indication of overall ToM. This is due to the presence of other complex components within ToM such as emotion and desires, which cannot be accounted for within this procedure. Subsequently, the debate of age continues, and the general consensus remains, driven by the false belief task, that ToM abilities develop around three to five years of age.

Perspective Taking

As was outlined in Chapter 1, perspective taking is a crucial component of ToM that significantly impacts everyday thought and action (Surtees, Apperly, & Samson, 2013), particularly in relation to predictions of what others are thinking and attending to. This process can be deconstructed further in terms of levels and different aspects of perspective taking. Both will be discussed in the following section in terms of their developmental assumptions.

Previous literature highlights a distinction within visual perspective taking in terms of levels. Level 1 visual perspective taking involves identification of whether another agent or model can or cannot see a target object, whereas, level 2 refers to whether the agent can or cannot perceive the target object differently (Flavell, Everett, Croft, & Flavell, 1981; Salatas & Flavell, 1976; Yaniv & Shatz, 1990). Additionally, the age at which an individual is able to successfully exhibit level 1 and level 2 perspective taking has also been examined, with level 1 being identified to develop earlier than level 2 (Moll & Tomasello, 2006; Surtees & Apperly, 2012).

Moll and Tomasello (2006) initially aimed to assess the age at which a child first engages with the ability to comprehend what another person sees may differ from the self, otherwise known as level 1 visual perspective taking. To achieve this, the authors developed a task in which children were asked to help the experimenter find a toy that was either visible to the experimenter and child, or was occluded by a barrier for the experimenter but not the child. The behaviour of the participant was recorded. It was found that when participants were asked to help the experimenter find the occluded object, 24-month-old children handed the object to the experimenter, but 12-month-old children did not. The authors argued that this difference in behaviour is driven by the children's visual perspective taking ability, which the 24-month-old children had developed, but the 12-month-old children had not. In other words, the authors were suggesting that the 24-month-old children were able to understand that the experimenter perceived the scene differently, and therefore was not able to perceive the target object, and thus handed them the toy, whereas the 12-month-old children had not developed this ability. Prior to this research it was thought that visual perspective taking developed at a much later stage in a child's development (e.g., Flavell, et al., 1981; Light & Nix, 1983; Masangkay et al., 1974; Piaget & Inhelder, 1956).

Perspective taking can also be distinguished in terms of visual and spatial perspective taking. Whereby visual perspective taking refers to whether individuals can distinguish that a target is seen or not, and spatial perspective taking refers to whether an individual is able to judge the target's positioning relative to the alternative viewpoint (Surtees & Apperly, 2012). Surtees, et al. (2013) examined the similarities and differences between these two components, particularly in relation to any influence that differing levels may also have. They found that level 2 visual perspective taking, and the understanding of angles in terms of spatial perspective taking, required complex cognitive processing, specifically related to mental rotation. In terms of age, it has been highlighted that a child aged between three and four years can distinguish between different spatial perspectives, for example, if an object is in front or behind the alternative viewpoint (Bialystok & Codd, 1987; Cox, 1981; Harris & Strommen, 1972), whereas older children can also distinguish the differing angles of the alternative viewpoints (Harris, 1972; Harris & Strommen, 1972). Additionally, Sodian, Thoemer and Metz (2007) highlighted that differences between rationalisation of perspective can also be influenced in regards to age. Specifically, Sodian, Thoemer and Metz (2007) identified that 14-month-old infants were able to rationalise that an individual was reaching for an object that was hidden from them, whereas 12-month-old infants did not have this ability. As can be seen, perspective taking develops throughout childhood with individual's ability progressively improving with age.

Spontaneous Perspective Taking

As was stated in Chapter 1, a number of authors have identified an extension of perspective taking, termed as 'spontaneous perspective taking', in which another's visual perspective is computed both 'rapidly' and 'spontaneously' (Samson, et al. 2010). In the following section the developmental assumption of this phenomenon will be outlined.

It has been argued that spontaneous visual perspective taking is not only developed throughout childhood, but can also to be linked to innate abilities by Surtees and Apperly (2012). Initially Surtees and Apperly (2012) assessed the ability to differentiate between egocentric and allocentric perspectives in accordance to age. It was found that all participants, regardless of age, processed the egocentric and allocentric perspectives, concurring with the RT differences leading to the conclusion of the spontaneous visual perspective taking effect. In other words, when shown a stimulus with an incongruent allocentric perspective, RTs increased due to egocentric interference (Surtees & Apperly, 2012). Consequently, this would suggest that spontaneous visual perspective taking, specifically in terms of egocentrism and automatic processes, are not developed throughout childhood. This conclusion is due to children in the experiment as young as six being influenced by the presence of an incongruent allocentric visual perspective. Conversely, this research only focused on one experimental procedure in the assessment of spontaneous visual perspective taking. Thus, further assessment in terms of spontaneous visual perspective taking and age is needed, specifically in terms the ambiguous number paradigm.

Epley, Morewedge, and Keysar (2004) assessed perspective taking in terms of the spontaneous assumption of visual perspectives as well as egocentrism and age. Using eye movement measures during a communication task the authors assessed the level of egocentrism within adult and child participants. Specifically, to see whether there are differences when participants were asked to reach for a certain object which position and visibility was manipulated to differ between the participant's and experimenter's visual perspective. The age of participants ranged from four to 12-year-olds. The authors found that both adults and children looked towards and identified the hidden object to the experimenter, which matched the experimenter's question. Therefore, if this object were chosen an egocentric response would have been given. However, adults were more likely to correct for this egocentric consideration and reach for another similar object which could also be seen by the experimenter, whereas, the younger participants were significantly more likely to reach for the hidden object, which resulted in an egocentric response. Consequently, Epley, et al. (2004) concluded that both children and adults rapidly and spontaneously compute the intention of the experimenter question egocentrically, and only differ in terms of the likelihood and speed of correcting this egocentric judgment. Thus, suggesting that the assumption of another's visual perspective is not developed with age, but instead the ability to correct oneself in terms of considering another visual perspective is improved with age. In other words, the authors have suggested that both adults and children have the same ability to respond egocentrically, but adults are able to correct and filter their response in terms of alternative perspectives. However, the age at which this ability to correct oneself to account for another perspective is still debatable. Furthermore, it can be argued that the experimental design that Epley, et al. (2004) used differed considerably to the current paradigms employed in spontaneous visual perspective taking literature. Thus, the reliability of whether the design specifically assesses spontaneous visual perspective taking is questionable.

Surtees, Butterfill and Apperly (2012) assessed differences in performance in terms of direct and indirect measures of visual perspective taking within a sample of children. The authors argued that the differences in performance could be due to differences in experimental designs that assess either direct or indirect measures of visual perspective taking. The authors suggest that direct measures assess a cognitive effortful ability to reason about another alternative viewpoint, whereas indirect measures assess the efficient less cognitively demanding account of ToM abilities. Using an adapted version of Samson et al. (2010) design, the authors were able to assess level 1 and 2 visual perspective taking using both direct and indirect measures in a sample of children aged six to 11-year-olds as well as adults. It was found that younger participants were able to make level 2 judgements using direct measures, but there was no evidence to support automatic indirect measures of level 2 judgements. Therefore, Surtees et al. (2012) support the notion of level 2 direct measures of perspective taking, which is developed throughout childhood. However, the age at which these abilities are developed is still debatable. In addition, as spontaneous visual perspective taking is still challenged as to whether it is an automatic process that resides within ToM or driven by other cognitive mechanisms, further investigation with children and infant participants is essential.

As can be seen the developmental trajectory of spontaneous visual perspective taking is still questionable. The current chapter will examine at what age the phenomenon is developed using adapted versions of the ambiguous number paradigm. Firstly, Experiment 6 will compare the responses of children and adult participants, using shapes as the ambiguous figures. Consequently, assessing the influence of ambiguity, validity, and age in respect to spontaneous visual perspective taking. It is predicted that if little to no difference is found in terms of age, Experiment 6 could suggest that the so-called spontaneous visual perspective taking notion may not be dependent upon the rapid assumption of an alternative visual perspective, routed within ToM, but may be due to other cognitive mechanisms not previously identified. Additionally, Experiment 7 again will examine age using an adapted ambiguous task. However, Experiment 7 will simplify the procedure and only examine ambiguity in respect to spontaneous visual perspective taking, using a face variation of the ambiguous figure. It is predicted that if spontaneous visual perspective taking is the rapid assumption of an alternative perspective, routed in ToM, the youngest participants will not be affected by the alternative perspective and only respond using the egocentric self perspective, whereas the alternative perspective will influence the older participants, which in turn will affect their responses during Ambiguous conditions.

4.2 Experiment 6 – The Ambiguous Shape Task for Adults and Children

4.2.1 Introduction

Experiment 6 was devised to examine whether spontaneous visual perspective taking is a process routed within ToM and is not dependent upon other cognitive shortcuts or vision attribution. This experiment used a sample of children and adult participants, for a cross comparison of performance between ages, thus assessing the developmental trajectories of the spontaneous visual perspective taking notion. As a meta-analysis of developmental research by Wellman, Cross and Watson (2001), found that the likelihood of a child correctly responding to a false belief task as a result of a fully formed ToM ability significantly improves between three and four years of age, a sample of, three, five, seven, and nine-year olds, as well as adults above the minimum age of 18 was obtained. These ages were gathered due to the assumption that the three-year-old participants should not have a fully developed ToM, whereas the seven-year-old, nine-year-old, and adult participants should. Therefore, is it hypothesised that the responses of the three-year-old participants should significantly differ to those of the seven-year-old, nine-year-old and adult participants, due to the development of a ToM. Additionally, with respect to the five-year-old participants, as the previous literature has also suggested that ToM is developed at five years of age (e.g., Wimmer & Perner 1983), half of the sample should perform in terms of the younger participants and half in terms of the older, due to individual differences in the speed of ToM development. However, if little to no difference is found in terms of age, the current work would challenge previous conclusions of the rapid assumption of another's visual perspective, routed within ToM, in favour of previously unidentified cognitive mechanisms.

Participants were required to manually react to ambiguous shapes presented, embedded within stimuli on a computer screen. The stimuli used possessed two important properties. Firstly, the shape could be interpreted the same (Unambiguous) or differently (Ambiguous) depending on the visual perspective adopted be that the egocentric self perspective or the depicted agent's allocentric perspective. Secondly, the agent in the stimulus was either looking towards the shape (Valid) or away from the shape (Invalid). This experiment therefore allowed the direct comparison of responses, as well as RT, of Unambiguous and Ambiguous conditions (when the agent and participants perspectives matched or did not match) as well as Valid and Invalid conditions. It is expected that if spontaneous visual perspective taking is the rapid assumption of an alternative visual perspective, routed in ToM, the responses of participants with a formed ToM ability will significantly differ to those without a formed ToM. Specifically, the older participants RT will be significantly shorter with an Unambiguous shape in comparison to the Ambiguous shape, but only during the Valid conditions, whereas, the younger participants will not.

4.2.2 Method

Participants

Ninety-nine participants (64 female) were recruited, 79 students from a local Primary school, and 20 University students. The sample of, three, five, seven, and nine-year-olds were obtained from a local Primary school and the sample of adults from the University of Essex online recruitment platform, known as 'SONA'. The distribution of gender for each age group can be viewed in Figure 4.1, and all reported normal or corrected to normal vision. As a means of reimbursement, the university sample received course credits, and the Primary school sample received stickers and a 'Young Scientist Award'.

Age Group	*Total*	*Female*	*Male*
3	*18*	*12*	*6*
5	*21*	*12*	*9*
7	*19*	*9*	*10*
9	*21*	*15*	*6*
18+	*20*	*16*	*4*

Figure 4.1: The distribution of female to male participants for each age group

In terms of the Primary school sample, informed consent was obtained from the head teacher after a full brief and outline of the experiment was discussed. In addition, consent forms were sent to the participant's parent/carers, using an opt-out procedure, ensuring that parents/carers were fully aware of the experiment and how the data was to be obtained and distributed. Additionally, the university sample received a consent form prior to the beginning of the experiment.

Stimuli and Apparatus

Six images were generated; three of these are shown in figure 4.2. The additional three followed the same pattern in factor manipulation. Two factors were manipulated in these images, validity of the trials depicted in terms of the agent's gaze direction and ambiguity of the shape. Valid trials (top and middle panel in figure 4.2) showed the agent gazing toward the target shape on the table, whereas Invalid trials (bottom panel in figure 4.2) showed the agent gazing off to the right-hand side of the image, away from the target shape. Additionally, for ambiguity of shape one of three shapes was used, either a circle, upward facing triangle or downward facing triangle. The circle is deemed as Unambiguous, as the perception of the shape is identical for the visual perspective that is adopted, be that from the viewer's egocentric perspective or the allocentric perspective of the agent. Alternatively, both triangles are deemed as Ambiguous as the perceptions of the triangles differ in terms of whether the viewer adopted their own egocentric visual perspective or the agent's allocentric visual perspective. Consequently, six trials were used, three Valid trials using; upwards-pointing triangle, downwards-pointing triangle and circle and three Invalid trials using; upwards-pointing triangle, downwards-pointing triangle and circle.

The presented images were 886 x 752 pixels in size, with a display ratio of 886 x 752 pixels. Additionally, the experiment was conducted using the SuperLab 5.0 desktop application, on a windows 10 laptop, with a screen size of 12 inches and resolution of 1366 x 768. RTs were recorded using an RB-844 cedrus.com response pad, which held visual representations for each of the shapes used in the experiment presented on different buttons on the response pad. The data were analysed using IBM SPSS (Version 21) and R Studio (R Core Team, 2016) statistical software.

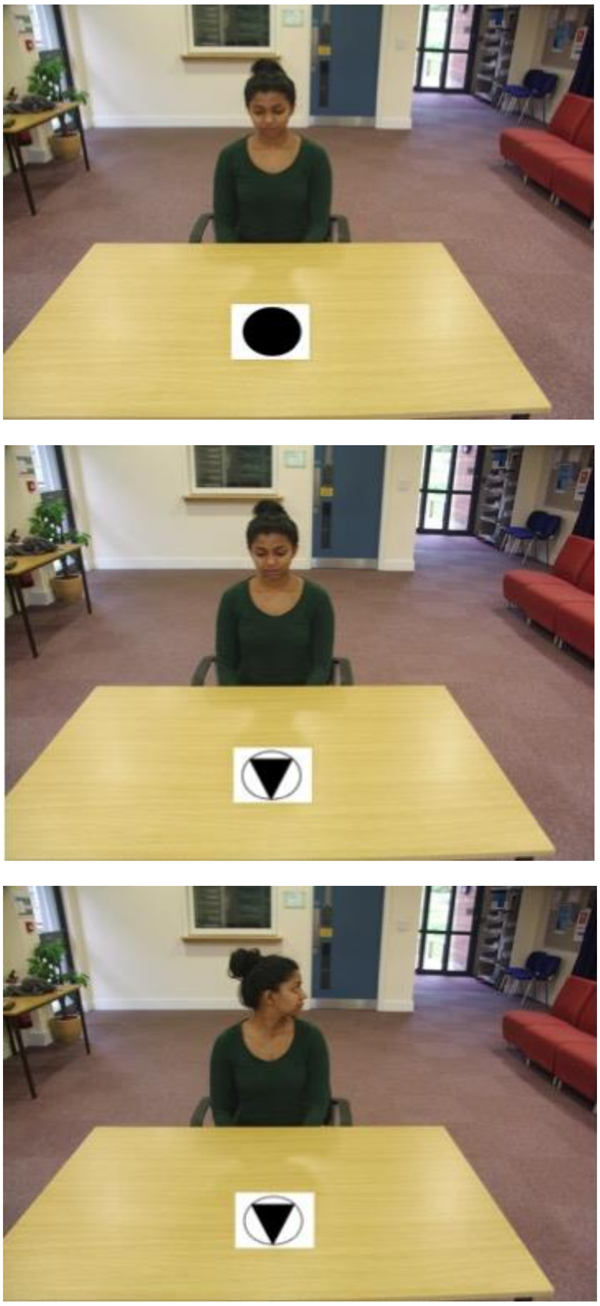

Figure 4.2: These images represent the critical conditions used in Experiment 6. The top panel is Unambiguous, Valid. This image is Unambiguous as the visual perspective of the shape is identical for the viewer as well as the agent, additionally; it is Valid as the agent is shown to be freely gazing at the shape. Alternatively, the middle image is Ambiguous Valid, as the agent is still freely gazing at the shape, but the shape can be perceived differently for the viewer and agent's visual perspective. Lastly, the bottom image is Invalid Ambiguous, as the agent's gaze is averted to the side of the image, away from the shape, and the visual perspective of the shape differs for the viewer and agent's visual perspective.

Design and Procedure

A 2 x 2 design was used with shape ambiguity (Ambiguous, Unambiguous) and validity (Valid, Invalid) as within-participant factors. In other words, the factors that were manipulated were the agent’s gaze direction and the shape presented. The target shape presented was either Unambiguous (circle) or Ambiguous (triangle) in terms of the agent and viewer’s visual perspective. In addition, in term of validity, the agent was either depicted as looking towards the shape (Valid), or away from the shape (Invalid). Furthermore, a between-participants factor was also accounted for in terms of age. A sample of participants for the following age groups of; three, five, seven, nine-year olds, as well as adults above the minimum age of 18 were sought to assess the developmental trajectory of spontaneous visual perspective taking.

Due to the age of the participant sample, an understanding check was administered prior to testing. This was achieved by presenting participants with two different variations of the experimental stimuli, in isolation. The images were presented in colour, high resolution laminated A5 photographs. Each participant was required to press the corresponding button on the RB-844 cedrus.com response pad. If the correct shape response was given, regardless of orientation, participants could proceed with the experiment. Participants’ verbal consent was then sought for the rest of the procedure.

Participant responses were recorded to identify the visual perspective adopted for each trial, as well as RT. Eighteen trials were used in total, three trials for each condition, thus ensuring that an equal distribution of shape orientation and model positioning were presented. Prior to the onset of each trial, participants were required to press the ‘GO!’ button on the response pad when the corresponding picture was presented on the screen. This response was not recorded and was only used to ensure that the participant was attending to the screen before the onset of the eighteen-recorded trials. At the end of each trial, after participant response, a blank screen was presented for 250ms signifying the end of the trial. A flow chart of this process is depicted in Figure 4.3. Once the experiment was complete, participants were reimbursed for their time.

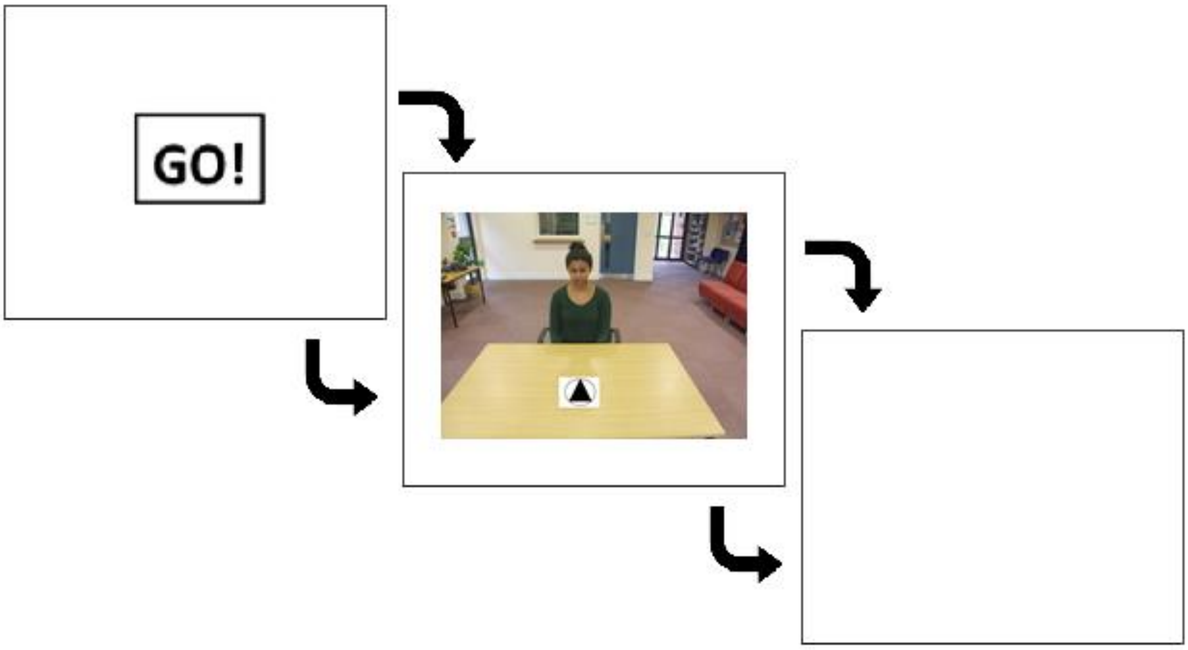

Figure 4.3: An example of the order for a Valid, Ambiguous trial, in which a participant pressed, "GO!" to signify the onset of the trial, then responded to the stimulus before a blank screen was displayed for 250ms indicated the end of the trial.

4.2.3 Results and Discussion

Data Preparation. The data were collated and reorganised using R Studio (R Core Team, 2016) and collapsed into a single data file. It was then coded to produce mean RTs for the two levels of each manipulated factor of validity and ambiguity. Any outliners that were 2 standard deviations above or below the mean were excluded from the overall analysis. The data were then filtered to ensure that only participants with scores for each level of each factor were used in the overall analysis.

Consequently, five participant's data were removed from analysis, three of which from the three-year-old sample and two from the nine-year-old sample. Therefore, a total of 15 three-year olds, 21 five-year olds, 19 seven-year olds, 19 nine-year olds and 20 adults were used in the following analysis. Additionally, a percentage score for correct self-perspective answers was also calculated for further analysis investigating error rate.

Reaction Time. Figure 4.4 depicts the overall mean RT for the four conditions. An ANOVA with ambiguity and validity as within-participants factors and age as a between-participant factor found a significant main effect of ambiguity, $F(1, 94) = 3.87$, $p < .05$, $n_p^2 = .04$, a significant main effect of validity, $F(1, 94) = 7.63$, $p < .007$, $n_p^2 = .08$, but there was no significant interaction, $F(1, 94) = 1.44$, $p < .23$, $n_p^2 = .02$. However, when factoring in age as a between subjects factor there was no significant effect of ambiguity, $F(1, 94) = .43$, $p < .79$, $n_p^2 = .02$, but there was an effect of validity, $F(1, 94) = 4.06$, $p < .004$, $n_p^2 = .15$, and no interaction between all factors, $F(1, 94) = 1.72$, $p < .15$, $n_p^2 = .07$. In other words, when assessing the overall effects of validity and ambiguity, participants were significantly faster to respond during Unambiguous Valid conditions when compared with Ambiguous Invalid. However, when accounting for age, there was only a significant effect of validity. Thus, in accordance to the different age groups, there was only a significant effect between Valid and Invalid conditions, and no significant difference in terms of the ambiguity of the presented shapes.

