The Development of Gaze Understanding

By

Catherine Marianne Sayer

School of Psychology

University of East Anglia

A thesis submitted in partial fulfilment of the requirements of the University of East Anglia for the degree of Doctor of Philosophy.

September 2024

Abstract

Introduction. Children follow others' gaze by their second year (e.g. Corkum & Moore, 1995; Moore & Corkum, 1998), yet many 3-year-olds cannot judge where someone is looking from eye direction alone (Doherty & Anderson 1999; Doherty et al., 2009). It has been claimed that humans have two distinct systems for gaze processing (Doherty 2006; Doherty et al., 2009), with the later developing system motivated by pre-school theory of mind changes.

Objectives. The current thesis explores the two-systems theory and investigates whether young children use a simpler conception of attention not requiring understanding of eye direction.

Methods. Experiment 1 measures 24- to 54-month-old's gaze judgement, theory of mind performance, and understanding of occlusion. Experiment 2 replicated this longitudinally from 33-to 48-months-old. Experiment 3 and 4 assesses the claim that System 1 remains distinct in adulthood. Faces filtered to remove edge or luminance information were presented in gaze orienting tasks to measure if System 1 can be identified by the visual information it processes.

Results. Young children's gaze judgement gradually improves in the preschool period. This development is not related to or predicted by theory of mind ability. Children's early understanding of occlusion does not differentiate between seeing and knowing. A more flexible understanding is associated with gaze judgement ability and predicted by Knowledge Access. Finally, faces retaining luminance or edge information produced equal gaze orienting in adults.

Conclusion. The results are consistent with the claim that children have two systems for processing gaze, but there is no evidence that System 2 development is motivated by children's increasing interest in others' minds. Prior to this, children likely understand seeing in terms of spatial and behavioural relations. However, the evidence suggests the two systems do not remain distinct in adulthood.

Table of Contents

List of Figures

List of Tables

Acknowledgements

Words cannot express my gratitude to Dr. Martin Doherty for his supervision of this thesis. In my Undergraduate research project, I first acknowledged that Martin continually challenged me to become a better researcher. Four years later and he is still challenging me. Thank you for nurturing my passion for research and I look forward to continuing our collaboration for years to come. I am also thankful to my second supervisor Dr. George Malcolm for his feedback and advice on studies and at review meetings.

This thesis would not have been possible without the schools and parents who gave permission for their children to take part in the studies. Special thanks to the families in the longitudinal study who returned for multiple visits and welcomed me to see their children grow into mini adults.

I have had the pleasure of working with many colleagues in the School of Psychology who have all been supportive of my academic journey. I would like to highlight my thanks to Dr. Teodora Gliga for being my academic Mum. You invited me into your lab group, baked amazing (and super healthy) snacks, and helped me grow the Pop-up Lab which is one of my proudest achievements from the PhD. It was through you I met Dr. Kelsey Frewin, who has been an absolute ray of sunshine and has taken time to train me on infant EEG and eye-tracking ready for future research. Additional thanks to Dr. Prerna Aneja and Dr. Alvaro Rodriguez for being my cheerleaders when it came to teaching experience.

Many thanks to Anna, Khadijah, and Lizzie for our weekly writing group which really helped keep me sane during the writing period, and to Reuben for sending great music when times were tough.

Finally, I am deeply grateful for my husband and son who have been there throughout this whole journey. They've seen me on the good and the bad days, but importantly never let me give up on this dream.

Declaration

I declare that the work contained in this thesis has not been submitted for any other award and that it is all my own work. I also confirm that this work fully acknowledges opinions, ideas and contributions from the work of others.

Any ethical clearance for the research presented in this thesis has been approved. Approval has been sought and granted by the School of Psychology Ethics Committee at the University of East Anglia.

Part of the work from Chapter 2 has been presented at the following conferences:

Oral Presentations

Sayer, C., Macfadyen, E., Doherty, M. J., (2023). "Knowing it and seeing it are two different things" unless you are a two-year-old. School of Psychology Postgraduate Conference Day, University of East Anglia, UK.

Sayer, C., Macfadyen, E., Doherty, M. J., (2024). Why young children don't understand seeing or knowing: the transition from engagement to theory of mind. European Society for Philosophy and Psychology, Université Grenoble-Alpes, Grenoble.