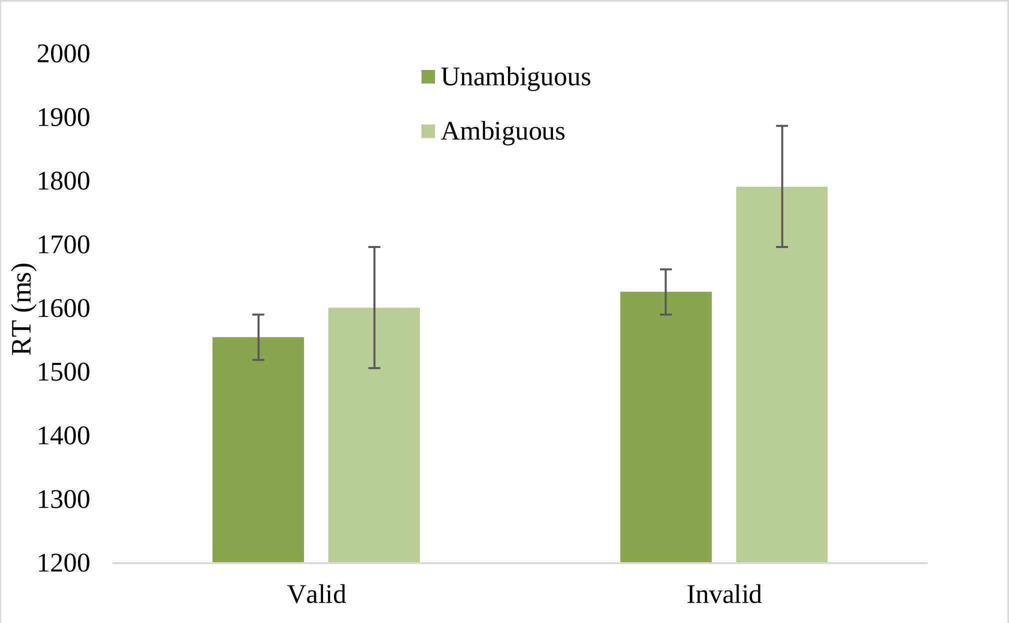

Figure 4.4: Mean RTs for each of the four conditions, for the complete dataset, collapsed across age groups, with standard error bars included

Error Rate. An ANOVA with ambiguity and validity as within-participant factors and age as a between subjects factor found a significant main effect of ambiguity, $F(1, 94) = 14.39$, $p < .001$, $n_p^2 = .13$, but no significant main effect of validity, $F(1, 94) = .38$, $p < .54$, $n_p^2 = .004$, there was also no significant interaction, $F(1, 94) = 1.11$, $p < .29$, $n_p^2 = .01$. When factoring in age as a between subjects factor there was no significant effect of ambiguity, $F(1, 94) = 1.79$, $p < .14$, $n_p^2 = .07$, validity, $F(1, 94) = 1.69$, $p < .16$, $n_p^2 = .07$, or interaction between all factors, $F(1, 94) = .93$, $p < .45$, $n_p^2 = .04$. In other words, when looking at the overall effects of validity and ambiguity in terms of error rate, participants were significantly more likely to produce errors during Ambiguous conditions, when compared with Unambiguous conditions. However, when accounting for age there was no significant effect of validity or ambiguity identified in terms of error rate.

Additional Analysis

As the overall results of Experiment 6 are inconclusive specifically in terms of the different effects of ambiguity and validity in the differing age groups, additional analyses were run on each age group in isolation. This was also due to the notion that significant effects embedded within each age group may have been unidentified due to significant noise and range within the data between age groups. For example, the three-year-olds were averaging around 2700ms, whereas adults were averaging around 840ms before a response was given. For a visual representation of mean RTs for each of the four conditions see Figure 4.5 for the children sample and Figure 9 for the adult sample.

Three Year-Olds

Reaction Time. Figure 4.5 (top left) depicts the mean RT for the four conditions. An ANOVA with ambiguity and validity as within-participant factors did not find a significant main effect of ambiguity, $F(1, 17) = .07$, $p < .79$, $n_p^2 = .004$, but a significant effect of validity was found, $F(1, 17) = 4.52$, $p < .04$, $n_p^2 = .21$. There was also no significant interaction, $F(1, 17) = 1.51$, $p < .24$, $n_p^2 = .08$. Consequently, three-year-olds were significantly faster to respond during Valid conditions in comparison with Invalid conditions.

Error Rate. An ANOVA with ambiguity and validity as within-participant factors found a significant main effect of ambiguity, $F(1, 17) = 5.32$, $p < .03$, $n_p^2 = .24$, but there was no significant main effect of validity, $F(1, 17) = 2.44$, $p < .14$, $n_p^2 = .13$. There was also no significant interaction, $F(1, 17) = 1.4$, $p < .25$, $n_p^2 = .08$. In other words, the three-year-olds were significantly more likely to make errors during Ambiguous conditions, when compared with Unambiguous conditions.

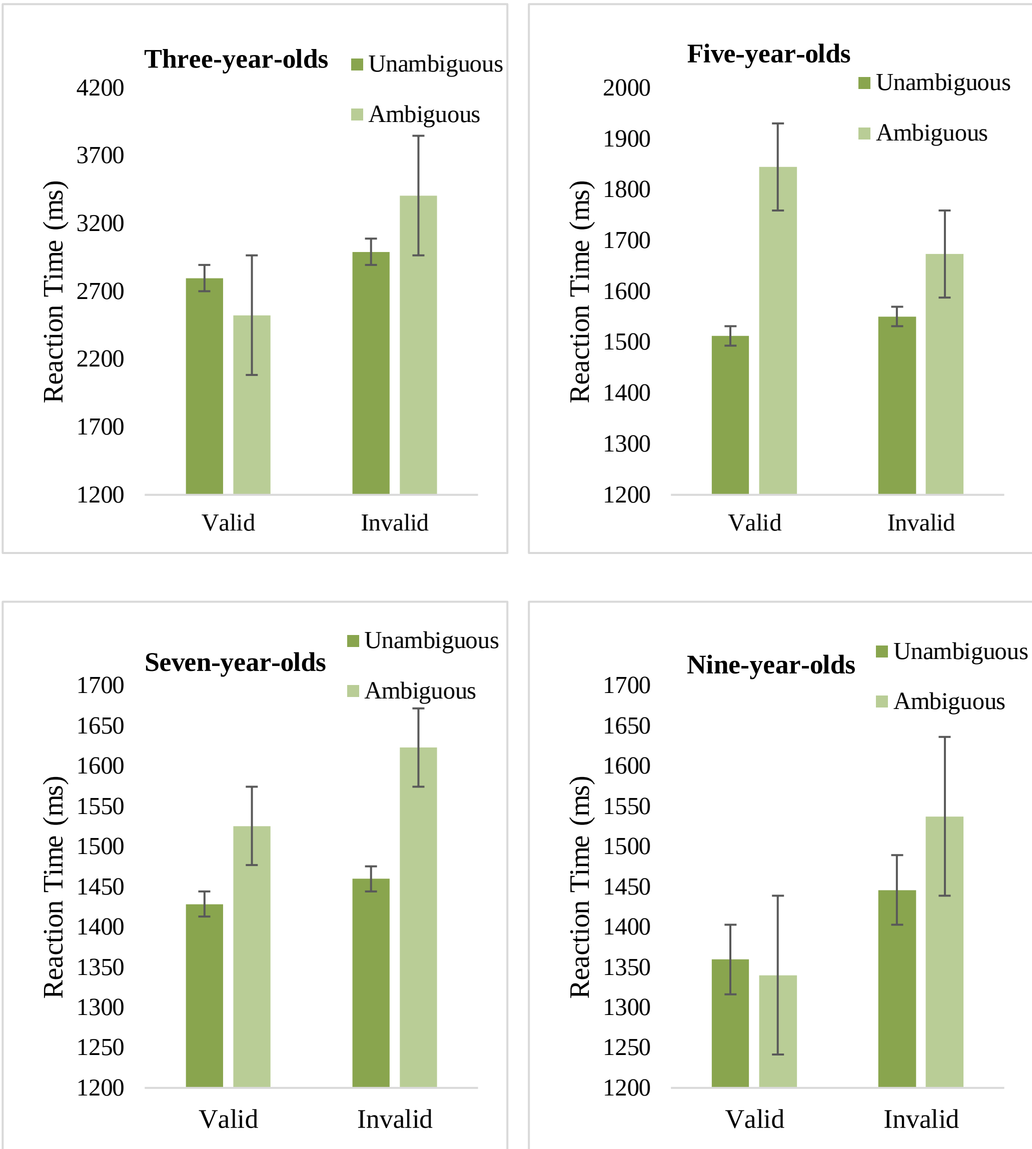

Figure 4.5: Mean RTs for the four conditions, for 3-year-olds (top left), 5-year-olds (top right), 7-year-olds (bottom left) and 9-year-olds (bottom right), with standard error bars included

Five Year-Olds

Reaction Time. Figure 4.5 (top right) depicts the mean RT for the four conditions. An ANOVA with ambiguity and validity as within-participant factors found a significant main effect of ambiguity, $F(1, 20) = 10.95$, $p < .003$, $n_p^2 = .35$, but no significant main effect of validity, $F(1, 20) = .1.79$, $p < .19$, $n_p^2 = .08$. There was also no significant interaction, $F(1, 20) = 2.76$, $p < .11$, $n_p^2 = .12$. Thus, five-year-old participants were significantly faster to respond during Unambiguous conditions when compared with Ambiguous conditions.

Error Rate. An ANOVA with ambiguity and validity as within-participant factors did not find a significant main effect of ambiguity, $F(1, 20) = 3.38$, $p < .08$, $n_p^2 = .15$, or validity, $F(1, 20) = .902$, $p < .35$, $n_p^2 = .04$. There was also no significant interaction, $F(1, 20) = .24$, $p < .63$, $n_p^2 = .01$. In terms of error rate, there was no significant difference in validity or ambiguity for the five-year-olds.

Seven Year-Olds

Reaction Time. Figure 4.5 (bottom left) depicts the mean RT for the four conditions. An ANOVA with ambiguity and validity as within-participant factors found a significant main effect of ambiguity, $F(1, 18) = 3.38$, $p < .01$, $n_p^2 = .05$, but no significant main effect of validity, $F(1, 18) = .93$, $p < .35$, $n_p^2 = .05$. There was also no significant interaction, $F(1, 18) = .25$, $p < .63$, $n_p^2 = .01$. In other words, within the seven-year-old sample, participants were significantly faster to respond during Unambiguous conditions when compared with Ambiguous conditions.

Error Rate. An ANOVA with ambiguity and validity as within-participant factors did not find a significant main effect of ambiguity, $F(1, 18) = .17$, $p < .69$, $n_p^2 = .009$, or of validity, $F(1, 18) = .14$, $p < .72$, $n_p^2 = .007$. There was also no significant interaction, $F(1, 18) = .19$, $p < .76$, $n_p^2 = .01$. Thus, in terms of error rate, there was no significant difference for validity or ambiguity for the seven-year-olds.

Nine Year-Olds

Reaction Time. Figure 4.5 (bottom right) depicts the mean RT for the four conditions. An ANOVA with ambiguity and validity as within-participant factors did not find a significant main effect of ambiguity, $F(1, 20) = .25$, $p < .62$, $n_p^2 = .01$, but a significant main effect of validity was found, $F(1, 20) = 5.87$, $p < .02$, $n_p^2 = .23$.

There was also no significant interaction found, $F(1, 20) = 1.23$, $p < .28$, $n_p^2 = .06$. Thus, meaning that nine-year-old participants were significantly faster to respond during Valid conditions when compared with Invalid conditions.

Error Rate. An ANOVA with ambiguity and validity as within-participant factors did not find a significant main effect of ambiguity, $F(1, 20) = 3.36$, $p < .08$, $n_p^2 = 1.44$, or of validity, $F(1, 20) = 1.00$, $p < .33$, $n_p^2 = .05$. There was also no significant interaction found, $F(1, 20) = 1.00$, $p < .33$, $n_p^2 = .05$. Consequently, in terms of error rate, there was no significant difference in terms of the factors of validity or ambiguity for the nine-year-olds.

Adults

Reaction Time. Figure 4.6 depicts the mean RT for the four conditions. An ANOVA with ambiguity and validity as within-participant factors found a significant main effect of ambiguity, $F(1, 19) = 4.28$, $p < .05$, $n_p^2 = .18$, but there was no significant effect of validity, $F(1, 19) = .99$, $p < .33$, $n_p^2 = .05$. There was also no significant interaction, $F(1, 19) = .06$, $p < .81$, $n_p^2 = .003$. In other words, adult participants were significantly faster to respond during Unambiguous conditions when compared with Ambiguous conditions.

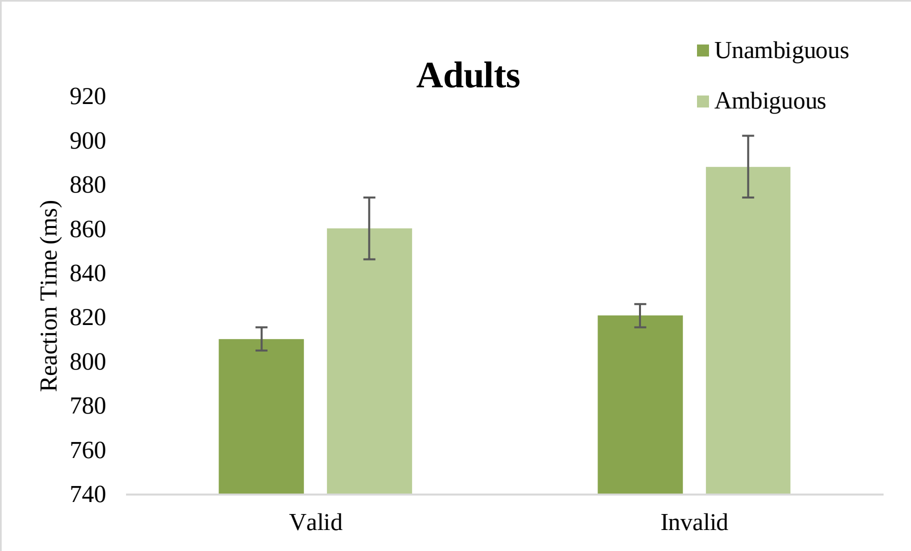

Figure 4.6: Mean RTs for each the four conditions, for the adult sample, with standard error bars included

Error Rate. An ANOVA with ambiguity and validity as within-participant factors found a significant main effect of ambiguity, $F(1, 19)$ =10.89, $p < .001$, n_p^2 = .36, but there was no significant main effect of validity, $F(1, 19)$ = .001, $p <$ .99, n_p^2 = .001. There was also no significant interaction, $F(1, 19)$ = .88, $p < .36$, n_p^2 = .04. Consequently, the adult participants were significantly more likely to make errors during Ambiguous conditions, when compared with Unambiguous conditions.

Overall irrespective of age a significant main effect of validity and ambiguity was found in terms of RT as well as a significant ambiguity effect within error rate. Specifically, participants were significantly faster to response during Valid Unambiguous conditions. Additionally, participants were significantly more error prone during Ambiguous conditions. However, when accounting for age in terms of RT, there was only a significant effect of validity. It could be argued that gaze cuing may have influenced this difference in RT between Valid and Invalid conditions. During Invalid conditions the agent is depicted as gazing towards the side of the image. This may unconsciously draw the attention of participants away from the centre of the image to the side of the stimulus, resulting in slower RTs. However, to assess this assumption further, the validity manipulation should be altered to account for the criticism of directed attention that could arise through gaze cuing.

Furthermore, when isolating the investigation of ambiguity and validity to different ages, the findings are inconclusive. Firstly, three-year-olds were faster to respond during Valid conditions when compared with Invalid, which may have been due to the three-year-olds being more susceptible to gaze cuing. Whereas, five and seven-year-olds were faster to respond during Unambiguous conditions when compared with Ambiguous, which may be due to the early development of the so-called spontaneous visual perspective taking phenomenon. However, nine-year-olds were reported faster to response during Valid conditions when compared with Invalid conditions, with no reported effect of ambiguity. Lastly, adults were more error prone as well as slower to respond to Ambiguous conditions when compared with Unambiguous conditions. Although gaze cuing could be a factor driving these findings, this cannot be concluded, as the results of this experiment are unclear.

As can be seen the results of Experiment 6 are inconclusive. Mixed effect of ambiguity and validity were found within the dataset. Consequently, Experiment 7 was devised. Experiment 7 focuses on investigating ambiguity in isolation within the spontaneous visual perspective taking notion, in respect to child development. This will limit the issues that gaze following may have over the findings obtained.

<u>4.3 Experiment 7 – The Ambiguous Face Task for Adults and Children</u>

4.3.1 Introduction

As Experiment 6 was inconclusive, Experiment 7 was adapted to simplify the manipulation of ambiguity, without the additional validity manipulation. The overall purpose of Experiment 7 was also to assess whether spontaneous visual perspective taking is a process routed within ToM using the same assumption that; if spontaneous visual perspective taking is routed within ToM then it should also be developed within childhood. As with Experiment 6 different ages were assessed. However, in order to (informally) increase Power, only three-year-olds, six-year-olds and adults were sought. Three-year-old participants were sought, as, during this stage in a child's development, ToM should not have developed. Consequently, the lack of spontaneous visual perspective taking should be easiest to identify when compared with older children. Additionally, six-year-old participants were sought as both the five and seven-year-old participants exhibited the same pattern in their responses within Experiment 6. Lastly, adults were sought to ensure that the six-year-olds were not performing significantly different to the adults, as both samples should be exhibiting similar response patterns if spontaneous visual perspective is indeed developed as part of ToM abilities in childhood.

During Experiment 7, participants were required to respond to an image presented on a table in front of them by pressing a matching button as quickly as possible. Again, the image could be interpreted differently depending upon the perspective that the participant assumes, his or her own, or the visual perspective of the experimenter. Participants were asked to complete three blocks of trials, one block with the experimenter completing the same task alongside the participant, next, with the experimenter adjacent to the participant, and lastly, responding with the experimenter's visual perspective, who again was positioned adjacent to the participant. This procedure allowed the comparison of Unambiguous and Ambiguous conditions, as well as assessing the effect of forced assumption of an alternative visual perspective within different ages. It is expected that the three-year-olds will not be affected by the presence of an allocentric visual perspective and only respond using the egocentric self-perspective, even during the forced assumption condition. Whereas it is expected that the six-year-olds and adults will be influenced by the presence of an allocentric visual perspective, which in turn will influence their RT during Ambiguous and forced assumption of Alternative visual perspective conditions.

4.3.2 Method

Participants

Forty-five new participants (29 female) were recruited, 30 participants from a local Primary school, containing a sample of 15 three-year-olds and 15 six-year-olds, and 15 University of Essex students recruited through an online recruitment platform, known as 'SONA'. The distribution of gender for each group can be viewed in Figure 4.7. All participants reported normal or corrected to normal vision. As a means of reimbursement, the university sample received course credits, and the Primary school participants received stickers and a 'Young Scientist Award'. Additionally, the consent procedures for Experiment 7 were a direct replication from those used in Experiment 6.

Age Group	*Total*	*Female*	*Male*
3	*15*	*9*	*6*
6	*15*	*7*	*8*
Adults	*15*	*13*	*2*

Figure 4.7: The distribution of female to male participants for each age group

Stimuli and Apparatus

Two images were used in Experiment 7, which can be observed in Figure 4.8. The images are identical, apart from the orientation; on the left the image depicts a smile, whereas on the right the smile has been rotated 180^{o} to depict a frown. Consequently, the figure is deemed to be Ambiguous as it can be interpreted differently depending on the orientation of the image.

Figure 4.8: The stimuli used in Experiment 7 depicting a smile (left) and frown (right)

The presented images were 886 x 752 pixels in size, with a display ratio of 886 x 752 pixels. Three trial types were used, Unambiguous, Ambiguous and Alternative. For Unambiguous trials the experimenter was sat alongside the participants directly looking at the image presented. This is deemed as Unambiguous as both the self and allocentric experimenter visual perspective are identical. Whereas for Ambiguous trials, the experimenter is positioned opposite to the participant completing the same task, and thus the self and allocentric visual perspective perceive different variations of the image. Lastly, for the alternative trials, the experiment was again positioned adjacent to the participant, but participants were instructed to assume the alternative visual perspective of the experimenter and respond accordingly, which was consequently different to what participant was viewing.

The experiment was conducted using the SuperLab 5.0 desktop app, on a windows 10 laptop in a tablet/flat screen position, with a screen size of 12 inches and resolution of 1366 x 768. Participant response and RT were recorded using an RB-844 cedrus.com response pad, which displayed visual representations for the smile and frown variation. Additionally, the data were analysed using IBM SPSS (Version 21) and R Studio (R Core Team, 2016) statistical software.

Design and Procedure

We employed a mixed method design with ambiguity (Unambiguous, Ambiguous, and Alternative) as a within-participant factor and age (three-year-olds, six-year-olds, adults) as a between-participant factor. In other words, the factors that were manipulated are the experimenters seated positioning for ambiguity, and the age of participants. For Unambiguous manipulations the experimenter was sat alongside the participant, sharing the same visual field of the stimuli, whereas during Ambiguous manipulations the experimenter was positioned adjacent to the participant, and therefore observed a different visual perspective of the stimuli. Lastly, for Alternative ambiguity manipulations, participants were instructed to respond explicitly with what the experimenter was viewing when positioned adjacent to them, thus instructing the participant to assume an alternative visual perspective. For age manipulations, a sample of three-year-olds, six-year-olds and adults over the minimum age of 18 were used.

Firstly, the same procedure for an understanding check that was used in Experiment 6 was used in Experiment 7. Once this was complete each participant’s verbal consent was sought prior to the onset of the rest of the procedure.

60 trials were used in total, 20 for each condition of Unambiguous, Ambiguous and Alternative. During each trial, a blank screen was first presented for

250ms, before the stimulus onset. Participants were required to press the matching button on the response pad as quickly as possible. Once a response was given another blank screen was presented for a further 250ms, signifying the end of the trial. The first two conditions of Unambiguous and Ambiguous were counterbalanced, but the Alternative condition was always presented last. This was to ensure that the Alternative condition did not influence the results of the Unambiguous and Ambiguous conditions.

For the Unambiguous conditions, the experimenter was positioned alongside the participant also completing the same task with a response pad. However, the responses of the experimenter were not recorded. During Ambiguous conditions the experimenter completed the same task with a response pad but adjacent to the participant. Consequently, the experimenter's visual perspective offered an alternative view of the stimuli to the participant. Lastly, prior to the Alternative condition, the participant was invited to the adjacent side of the table, to where they had previously been positioned, to view the stimuli. This ensured that they were able to understand that the view of the image is different depending on positioning. Each participant was then required to sit back in their original seat and complete the final task, by responding to the presented stimuli with what the experimenter could see from their position adjacent to them. Reminder prompts for the participant were given after every five trials in the form of the following statement, "*remember you are pressing what I can see*". Again, in the alternative conditions the experimenter was preforming the same task alongside the participant, as was the procedure for the other conditions. See Figure 4.9 for the positioning of the experimenter and participant for Experiment 7. Once the experiment was complete, participants were reimbursed for their time; ensuring attention was focused on the experiment throughout.

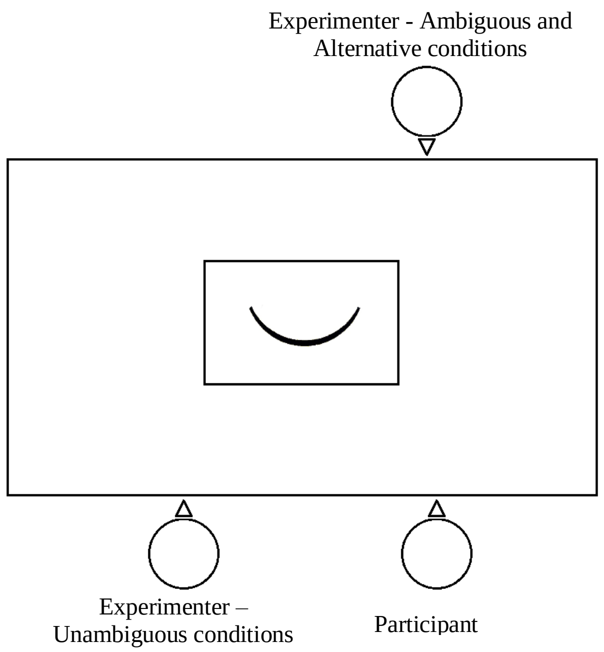

Figure 4.9: The positioning of the experimenter and participant during Experiment 7

4.3.3 Results and Discussion

Data Preparation. As with Experiment 6 the data were collated and reorganised using R Studio (R Core Team, 2016) and collapsed into a single data file. The data were then coded producing mean RTs for Ambiguous, Unambiguous and Alternative conditions. Any outliners that were 2 standard deviations above or below the mean were excluded from the overall analysis. Additionally, percentages score for correct self-perspective responses were also calculated for Ambiguous, and Unambiguous conditions and the alternative perspective for Alternative conditions for analysis investigating error rate. Three participants from the three-year-old sample were excluded from analysis due to a 90% error rate, signifying lack of understanding for the task.

Reaction Time. Figure 4.10 depicts the mean RT for the three conditions of ambiguity in the total sample, as well as the individual samples of three-year-olds, six-year-olds and adults above the minimum age of 18. An ANOVA with ambiguity as a within-participant factor and age as a between-participants factor did not find a significant main effect of ambiguity, $F(2, 78) = .32, p = .73, n_p^2 = .008$, but a significant interaction between ambiguity and age was found, $F(4, 78) = 3.46$ $p < .01$, $n_p^2 = .15$. Post Hoc Tukey's HSD tests identified that the three-year-olds were significantly different to both the six-year-old ($p < .001$) and adult ($p < .001$) samples, but there was no significant difference between the six-year-old and adult samples ($p = .14$). In other words, there was no significant effect of ambiguity within the entire sample. Yet when factoring in age, the three-year-old sample significantly differed in RT to the ambiguity conditions in comparison to the six-year-olds and adults, which did not significantly differ from each other. The effects of ambiguity in respect to the different samples will be extended within the *Additional Analysis* section.

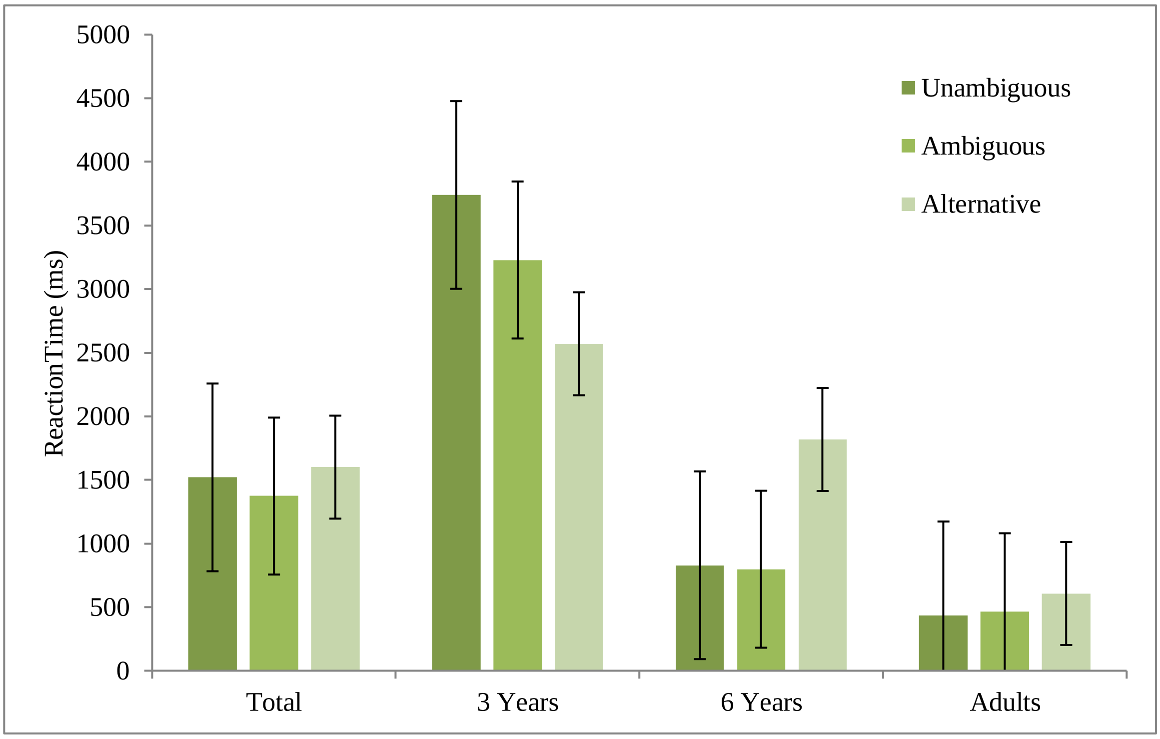

Figure 4.10: Mean RTs for different conditions of ambiguity for the total sample as well as the individual samples of 3-year-olds, 6-year-olds and adults, with standard error bars included

Error Rate. An ANOVA with ambiguity as a within-participant factor and age as a between-participants factor found a significant main effect of ambiguity, $F(2, 78) = 13.44$, $p < .001$, $n_p^2 = .26$, and a significant interaction between ambiguity and age, $F(4, 78) = 2.58$, $p < .04$, $n_p^2 = .12$. Post Hoc Tukey's HSD tests identified that in term of error rate, the three samples were all statistically different from each other (all *p*s < .001).

Consequently, three paired samples *t*-tests were used to form post-hoc comparisons between the three ambiguity conditions within the entire sample. The first paired samples *t*-test indicated that there was no significant difference in error rate between Unambiguous (M=.80, SD=.12) and Ambiguous (M=.78, SD=.20) conditions; $t(1, 41) = .94$, $p = .35$, $d = .17$, $BF_{10} = 0.25$, and is further supported by a Bayesian *t*-test which found the data to be 4 times more likely under the null than the alternative hypothesis. The second paired samples *t*-test indicated that there was a significant difference in error rate between Alternative (M=.62, SD=.31) and Unambiguous (M=.80, SD=.12) conditions, $t(1, 41) = -3.67$, $p < .001$, $d = .58$, $BF_{10} = 43.6$, and is further supported by a Bayesian *t*-test which found the data to be 43.6 times more likely under the alternative hypothesis. The last paired samples *t*-test indicated that there was a significant difference in error rate between Ambiguous (M=.78, SD=.20) and Alternative (M=.62, SD=.31) conditions; $t(1, 41) = 3.56$, $p < .001$, $d = .55$, $BF_{10} = 31.21$. and is further supported by a Bayesian *t*-test which found the data to be 31.2 times more likely under the alternative hypothesis.

These results suggest that ambiguity and age significantly impact error rate. In terms of ambiguity, it was found that forcing the assumption of the alternative visual perspective for participants was the most influential ambiguity manipulation at producing errors. Additionally, as post hoc Tukey tests identified that all ages differed from each other, further analysis upon the specific response pattern for error rate is needed within each sample in isolation.

Additional Analysis

As the initial analysis of ambiguity and age has produced interesting results in terms of RT and error rate, each of the different samples will now be further analysed in terms of each individual patterns of ambiguity.

Three Years-Olds

Reaction Time. Figure 4.10 depicts the mean RT for the three conditions of ambiguity in the three-year-old sample. A one-way ANOVA with ambiguity as a within-participant factor did not find a significant effect of ambiguity, $F(2, 22) = 1.67$, $p = .21$, $n_p^2 = .13$. Thus, in terms of RT, there was no significant difference between the different conditions of ambiguity for the sample of three-year-olds.

Error Rate. A one-way ANOVA with ambiguity as a within-participant factor found a significant effect of ambiguity, $F(2, 22) = 3.98$, $p < .03$, $n_p^2 = .27$. Therefore, three paired samples *t*-tests were used to form post-hoc comparisons between the three ambiguity conditions. The first paired samples *t*-test indicated that there was no significant difference in error rate between Unambiguous (M=.56, SD=.18) and Ambiguous (M=.62, SD=.18) conditions; $t(1, 11) = -.68$, $p = .51$, $d = .23$, $BF_{10} = 0.35$, and is further supported by a Bayesian *t*-test which found the data to be 2.9 times more likely under the null than the alternative hypothesis. The second paired samples *t*-test indicated that there was also no significant difference in error rate between Alternative (M=.39, SD=.23) and Unambiguous (M=.56, SD=.18) conditions, $t(1, 11) = -1.92$, $p = .08$, $d = .56$, $BF_{10} = 1.15$, and is further supported by a Bayesian *t*-test which found the data to be 1.2 times more likely under the alternative hypothesis. The last paired samples *t*-test indicated that there was a significant difference in error rate between Ambiguous (M=.62, SD=.18) and Alternative (M=.39, SD=.23) conditions; $t(1, 11) = 2.18$, $p < .05$, $d = .64$, $BF_{10} = 1.6$, and is further supported by a Bayesian *t*-test which found the data to be 1.6 times more likely under the alternative hypothesis. In other words, in terms of error rate for the three-year-old sample, the only noteworthy finding was the significant difference between Ambiguous and Alternative conditions of ambiguity.

Six Year-Olds

Reaction Time. Figure 4.10 depicts the mean RT for the three conditions of ambiguity in the six-year-old sample. A one-way ANOVA with ambiguity as a within-participant factor did not find a significant effect of ambiguity, $F(2, 28) = 3.18$, $p = .06$, $n_p^2 = .19$. Thus, in terms of RT, there was no significant difference between the different conditions of ambiguity for the sample of six-year-olds.

Error Rate. A one-way ANOVA with ambiguity as a within-participant factor found a significant effect of ambiguity, $F(2, 28) = 8.01$, $p < .002$, $n_p^2 = .36$. Therefore, three paired samples *t*-tests were used to form post-hoc comparisons between the three ambiguity conditions. The first paired samples *t*-test indicated that there was no significant difference in error rate between Unambiguous (M=.84, SD=.16) and Ambiguous (M=.76, SD=.21) conditions; $t(1, 14) = 1.84$, $p = .09$, $d = .47$, $BF_{10} = 1.01$, and is further supported by a Bayesian *t*-test which found the data to be 1 times more likely under the alternative hypothesis. The second paired samples *t*-test indicated that there was a significant difference in error rate between Alternative (M=.53, SD=.32) and Unambiguous (M=.84, SD=.16) conditions, $t(1, 14) = -3.17$, $p < .007$, $d = .81$, $BF_{10} = 7.6$, and is further supported by a Bayesian *t*-test which found the data to be 7.6 times more likely under the alternative hypothesis. The last paired samples *t*-test indicated that there was a significant different in error rate between

Ambiguous (M=.76, SD=.21) and Alternative (M=.53, SD=.32) conditions; $t(1, 14) = 2.57$, $p < .02$, $d = .66$, $BF_{10} = 2.92$, and is further supported by a Bayesian *t*-test which found the data to be 3 times more likely under the alternative hypothesis. In other words, when comparing error rate between different ambiguity conditions in the six-year-old sample, Alternative conditions were significantly different from both the Ambiguous and Unambiguous conditions. This could suggest that the six-year-olds were beginning to comprehend the allocentric visual perspective and are therefore more likely to be influenced by spontaneous visual perspective taking.

Adults

Reaction Time. Figure 4.10 depicts the mean RT for the three conditions of ambiguity in the adult sample. A one-way ANOVA with ambiguity as a within-participant factor found a significant effect of ambiguity, $F(2, 28) = 28.43$, $p < .001$, $n_p^2 = .67$. Therefore, three paired samples *t*-tests were used to form post-hoc comparisons between the three ambiguity samples. The first paired samples *t*-test indicated that there was no significant difference in RT between Unambiguous (M=434.98, SD=68.02) and Ambiguous (M=463.35, SD=84.20) conditions; $t(1, 14) = -1.87$, $p = .08$, $d = .48$, $BF_{10} = 1.06$, and is further supported by a Bayesian *t*-test which found the data to be 1 times more likely under the alternative hypothesis. The second paired samples *t*-test indicated that there was a significant difference in RT between Alternative (M=607.14, SD=126.78) and Unambiguous (M=434.98, SD=68.02) conditions, $t(1, 14) = 5.92$, $p <.001$, $d = 1.53$, $BF_{10} = 677.81$, and is further supported by a Bayesian *t*-test which found the data to be 677 times more likely under the alternative hypothesis. The last paired samples t-test indicated that there was a significant difference in RT between Ambiguous (M=463.35, SD=84.20) and Alternative (M=607.14, SD=126.78) conditions; $t(1, 14) = -5.38$, $p < .001$, $d = 1.39$,

BF_{10} = 294.21, and is further supported by a Bayesian *t*-test which found the data to be 294 times more likely under the alternative hypothesis. As can be seen, for the adult sample, the Alternative condition was significantly different to Ambiguous and Unambiguous conditions in terms of RT. This could be due to the extra effort required of participants to consciously switch between visual perspectives before responding.

Error Rate. A one-way ANOVA with ambiguity as a within-participant factor did not find a significant effect of ambiguity, $F(2, 28) = 2.53$, $p = .09$, $n_p^2 = .15$. Hence, in terms of error rate, there was no significant difference of ambiguity manipulations for the adult sample.

Overall, this experiment has begun to highlight age differences when investigating ambiguity as a manipulation of the spontaneous visual perspective taking notion. In terms of the overall RT, there was no significant main effect of ambiguity, yet, when factoring in age, the three-year-old sample significantly differed in comparison to both the six-year-old and adult samples, which did not significantly differ from each other. Whereas, in terms of error rate ambiguity and age were found to be significant overall, with conditions in which participants were asked to assume the allocentric alternative visual perspective producing the most errors. Additionally, as post hoc tests identified that all ages differed from each other, isolated analysis of the different samples were undertaken. Within this additional analysis the most noteworthy finding was that the adult sample was not found to be influenced significantly by ambiguity in terms of error rate, whereas both the three and six-year-old samples were. However, the lack of a significant main effect of ambiguity in the initial analysis, as well as in the RT analysis of adult participants, questions whether the ambiguous figure implemented was an efficient manipulation of spontaneous visual perspective taking. As the spontaneous visual perspective taking assumption clearly states that RT should be significantly shorter during Unambiguous conditions, in comparison to Ambiguous and Alternative conditions.

4.4 General Discussion

The principal rationale for the present chapter was to assess whether the so-called spontaneous visual perspective taking notion is reliant upon ToM processes, as previous literature has claimed. If this assumption is accurate, a clear difference in performance should be distinguishable between the different ages. Specifically, three and five-year-olds should significantly differ in terms of both their RT and error rate when compared with older children and adults (e.g., Wimmer & Perner 1983; Wellman, et al. 2001). However, the results of Experiment 6 are inconclusive, with different patterns of ambiguity and validity being found between ages, which did not follow a progressively improved performance with increased age. In consequence it was suggested that the findings of Experiment 6 could be due to extraneous variables associated with the design of the experiment that did not lead to the successful assessment of visual perspective taking ability. Thus, Experiment 7 was devised.

Experiment 7 reduced the complexity of the design used in Experiment 6 by eliminating the additional factor of validity, and subsequently focused solely upon measuring differences in RT and error rate under conditions of ambiguity. This is due to the notion that manipulating ambiguity has been reported to be a reliable method in the assessment of spontaneous visual perspective taking (e.g., Surtees et al., 2012; Surtees et al., 2016; Zhao, Cusimano, & Malle, 2015). In terms of RT, Experiment 7 did not find a significant main effect of ambiguity, however when factoring in age, the three-year-old sample significantly differed in comparison to both the six-year-old and adult sample, which did not significantly differ from each other. Whereas, in terms of error rate ambiguity and age were found to be significant overall, with the condition in which participants were asked to assume the alternative visual perspective producing the most errors. However, the lack of a significant main effect of ambiguity in the initial analysis, as well as in the RT analysis of the adult sample, highlights the need to question whether the ambiguous figure used in Experiment 7 was also not an efficient measure of investigating spontaneous visual perspective taking.

In contrast to the current chapter, previous literature used ambiguous numbers such as '*6/9'* in the assessment of ambiguity as a measurement for spontaneous visual perspective taking (Surtees et al., 2012; Surtees et al., 2016; Zhao et al., 2015). Thus, it could be criticised that the ambiguous figures used in the current work may not be as effective at producing the RT differences, in comparison to ambiguous numbers, that has led previous authors to conclude in terms of spontaneous visual perspective taking. As the spontaneous visual perspective taking assumption is clear, when using the ambiguous method, RT should be significantly shorter during Unambiguous conditions, in comparison to Ambiguous and Alternative conditions. The rationale behind using the shape and face variation instead of an ambiguous number was due to the concept of reducing the cognitive effort that is arguably associated with reading a number that the younger participants may have struggled with. Future work would benefit from piloting the standard '*6/9'* variation of ambiguity instead of the face variation in a direct replication of Experiment 7. If the younger participants can understand this, and are able to complete the task, then the additional factor of validity could then also be revisited.

Another issue that could be influencing the results of the current chapter is sample size. Experiment 6 used a sample size of around 20 participants per age group, and Experiment 7 used 15 participants per age group. Unfortunately, due to the nature of this work, gaining access to larger participant samples is difficult, particularly when accounting for year groups and student intake.

Additionally, automaticity and attention are two concerns associated with the literature investigating spontaneous visual perspective taking. It has been debated that measuring a spontaneous process can be challenging, as the investigation should not draw the attention of participants or promote conscious control over the phenomena (Moor & Houwer, 2006). In the current chapter instigating a game-like procedure counteracted this issue. This was particularly beneficial for the diverted but continued attention on the task for the younger participants. However, it also increased the length of the procedure, which may have increased conscious control over the phenomenon for the older participants. Nevertheless, the debate continues as it can still be argued that using any form of repetitive trials in the evaluation of an automatic process could detract from an efficient assessment. Zhao et al. (2015) have provided an example of one procedure that can be used to assess spontaneous visual perspective taking, absent of repetitive trials. In their research, participants were required to complete one trial with a single response. Although, this single response procedure has logistical issues regarding large participant numbers, and the inability to provide an understanding check prior to the onset of the experimental procedure, this is one possible avenue for future work to investigate.

Overall, the current chapter was unable to determine whether spontaneous visual perspective taking is routed within ToM and perspective taking abilities. Experiment 6 did provide initial insight showing that differences within ability in respect to age is apparent, which may be due to cognitive development. Plus, Experiment 7 supported Experiment 6 by signifying that adults significantly differed in performance to the children. However, despite these findings, further investigation is essential in the developmental understanding of spontaneous visual perspective taking.

Chapter Five:

Perceived Ownership

5.1 Chapter Overview

So far, the spontaneous visual perspective taking theory has been examined closely in relation to its theoretical framework and its central assumptions. The following chapter will take an alternative approach and assess whether manipulating aspects of an individual's self-concept will influence the spontaneous visual perspective taking response pattern.

The 'self' is a complex construct that has been examined extensively in terms of cognitive psychology. For the purpose of the current work, the 'minimal' sense of self (Gallagher 2000) will be adopted, which emphasises the ability to distinguish the self from the external environment (Boyer, Robbins & Jack, 2005; Gallagher 2000; Humphrey, 2000; Neisser, 1988) enabling actions and processes to be executed in relation to the body and external world. This distinction between the external world and the self has often been related back to the individual's capacity of performing actions in terms of affordances (Gibson, 1979). Affordances are used to describe the relationship between the external world and an individual's features (e.g., Stoffregen, 2003; Plumert et al., 2004). Understanding this relationship and what influences the likelihood of an action from an allocentric frame of reference upon an object can influence the behaviour of the egocentric frame of reference, in other words, the self. This relationship is also influenced by and actively influences the promotion of ownership over objects within the external world (Fasig, 2000; Ross, 1996). Ownership can be defined as an association between the self and an object (Fiske & Taylor, 1991). On a daily basis the self is exposed to vast quantities of information, including input relating to objects. Consequently, it is imperative that the self can distinguish between objects that are important or owned by the self and those that are not. As the level of attention to objects that are important to the self can be increased and reduce the level of attention to those that are not important. This difference in encoding has often been referred to as the self-reference effect (Beggan, 1992; Belk, 1988, 1991) and has also been found to influence memory (e.g., Cunningham, Turk, Macdonald, & Macrae, 2008).

Currently, spontaneous visual perspective taking has not been examined relative to the influence of other aspects of social interaction or cognition. Consequently, the current chapter will evaluate whether ownership, in terms of the self-reference effect, will in turn influence the spontaneous visual perspective taking theory. It is hypothesised that if spontaneous visual perspective taking is truly spontaneous, (i.e., it is a fast, absent of conscious control, and is stimulus driven) then other constructs, such as ownership in this case, should not influence the phenomenon. However, if ownership in accordance to the self-reference effect were found to influence spontaneous visual perspective taking, the current work would challenge the specific spontaneous assumption of the visual perspective taking theory.

Subsequently, during Experiment 8 distance will be used to manipulate the perceived ownership for the participant over the ambiguous number. This experiment will use a variation of the ambiguous number paradigm that purposely assesses the '*spontaneous*' element of the visual perspective taking notion by utilising a single response from each participant, under either a Near or Far condition. Additionally, Experiment 9 will also use the single response procedure in the investigation of ownership. However, differences within the stimuli will be used to manipulate perceived ownership via distance. Specifically, the stimulus that the ambiguous number will be presented in will either be photographed through an open or closed door, creating a sense of whether the viewer is part of the scene or outside or it. Lastly, Experiment 10 will again assess perceived ownership using the stimuli differences employed in Experiment 9, however instead of the single response procedure, the same procedure that was used in Experiment 1 and 2 that used standard RT measures, will be employed in the assessment of ownership.

5.2 Experiment 8 – Distance and the Single Response Method

5.2.1 Introduction

Firstly, when manipulating an automatic phenomenon, the issue of attention needs to be addressed. In other words, the procedure must not draw the attention of participants to the manipulation, or promote conscious control over the phenomenon, as this would therefore hinder the examination of an automatic process. To reiterate, the manipulation should be hidden to participants, thus ensuring that it remains an automatic process. Consequently, in Experiment 8 a novel procedure utilising a single response was employed. Participants were asked one question about a stimulus in which ambiguous information was presented. Depending on participant response, the visual perspective adopted can be identified. For example, if the participant responds with what they see, an egocentric visual perspective response has occurred, whereas if the participant responds with what the embedded agent can see, the allocentric visual perspective is assumed to have been computed. Additionally, to examine the influence of ownership within this experimental design, the distance between the participant and the stimulus was manipulated. It has previously been found that when asked to lift and place an object on a table, participants placed objects that were owned by the participant closer to themselves, in comparison to objects that were owned by another (Constable, Kritikos & Bayliss, 2011). Consequently, it was assumed that a stimulus that is presented close to the participant would increase the sense of ownership the participant has over the stimulus, compared with a stimulus that is presented further away. In Experiment 8, participants were presented with the stimulus either 60cm, or 5.5 meters away whilst simultaneously being asked, '*what is the number on the table?*' It is expected that if ownership has an influence on spontaneous visual perspective taking, participants will be significantly more likely to respond with an egocentric perspective response when the stimulus is presented closer to them, in comparison to when the stimulus is presented further away. However, this will challenge the spontaneous claim of the visual perspective taking theory, as self-constructs, such as ownership, should not influence the spontaneous visual perspective taking theory, if it is indeed a spontaneous phenomenon.

5.2.2 Method

Participants

One hundred participants were obtained through opportunity sampling within the surrounding community. All reporting normal or corrected to normal vision and consisted of 26 males (74 females) all above the age of 18.

Stimuli and Apparatus

The stimulus is shown in Figure 5.1. As the image shows, the scene depicts an agent sat at a table with an ambiguous number presented in front of them. This number can be interpreted as '*98*' form the viewer's egocentric visual perspective, but '*86*' from the agent's allocentric visual perspective. For both the Near and Far condition, the same stimulus was used and presented in high-resolution colour and laminated on A4 photographic paper. For Near conditions, the stimulus was presented 60cm away from the participant and for Far conditions 5.5 meters.

Figure 5.1: This image was used for both conditions. From the agent's allocentric visual perspective the number is read as '86', whereas from the viewer's egocentric visual perspective the number is read as '98'. Consequently, participant response to the question, 'what is the number on the table?' can be used to determine whether participants adopt the egocentric or allocentric visual perspective.

Design and Procedure

A single-factor design was employed in which the frequency of responses from each visual perspective (egocentric or allocentric) was taken. After providing consent participants were randomly allocated to one condition (either Near, with the stimulus being presented 60cm away, or Far 5.5meters away), using a dice roll. Participants were tested individually and told that they would be shown a photograph that includes a number placed on a table. They were then presented with the photograph and asked, "*What is the number on the table?"* The experimenter recorded the participants' gender and response before providing a debrief and demonstration of the other condition.

5.2.3 Results and Discussion

Percentage (and absolute) responses for egocentric and allocentric visual perspective across the two conditions of Near and Far can be viewed in Figure 5.2. If a participant responded with '*98*' it was deemed as an egocentric visual perspective response, as this is how the ambiguous number is interpreted from the viewer's perspective. Alternatively, if a participant responded with '*86*' it was deemed as an allocentric visual perspective response, as this is how the agent in the stimulus perceives the ambiguous number.