Poster Presentations

Sayer, C., Macfadyen, E., Doherty, M. J., (2023). "Seeing and knowing are two different things" unless you are a two-year-old. Lancaster Conference on Infant and Early Child Development, Lancaster University, UK.

Sayer, C., Macfadyen, E., Doherty, M. J., (2024). Pre-schooler's understanding of attentional relationships: The transition from engagement to eye-direction. Budapest CEU Conference on Cognitive Development. Central European University, Budapest.

Sayer, C., Macfadyen, E., Doherty, M. J., (2024). Pre-schooler's understanding of attentional relationships: The transition from engagement to eye-direction. European Society for Philosophy and Psychology, Université Grenoble-Alpes, Grenoble.

Chapter 1

Developmentally Distinct Systems for Processing Gaze: Relations to Theory of Mind

There is a consensus view that infants can think about others' visual attention (Butler et al., 2000; Corkum & Moore, 1998), which has frequently been suggested as a precursor to understanding belief (Goméz, 1996). As Goméz (1996) put it, understanding that someone sees something is "… an early and simple way to know what is in the other's mind, because the contents of the other's mind – the object looked at – is in front of the beholder's eyes…" (p.334). Thus, tracking and interpreting eye gaze is central to human social behaviour. On the contrary, I discuss the claim that children do not understand others' gaze direction until four years, with its development motivated by the emergence of uniquely human theory of mind abilities (Doherty et al., 2009).

Gaze processing has been extensively studied in non-human primates, human infants, preschoolers, and adults, with the mastery of gaze following used as the determining milestone for gaze understanding. That being, once one can follow gaze then one can use it referentially. The literature consistently shows that we are sensitive to gaze from birth (Farroni et al., 2002) and can follow the head or eye-direction of others from at least the second year (Corkum & Moore, 1995), and perhaps earlier (D'Entremont et al., 1997). In contrast, many 3-year-olds are unable to judge where someone is looking from their eye-direction alone (Doherty & Anderson 1999; Doherty et al., 2009). Whilst these children can follow gaze, they cannot make judgements about where someone is looking. In other words, they can direct their attention toward the object of others' attention, but do not necessarily understand the attentional relation between the observer and the object.

One way to explain the relationship between gaze following in infancy and preschoolers' novel ability to judge gaze direction would be to argue that the two abilities share a core competency. This gaze understanding is initially implicit and not available for verbal report, later becoming explicit. Evidence from comparative (e.g. Anderson & Mitchell, 1999; Povinelli & Eddy, 1996a; 1996b) and infant (e.g. Corkum & Moore, 1995; Brooks & Meltzoff, 2002; Doherty et al., 2009) research support the opposite possibility, that the two abilities are quite distinct and not directly related developmentally. I will be suggesting that the ability to make explicit judgments of eye direction, the hallmark of the later system of gaze understanding, relates to theory of mind changes at approximately 4 years. Prior to this, children have a simpler conception of attention that does not require any mental state concepts.

Two-Systems Theory of Gaze Processing

Humans respond to gaze direction in two ways. The first (A) is by using others' gaze cue to locate objects of attention. The second (B) is using others' gaze cues to think about the attentional relationship between observer and object. In other words, A allows thoughts of the form "there is an apple", whereas B allows thoughts of the form "she is looking at an apple".

From this, Doherty (2006) hypothesised there are two distinct gaze processing systems. One system that is innate, fast and approximate (System 1), and another learned and computationally complex one (System 2). System 1 serves function A, thought to use gaze-following to direct attention to the locus of another's attention for things attended to by others are often important to be aware of (potential danger, food, etc.). It uses a variety of cues such as body posture, head-and-eye, or eye direction, and is available early in development, to monkeys, apes, and infants. Although this system redirects one's own attention, it does not provide understanding the attention of others.

Attention, as conceived of by adults, concerns representational mental states: things internal to the mind, that are about something external (usually) to the mind. Because System 1 is about objects and not attentional relationships, it cannot be identified with Theory of Mind, infant or otherwise, but System 2 can. Whilst it employs the same attentional input cues as 1, it serves function B. It yields not simply an object of attention, but also the relation between the observer and object. Theory of mind also goes through contemporaneous development as gaze judgements after 3 years of age. Thus, System 2 seems only available to humans beyond infancy, possibly due to increasing interest in others' mental states.