	Egocentric Perspective	*Allocentric Perspective*
Near	*90 (45)*	*10 (5)*
Far	*98 (49)*	*2 (1)*

Figure 5.2: Percentage (and absolute) egocentric and allocentric visual perspective responses across the two conditions

Using the raw frequency of responses, the proportion of egocentric and allocentric visual perspective responses were examined to see if they differed depending on the manipulation of ownership, in the form of distance (i.e., during Near and Far conditions). This analysis did not find a significant difference in proportions, $X^2 = 2.84$, $p = .09$, Cramer's Phi = .09. An additional analysis was run to assess whether gender influenced response, this again did not find a significant difference in proportions, $X^2 = 1.91$, $p = .17$, Cramer's Phi = .17.

It was predicted that presenting the stimulus near to the participant would increase the degree of perceived ownership the participant has over the ambiguous number, compared with presenting the stimulus further away. Thus, it was predicted that the frequency of perspective adopted would be affected. This was not identified, instead no significant result was found in terms of the frequency of visual perspective adopted when distance was manipulated.

It could be argued that the experimental procedure employed, in other words the single response method was driving the non-significant results. This could be due to the possibility that the single response method was unable to obtain significant power to identify the overall effect of spontaneous visual perspective taking. However, the rationale behind the single response method remains clear to specifically assess the spontaneous claim of the so-called spontaneous visual perspective taking phenomenon. Consequently, in the following experiment, the use of the single response method will again be used to examine ownership. Experiment 9 will present stimuli in which the agent and ambiguous number, and the participant's depicted position are either in the Same-Room or a Different-Room. It is predicted that the Same-Room condition will increase the perceived ownership the agent has over the ambiguous number. In other words, if ownership contributes towards the visual perspective taking phenomenon then the allocentric responses should be greater in the Same-Room condition.

5.3 Experiment 9 – Same or Different-Room and the Single Response Method

5.3.1 Introduction

As previously outlined, Constable, et al. (2011) found that when asked to lift and place an object on a table, participants placed objects that were owned by the participant closer to themselves, in comparison to objects that were owned by another. Therefore, in the present experiment ownership was manipulated by presenting the agent and ambiguous number either in the Same or a Different-Room as the participant's suggested position. This was achieved by the camera focusing on a table inside a room through an open or closed door (using the window on the door to show that the door was closed). The stimulus was also manipulated to include an agent, and therefore alternative allocentric visual perspective, or not. Subsequently, using a variation of the single response procedure of Experiment 8, Experiment 9 examined the frequency of adopted perspective for four conditions with an open or closed door together with the presence or absence of an agent.

As predicted during Experiment 8, it is expected that if ownership influences spontaneous visual perspective taking, participants will be significantly more likely to respond with an egocentric response when the ambiguous number is least challenged to be owned by the agent (i.e., when the door is open, absent of an alternative allocentric visual perspective). In addition, it is predicted that the stimulus displaying an agent photographed through a closed door should produce the greatest frequency of allocentric responses, as the scene strongly suggests that the ambiguous number does not 'belong' to the participant, but to the agent. However, if these findings are obtained, this experiment will challenge the spontaneous claim of the visual perspective taking theory, as self-constructs, such as ownership, should have no influence over the visual perspective taking theory, if it is indeed a spontaneous phenomenon.

5.3.2 Method

Participants

Two hundred and forty participants were obtained through opportunity sampling within the surrounding community. All reporting normal or corrected to normal vision and consisted of 85 males (155 females) all above the age of 18.

Stimuli and Apparatus

The four stimuli used for each condition are shown in Figure 5.3. Two conditions depict an agent sat at a table looking at the ambiguous number presented in front of them. Alternatively, the remaining two conditions were absent of the agent, with the same ambiguous number presented. Additionally, two of the stimuli, one with an agent and one without, were taken through an open door, hereafter deemed as Same-Room, and the remaining two through a closed door, hereafter deemed as Different-Room, using the glass in the door to showcase this manipulation.

As with Experiment 8, Experiment 9 used the same ambiguous number for all conditions, i.e., ‘*98*’. For all conditions, the stimulus was presented in colour, high resolution and laminated on A4 paper photographic paper, with approximately 60cm viewing distance.

Figure 5.3: The four conditions used in Experiment 9. Top left, Same-Room Agent Absent, promotes the viewer's perceived ownership of the ambiguous number as the viewer is portrayed as inside the same room as the ambiguous number and there is no allocentric visual perspective challenging ownership. Top right, Different-Room Agent Absent, challenges the viewer's ownership over the ambiguous number by suggesting that they are outside of the room. Bottom left, Same-Room Agent Present challenges the viewer's ownership by presenting an allocentric visual perspective. Bottom right, Different-Room Agent Present reduces the viewer's ownership by the presenting an allocentric visual perspective and suggesting that the viewer resides in a different room.

Design and Procedure

As with Experiment 8, a single-factor design in which the frequency of responses from each visual perspective (egocentric or allocentric) was taken. After providing consent participants were randomly allocated to one of the four conditions, using a dice roll. Participants were tested individually and told that they would be shown a photograph that included a number placed on a table. They were then presented with the photograph and asked, "*What is the number on the table?*" The experimenter recorded the participants' gender and response before providing a debrief, including exposure to the alternative conditions.

5.3.3 Results and Discussion

The percentage (and absolute) responses for egocentric and allocentric visual perspective responses across the four conditions of Same-Room Agent Absent, Different-Room Agent Absent, Same-Room Agent Present, and Different-Room Agent Present can be viewed in Figure 5.4. A participant response of '*98*' was deemed as an egocentric visual perspective response, whereas, a response of '*86*' was interpreted as an allocentric visual perspective response.

	Egocentric Perspective	*Allocentric Perspective*
Same-Room Agent Absent	*85 (51)*	*15 (9)*
Different-Room Agent Absent	*88 (53)*	*12 (7)*
Same-Room Agent Present	*77 (46)*	*23 (16)*
Different-Room Agent Present	*85 (51)*	*15 (9)*

Figure 5.4: Percentage (and absolute) egocentric and allocentric visual perspective responses across the four conditions

Using the raw frequencies of responses, the proportion of egocentric and allocentric visual perspective responses were analysed to see if they differed depending on the manipulated ownership over the ambiguous number. This analysis did not find a significant difference in proportions, X^2 = 3.28, p = .35, Cramer's Phi = .12. Additional analyses were run to assess each condition's specific difference to the baseline condition of Same-Room Agent Absent. In terms of the difference between Same-Room Agent Absent, and Different-Room Agent Absent, there was no significant difference between proportions, X^2 = .29, p = .59, Cramer's Phi = .05. The proportional difference between Same-Room Agent Absent, and Same-Room Agent Present, was also not significant, X^2 =1.35, p = .25, Cramer's Phi = .11. Lastly, there was no difference between Same-Room Agent Absent, and Different-Room Agent Present, as the frequency of responses was identical between the two conditions.

It was predicted that conditions with an allocentric visual perspective and the suggested participant position residing outside of the room as the ambiguous number would challenge the perceived ownership over the ambiguous number significantly. However, this was not found. Instead, it was found that participants predominately adopted the egocentric visual perspective for all of the four conditions that manipulated perceived ownership over the ambiguous number.

As was suggested by Experiment 8, the single response procedure could be driving the non-significant finding. In other words, the single response obtained may not have been able to access sufficient power to identify significance, even with the increased sample size that Experiment 9 used in comparison with Experiment 8. Consequently, Experiment 10 will examine perceived ownership using an alternative approach. Instead of the single response method, standard RT measures using multiple trials will be employed. In other words, the procedure that was used by Experiment 1 and 2 will be used in the assessment of perceived ownership and spontaneous visual perspective taking using the Same, Different-Room manipulation.

5.4 Experiment 10 – Same or Different-Room and Reaction Time Measures

5.4.1 Introduction

Experiments 8 and 9 have been unsuccessful in the investigation of perceived ownership and the spontaneous visual perspective taking theory using the single response method. Consequently, Experiment 10 will use standard RT measures alongside perceived ownership using a variant of the method employed in Experiments 1 and 2. To reiterate, RT will be compared under different manipulations of perceived ownership, alongside the standard ambiguity manipulation. Consequently, the stimuli used in Experiment 9, depicting an ambiguous number under different ownership manipulations were adapted to include both an Ambiguous (i.e., '*68*' and '*89*') and Unambiguous number (i.e., '*69*' and '*88*') variations. It is expected that the findings of Experiment 1 and 2 will be replicated with RTs being significantly shorter during Unambiguous conditions when compared with Ambiguous trials. However, if spontaneous visual perspective taking is truly spontaneous, manipulated ownership should not modulate this effect. However, if RT were found to be shorter during conditions when ownership over the ambiguous number is emphasised for the participant and not the agent, the self-reference effect as previously outlined in *Section 5.1 Chapter Overview* would be supported. Thus, providing evidence to refute the spontaneous claim of the spontaneous visual perspective taking theory.

5.4.2 Method

Participants

Thirty-seven participants were recruited with a mean age of 19.22 (*SD* = 3.63, range = 18-40), with 30 of the sample being female (seven male). Two participants identified themselves as left-handed whilst the remaining 35 identified themselves as right-handed. All participants reported normal, or corrected to normal vision, and were recruited through the University of Essex online volunteer portal known as 'SONA', with participants being reimbursed for their time.

Stimuli and Apparatus

Eight images were used in this experiment. Four of these are shown in Figure 5.5. The other four were identical but with the alternative '*68*' and '*69*' ambiguous number pair. As was with Experiment 1 and 2, the two ambiguous number pairs of '*88*' with '*89'*, and '*68*' with '*69*' were used to counterbalance ambiguity of response to both the left and right-hand, whilst simultaneously accounting for the left-to-right ascending order of the numerals.

Additionally, the presented images were 3264 x 2448 pixels in size, with a display ratio of 816 x 612 pixels. The agent in each stimulus was always presented directly looking at the number on the table. However, for half of the stimuli, the display was positioned through an open door, with no obstruction of the view for the participant, and the remaining stimuli were positioned through a closed door, using a windowpane in the door to view the scene. Consequently, for half of the stimuli the participant's view of the image was obstructed but they were still able to see the ambiguous number and agent. This manipulation aimed to influence the overall sense of ownership over the ambiguous number pairs for participants.

The experiment used the SuperLab 5.0 desktop app, on an Apple Mac with a screen size of 27-inch with the Apple iMac running on version 10.116 with 5120 x 2880 display dimensions. Participant responses and RTs were recorded using a standard keyboard, and the data were analysed using IBM SPSS Statistics (Version 21) computer software.

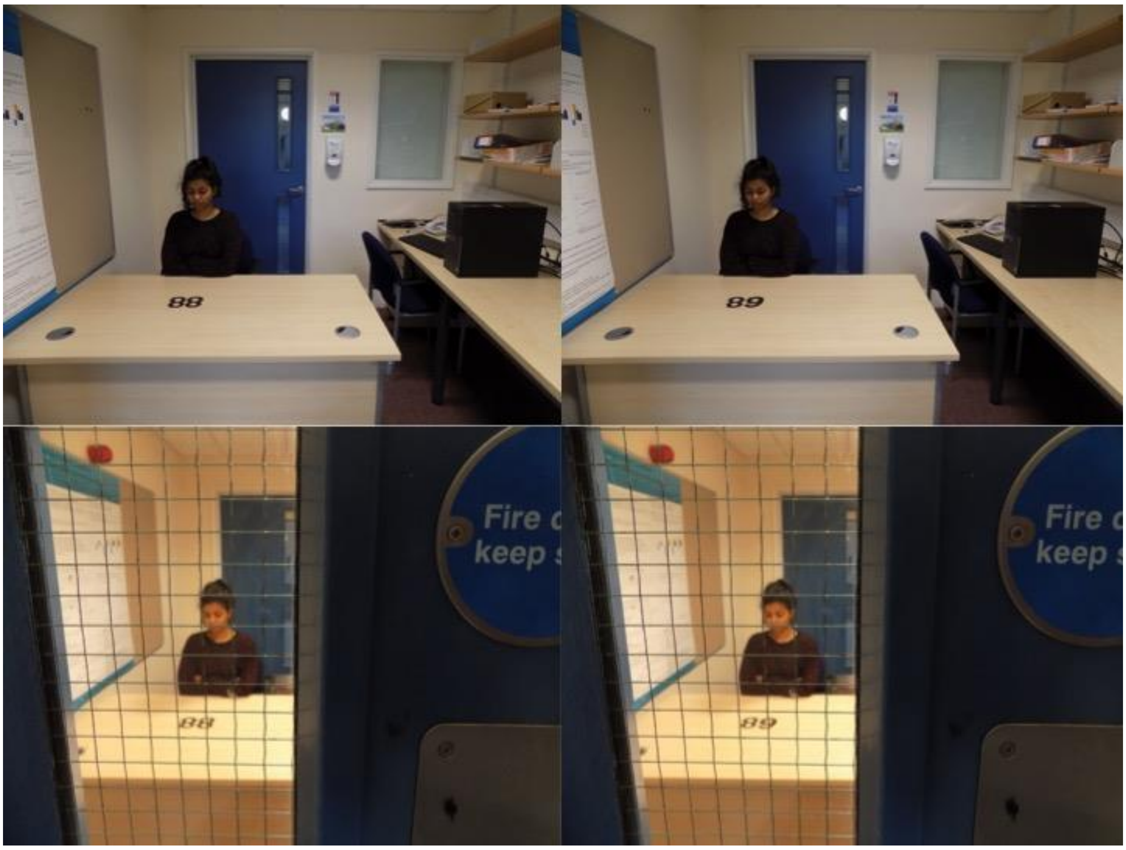

Figure 5.5: The top images represent the critical conditions in the basic ambiguous number paradigm. The left number is Unambiguous in that it is perceived as the same for the viewer and the agent visual perspective (i.e., '88'). In contrast, the number on the right is Ambiguous; it is '89' for the viewer visual perspective but '68' for the agent visual perspective. Additionally, the top images show the Same-Room variation and can be freely viewed through an open door, thus the level of ownership can be questioned. Whereas the bottom images present the Different-Room variation and are obstructed as they are taken through a closed door. Therefore, the perceived ownership of the ambiguous number is strongest for the agent.

Design and Procedure

A 2 x 2 repeated measures design was employed, with the first factor manipulating ambiguity (Unambiguous, Ambiguous) and the second manipulating perceived ownership (Same-Room, Different-Room). In other words, ambiguity relates to the view of the ambiguous number, i.e., whether it can be perceived as the same from each visual perspective (Unambiguous) or differently (Ambiguous). Whereas the level of perceived ownership relates to whether the door in the stimulus was either open, insinuating the participants as being inside the Same-Room as the ambiguous number, or shut, insinuating the participants as being part of a Different-Room.

As with Experiment 1 and 2, participants were required to respond to two numbers displayed on a computer monitor, which were embedded in different stimuli, using a standard computer keyboard. Participants began with their hands positioned upon the two response keys at the start of each block. The '*Z*' key was used to indicate a left-hand response corresponding to the numbers '*68'* or '*88*', and the '*M*' key was used for right-hand response, corresponding to the numbers '*69*' or '*89*'.

During each block the stimulus that depicted the different conditions, be that Same-Room or Different-Room remained the same throughout. At the start and end of each trial, the same stimulus was also displayed absent of a target, so that each trial was easily differentiated. 250 milliseconds after the onset of the neutral image, the target was presented. The trial ended only after the participant responded using the respective key. At the end of the trial the neutral image was again shown for 250 milliseconds prior to the onset of the next trial.

Four blocks were used in total. Two blocks used the '*68'* and '*69'* ambiguous number pair, with a Same and Different-Room variation, which was repeated with the second ambiguous number pair of '*88*' and '*89*'. Each block presenting 96 trials resulting in a total number of 384 trials altogether.

5.4.3 Results and Discussion

Data Preparation. As with previous experiments, the data were collated and reorganised using Microsoft Excel and collapsed into a single data file. The data were then coded, producing mean RTs for the different factors of ambiguity and perceived ownership. Any outliners that were 2 standard deviations above or below the mean were excluded from the overall analysis. In addition, percentages for correct egocentric visual perspective responses were calculated for each condition so that analysis of error rate could take place.

Reaction Time. Figure 5.6 illustrates the mean RT for the four conditions. A repeated measures ANOVA with ambiguity and perceived ownership as within-participant factors found a significant main effect of ambiguity, $F(1, 36) = 15.32$, $p < .001$, $n_p^2 = .29$, but no significant main effect of ownership, $F(1, 36) = 3.21$, $p = .08$, $n_p^2 = .08$. There was also no significant interaction, $F(1, 36) = 2.37$, $p = .13$, $n_p^2 = .06$. As can be seen, participants were significantly faster to respond during Unambiguous conditions, in comparison to Ambiguous conditions, regardless of perceived ownership.

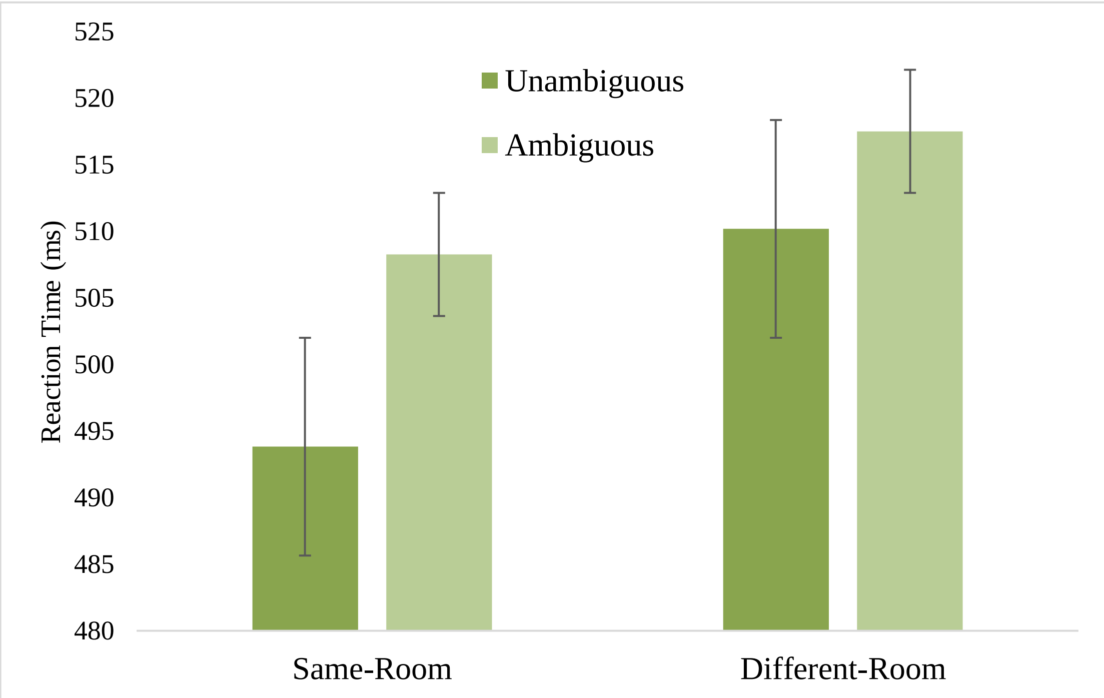

Figure 5.6: Mean RTs for each of the four conditions with standard error bars included

Error Rate. A repeated measures ANOVA with ambiguity and perceived ownership as within-participant factors did not find a significant main effect of ambiguity, $F(1, 36) = 1.4$, $p = .24$, $n_p^2 = .04$, or of perceived ownership, $F(1, 36) =$

.26, p =.61, n_p^2 = .007. There was also no significant interaction, $F(1, 36) = .49$, p = .49, n_p^2 = .01. Thus, there was no significant finding in terms of error rate.

To reiterate, the significant ambiguity finding is suggestive that participants were assuming the visual perspective of the agent, as previous literature has suggested. This is supported by RTs being significantly shorter during Unambiguous trials, when the visual perspective of the participant and agent are identical, in comparison to Ambiguous trials where the visual perspectives differ. Interestingly, the same pattern in RT was found irrespective of perceived ownership, be that the Same or Different-Room variations. Subsequently, Experiment 10 does not support the idea that manipulating perceived ownership influences the likelihood of participants spontaneously assuming the visual perspective of the agent.

Further Analysis

As with Experiments 1 and 2, further analysis was undertaken to assess whether the trend occurs in both of ambiguous number pairs.

'68' – '69' Ambiguous Number Pair

Reaction Time. Figure 5.7 illustrates the mean RT for the four conditions, using the '*68*' and '*69*' ambiguous number pair. A repeated measures ANOVA with ambiguity and perceived ownership as within-participant factors found a significant main effect of ambiguity, $F(1, 36) = 22.87$, $p < .001$, $n_p^2 = .39$, but there was no significant main effect of perceived ownership, $F(1, 36) = 1.06$, $p = .31$, $n_p^2 = .03$. Interestingly there was a significant interaction, $F(1, 36) = 5.64$, $p = .02$, $n_p^2 = .14$.

To examine the significant interaction three planned comparison t-tests were conducted. Firstly, there was a significant difference between Same-Room, Unambiguous (M=519.74, SD=56.09) and Same-Room, Ambiguous (M=502.22, SD=62.28) conditions; $t(1, 36) = 3.22$, $p = .003$, $d = .53$, $BF_{10} = 12.83$. Consequently, it was highlighted that participants were significantly faster to respond to ‘*68*’ in comparison to ‘*69*’ within the Same-Room condition. This is further supported by a Bayesian *t*-test which found the data to be 13 times more likely under the alternative hypothesis. There was also a significant difference between Different-Room, Unambiguous (M=539.57 SD=93.77) and Different-Room, Ambiguous (M=510.22, SD=88.75) conditions; $t(1, 36) = 5.29$, $p < .001$, $d = .87$, $BF_{10} = 3221.47$. Again, it was highlighted that participants were significantly faster to respond to ‘*68*’ in comparison to ‘*69*’ during Different-Room conditions. This is further supported by a Bayesian *t*-test which found the data to be 3221 times more likely under the alternative hypothesis. Interestingly this ambiguity effect was found to be greatest during the Different-Room variations. Additionally, there was no significant difference between Same-Room, Ambiguous (M=502.22 SD=62.28) and Different-Room, Ambiguous (M=510.22, SD=88.75) conditions; $t(1, 36) = -0.56$, $p = .56$, $d = -.09$, $BF_{10} = 0.21$, which was further supported by a Bayesian *t*-test that found the data to be 4.9 times more likely under the null than the alternative hypothesis.

Consequently, it was highlighted that participants were significantly faster to respond during trials using the ‘*68*’ numeral compared with ‘*69*’. However, most importantly there was no significant difference in RT between Ambiguous Same-Room and Ambiguous Different-Room conditions. Yet, in terms of the overall effect of perceived ownership, RT was not significantly influenced.

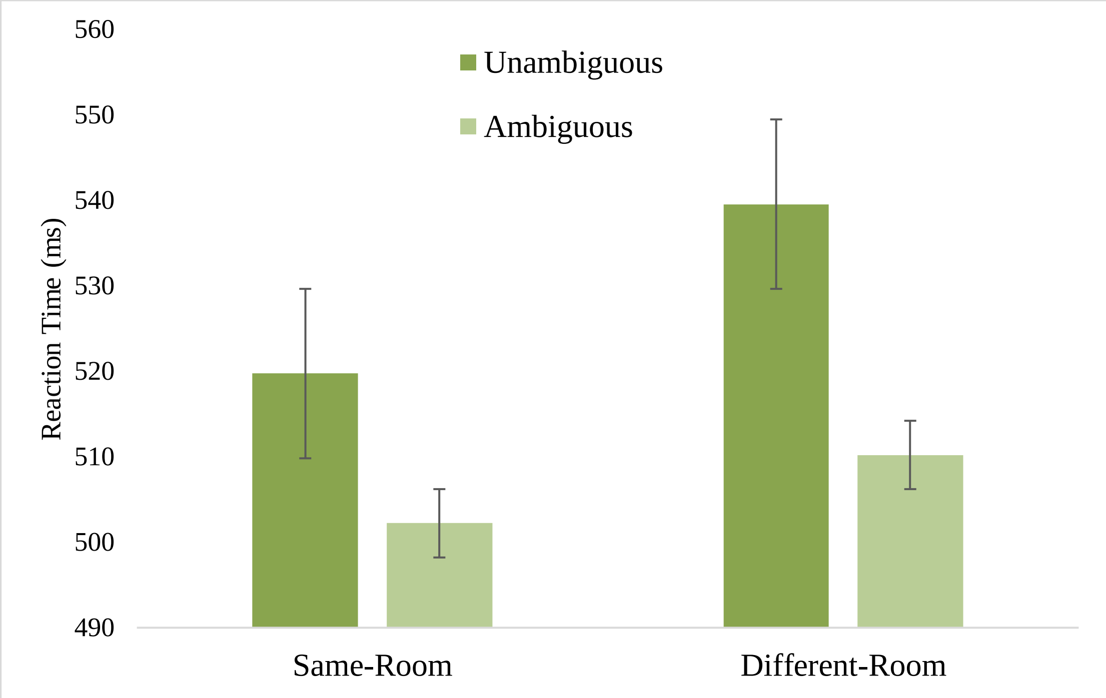

Figure 5.7: Mean RTs for each of the four conditions for the '68' and '69' ambiguous number pair, with standard error bars included

Error Rate. A repeated measures ANOVA with ambiguity and perceived ownership as within-participant factors did not find a significant main effect of ambiguity, $F(1, 36) = .24$, $p = .63$, $n_p^2 = .007$, or of perceived ownership, $F(1, 36) = .26$, $p = 1.89$, $n_p^2 = .05$. There was also no significant interaction, $F(1, 36) = .13$, $p = .72$, $n_p^2 = .004$. In sum, there was no significant finding in terms of error rate when isolating the '*68*' and '*69*' ambiguous number pair.

'88' – '89' Ambiguous Number Pair

Reaction Time. Figure 5.8 illustrates the mean RT for the four conditions, using the '*88*' and '*89*' ambiguous number pair. A repeated measures ANOVA with ambiguity and perceived ownership as within-participant factors found a significant main effect of ambiguity, $F(1, 36) = 64.41$, $p < .001$, $n_p^2 = .64$, but no significant main effect of perceived ownership, $F(1, 36) = 3.25$, $p = .08$, $n_p^2 = .08$. There was also no significant interaction, $F(1, 36) = .10$, $p = .75$, $n_p^2 = .003$. Therefore, participants were significantly faster to respond during trials that used the '*88*' numeral in comparison to the '*89*', irrespective of the perceived level of ownership.

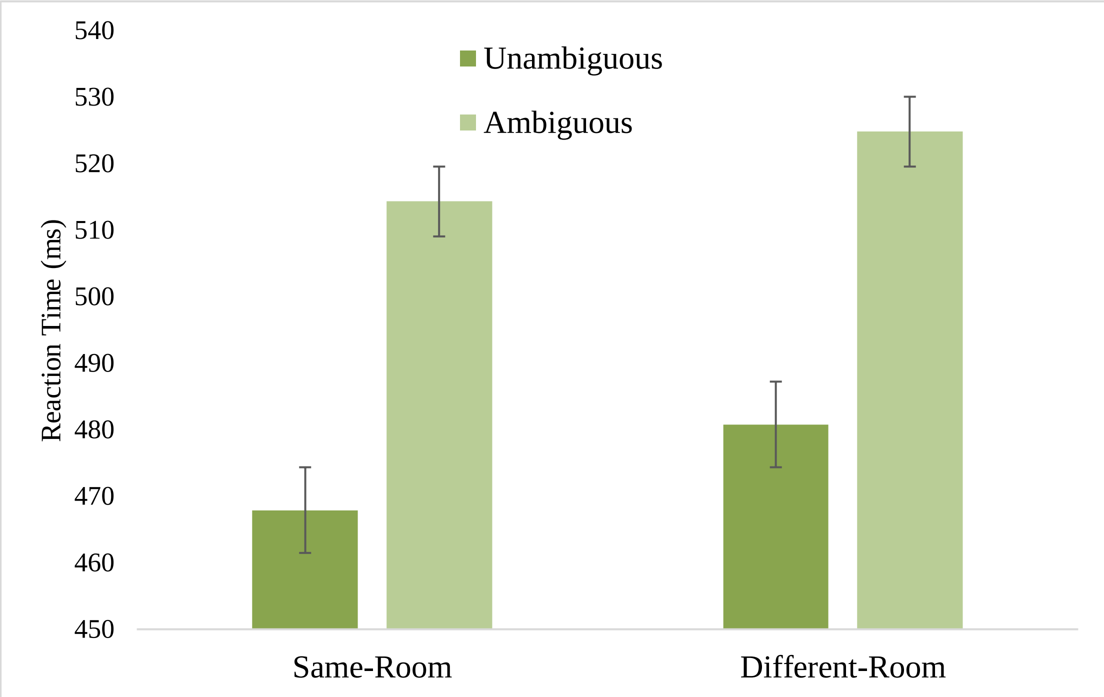

Figure 5.8: Mean RTs for each of the four conditions for the '88' and '89' ambiguous number pair, with standard error bars included

Error Rate. A repeated measures ANOVA with ambiguity and perceived ownership as within-participant factors did not find a significant main effect of ambiguity, $F(1, 36) = 2.49$, $p = .12$, $n_p^2 = .07$, or of perceived ownership, $F(1, 36) = .63$, $p = .44$, $n_p^2 = .02$. There was also no significant interaction, $F(1, 36) = .37$, $p = .55$, $n_p^2 = .01$. Thus, there was no significant finding in terms of error rate when isolating the '*88*' and '*89*' ambiguous number pair.

The results of these additional analyses have found that in regard to the '*88*' and '*89*' ambiguous number pair, the same patterns in the data were exhibited. However, when isolating the '*68*' and '*69*' ambiguous number pair the opposite effect was found. RTs were shortest during Ambiguous trials in comparison to Unambiguous trials, irrespective of the perceived ownership. It can be determined that this effect was weaker than the ambiguity effect identified in the '*88'* and '*89*' pair, as the overall ambiguity effect overrode the effect identified in the '*68*' and '*69*' ambiguous number pair when the dataset was investigated in its entirety. One possible explanation of this counter-effect is the influence of spatial hand mapping. Participants may be faster to response for either the left or right-hand responses overall. This validates the use of the two ambiguous numbers pairs and ensures that any effects of spatial hand mapping are counterbalanced, and reduces the chances that spatial hand mapping will mask any ambiguity or ownership effects that may be present.

Overall, the present experiment replicated the ambiguous number effect; RTs were significantly shorter during Unambiguous trials in comparison to Ambiguous trials. However, this was irrespective of perceived ownership the participant has over the ambiguous number. Consequently, the results of Experiment 10 do not support the concept that manipulations to perceived ownership over an ambiguous number influences the spontaneous visual perspective taking theory.

5.5 General Discussion

Overall, Experiment 8 predicted that if ownership influences the spontaneous visual perspective taking effect, participants would be significantly more likely to respond with an egocentric visual perspective response when the stimulus was presented closer to them. However, this would challenge the spontaneous claim of the spontaneous visual perspective taking theory, as self-constructs such as ownership, should not influence the visual perspective taking theory, if it is indeed a spontaneous phenomenon. Nonetheless, this finding was not identified. Instead, no significant result was found in regard to the frequency of perspective adopted in respect to the distance between the stimulus and participant. Next, Experiment 9 predicted that if perceived levels of ownership influence the spontaneous perspective taking theory, participants would be significantly more likely to respond with an egocentric visual perspective response when the ambiguous number is least challenged to be owned by the agent. In Experiment 9, this was when the door was open, and absent of an alternative perspective. Additionally, it was predicted that the stimulus displaying an agent and photographed through a closed door would produce the greatest frequency of allocentric visual perspective responses, as the scene strongly suggested that the ambiguous number does not 'belong' to the participant. However, Experiment 9 found that participant responses were not affected by manipulations to the perceived ownership.

It is worth noting that the method adopted by Experiment 8 and 9 can be argued to impact the likelihood of identifying a significant effect. In other words, the single response method may be unable to access sufficient power in order to identify an overall effect of spontaneous visual perspective taking. However, the rationale behind the single response method is clear in its specific assessment of investigating the spontaneity element of the so-called spontaneous visual perspective taking phenomenon.

Additionally, using a traditional RT based approach alongside the ambiguous number paradigm; Experiment 10 found that RTs were significantly shorter during Unambiguous conditions in comparison to Ambiguous conditions; irrespective of the perceived level of ownership over the ambiguous number. Consequently, again the results of Experiment 10 do not support the idea that perceived ownership over an ambiguous number influences the likelihood of participants spontaneously assuming the visual perspective of an agent.

However, it should also be acknowledged that Experiment 9 and 10 can be criticised as the Different-Room stimuli are harder to see and perceive through the windowpane in the door. Therefore, the effects identified may be a result of the participants impaired visibility and not due to the experimental manipulations. This is due to the focus of the camera being aimed at the door, not at the scene within the room. However, this positioned was used in order to maximise the emphasis on the participant being in a different room to the ambiguous number and agent. In order to assess whether this stimulus characteristic is influencing the results of Experiment 9 and 10, future experiments using the Same, Different-Room manipulation should superimpose the door, absent of the windowpane onto the Same-Room variation. To reiterate, the same stimulus would be used, however, the Different-Room variation would have the overlay of the door without the windowpane hindering the observers view of the scene.

All things considered; the current chapter addressed the question of whether manipulating ownership over ambiguous information impacts the spontaneous visual perspective taking theory. Although the present experiments have been unable to report an effect of manipulated ownership, the chapter has advanced the spontaneous visual perspective taking methodological. Specifically, this chapter has begun to challenge what exactly it means to assess a spontaneous process and how the experimental design can influence the onset of a spontaneous phenomenon. It was highlighted that spontaneous investigations should ensure that attention is not drawn to the task or promote conscious control over the phenomenon. This consequently led to the development of the single response method. Thus, this procedure will be revisited in *Chapter 6: Attribution of Vision and Knowledge in 'Spontaneous Perspective Taking'* which will again examine the application of occluding barriers to the spontaneous visual perspective taking theory. Furthermore, an alternative explanation that can be used to account for the so-called spontaneous visual perspective taking effects will also be introduced.

Chapter Six:

Attribution of Vision and Knowledge in 'Spontaneous Perspective Taking'

6.1 Chapter Overview

The present thesis has so far described ten experiments that have examined the spontaneous visual perspective taking theory. Overall, the data do not concur with the notion that humans spontaneously represent the viewpoint of others. After a further attempt to show an effect of another person's visual perspective (Experiment 11), this final empirical chapter will present an alternative account to the visual perspective taking explanation. This explanation will then be tested in two final experiments (Experiments 12 and 13).

Experiment 11 examined the spontaneous visual perspective taking theory by incorporating the (single response) ambiguous number paradigm (i.e., Surtees, Samson, & Apperly 2016; see also Surtees, Butterfill, & Apperly, 2012; Zhao, Cusimano & Malle, 2015). Recall that participants are presented with a display, including a number that is ambiguous in the sense that it is different depending on which position it is viewed from. For instance, the number positioned on the table in Figure 6.1 is '*9*' from the viewpoint of the participant but '*6*' from the viewpoint of the agent. Surtees, et al. (2016) found that RTs to determine the identity of the number were longer under this situation of ambiguity compared to when the number was positioned such that its identity was the same for both the participant and agent. A further application of the ambiguous number technique was undertaken by Zhao et al. (2015). As reproduced in Figure 6.1, Zhao, et al. (2015) presented participants with images that included a person, facing the observer, looking at a number placed on a table. Observers were simply asked *"What number is on the table?"*.

Figure 6.1: The ambiguous number paradigm.

One particularly interesting aspect of this design is that it is a direct measure of potential spontaneous visual perspective taking, as opposed to the indirect attentional measures generated via RTs in the other paradigms. Zhao et al. found that 42% of observers judged the number from the viewpoint of the actor in the display (i.e., stated '*6*'). This contrasted with responses when no actor was present, in which case all observers judged the number from their own egocentric visual perspective. The fact that data suggestive of spontaneous visual perspective taking has come from a very different paradigm to other methods does provide convergent evidence for the theory. The second aim of the present chapter was to assess an alternative explanation as to why spontaneous visual perspective taking effects occur.

In the present Experiment 11, a close replication of the Zhao et al. procedure was undertaken, and a condition was added in which the model could not see the number due to the position of a physical barrier. Specifically, she held a newspaper occluding its view. In Experiments 12 and 13 we manipulated this number judgement paradigm to test the alternative account of the phenomenon.

6.2 Experiment 11 – Occluding Barriers and the Single Response Method

6.2.1 Introduction

In the present experiment we adopted the basic procedure employed by Zhao et al. (2015; Experiment 1) in which a model/actor sits at a table facing a number placed upon it. Participants are asked what the number is. In our experiment we presented (independent) participants with one of three conditions of an office scene (see Figure 6.2). In the 'Gaze' condition, the actor looks directly at the number. In the 'Barrier' condition, the actor looks towards the number but cannot see it because a newspaper she is holding occludes it. As with Zhao et al. we also included a no-actor condition as a Baseline.

6.2.2 Method

Participants

One hundred and fifty-one participants took part, obtained through opportunity sampling within the surrounding community, all reporting normal or corrected to normal vision. The sample consistent of 33 males, (118 females) all above the age of 18.

Stimuli and Apparatus

The stimuli are shown in Figure 6.2. As the images show, the scene was designed to look like a 'hot desk' located at one end of a rectangular room. One number was always used, being either '*1801*' or '*1081*' depending upon the position it could be viewed. The images were presented in colour, in high resolution, and laminated on A4 paper.

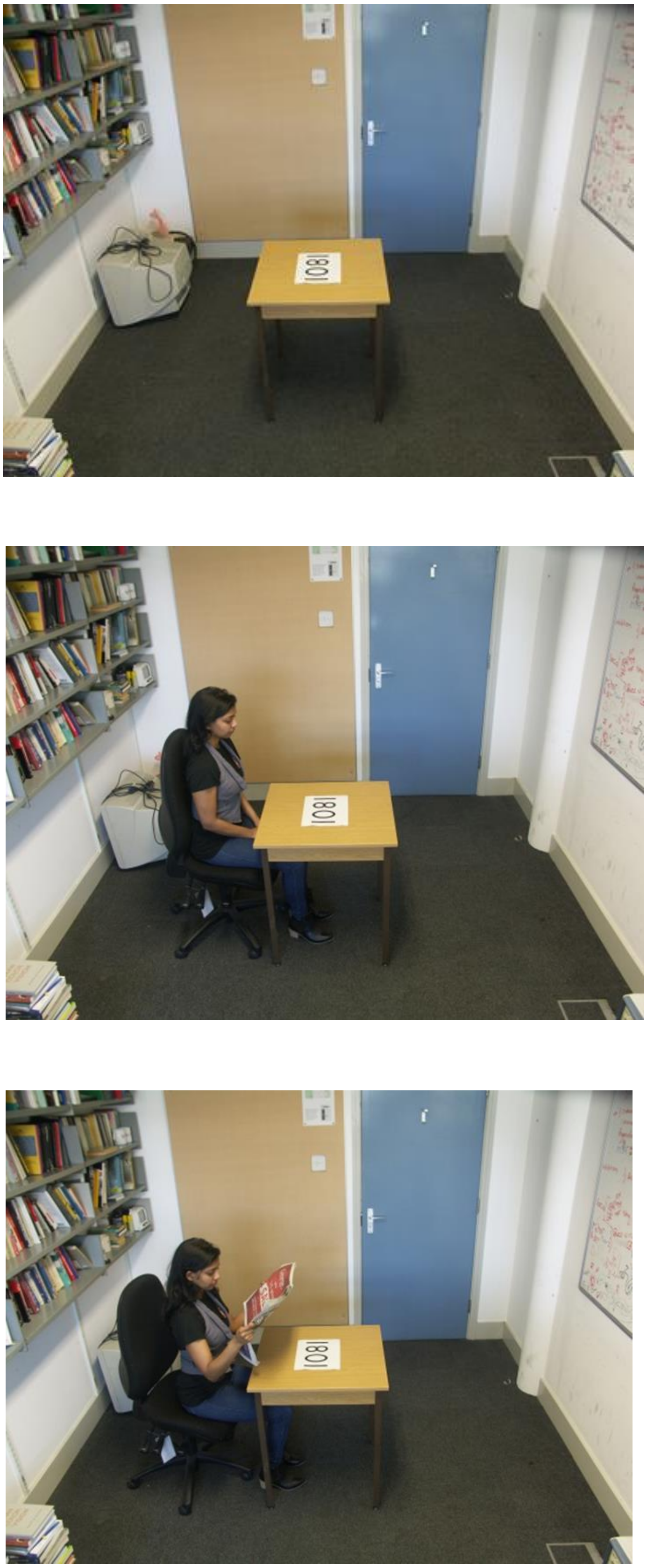

Figure 6.2: Stimuli used in Experiment 11. The upper panel shows the no-actor Baseline, the middle panel shows the Gaze condition, and the lower panel shows the Barrier condition.

Design and Procedure

A single-factor design was employed in which the frequency of responses from each visual perspective (left and right of the table) was taken. After providing verbal consent participants were randomly allocated to one condition, using a dice roll. Participants were tested individually and told that they would be handed a photograph that includes a number placed on a table. They were then presented with the photograph and asked, "*What is the four-digit number on the table?*" The experimenter recorded the participants' gender and response before providing a debrief.

6.2.3 Results and Discussion

The response of one (female) participant was omitted due to a time delay of more than two minutes before a response was given. Figure 6.3 presents responses, converted to percentages, pertaining to the left and right perspective of the number. If a participant responded "*1801*", this was considered as a response from the left perspective.

	Left Perspective	*Right Perspective*
Baseline	*52 (26)*	*48 (24)*
Gaze	*86 (43)*	*14 (7)*
Barrier	*78 (39)*	*22 (11)*

Figure 6.3: Percentage (and absolute) left and right perspective responses across the three stimulus conditions

Using the raw frequencies, we examined whether the proportions of left and right perspective responses were different across the three stimulus conditions (i.e., Baseline, Gaze and Barrier). This analysis found a significant difference in proportions, X^2 = 15.68, p < .001, Cramer's Phi = .323. An additional analysis found that the proportions of left and right perspective responses for the Gaze and Barrier conditions were not significantly different from each other $X^2 = 1.1$, $p > .29$.

These results show that when no actor was present, the number was highly ambiguous in that approximately half of the participants considered it from the left (i.e., *1801*) and half from the right (i.e., *1081*). This can be contrasted with the Gaze condition in which 86% of participants considered the number from the left. This on its own is highly suggestive of spontaneous visual perspective taking since this position corresponded with the viewing position of the actor. However, when the actor could not see the number, 78% of observers again judged the number from the leftward perspective. Since the number was not visible to the agent this cannot be due to the actor’s visual perspective being taken. These results therefore concur with those of previous experiments using the barrier method (Cole et al. 2015; 2016); a perspective-taking-like pattern of data has been found even when the model in the display cannot see the same critical stimuli as the participant. This in turn does not support the spontaneous visual perspective taking theory.

Although one does not need a theory as to why a phenomenon does *not* occur, there is still the issue of why visual perspective-taking-like data are observed in the present Experiment 11 and in previous studies. In Experiments 12 and 13 and further in *6.5 General Discussion,* we sketch out an alternative theory and test one of its assumptions.

Additionally, although she could not see the number, the actor in the gaze condition of Experiment 11 had an effect on responses; results were clearly driven by her presence. To explain this and other visual perspective taking-like effects, we suggest a model in which two distinct stages/processes operate to generate the phenomenon. First, when an observer is presented with the displays typical of spontaneous visual perspective taking studies, the seeing agent *acts as a reference point*, or *anchor*. Information gained from this object/position is then used for the second stage. In this latter stage, a mental rotation process occurs, or something akin to mental rotation, in which the observer orients to the approximate position of the agent and assumes the direction of the agent. The critical stimuli are then represented from this position/side. As a consequence, this representation is afforded primacy over alternative representations, leading to, for instance, facilitated RTs. One of the central assumptions of this explanation is that *any stimulus acting as the directional reference point* will induce the same effect as when an agent does. This was tested in Experiment 12 and 13.

<u>6.3 Experiment 12 – A Test of the Reference Point Theory</u>

6.3.1 Introduction

The basic procedure of Experiment 11 was again employed for Experiment 12. i.e., participants were presented with a photograph and asked to decide what number was placed on the table. The same Baseline was used and compared with two other conditions both of which included directional reference points (see Figure 6.4). In one of these conditions, the table included items typical of a hot desk. In the other condition, the table abutted one of the room's walls. Thus, the hot desk items provide the reference point in the former condition, and the room itself may be said to provide the reference point in the latter.

6.3.2 Method

All aspects of the Method were as described for Experiment 11 with the following exception. A new sample of 151 participants (84 females) took part.

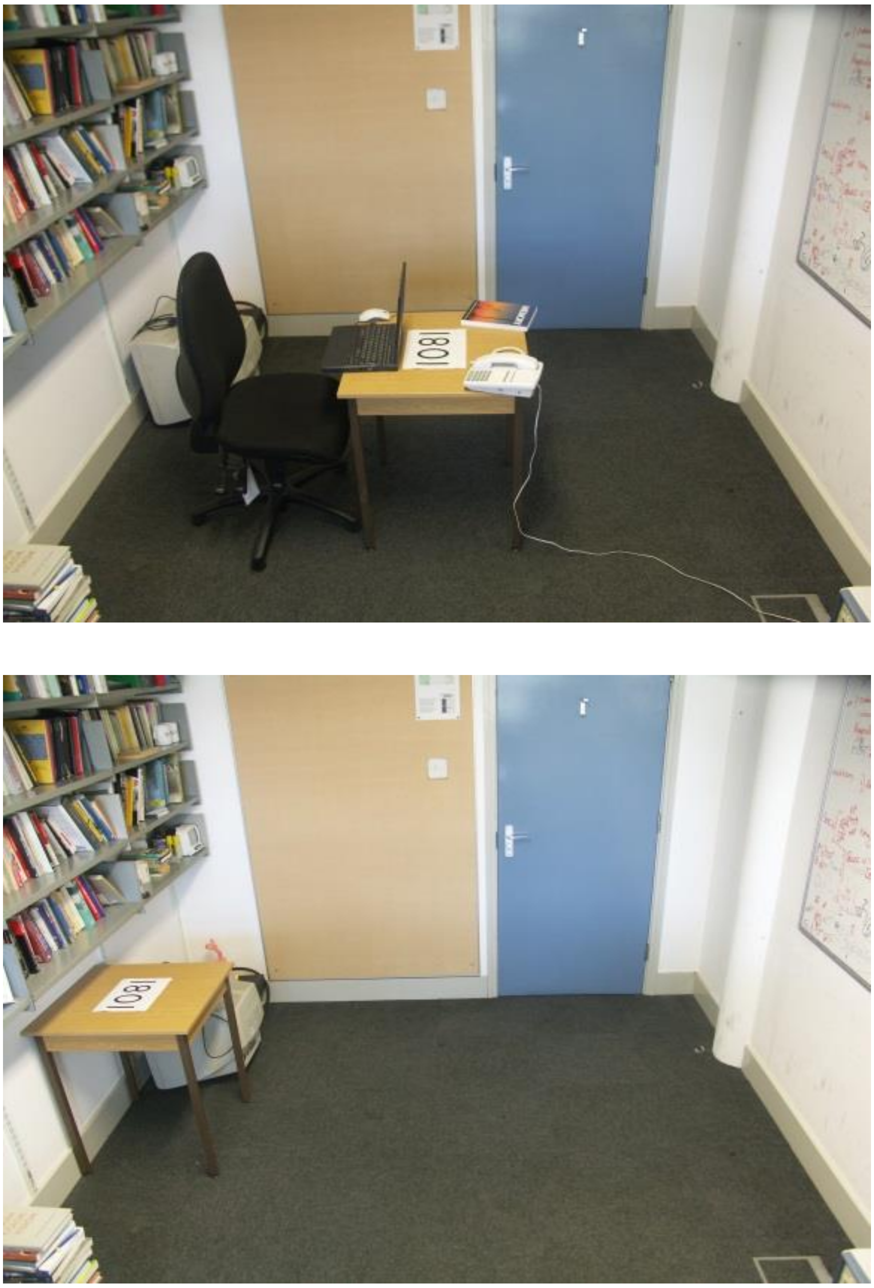

Figure 6.4: Stimuli used in Experiment 12. We also included the same Baseline condition as was used in Experiment 11, but not shown here.