Morphology of the Human Eye

Human eyes are unique. Kobayashi and Kohshima's (2001) landmark study demonstrates this in detail. Like most other primates, humans have forward-facing eyes and dense central fovea, meaning that our vision has good depth perception and high resolution. In other primates, determining locus of attention can be done with reasonable accuracy using head direction. In comparison, human head turns take more time and energy. Thus, humans (and larger primates) benefit from moving their eyes independently of their head. This is supported by Kobayashi and Kohshima's (2001) finding that humans move their eyes without a corresponding head turn 60% of the time, but our closest relative the chimpanzee only 25% of the time.

Compared to head direction, eye direction is a harder cue to interpret. It is reliant on iris-sclera contrast; the higher the contrast between the iris and sclera, the easier eye-direction is to interpret (Emery, 2000; Kobayashi & Kohshima, 1997; Yorzinski et al., 2021). Primates have sclera that is typically pigmented the same colour as the iris and face. Humans, however, have an extensive white sclera, creating a high iris-sclera contrast. It is plausible that differences in iris-sclera contrast are an adaption to meet the communicative needs of the species. After all, *"the eyes not only look, but are looked at*" (Gibson & Pick, 1963, p. 386).

Primates often view direct gaze as a signal of aggression or submission (Andrew, 1963; Coss, 1970). Thus, the ability to camouflage gaze confers a survival benefit, by reducing the risk of conflict. It also allows for discreet collection of information, such as location of food and spotting of predators (Kobayashi & Hashiya, 2011). To test this claim, Mearing et al., (2022) analysed the scleral brightness of 108 primate species and correlated this with the percentage of deaths per species attributable to conspecific lethal aggression (Gómez, et al., 2016). They found primate scleral brightness to be negatively associated with conspecific lethal violence. As such, primates having a low iris-sclera contrast may serve a camouflaging function.

The high contrast in humans' functions as *gaze-signalling*. Direct gaze in humans can signal a multitude of meanings such as intent, sincerity, and sharing of information (Coss, 1970; Tomkins, 1963). Averted gaze can also inform us about another's attention (Carlin & Calder, 2013), and responding to gaze is thought to play a vital role in the early development of social cues (Jones & Klin, 2013). Thus, discriminating eye direction is important for detecting social cues. In support of this, Kobayashi and Hashiya (2011) found that eye morphology correlated with the social complexity (group size) of the species. When living in groups, grooming is important for maintaining social relationships (Muroyama, 1991). However, as the group size gets larger, directing contact-based grooming or vocalisations to the right being becomes costly in time and energy. Thus, Kobayashi and Hashiya (2011) propose to overcome this, non-contact methods of social grooming such as gaze become necessary. This is because gaze direction allows one to determine if communicative signals such as vocalisations and facial expression are directed at oneself. Kobayashi and Hashiya (2011) refer to this as *"gaze-grooming"*. Greater exaggeration of sclera contrast allows for better gaze discrimination, conferring greater success as gaze-grooming in social groups. With humans living in over double the group size of that of non-human primates (Kobayashi & Hashiya, 2011), the exaggerated morphology of the human eye may be a specific adaptation for social evolution seen in humans.

Primate Gaze Following

Some non-human primates can detect gaze. Prosimians, such as lemurs, cannot follow human gaze (Anderson & Mitchell, 1999), but can follow the gaze of conspecifics (Ruiz et al., 2008; Shepherd & Platt, 2007). Monkeys such as macaques, will follow human or conspecific head direction (Anderson & Mitchell, 1999; Emery et al., 1997). Chimpanzees can follow human head and eye direction, and to a slightly greater extent if they are human-reared (Call et al., 2000). They also expect gaze to terminate at an object. Chimpanzees have been shown to follow gaze behind various types of barriers including doors and opaque windows, but not through transparent windows (Bräuer et al., 2005; Okamoto-Barth et al., 2007). Notably, Povinelli and Eddy (1996a) placed an opaque screen between and to one side of the experimenter, who looked to a position on the screen invisible to the chimpanzee. Simply following the line of the experimenter's gaze (through the opaque screen) would lead chimpanzees to turn to look behind them. Instead, they moved to one side, presumably to see what on the other side of the screen the experimenter was looking at.