6.3.3 Results and Discussion

One female participant's response was again omitted as a result of a time delay exceeding two minutes. Figure 6.5 presents percentage responses pertaining to left and right visual perspectives of the number.

	Left Perspective	*Right Perspective*
Baseline	*54 (27)*	*46 (23)*
Hot Desk	*76 (38)*	*24 (12)*
Wall	*6 (3)*	*94 (47)*

Figure 6.5: Percentage (and absolute) left and right perspective responses across the three stimulus conditions

Using the raw frequencies, we again examined whether the proportions of left and right perspective responses were different across our stimulus conditions (i.e., Baseline, Hot Desk and Wall). This revealed a significant difference, $X^2 = 51.7$, $p < .001$, Cramer's Phi = .59. Additional planned analyses found that the proportions for the Baseline and Hot Desk conditions were also significantly different from each other $X^2 = 5.32$, $p < .05$, Cramer's Phi = .23, as was the difference between the Baseline and Wall conditions, $X^2 = 27.4$, $p < .001$, Cramer's Phi = .52.

As with Experiment 11, the results from the Baseline condition show that the number is ambiguous. However, this ambiguity is reduced in the Hot Desk condition where a significant number of observers considered the number's visual perspective from the same direction that the Hot Desk items suggest (i.e., the left). The same effect occurred in the Wall condition where 94% of responses considered the number to be viewed from the central room position rather than wall viewpoint. This supports the central assumption of reference point theory; any stimulus that acts as the directional reference point will induce the same effect as when an agent is present.

6.4. Experiment 13 – Spontaneous Perspective Taking or Reference Point Theory?

6.4.1 Introduction

Participants in Experiments 11 and 12 were effectively asked to choose between two equally likely allocentric visual perspectives. To put another way, the change in viewpoint that was required always corresponded to approximately 90°. This can be contrasted with the agent's position in Figure 6.1 which more strongly juxtaposes her view with our own, corresponding to a difference of 180°. The data from Experiments 11 and 12 could therefore be an artefact of the scene.

In Experiment 13, we designed the stimuli such that the choice was effectively between an egocentric and allocentric viewpoint. We presented participants with one of two images and, as with Experiments 11 and 12, asked them to indicate what number is on the table. One of the images showed a human agent sitting at a table facing the participant; the other again replaced the agent with a hot desk set-up (see Figure 6.6). Experiment 13 can therefore also be considered as a conceptual replication of Experiment 12 using different stimuli. The visual perspective taking theory suggests that participants should only take the non-egocentric viewpoint when the agent is present. In contrast, the reference point theory predicts that the non-egocentric viewpoint will be taken in both the Agent-Present and Hot Desk conditions.

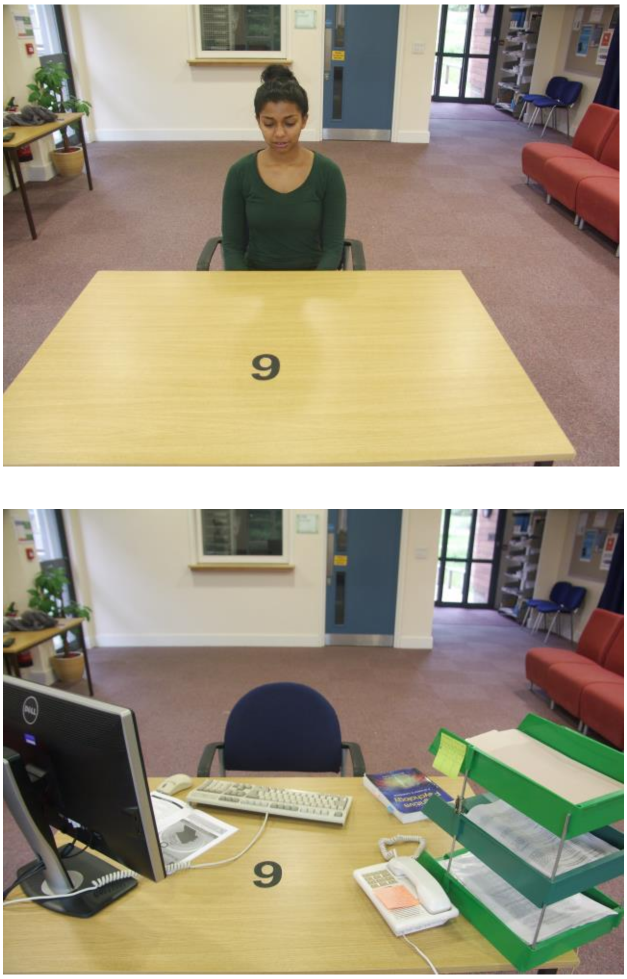

Figure 6.6: Stimuli used in Experiment 13. The upper panel shows the Agent-Present condition, the lower shows the Hot Desk condition.

6.4.2 Method

All aspects of the method were as described previously with the following exceptions. Eighty-six participants were recruited and randomly assigned to one of two conditions. Forty-five were presented with the Agent-Present stimulus, and 41 with the Hot Desk stimulus. Unlike the laminated photographs of Experiments 11 and 12, the images were presented via an Apple iMac computer linked to a Dell 7609 projector and participants tested in groups of approximately 12. Participants were simply asked to write down the number placed on the table.

6.4.3 Results and Discussion

	Allocentric	*Egocentric*
Agent-Present	*71 (29)*	*29 (12)*
Hot Desk	*73 (33)*	*27 (12)*

Figure 6.7: Percentage (and absolute responses) indicating "6" (i.e., allocentric visual perspective) and "9" (i.e., egocentric visual perspective) in the two conditions.

Analysis of the raw frequencies showed that the proportion of responses pertaining to the allocentric visual perspective did not differ across the Agent-Present and Hot Desk conditions, X^2 =.07, p >.7. Indeed, the latter condition generated slightly more responses from the non-egocentric viewpoint. These data therefore concur with the prediction made by the reference point theory.

One further aspect of these data worthy of note is the fact that overall non-egocentric 'visual perspective taking' was greater than that reported by Zhao et al. (2015; 42%). This itself could be due to the processing of reference points; there were simply more in the current stimuli (e.g., room, phone, books) than were present in the impoverished displays of Zhao et al.

<u>6.5 General Discussion</u>

The present chapter examined the claim that humans spontaneously compute the visual perspective of others. Experiment 11 replicated the findings of Zhao et al. (2015) showing that observers' judgement of an ambiguous number is consistent with the visual perspective of a human agent also present in the display, suggestive of visual perspective taking. However, we found the same effect even when the agent's view of the number was obstructed by another object. This does not therefore support the theory that this particular type of ToM is spontaneously computed. Thus, the present method represents a further paradigm, in addition to gaze following (Cole et al., 2015) and the dot-perspective task (Cole et al., 2016), to show visual-perspective-taking-like data even though the agent in the display could not see the critical stimuli.

We additionally presented an alternative explanation that can account for previous results and the data obtained in the present Experiment 11. We suggest that an agent in a scene acts as a reference point, which cues the observer, via mental rotation, to view the critical stimuli from this position, and in the direction to which the agent faces. This results in RTs being reduced and, as with the present chapter, ambiguous stimuli becoming less ambiguous. One might be tempted to counter that this process is itself a spontaneous perspective taking process; since such perspective taking must also involve a mental rotation process that assumes the position and vision of the agent. However, the fact that the same data occur when the agent's vision of the critical stimuli is obscured, reveals that the observer's perspective, resulting from mental rotation, cannot be from the agent's perspective. This is why the alternative theory suggests that the *approximate position* of the agent is assumed. Direction of sight is indeed computed, and this computation is certainly based on information from the agent, but it is not for its actual visual perspective. Furthermore, computing a ‘direction of sight’ from the position of a stimulus does not need to involve computation of the stimulus’ actual visual perspective. For instance, the word ‘left’ clearly has no perspective, it is a word not a seeing agent, but the word is likely to induce a viewpoint, via mental rotation, in the observer. Put simply, our reference point notion posits that an alternative representation is computed, based on the agent, but it is not computation of the agent’s actual visual perspective. This notion is supported by the results from the present Experiments 12 and 13 in which the human agent was replaced by a reference point (i.e., a ‘hot desk’) that informed the observer as to the direction in which the ambiguous stimulus should be viewed. We found the same pattern of data both when the human agent was present and when it was not.

We additionally argue that a further process occurs in this alternative model, a process concerning attribution of knowledge. The observer, implicitly or explicitly, will generate knowledge as to what the agent knows about the critical stimuli. Specifically, the observer will likely assume that the agent has at some point seen the stimuli. If one considers the Barrier condition of Experiment 11 (Figure 6.2, lower panel), the agent reading the newspaper is very likely to have seen the number on the table. The observer thus knows what the agent knows. This of course *is* a ToM attribution but importantly it is not one that computes the agent’s visual perspective. One can add that this knowledge-based interpretation also applies to what an agent *potentially* knows about its environment; what it might see in the next few moments.

It is worth emphasising that our reinterpretation of the typical visual perspective taking results is different to the standard interpretation in a subtle but crucially different way. Following Samson et al. (2010), the standard argument is that an agent’s *visual perspective* is computed. Of course, the very term, visual perspective taking, suggests that the *computation of a percept* occurs. This is why the phenomenon is referred to as ‘perspective taking’ and not, for instance, ‘position taking’; it is a reference to an agent’s perception, i.e., what they actually see. This is often made explicit within the field. For example, Moll and Kadipasaoglu (2013) refer to representing “others’ as well as one’s own “snapshot” perspectives in a literal, i.e., optical sense of the term”. Similarly, Capozzi, Cavallo, Furlanetto, and Bechio (2014) stated that “in simple perspective-taking tasks, one’s own and others’ *visual experience* influence each other” (italics added). Our alternative account, in contrast, emphasises what an agent knows about the world as opposed to what they see. It is therefore based on an attribution processes, the same process that lead to the modulation of gaze following when a participant is told that the agent cannot see (i.e., Nuku & Bekkering, 2008).

Although we have provided support for the reference point theory, we do acknowledge that mimicking the data that occur when an agent is present, with the use of a reference point (i.e., a hot desk) may be itself problematic. This pertains to a general issue that is common, particularly in visual cognition work. In order to argue that a particular phenomenon includes an attention orienting component, authors sometimes employ the experimental strategy of showing the same effect (i.e., mimicking the data) when the critical stimulus is replaced by an attentional cue. For example, Doneva and Cole (2014) observed the same pattern of data that occur in a common joint action phenomenon (‘social inhibition of return’; Welsh et al. 2005), thought by some to be due to imitation, when a co-actor was absent and replaced by an attention capturing cue. The authors therefore concluded that the basic joint action phenomenon is due to attention shifts induced by the observation of a limb movement. Making a similar argument, Dolk, Hommel, Prinz, and Liepelt (2013) found ‘joint’ Simon effect-like data when a co-actor was replaced by an attention capturing ticking clock. This mimicking-the-data strategy has also been employed previously within the visual perspective taking field. Santiesteban, et al. (2014) found that an arrow, replacing the avatar, also generates a consistency effect in the dot perspective paradigm. The authors concluded that a non-mental state attribution process was therefore driving the effect that occurs when the avatar is present. However, as Cole et al. (2017) point out, mimicking the typical results that occur in a paradigm by introducing process/stimulus x does not mean that those typical results are due to x. The typical results and the mimicked results may be due to different processes. This issue could therefore be applied to the current work. The results obtained in the Agent-Present and Hot Desk conditions may be due to different mechanisms that happen to generate the same data, with the former being due to visual perspective taking. It is for this reason that Cole et al. (2016) introduced the barrier/occlusion method; one should expect visual perspective taking-like data *to be absent* when the avatar cannot see the critical stimuli.

It is also worth noting that although the preliminary evidence supporting this alternative theory has used the ambiguous number paradigm, the theory is also applicable to the dot perspective task. This is due the fundamental aims of spontaneous visual perspective taking remaining the same. Interesting, this alternative theory could be one explanation as to why non-human orientation cues were able to produce the same pattern in consistency that the spontaneous visual perspective taking theory literature suggests (e.g., Santiesteban, et al. 2013, Nielsen et al. 2015). However, further research specifically assessing this alternative theory in relation to the dot perspective task is needed to solidify whether this alternative account can conclusively be applied.

Finally, another issue in the spontaneous visual perspective taking debate is the question of when a process should be considered as being spontaneous. Although spontaneity is most often thought of as a dichotomous phenomenon (i.e., spontaneous or not spontaneous), Logan (1985) made the point that it may be more expedient to view it as an evolving continuum, with processes continuously fluctuating upon it (Logan, 1985). Future researchers may want to examine where on the continuum of spontaneous processing ‘spontaneous’ visual perspective taking lies.

Chapter Seven:

General Discussion

7.1 Overall Findings and Implications

Initially, Chapter Two examined the spontaneous visual perspective taking theory using the combination of the dot perspective and ambiguous number paradigms. At first, Experiments 1 and 2 identified the spontaneous visual perspective taking phenomenon during conditions in which the process should not have been apparent. As previous literature suggests, if this process is in fact the *rapid* and *spontaneous* assumption of another individual's visual perspective, the effect should not have been present when the allocentric perspective's gaze was averted from the ambiguous number (Experiment 1), or obstructed by an occluding barrier (Experiment 2). However, this was not the case. Similar patterns in the data were found, irrespective of the two manipulations. Thus, Experiments 1 and 2 challenge the spontaneous visual perspective taking assumption. Additionally, in Experiment 3, a variant of Gardner et al. (2018) Rubbernecking manipulation was employed together with consistency and visibility manipulations using the dot perspective paradigm. This was due to the concept that directional cueing has been found to be strongest when the head of an avatar is not aligned with the body. However, Experiment 3 did not find any significance regarding these factors within the spontaneous visual perspective taking theory. Consequently, Chapter Two, driven by the results of Experiment 1 and 2, has begun to question the limitations of the spontaneous visual perspective taking notion.

Chapter Three increased the sensitivity of the spontaneous visual perspective taking theory by applying eye movement measures, whilst assessing whether the phenomenon is an extension of ToM by correlating the notion with self-reported measures of empathy. Consequently, both Experiments 4 and 5 assessed SOA, validity, and self-reported measures of empathy using eye movement measures.

However, Experiment 4 also examined the importance of human and non-human orientation cues, whilst Experiment 5 investigated the impact of occluding barriers. As expected, and justifying the use of eye movements, Experiment 4 found that RTs were shorter for fixation upon the target within the Overt task, in comparison to manual responses in the Covert task and were significantly shorter during Valid conditions in comparison to Invalid conditions for both tasks. Moreover, SOA induced a significant difference with 300ms SOA RTs being significantly shorter than 80ms. There was however no correlation between self-reported measures of empathy and susceptibility to spontaneously assume an allocentric visual perspective, and no difference in terms of the human and non-human directional cues. Consequently, the effects found in Experiment 4 are unlikely to be associated with the spontaneous assumption of an allocentric visual perspective, as the same effects were found in conditions absent of a humanised directional cue and with a greater effect at longer SOAs. Furthermore, Experiment 5 found that the agent's gaze shifted the attention of the observer accordingly. Importantly however, this effect was not abolished by the presence of an occluding barrier. Therefore, Experiment 5 challenges the spontaneous visual perspective taking theory since an individual cannot assume the visual perspective of another when that individual cannot see the target. In addition, as with Experiment 4, there was no correlation between self-reported measures of empathy and susceptibility to spontaneously assume an allocentric visual perspective within Experiment 5. Thus, Chapter Three again challenges the spontaneous visual perspective taking theory, as the same patterns in the data were observed in conditions without a humanised agent (Experiment 4), regardless of an occluding barrier (Experiment 5), and did not correlate with self-reported measures of empathy.

Chapter Four assessed the developmental trajectory of spontaneous visual perspective taking using adapted variations of the ambiguous number paradigm, as if this process is routed in ToM it should therefore be developed in childhood also. For Experiment 6, manipulations of ambiguity and validity were observed within a sample of three, five, seven, and nine-year-old children as well as an adult sample. Although the experiment did provide insight to show that differences within ability with respect to age are present, which could be due to ToM development, the results were largely inconclusive. Consequently, Experiment 7 reduced the complexity of the experimental design in order for age and ambiguity to be investigated in isolation, with the experimenter acting as the allocentric visual perspective. It was found that adult responses were significantly different in comparison to the sample of three and six-year-old children. Yet, this was not the case for both accuracy and RT. Taken together, the results of Experiment 6 and 7 are suggestive, but not definitive, as there was no overall effect of ambiguity when factoring in age in Experiment 6, and Experiment 7 did not identify a clear relationship for the adult sample in terms of RT, only in terms of error rate. Consequently, no overall conclusion in regards to the developmental trajectory of spontaneous visual perspective taking can be assumed.

Chapter Five assessed whether manipulating the perceived ownership over ambiguous information impacts spontaneous visual perspective taking using the ambiguous number paradigm. To modulate ownership two manipulations were devised. For Experiment 8 the physical distance between the ambiguous number and participant was manipulated, whereas Experiments 9 and 10 manipulated distance by altering the depth and location of stimuli. Specifically, the participant's suggested positioning was portrayed as being either inside or outside of the room that the ambiguous number was presented in. Additionally, Experiments 8 and 9 used a novel single response procedure, in which participants gave one response by reading the ambiguous number, while Experiment 10 used traditional behavioural RT and error rate measures. There were no significant results found in frequency of adopted perspective in Experiment 8 and 9. However, Experiment 10 showed that RTs were significantly shorter during Unambiguous conditions in comparison to Ambiguous conditions, irrespective of perceived ownership. Thus, Experiment 10 does not support the idea that perceived ownership for the ambiguous number influences the likelihood of participants spontaneously assuming the visual perspective of an agent. It is worth noting that although Chapter Five has been unable to report robust effects of ownership it has furthered the procedural experimentation of spontaneous visual perspective taking. Specifically, Chapter Five began to challenge how to measure a spontaneous process efficiently, absent of experimental interference. As a consequence, the use of the single response procedure was examined further during Chapter Six.

The last empirical chapter proposed an alternative theory for the spontaneous visual perspective taking effect. Specifically, that a human agent acts as a reference point that leads to facilitated responses. Again, Experiments 11, 12, and 13 employed a single response procedure. However, variations were made to the stimuli including the addition of non-human orientation stimuli instead of a human agent. Experiment 11 again found that the presence of an occluding barrier did not hinder the visual perspective taking-like effect. Experiment 12 found that the Hot Desk produced the same effects as an agent in favouring an alternative non-human orientation, which was also found in Experiment 13, with an additional Hot Desk set up. Thus, Chapter Six again challenges the spontaneous visual perspective taking theory.

Overall the spontaneous visual perspective taking theory has been challenged significantly by the current work. In the following sections, what it explicitly means to take another's visual perspective will be discussed in conjunction with the difference between perspective and perception as well as out-of-body experiences and mental imagery. Next, the reference point theory will be further defined before the schema theory of visual perspective taking is introduced. Lastly suggestions for future research will be outlined before the overall conclusion for the current work is drawn.

7.2 What does it mean to assume another's perspective?

A plethora of literature has been published examining the relevance and fundamental components of perspective taking, from assessing the levels of perspective taking (e.g., Salatas & Flavell, 1976; Flavell, Everett, Croft & Flavell, 1981; Yaniv & Shatz, 1990), to spatial differences (e.g., Harris & Strommen, 1972; Cox, 1981; Bialystok & Codd, 1987; Surtees, Apperly & Samson, 2013), as well as the developmental trajectories (e.g., Perner, 1991; Wimmer & Perner, 1983; Deutsch & Pechmann, 1982; Sonnenschein & Whitehurst, 1984). However, despite these differences in the literature, the overall consensus is that perspective taking is the ability to assume another individual's point of view, including visual, spatial, and emotional knowledge that may differ from the self. Yet, it remains unclear how exactly the self is portrayed in this sense, and how this process of assuming another's perspective is achieved. Consequently, if one begins to question exactly '*what does it mean to take someone's perspective?*' a number of other questions arise. How is this process achieved? How is the self perceived in this context? When an individual assumes the perspective of another, what exactly is being attended to? How can we be certain of intentions and beliefs of another individual? In the following section I will ask, *what does it mean to assume another's visual perspective*? This will be addressed in reference to published literature as well as suggestions for future work that could be used to explore this question.

7.2.1 Perspective Taking or Embodied Perception?

The literature surrounding perspective taking has contributed to a number of key advances. For example, different levels of visual perspective have been identified, with level 1 involving the identification of whether another agent can or cannot see a target and level 2 referring to whether an individual can identify that the visual image of a target is perceived differently (Flavell, Everett, Croft & Flavell, 1981; Salatas & Flavell, 1976; Yaniv & Shatz, 1990). Yet, how exactly the allocentric perspective is adopted, is still unclear. Proposals have been made in terms of a *literal assumption* of an alternative visual perspective, using key terms such as '*visual experience'* (Capozzi, Cavallo, Furlanetto, & Bechio, 2014), '*snapshot*' (Moll & Kadipassaoglu, 2013), and '*imaging*' (Erle & Topolinski, 2017). Thus, in terms of the current literature it would seem, and indeed stated, that visual perspective taking is the assumption of another's *visual experience*. This clearly suggests that the individual is somewhat able to see through the embodied percept of another. Furthermore, how could one possibly judge what another individual is attending to? For example, one may be focused and gazing directly at a target, yet your attention may be diverted to the side of the target, or at a different depth. To reiterate, if one was to view an agent looking through a window at a child playing in the centre of a room, one may argue that the agent is attending to the child, as the gaze direction would suggest. However, this may not be the case. The individual may be focusing on imperfections in the windowpane, at the floor, or the walls of the room etc. Consequently, in terms of the visual experience in which the perspective taking theory assumes, we cannot create a fully formed "snapshot" of the scene, as we are unable to fully know, or guess the agent's intentions.

Reflecting this criticism back to spontaneous visual perspective taking, how can one assume the visual perspective of another both spontaneously and rapidly, absent of conscious thought, in order to process scenes with ambiguous information or alternative perspectives? The very nature of this phenomenon suggests that an embodied computation of an alternative view is automatic, requiring minimal cognitive effort, which when reflected back to the above, cannot be the case. The processing of the scene, in terms of visual experience, will require a high level of conscious effort in reference to mental states and intentions of the allocentric perspective, which perspective taking is dependent upon. In the following section, this criticism will be applied to out-of-body experiences (OBE), as there are similar assumptions between this phenomenon and the suggested embodiment of perspective taking.

7.2.2 Out-of-Body Experiences

Individuals that experience OBE report their sense of self leaving their physical body spontaneously and have no control over the relocation of the self to a different position in space, which may be above or beside the body, or in a different location altogether (Blackmore, 1982; Cook & Irwin, 1983; Eastman, 1962; Irwin, 1985). The causes of this spontaneous experience have been heavily debated, yet for the purpose of this theoretical discussion the causes are unimportant, instead the similarities between OBE and spontaneous visual perspective taking will be focused upon.

Although it may seem peculiar to compare OBE alongside spontaneous visual perspective taking, the two experiences have a common factor. If we take the assumption of perspective taking as the embodied process of another's visual perspective, the self must be extracted somehow from the physical body, as is the case of those who experience the extraction of the self for an OBE to occur. This idea is supported by the work of Kessler and Braithwaite (2016) and Braithwaite and Dent (2011) who discussed the similarities between OBE and visuo-spatial perspective taking.

This issue of veridical perception has been extensively assessed within OBE literature, however, most famously, and most applicable in terms of this theoretical discussion is the experimentation of 'Miss Z' by Tart (1968). A five-digit number was placed on a shelf out of sight and reach of Miss Z. Miss Z's EEG pattern was recorded whilst she slept, to assess OBE. Most interestingly, Tart reported that Miss Z was able to correctly recall the number that was placed out of sight and reach from her, as a result of an OBE. The assumption was that Miss Z was able to extract herself from her physical body while she slept, and relocated herself to a new location in the room that allowed her to see the number. Clearly, this could not have occurred. Thus, research that assesses whether visual perspective taking is an embodied process, like that of OBE, may wish to incorporate a hidden element into the methodology in which the agent or avatar can see the target, but the participant cannot.

The OBE literature also documents the phenomenon of heautoscopy, in which an individual, experiences being in two places at once, be that within the bounds of the physical body, as well as the extracted self-location (Blanke & Mohr, 2005). This is similar to what has been claimed in the visual perspective taking literature. An individual may simultaneously be assuming the embodied allocentric perspective, as well as the egocentric self perspective. Thus, how the embodiment of the allocentric perspective is processed, which is the focus in question, is still not accounted for. The point is therefore somewhat moot; understanding how, and whether, individuals do indeed embody the visual perspective of another within spontaneous visual perspective taking is still required.

7.3 Mental Imagery and Perspective Taking

Recall from Chapter One the debate surrounding mental models (Craik, 1943), mental imagery (Craik, 1943; Shepard and Metzler 1971; Pylshyn, 1973; Just & Carpenter, 1976; Hochberg & Gellman, 1977; Dennett, 1991; Kosslyn, 1994), and mental rotation (Shepard and Metzler 1971; Just & Carpenter, 1976; Hochberg & Gellman, 1977). In summary, the debate questions whether mental models are processed in terms of visual representation (Kosslyn, 1994) or are instead processed using prior experience as well as pre-existing knowledge (Pylshyn, 1973). This in turn affects mental rotation as to whether a visual picture presented within the mind's eye is virtually rotated (Shepard and Metzler 1971; Just & Carpenter, 1976; Hochberg & Gellman, 1977; Dennett, 1991; Kosslyn, 1994), or instead, knowledge of the image is used in the rotation (Pylshyn, 1973). Thus, it was stated in Chapter One that in terms of the spontaneous visual perspective taking notion, individuals may not in fact be embodying the allocentric visual perspective displayed in the images as implied by previous literature; but instead maybe forming a mental model of the scene (Craik, 1943). Thus, RT differences could be dependent upon the processing of the scene via mental models and not be associated with the counter correction of the spontaneous assumption of an allocentric visual perspective. However, the debate of whether these mental models are visual pictorial representations (Dennett, 1991; Kosslyn, 1994), or are simply by-products of previous experience and knowledge (Pylshyn, 1973) is still on-going. Consequently, the spontaneous visual perspective taking theory can also be related to the mental imagery debate. For example, does an individual use what they 'see' in their mind's eye in which to response within the paradigms used to examine spontaneous visual perspective taking, or do they simply 'know' what is, can, and cannot be seen? The visual perspective taking literature suggests the former with its reference to "visual experience" and to "literal snapshots" of a scene. To reiterate further, recall the stimuli used in Experiment 2, (see Figure 7.1) in which the agent cannot see the target number. When interpreting the results of this experiment using past knowledge and experience of the agent, observers of the image may simply know

Figure: 7.1 The Non-Visible stimuli used in Experiment 2, with the Unambiguous number of the left and Ambiguous number on the right

that moving the newspaper from the field of view would allow the agent to see the target, through past experience. In addition, the observer may unconsciously assume that the agent is aware of the nature of the experiment and target placement, thus they may be unconsciously altering their response pattern accordingly. This is one possible explanation of why the same patterns in the data were found in Experiment 2 for Ambiguous conditions regardless of the occluding barrier. Alternatively, when applying the pictorial assumption of the mental imagery debate, differences in RT may be due to the transformations required of the mental image held within the observer's mind's eye. As previously stated, Shepard and Metzler (1971), Just and Carpenter (1976), and Hochberg and Gellman (1977) all support this claim, as they found that RT could be progressively influenced with the increased number of mental rotations required. Additionally, Michelon and Zacks (2006) and Surtees, et al. (2013) also investigated the influence of distance and angles and found that RT increased with an increased angle between the agent and participant. Nonetheless, as this explanation is based on speculation, future work should assess whether this assumption can be applied to other assessments of the spontaneous visual perspective taking notion. However, in order to test this assumption, replications of the dot perspective and ambiguous number paradigms will need to be implemented alongside an additional factor that manipulates the level of knowledge that the agent or avatar has over the targets. For example, the avatar in the dot perspective paradigm could be seen to look around the room, prior to the presentation of the targets, as could the agent in the ambiguous number paradigm. It would then be definitive that observer knows that the agent or avatars visual perspective, regarding the targets, differ from their own. In other words, the observer knows what the agent knows.

7.4 Reference Point Theory

In Chapter Six, the reference point theory was introduced. This theory suggests that the agent, be that the avatar in the dot perspective paradigm or the model in the ambiguous number paradigm, acts as a reference point in which the observer is cued, via mental rotation, to view the target stimulus from the agent or model's positioning. This therefore impacts RT in the dot perspective paradigm, and the significance of ambiguity in the ambiguous number paradigm. Bryant, Tversky, and Franklin (1992) can be seen as support for this claim as they found that when instructed to adopt the perspective of an inanimate object, observers are able to perform this easily. It is worth noting that one could argue that the reference point theory is using spontaneous visual perspective taking during the mental rotation of the agent's visual field. However, the *actual* visual perspective of the agent is not computed or assumed in the reference point theory. Indeed, the visual positioning and line of sight of the agent is acknowledged in terms of the spatial arrangement of the stimulus, but other inanimate objects can be used to orientate the stimulus also, which is not accounted for by the spontaneous visual perspective taking theory. Therefore, a clear distinction between the two theories is that within the reference point theory the assumption of the agent's visual perspective is not embodied, whereas the spontaneous visual perspective taking theory suggests that it is.

Interestingly, the reference point notion can account for unexpected spontaneous visual perspective taking-like findings. For example, Experiment 1 found the spontaneous visual perspective taking-like differences during conditions in which the agent was not directly looking at the stimulus. This can be explained as the agent acting as a reference point in which to orientate the image. The offset of the agent's head does not therefore detract from the observer mentally rotating the image in accordance to the agent position, thus influencing RT. Additionally, the influence of the barrier method in Experiments 2, 5, 11, and 12 can also be explained using the reference point idea. The presence of the occluding barrier does not inhibit the processing of the agent's line of sight; thus, RT is still affected through the orientating of the image. In fact, most interestingly, in Experiments 12 and 13 we found the same pattern in the data both when the agent was present and when it was not. Thus, supporting the reference point theory, and countering the assumption that ToM processes were influencing the results of these experiments.

Attribution of knowledge, as discussed above, may also contribute towards the reference point theory. As observers view the scene, they will generate knowledge, be that explicitly in terms of what is expected of them in a specific task, but also implicitly in terms of what an agent may or may not have seen, as well as other a priori assumptions about the scene. Hence, unintentionally observers may judge whether the agent has previously seen or is currently seeing the target. This includes conditions in which the agent is freely viewing the target, but also during conditions in which the target may be barred from view. Thus, the observer may generate the knowledge that the agent knows that the target is behind the barrier. It is worth noting that this is an extension of ToM ability, as the observer is making a judgement about the mental state of the agent. However, it is not the computation of the agent's visual perspective in a literal sense, which spontaneous visual perspective taking proposes. Therefore, the attribution of this knowledge, which may be unconsciously processed, can influence RT as well as the response given, contributing towards the conclusions of spontaneous visual perspective taking.

Reflecting the alternative reference point notion back to the literature associated with the spontaneous visual perspective taking theory, several different cues may have contributed to orientations of the stimuli. In terms of the dot perspective paradigm, the room, and any inanimate objects, be that arrows or avatars, all may have acted as orientation cues in which the observer can position and orientate the image. This is also the case for reflexive gaze following, joint action, and the ambiguous number paradigm. However, within these paradigms the chances of additional orientation cues are more apparent. For example, in the Zhao et al. (2015) ambiguous number experiment, the agent is sat at a table, which again can be used as an orientation stimulus to anchor and transform the image. Nonetheless, to further examine this theory, future experiments should be encouraged to use inanimate objects as orientation cues, to see whether the same spontaneous visual perspective taking-like data effects can be found. This is particularly crucial for experimentation using the dot perspective paradigm, as the preliminary evidence for the alternative reference point theory has utilised the ambiguous number paradigm. Therefore, in order to conclusively associate the alternative reference point theory to the dot perspective task, further experimentation is essential. Additionally, it would be beneficial to assess whether the SOA differences that are apparent within spontaneous visual perspective taking literature are also present when using inanimate objects in the reference point theory.

7.5 Does the Schema Theory apply?

As gaze cuing and spontaneous visual perspective taking have similar principles within their theoretical framework, (i.e., both are dependent upon ToM and have been argued to be automatic) alternative theories that can be applied to gaze cuing could also be applied to the spontaneous visual perspective taking phenomenon. Cole et al. (2015) introduced a novel theory of gaze cuing, defined as the schema theory. This theory states that specific goals, be that cueing attention, are learnt through sequences of action known as schemas. Once these actions are learnt, they are adapted from top down processes, requiring conscious control, into bottom up actions, which are executed automatically (Kahneman, 1973). In this sense, Cole et al. (2015) proposed that the repetition of observed gaze direction enables the development of a gaze cuing schema, with the primary outcome of the schema being the *rapid orientating of spatial attention to the gaze-at location* (Cole et al. 2015). This therefore alters RT in relation to consistency of gaze direction and target placement, leading to gaze cuing.

The schema theory can also be applied to spontaneous visual perspective taking. When individuals are presented with a scene that have an implied directed attention, be that with gaze, arrows, or non-human orientation stimuli, the individual is encouraged to develop a directional schema. This schema maps the stimulus or scene spatially, and thus is used to rapidly orientate the image or scene. Consequently, once developed, the schema can be applied to conditions with similar arrangements and in turn alter RT in accordance to experimental manipulations. It is worth noting how the schema theory of visual perspective taking and the spontaneous visual perspective taking assumption differ. Most importantly, spontaneous visual perspective taking claims that an alternative visual perspective is rapidly and spontaneous assumed, implying a forced embodiment of the alternative visual perspective. Whereas, the schema theory of perspective taking proposes a response pattern is learnt, that relies on mapping the spatial layout of the scene thus orientating the image. To reiterate, the spontaneous visual perspective taking notion implies an embodied process, whereas the schema theory of visual perspective taking emphasises orientation of the scene, absent of embodiment.

Alongside the initial outline of the schema theory, Cole et al. (2015) also highlighted a number of predictions about when the gaze cuing effect should and should not be observed. Three of these predictions, which are non-specific of gaze cuing, will now be applied to the schema theory of visual perspective taking. First, the formation of the gaze cuing schema is dependent upon the strength of the cuing direction (Cole et al. 2015). In other words, the motivation for ignoring the gaze direction and the strength to which the gaze implies, or is ambiguous at promoting a direction of attention, will directly impact the formation of a gaze cuing schema. This is also true for the schema theory of visual perspective taking. For the greater the emphasis of a particular direction should directly correlate towards the greater tendency of the spontaneous visual perspective taking-like effect being found. For example, refer back to Zhao et al. (2015), 42% of participants reported the ambiguous number from the visual perspective of the model within the gaze condition, yet 46% during the gazing and reaching condition. The authors suggested that interaction with the ambiguous number further impacts the likelihood of a spontaneous visual perspective taking response. However, in terms of the schema theory of visual perspective taking, enhancing the cued direction in any form will increase the strength of a schema formation. Thus, the actions of reaching alongside the gaze direction combine to the overall strength of the cued direction. Additionally, Gardner et al.

(2018) identified a greater spontaneous visual perspective taking-like effect when an avatar's head was averted to the side, towards the target, with the torso facing the observer. Although it could be argued that ToM encompasses social elements that the averted gaze condition may have incorporated, thus encouraging the need to understand why the avatar is looking to the side. The spontaneous visual perspective taking theory does not integrate this into its overall concept. Nevertheless, when applying the schema theory of visual perspective taking to Gardner et al. (2018), one might conclude that the averted gaze condition has a greater cuing strength in comparison to the standard avatar positioning, and thus contributes towards the formation of a stronger directional cuing schema, absent of avatar embodiment.

Cole et al. (2015) also predicted that the cuing effect, in accordance to the schema theory, would have a clear developmental trajectory with young children being slow and predominately using conscious control within the associated tasks. Additionally, as the child grows the effect should become progressively faster with reduced levels of conscious control. Applying this prediction towards the spontaneous visual perspective taking effect, both the schema and spontaneous visual perspective taking theory agree, the effect is indeed developed throughout childhood, and this is supported by a number of experiments (Harris, 1972; Harris & Strommen, 1972; Cox, 1981; Bialystok & Codd, 1987; Epley, Morewedge & Keysar, 2004; Moll & Tomasello, 2006; Sodian, Thoemer & Metz, 2007; Surtees & Apperly, 2012; Surtees, Butterfill & Apperly, 2012). Yet the age at which the phenomenon is first developed is still debated.

Lastly, Cole et al. (2015) predicted that imposing cognitive load and/or incorporating ego depletion will modulate the development of an associated schema. Thus, it is hypothesised that in terms of the schema theory of visual perspective taking, increasing the cognitive load with distraction tasks, and progressive ego depletion, will modulate the development of the directional schema. However, there is currently limited exploration of the influence of cognitive load or distraction tasks within the spontaneous visual perspective taking phenomenon. It is worth noting that exploring the influence of cognitive load and distractor tasks in future work would benefit the literature in two ways. Firstly, it would allow the exploration of a key prediction of the schema theory of visual perspective taking, but it would also examine of the automaticity assumption of visual perspective taking. There have been many different theories developed in the attempt to categorise and define automatic processes (refer to Chapter One for a brief review of three of these theories), which have differing claims over the essential characteristics of a truly automatic process. These characteristics have suggested that automatic processes reside on a continuum (gradual theory, Logan, 1985), are dependent upon preconditions (conditioned approach, Bragh, 1992), and are dependent upon conscious goals (goal dependency, Bragh and Gollwitzer 1994), to highlight a few. Consequently, this investigation of whether cognitive load and/or ego depletion affects the notion, will either highlight that the process is conscious, thus supporting the schema theory of visual perspective taking. Or it will highlight that the notion is automatic and exhibited regardless, thus supporting the spontaneous visual perspective taking notion and refuting the schema claim.

7.6 Future Directions

In this following section, future experiments will be proposed that will further the understanding of spontaneous visual perspective taking in relation to the criticisms and alternative theories previously highlighted.

Initially, investigating to what extent the perspective of another is assumed should be considered. Specifically, clarification in terms of whether the spontaneous visual perspective taking phenomenon is an embodied process, is essential. This could be achieved through replications of the dot perspective and ambiguous number paradigms with alterations to the target presentation. During the conditions in which the participant is required to assume the perspective of the agent or avatar, *the target should only be detectable by the agent or avatar only, and not by the participant.* For example, one can imagine an ambiguous figure that can only be (clearly) resolved when seen from the avatar's position. In contrast, when seen from the position of the participant, this figure appears as a blob. This is a similar approach to the methodology of Tart (1968) in their experimentation on OBE. Although this seems peculiar as an approach, a clear identification of whether the alternative visual perspective is embodied will be identifiable, as pervious research cannot avoid the criticism that participants are not implicitly aware of what the avatar or agent can and cannot see, which the proposed experiment will avoid. The suggested experiment will consequently be unable to be criticised in terms of conscious thought and knowledge over the targets and potential targets. It is predicted that if the spontaneous visual perspective taking notion is in fact the embodiment of the alternative perspective, participants will be able to correctly respond in terms of the targets detected only by the alternative visual perspective. However, if spontaneous perspective taking is not the embodiment of the alternative visual perspective, it is predicted that participants will not be able to correctly respond in terms of the targets detected only by the alternative visual perspective.

Furthermore, in regards to the mental imagery debate, classification as to whether spontaneous visual perspective taking relies upon pictorial representations of mental models or is in fact reliant upon the processing of past experiences and preconceived knowledge over the scene, is required. To assess this, future work should aim to manipulate the level of knowledge that participants have over the scene. This can be achieved using both the dot perspective and ambiguous number paradigms initially, but it could also be applied to reflexive gaze following and co-operative tasks where spontaneous visual perspective taking may not be the sole focus. To manipulate the level of knowledge that the participant has over the image, the agent, or avatar being presented, should be seen to first look around the room, which may or may not have the target information already placed within it. Then the targets could move, appear, or be removed. Consequently, it would be definitive that the participants know that the avatar or agent's knowledge of the targets is different to their own. It is predicted that if differences in avatar and participant knowledge over the image is crucial to the processing of the scene, then manipulating knowledge will dramatically affect the RT response pattern that has been identified in previous literature.

Additionally, to assess whether the reference point theory can be used as an explanation for spontaneous visual perspective taking, future work should examine whether inanimate objects orientate images to the same effect as agents or avatars. For example, the investigation of inanimate orientation stimuli such as the desk scenes in Experiments 12 and 13, as well as arrows in Experiment 4 could be extended. Moreover, it would be interesting to see whether other inanimate objects with emphasised orientations, that have not been previously assessed, can be used to influence the orientation of the scene and act as anchors. For example, weapons, automobiles, and everyday electronics. Using these objects, simple replications of the dot perspective and ambiguous number paradigm can be applied, using both the standard RT method with multiple trials, as well as the single response method that was used in Experiments 8, 9, 11, 12, and 13. It is predicted that if the reference point theory partly contributes to the phenomenon, and participants are using the orientation stimuli as a means to anchor and rotate an image, and are not rapidly and spontaneously assuming the visual perspective of another, similar patterns in the data will be identified with any non-human orientation stimuli.

Lastly, in terms of furthering the schema theory of visual perspective taking, cognitive load and ego depletion should be examined. Incorporating standardised distraction tasks into the dot perspective or ambiguous number paradigm is one way in which to assess the influence of cognitive load and ego depletion within the spontaneous visual perspective taking phenomenon. Visual and auditory distraction (see e.g., Mastroberardino & Vredeveldt, 2014; Kyriakidou et al. 2014), social tasks (see e.g., Baddeley et al., 1975; Logie et al., 1990; Craik, 2014), and stroop tasks (see e.g., Lautenbach, Laborde, Putman, Angelidis, & Raab, 2016; Kalanthroff, Henik, Derakshan, & Usher, 2016) are all examples of distraction tasks that could be used in the application of cognitive load or ego depletion within spontaneous visual perspective taking. If it is shown that cognitive load and/or ego depletion effects or abolishes the suggested spontaneous visual perspective taking-like data trends, the schema theory of visual perspective taking would be supported. However, if it were found that individuals do indeed rapidly and spontaneously assume the visual perspective of another, regardless of increased cognitive load or ego depletion, the schema theory of visual perspective taking would be refuted.

7.7 Conclusion

The current work examined the spontaneous visual perspective taking notion using a number of different methodologies and underlying theoretical rationales. Although the phenomenon was identified during replications of the central methodological paradigms, similar effects were also observed during conditions in which the notion should have been abolished (e.g., with an occluding barrier or a non-human orientation stimulus). Consequently, an alternative theory that explains spontaneous visual perspective taking-like effects was introduced, defined as the reference point theory. Lastly, a number of theoretical issues regarding the notion were discussed, including the issues of what is means to take another individual's perspective as well as OBEs, mental imagery, and the schema theory of visual perspective taking. To conclude, it is clear that the spontaneous visual perspective taking-like effects are apparent. Yet, the claim that this notion is solely caused by individuals rapidly and spontaneously assuming the visual perspective of another is reductionist. Instead, it is clear to state that other mechanisms, such as the suggested reference point theory, gaze cuing application, schema theory of visual perspective taking, as well as many other aspects of cognition, all contribute towards the previously suggested spontaneous visual perspective taking-like effect. Thus, further experimentation within this field is essential.

References

Allison T, Puce A, McCarthy G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences*. *4*, 267–278.

Amorim, M. A., Glasauer, S., Corpinot, K., & Berthoz, A. (1997). Updating an object's orientation and location during nonvisual navigation: A comparison between two processing modes. *Perception & Psychophysics, 59,* 404-418.

Amorim, M. A., & Stucchi, N. (1997). Viewer- and object-centered mental explorations of an imagined environment are not equivalent. *Cognitive Brain Research*, *5,* 229-239.

Amorim, M. A., Trumbore, B., & Chogyen, P. L. (2000). Cognitive repositioning inside a desktop VE: The constraints introduced by first versus third-person imagery and mental representation richness. *Presence: Teleoperators & Virtual Environments*, *9*, 165-186.

Apperly, I. (2011). Mindreaders: The cognitive basis of. *Theory of Mind*, 1093-1108.

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, *116*(4), 953.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The Psychology of learning and motivation: Advances in research and theory, Vol 2, (pp. 89 - 195). New York: Academic Press.*

Baddeley A. D., Grant S., Wight E., Thomson N. (1975). "Imagery and visual working memory," *In Attention and Performance V, eds Rabbitt P. M. A., Dornic S., editors.* (London: Academic Press), 205–217

Baker, L. J., Levin, D. T., & Saylor, M. M. (2016). The extent of default visual perspective taking in complex layouts. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(4), 508.

Baron-Cohen, S. (1997). *Mindblindness: An essay on autism and theory of mind*. MIT press.

Baron-Cohen S. The eye direction detector (EDD) and the shared attention mechanism (SAM): Two cases for evolutionary psychology. In: Moore C, Bialystok, E., & Codd, J. (1987). *Children's interpretations of ambiguous spatial descriptions*. British Journal of Developmental Psychology, *5*, 205–211.

Bargh, J. A. (1992). The ecology of automaticity: Toward establishing the conditions needed to produce automatic processing effects. *American Journal of Psychology*, *105*, 181–199.

Bargh, J. A., & Gollwitzer, P. M. (1994). Environmental control of goal-directed action: Automatic and strategic contingencies between situations and behaviour. In W. D. Spaulding (Ed.), *Nebraska Symposium* (Vol. 41, pp. 71–124). Lincoln: University of Nebraska Press.

Beggan, J. K. (1992). On the social nature of nonsocial perception: The mere ownership effect. *Journal of Personality and Social Psychology*, *62*, 229–237.

Belk, R. W. (1988). Possessions and the extended self. *Journal of Consumer Research*, *15*, 139–168.

Belk, R. W. (1991). The ineluctable mysteries of possessions. *Journal of Social Behaviour and Personality*, *6*, 17–55.

Bialystok, E., & Codd, J. (1987). Children's interpretations of ambiguous spatial descriptions. *British Journal of Developmental Psychology*, *5*(3), 205-211.

Birch, S. A. J., & Bloom, P. (2004). Understanding children's and adults' limitations in mental state reasoning. *Trends in Cognitive Sciences*, *8,* 255-260.

Blackmore, S. J. (1982). Beyond the body: *An investigation of out-of-the-body experiences*. Chicago: Academy Chicago.

Blanke, O., & Mohr, C. (2005). Out-of-body experience, heautoscopy, and autoscopic hallucination of neurological origin: Implications for neurocognitive mechanisms of corporeal awareness and self-consciousness. *Brain Research Reviews*, *50*(1), 184–199.

Boyer, P., Robbins, P., & Jack, A. I. (2005). Varieties of self-systems worth having. *Consciousness and Cognition*, *14*, 647–660.

Braithwaite, J. J., & Dent, K. (2011). New perspectives on perspective-taking mechanisms and the out-of-body experience. *Cortex, 47*(5), 628-632.

Brown, T. L., Gore, C. L., & Carr, T. H. (2002). Visual attention and word recognition in Stroop color naming: Is word recognition automatic? *Journal of Experimental Psychology: General*, *131*, 220 – 240.

Bryant, D. J., Tversky, B., & Franklin, N. (1992). Internal and external spatial frameworks for representing described scenes. *Journal of Memory and Language*, *31*(1), 74-98.

Bukowski, H., Hietanen, J. K., & Samson, D. (2015). From gaze cueing to perspective taking: Revisiting the claim that we automatically compute where or what other people are looking at. *Visual Cognition*, *23*(8), 1020-1042.

Butterfill, S. A., & Apperly, I. A. (2013). How to construct a minimal theory of mind. *Mind & Language*, *28*(5), 378-384.

Calder, A. J., Lawrence, A. D., Keane, J., Scott, S. K., Owen, A. M., Christoffels, I., & Young, A. W. (2002). Reading the mind from eye gaze. *Neuropsychologia*, *40*, 1129–1138.

Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences*, *12*(5), 187-192.

Capozzi, F., Cavallo, A., Furlanetto, T., & Becchio, C. (2014). Altercentric intrusions from multiple perspectives: Beyond dyads. *PloS one*, *9*(12), e114210.

Caron, A. J., Butler, S. C., & Brooks, R. (2002). Gaze following at 12 and 14 months: Do eyes matter. *British Journal of Developmental Psychology*, *20*, 225–239.

Carpendale, J., & Lewis, C. (2006). *How children develop social understanding*. Blackwell Publishing.

Clark, H.H. (1992). *Arenas of language use.* Chicago: University of Chicago Press.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition.* Washington, DC: American Psychological Association.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. New York, NY: Routledge Academic.

Cohen, J. D., Servan-Schreiber, D., & McClelland, J. L. (1992). A parallel distributed processing approach to automaticity. *American Journal of Psychology*, *105*, 239 – 269.

Constable, M. D., Kritikos, A., & Bayliss, A. P. (2011). Grasping the concept of personal property. *Cognition*, *119*(3), 430-437.

Conway, J. R., Lee, D., Ojaghi, M., Catmur, C., & Bird, G. (2017). Submentalizing or mentalizing in a Level 1 perspective-taking task: A cloak and goggles test. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(3), 454.

Cook, A. M., & Irwin, H. J. (1983). Visuospatial skills and the out-of-body experience. *The Journal of Parapsychology*, *47*(1), 23–36.

Cooney, S.M.; Brady, N.; Ryan, K. (2017). Spatial orienting of attention to social cues is modulated by cue type and gender of viewer. *Experimental Brain Research*, *235*, 1481–1490.

Cole, G. G., Atkinson, M., Le, A. T., & Smith, D. T. (2016). Do humans spontaneously take the perspective of others?. *Acta psychologica*, *164*, 165-168.

Cole, G. G. & Millett, A. C. (2019). The closing of the Theory of Mind: A critique of perspective taking, *Psychonomic Bulletin and Review*, 1-16.

Cole, G. G., Smith, D., & Atkinson, M. (2015). Mental state attribution and the gaze cueing effect. *Attention, Perception & Psychophysics*, *77*, 1105-1115.

Cosijn, E., & Ingwersen, P. (2000). Dimensions of relevance. *Information Processing & Management*, *36*(4), 533-550.

Cox, M. V. (1981). Interpretation of the spatial prepositions in front of and behind. *International Journal of Behavioural Development*, *4*, 359–368.

Craik, K. (1943). *The nature of explanation*. Cambridge, England: Cambridge University Press.

Craik, F. I. (2014). Effects of distraction on memory and cognition: a commentary. *Frontiers in Psychology*, *5*, 841.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671-684.

Creem, S. H., Downs, T. H., Wraga, M., Proffitt, D. R., & Downs, J. H., III (2001). An fMRI study of imagined self-rotation. *Cognitive, Affective, & Behavioural Neuroscience*, *1*, 239-249.

Cunningham, S. J., Turk, D. J., Macdonald, L. M., & Macrae, C. N. (2008). Yours or mine? Ownership and memory. *Consciousness and Cognition*, *17*(1), 312-318.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*(1), 113-126.

Decety, J., & Michalska, K. J. (2010). Neurodevelopmental changes in the circuits underlying empathy and sympathy from childhood to adulthood. *Developmental Science*, *13*(6), 886-899.

Dennett, D. C. (1991). *Consciousness Explained*. Boston: Little, Brown & Company.

Deutsch, W., & Pechmann, T. (1982). Social interaction and the development of definite descriptions. *Cognition*, *11*, 159–184.

Doherty, M. (2008). *Theory of mind: How children understand others' thoughts and feelings*. Psychology Press.

Driver IV, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition*, *6*(5), 509-540.

Duran, N. D., Dale, R., & Kreuz, R. J. (2011). Listeners invest in an assumed other's perspective despite cognitive cost. *Cognition*, *121*(1), 22–40.

Eastman, M. (1962). *Out-of-the-body experiences*. Paper presented at the Proceedings of the Society for Psychical Research.

Emery, N. J., & Clayton, N. S. (2009). Comparative social cognition. *Annual Review of Psychology*, *60*, 87-113.

Erle, T. M., & Topolinski, S. (2017). The grounded nature of psychological perspective-taking. *Journal of Personality and Social Psychology*, *112*(5), 683.

Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology*, *40*(6), 760–768.

Farroni, T., Mansfield, E. M., Lai, C., & Johnson, M. H. (2003). Infants perceiving and acting on the eyes: Tests of an evolutionary hypothesis. *Journal of Experimental Child Psychology*, *85*(3), 199-212.

Fasig, L. G. (2000). Toddlers' understanding of ownership: Implications for self-concept development. *Social Development, 9*, 370–382.

Flavell, J. H. (1986). The development of children's knowledge about the appearance-reality distinction. *American Psychologist*, *41*, 418– 425.

Flavell, J. H. (1999). Cognitive development: children's knowledge about the mind. *Annual Review of Psychology, 50*, 21-45.

Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young childrens knowledge about visual-perception – further evidence for the level 1- level 2 distinction. *Developmental Psychology*, *17,* 99–103.

Fischer, M. H., Prinz, J., & Lotz, K. (2008). Grasp cueing shows obligatory attention to action goals. *The Quarterly Journal of Experimental Psychology*, *61*(6), 860-868.

Fiske, S. T., & Taylor, S. E. (1991). *Social Cognition*. (2nd Ed.). New York: McGraw-Hill.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 1030.

Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review*, *5*(3), 490-495.

Frith, C. D. (2004). Schizophrenia and theory of mind. *Psychological Medicine*, *34*(3), 385-389.

Frith, C., D., & Frith, U. (2005). Theory of Mind. *Current Biology*, *15*(17), 644-645.

Frith, C. D., & Frith, U. (2006). The Neural Basis of Mentalizing. *Neuron*, *50*(4), 531–534.

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological Bulletin*, *133*(4), 694.

Frischen, A., Loach, D., & Tipper, S. P. (2009). Seeing the world through another person's eyes: Simulating selective attention via action observation. *Cognition*, *11*(2), 212-218.

Furlanetto, T., Cavallo, A., Manera, V., Tversky B., & Becchio C. (2013). Through your eyes: incongruence of gaze and action increases spontaneous perspective taking. *Frontiers in Human Neuroscience, 7*, 455, 1-5.

Furlanetto, T., Becchio, C., Samson, D., & Apperly, I. (2016). Altercentric interference in level 1 visual perspective taking reflects the ascription of mental states, not submentalizing. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(2), 158.

Gardner, M., Bileviciute, A., & Edmonds, C. (2018). Implicit mentalising during level-1 visual perspective-taking indicated by dissociation with attention orienting. *Vision, 2*(1), 3.

Gallagher, S. (2000). Philosophical conceptions of the self: Implications for cognitive science. *Trends in Cognitive Sciences, 4,* 14–21.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin.

Goldstein, T. R., & Winner, E. (2012). Enhancing empathy and Theory of Mind. *Journal of Cognition and Development*, *13*(1), 19-37.

Graf, R. (1994). Self-rotation and spatial reference: The psychology of partner-centred localisations.

Graziano, M. S., & Gross, C. G. (1994). Mapping space with neurons. *Current Directions in Psychological Science, 3*(5), 164-167.

Harris, L. J. (1972). Discrimination of left and right, and development of the logic of relations. *Merrill-Palmer Quarterly*, *18*(4), 307–320.

Harris, L. J., & Strommen, E. A. (1972). Role of front-back features in children's front, back, and beside placements of objects. *Merrill-Palmer Quarterly of Behaviour and Development, 18,* 259–271.

Hart, R. A., & Moore, G. T. (1973). *The development of spatial cognition: A review*. Aldine Transaction.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, *108*, 356 – 388.

Happé, F., & Frith, U. (1995). Theory of Mind in autism. In, *Learning and cognition in autism* (pp. 177-197). Springer US.

Heyes, C. M. (1998). Theory of mind in nonhuman primates. *Behavioural & Brain Science, 330*, 1830-1834.

Heyes, C. (2014). Submentalizing: I Am Not Really Reading Your Mind. *Perspectives on Psychological Science*, *9*(2), 131-143

Hietanen, J.K. (2002). Social attention orienting integrates visual information from head and body orientation. *Psychological Research*, *66,* 174–179.

Hochberg, J., & Gellman, L. (1977). The effect of landmark features on mental rotation times. *Memory and Cognition*, *5*(1), 23-26.

Humphrey, N. (2000). How to solve the mind-body problem. *Journal of Consciousness Studies*, *7,* 5–20.

Huttenlocher, J., & Presson, C. C. (1973). Mental rotation and the perspective problem. *Cognitive Psychology*, *4*(2), 277-299.

Irwin, H. J. (1985). *Flight of mind: A psychological study of the out-of-body experience*. Metuchen, NJ: Scarecrow Press.

Johnson-Laird, P. N., & Bethell-Fox, C. E. (1978). Memory for questions and amount of processing. *Memory & Cognition*, *6*(5), 496-501.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, *8*(4), 441-480.

Just, M. A., & Carpenter, P. A. (1980). A Theory of Reading: From Eye Fixations to Comprehension. *Psychological Review*, *87*, 329-354.

Juurmaa, J., & Lehtinen-Railo, S. (1994). Visual experience and access to spatial knowledge. *Journal of Visual Impairment & Blindness, 88*, 157-170.

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Englewood Cliffs, NJ: Prentice-Hall.

Kalanthroff, E., Henik, A., Derakshan, N., & Usher, M. (2016). Anxiety, emotional distraction, and attentional control in the Stroop task. *Emotion*, *16*(3), 293.

Keehner, M., Guerin, S. A., Miller, M. B., Turk, D. J., & Hegarty, M. (2006). Modulation of neural activity by angle of rotation during imagined spatial transformations. *Neuroimage*, *33*(1), 391-398.

Kessler, K. (2000). Spatial cognition and verbal localisations: A connectionist model for the interpretation of spatial prepositions. *Wiesbaden: Deutscher Universitäts-Verlag*.

Kessler, K., & Braithwaite, J. J. (2016). Deliberate and spontaneous sensations of disembodiment: capacity or flaw?. *Cognitive Neuropsychiatry*, *21*(5), 412-428.

Kessler, K., & Rutherford, H. (2010). The two forms of visuo-spatial perspective taking are differently embodied and subserve different spatial prepositions. *Frontiers in Psychology*, *1*, 213.

Kessler, K., & Thomson, L. A. (2010). The embodied nature of spatial perspective taking: embodied transformation versus sensorimotor interference. *Cognition*, *114*(1), 72-88.

Kessler, K., & Wang, H. (2012). Spatial perspective taking is an embodied process, but not for everyone in the same way: differences predicted by sex and social skills score. *Spatial Cognition & Computation*, *12*(2-3), 133-158.

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, *89*(1), 25-41.

Kosslyn, S. M. (1994). *Image and Brain: The resolution of the imagery debate*. Cambridge. MA: MIT Press.

Kosslyn, S. M., & Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery? *Psychological Bulletin, 129*(5), 723.

Kozhevnikov, M., & Hegarty, M. (2001). A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory & Cognition*, *29*(5), 745-756.

Kozhevnikov, M., Motes, M. A., Rasch, B., & Blajenkova, O. (2006). Perspective-taking vs. mental rotation transformations and how they predict spatial navigation performance. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *20*(3), 397-417.

Kyriakidou M., Blades M., Carroll D. (2014). Inconsistent findings for the eyes closed effect in children: the implications for interviewing child witnesses. *Frontiers in Psychology*. *5*, 448.

Kuhn, G., Vacaityte, I., D'Souza, A. D., Millett, A. C., & Cole, G. G. (2018). Mental states modulate gaze following, but not automatically. *Cognition*, *180*, 1-9.

Laberge, D. & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*, 293 – 323.

Langton, S. R. H., & Bruce, V. (1999). Reflexive visual orienting in response to the social attention of others. *Visual Cognition, 6(5),* 541-567.

Langton, S. R. H., & Bruce, V. (2000). You must see the point: Automatic processing of cues to the direction of social attention. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 747–757.

Lautenbach, F., Laborde, S. J. P., Putman, P., Angelidis, A., & Raab, M. (2016). Attentional distraction by negative sports words in athletes under low-and high-pressure conditions: Evidence from the sport emotional Stroop task. *Sport, Exercise, and Performance Psychology, 5*(4), 296.

Leslie, A. M. (1987). Pretense and representation: The origins of" theory of mind.". *Psychological Review, 94*(4), 412.

Levine, M., Jankovic, I. N., & Palij, M. (1982). Principles of spatial problem solving. *Journal of Experimental Psychology: General*, *111*(2), 157.

Light, P., & Nix, C. (1983). Own view versus good view in a perspective-taking task. *Child Development, 54*(2), 480-483.

Logan, G. D. (1985). Skill and automaticity: Relations, implications, and future directions. *Canadian Journal of Psychology, 39*, 367–386.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*, 492 – 527.

Logie R. H., Zucco G. M., Baddeley A. D. (1990). Interference with visual short-term memory. *Acta Psychologica*. *75,* 55–74

Mani, K., & Johnson-Laird, P. N. (1982). The mental representation of spatial descriptions. *Memory & Cognition*, *10*(2), 181-187.

Mattan, B. D., Rotshtein, P., & Quinn, K. A. (2016). Empathy and visual perspective-taking performance. *Cognitive Neuroscience*, *7*(1-4), 170-181.

Masangkay, Z. S., McCluskey, K. A., McIntyre, C. W., Sims-Knight, J., Vaughn, B. E., & Flavell, J. H. (1974). The early development of inferences about the visual precepts of others. *Child Development, 45*, 357–366.

Mastroberardino S., Vredeveldt A. (2014). Eye-closure increases children's memory accuracy for visual material. *Frontiers in Psychology*. *5*, 241.

McCann, R. S., Folk, C. L., & Johnston, J. C. (1992). The role of spatial attention in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 1015 – 1029.

Melztoff, A. M., & Brooks, R. (2008). Self-experience as a mechanism for learning about others: A training study in social cognition. *Developmental Psychology, 44*, 1257-1265.

Michelon, P., & Zacks, J. M. (2006). Two kinds of visual perspective taking. *Attention, Perception, & Psychophysics*, *68*(2), 327-337.

Millett, A. C., D'Souza, A. D., & Cole, G. G. (2019). Attribution of vision and knowledge in 'spontaneous perspective taking'. *Psychological Research*, 1-8.

Moll, H., & Kadipasaoglu, D. (2013). The primacy of social over visual perspective-taking. *Frontiers in Human Neuroscience, 7*, 558.

Moll, H., & Tomasello, M. (2006). Level I perspective-taking at 24 months of age. *British Journal of Developmental Psychology, 24*, 603–613.

Moore, C., & Dunham, P. J., (Eds.) (1995). *Joint attention: Its origins and role in development*. Hove, United Kingdom: Erlbaum.

Moors, A., & De Houwer, J. (2006). Automaticity: a theoretical and conceptual analysis. *Psychological Bulletin*, *132*(2), 297.

Morgan, E. J., Freeth, M., & Smith, D. T. (2018). Mental State Attributions Mediate the Gaze Cueing Effect. *Vision*, *2*(1), 11.

Mou, W., McNamara, T. P., Valiquette, C. M., & Rump, B. (2004). Allocentric and egocentric updating of spatial memories. *Journal of experimental psychology: Learning, Memory, and Cognition*, *30*(1), 142.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles on inhibitionless spreading activation and limited –capacity attention. *Journal of Experimental Psychology: General, 106*, 226 – 254.

Neisser, U. (1988). Five kinds of self-knowledge. *Philosophical Psychology, 1*, 35–39.

Nielsen, M. K., Slade, L., Levy, J. P., & Holmes, A. (2015). Inclined to see it your way: Do altercentric intrusion effects in visual perspective taking effects reflect an industrially social process? *Quarterly Journal of Experimental Psychology*, *68*, 1931-1951.

Nuku, P., & Bekkering, H. (2008). Joint attention: Inferring what others perceive (and don't perceive). *Consciousness and Cognition*, *17*, 339–349.

O'Craven, K. M., & Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of Cognitive Neuroscience*, *12*(6), 1013-1023.

Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: Bradford Books/MIT Press.

Perner, J., & Ruffman, T. (2005). Infants' insight into the mind: How deep? *Science*, *308*, 214-216.

Perrett, D. I., Hietanen, J. K., Oram, M. W., Benson PJ. (1992). Organisation and functions of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society of London, Series B*. *335*, 23–30.

Pierno, A. C., Becchio, C., Wall, M. B., Smith, A. T., Turella, L., and Castiello, U. (2006). When gaze turns into grasp. *Journal of Cognitive Neuroscience*, *18*, 2130-2137.

Plumert, J. M., Kearney, J. K., & Cremer, J. F. (2004). Children's perception of gap affordances: bicycling across traffic ‐ filled intersections in an immersive virtual environment. *Child Development*, *75*(4), 1243-1253.

Pomianowska, I.; Germeys, F.; Verfaillie, K.; Newell, F.N. (2012). The role of social cues in the deployment of spatial attention: Head-body relationships automatically activate directional spatial codes in a Simon task. *Frontiers in Integrative Neuroscience*, *6*, 4.

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and Performance X* (pp.531–556). Hillsdale, NJ: Erlbaum.

Posner, M. I. & Snyder, C. R. R. (1975). Attention and cognitive Control. In R. S. Solso (Ed.), *Information processing and cognition: The Loyola Symposium (pp. 55-85). Hillsdale, NJ: Erlbaum.*

Piaget, J., & Inhelder, B. (1956). *The child's conception of space*. London: Routledge.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind?. *Behavioral and Brain Sciences, 1*(4), 515-526.

Presson, C. C., & Montello, D. R. (1994). Updating after rotational and translational body movements: Coordinate structure of perspective space. *Perception*, *23,* 1447-1455.

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin, 80,* 1-24.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of*

Experimental Psychology: Human Perception and Performance, 18(3), 849-860.

Reynolds, M., & Besner, D. (2006). Reading aloud is not automatic: processing capacity is required to generate a phonological code from print. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(6), 1303.

Ross, H. S. (1996). Negotiating principles of entitlement in sibling property disputes. *Developmental Psychology, 32*, 90–101.

Santiesteban, I., Catmur, C., Coughlan Hopkins, S., Bird, G., & Heyes, C. (2013). Avatars and Arrows: Implicit Mentalizing or Domain-General Processing? *Journal of Experimental Psychology: Human Perception and Performance, 36*, 1255-1266.

Santiesteban, I., Shah, P., White, S., Bird, G. & Heyes, C. (2015). Mentalizing or submentalizing in a communication task? Evidence from autism and a cameral control. *Psychonomic Bulletin and Review, 22*, 844-849.

Salatas, H., & Flavell, J. H. (1976). Behavioral and metamnemonic indicators of strategic behaviors under" remember" instructions in first grade. *Child Development*, 81-89.

Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. (2010). Seeing it their way: Evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 1255-1266.

Saracevic, T. (1996, October). Relevance reconsidered. In *Proceedings of the second conference on conceptions of library and information science* (CoLIS 2) (pp. 201-218). New York: ACM.

Schamber, L., Eisenberg, M. B., & Nilan, M. S. (1990). A re-examination of relevance: toward a dynamic, situational definition. *Information Processing & Management*, *26*(6), 755-776.

Schneider, W., & Shiffrin, R.M. (1977). Controlled and automatic human information processing. I. Detection, search and attention. *Psychological Review*, *84*, 1 – 66.

Senju, A., Southgate, V., Snape, C., Leonard, M., & Csibra, G. (2016). Do 18-month-olds really attribute states to others? A critical test. *Psychological Science, 22*(7), 878-880.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three dimensional objects. *Science*, *171*, 701-703.

Shelton, A. L., & McNamara, T. P. (1997). Multiple views of spatial memory. *Psychonomic Bulletin & Review*, *4*(1), 102-106.

Siegal, M., & Beattie, K. (1991). Where to look first for children's understanding of false beliefs. *Cognition*, *38*, 1–12.

Sonnenschein, S., & Whitehurst, G. J. (1984). Developing referential communication: A hierarchy of skills. *Child Development*, *55*, 1936– 1945.

Stephenson, B., & Wicklund, R. A. (1983). Self-directed attention and taking the other's perspective. *Journal of Experimental Social Psychology*, *19,* 58–77.

Stoffregen, T. A. (2003). Affordances as properties of the animal-environment system. *Ecological Psychology*, *15*(2), 115-134.

Surtees, A. D. R., & Apperly, I. A. (2012). Egocentrism and automatic perspective taking in children and adults. *Child Development*, *83*, 452–460.

Surtees, A. D. R., Butterfill, S. A., & Apperly, I. A. (2012). Direct and indirect measures of Level-2 perspective-taking in children and adults. *British Journal of Developmental Psychology*, *30*(1), 75–86.

Surtees, A., Apperly, I., & Samson, D. (2013). Similarities and differences in visual and spatial perspective-taking processes. *Cognition*, *129*(2), 426-438.

Surtees, A., Samson, D., & Apperly, I. (2016a). Unintentional perspective-taking calculates whether something is seen, but not how it is seen. *Cognition*, *148*, 97-105.

Surtees, A., Apperly, I., & Samson, D. (2016b). I've got your number: Spontaneous perspective-taking in an interactive task. *Cognition*, *150*, 43-52.

Tart, C.T. (1968). A psychophysiological study of out-of-the body experiences. *Journal of the American Society for Psychical Research*, *62,* 3–27.

Teufel, C., Alexis, D. M., Clayton, N. S., & Davis, G. (2010). Mental-state attribution drives rapid, reflexive gaze following. *Attention, Perception, & Psychophysics*, *72*(3), 695-705.

Teufel, C., Alexis, D. M., Todd, H., Lawrance-Owen, A. J., Clayton, N. S., & Davis, G. (2009). Social cognition modulates the sensory coding of observed gaze direction. *Current Biology*, *19*, 1274– 1277.

Teufel, C., Fletcher, P. C., & Davis, G. (2010). Seeing other minds: Attributed mental states influence perception. *Trends in Cognitive Sciences, 14,* 376–382.

Tomasello, M. (2008). *Origins of human communication*. Cambridge: MIT Press.

Tversky, B., Lee, P., & Mainwaring, S. (1999). Why do speakers mix perspectives?. *Spatial Cognition and Computation*, *1*(4), 399-412.

Waytz, A., Gray, K., Epley, N., & Wegner, D. M. (2010). Causes and consequences of mind perception. *Trends in Cognitive Sciences, 14*(8), 383-388.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72*(3), 655-684.

Wellman, H. M., Hollander, M., & Schult, C. A. (1996). Young children's understanding of thought-bubbles and of thoughts. *Child Development*, *67*, 768–788.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 103–128.

Wraga, M., Creem, S. H., & Proffitt, D. R. (1999). The influence of spatial reference frames on imagined object-and viewer rotations. *Acta Psychologica*, *102*(2-3), 247-264.

Wraga, M., Creem, S. H., & Proffitt, D. R. (2000). Updating displays after imagined object and viewer rotations. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *26*, 151-168.

Wraga, M., Shephard, J. M., Church, J. A., Inati, S., & Kosslyn, S. M. (2005). Imagined rotations of self versus objects: an fMRI study. *Neuropsychologia*, *43*(9), 1351-1361.

Yaniv, I., & Shatz, M. (1990). Heuristics of reasoning and analogy in children's visual perspective taking. *Child Development*, *61*(5), 1491-1501.

Xu, B., Tanaka, J., & Mineault, K. (2012). The head turn cueing effect is sustained at longer SOA's in the presence of an object distractor. *Journal of Vision*, *12*(9), 396– 400.

Zacks, J. M., & Michelon, P. (2005). Transformations of visuospatial images. *Behavioral and Cognitive Neuroscience Reviews*, *4*(2), 96-118.

Zacks, J. M., Mires, J., Tversky, B., & Hazeltine, E. (2000). Mental spatial transformations of objects and perspective. *Spatial Cognition & Computation*, *2*, 315-332.

Zacks, J. M., Vettel, J. M., & Michelon, P. (2003). Imagined viewer and object rotations dissociated with event-related fMRI. *Journal of Cognitive Neuroscience*, *15*, 1002-1018.

Zhao, X., Cusimano, C. J., & Malle, B. F. (2015). In Search of Triggering Conditions for Spontaneous Visual Perspective Taking. *In CogSci*.

Zwickel, J. (2009). Agency attribution and visuo-spatial perspective taking. *Psychonomic Bulletin & Review*, *16,* 1089-1093.

Zwickel, J. & Muller, H. J. (2010) Observing fearful faces leads to visuo-spatial perspective taking. *Cognition*, *117*, 101-105.