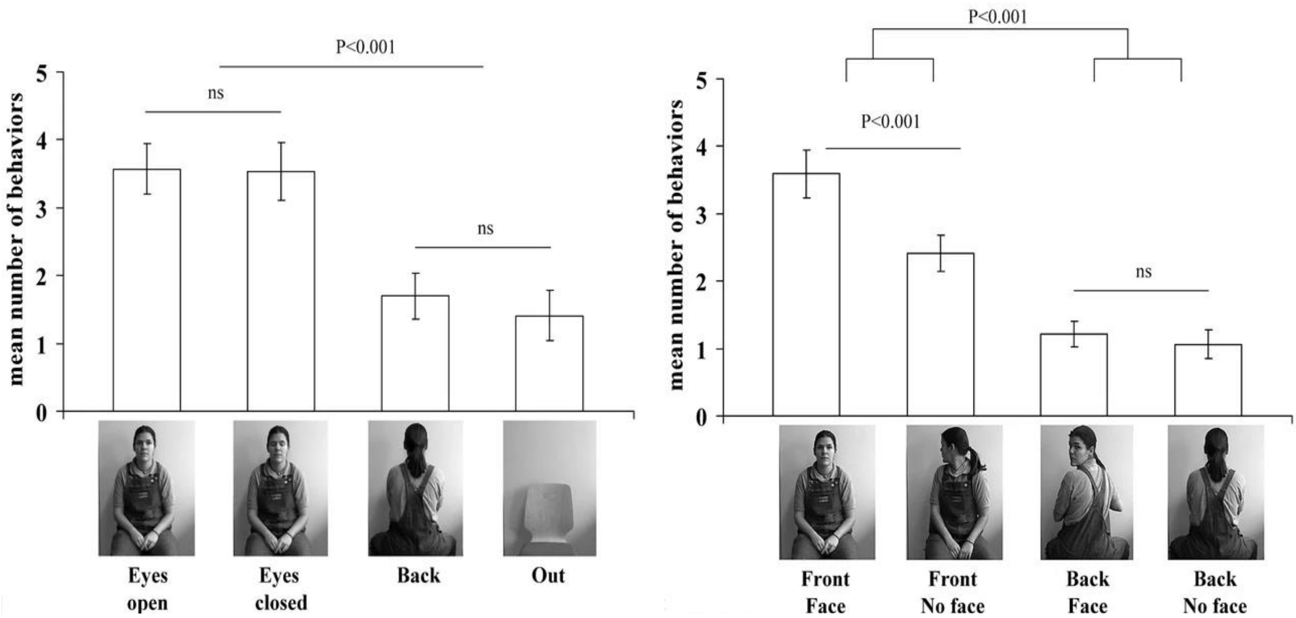
These abilities show that monkeys and chimpanzees understand something about gaze. Chimpanzees can follow line of sight and expect it to terminate at objects. This suggests that chimpanzees are not simply orienting their attention but appear to represent gaze as a spatial relationship between the eyes and an object. However, they do not seem to understand this relationship in any psychological way. For example, Anderson et al. (1996; Hauser et al., 2007) conducted a study in which the experimenter turned his head and eyes to look at one of two covered locations. Macaques and capuchins never learned to use this cue to find the food reward in the looked-to box. Chimpanzees are subject to a similar limitation. The evidence is mixed concerning whether they can learn to prefer the gazed at location in Anderson et al.'s (1996) paradigm. Itakura and Tanaka (1998) found that their chimpanzees could, even when the experimenter only glanced without moving his head. Povinelli et al. (1999) found that their chimpanzees followed head direction, but not eye direction. However, they responded similarly whether the experimenter's head and eyes were directed to the target, or 50cm above the target. Adult humans would interpret this as attention elsewhere, and 3-year-old children, unlike the chimpanzees, only respond to the head and eyes cue when it was directed at the target, not when it was directed above it. Thus, it may be that in both studies chimpanzees responded correctly because they are aware there is a relationship between observer and object, but do not understand the significance of it. In other words, it should be questioned whether chimpanzees understand seeing as a form of attention or simply an orienting mechanism.

To test this, Povinelli and Eddy (1996b) took advantage of chimpanzee begging behaviour. Chimpanzees were offered the chance to beg from one of two humans. One human could see the chimpanzee. The other would be either facing the other way, blindfolded, have a bucket on her head or hands covering the eyes. If chimpanzees understood something about visual attention, then they should beg from the experimenter who could see them. The chimpanzees typically showed no preference for the experimenter who could see them. The only cue that chimpanzees systematically preferred was whole body facing them compared to whole body facing away. Thus, chimpanzees may not understand seeing but are merely sensitive to body orientation.

However, Povinelli and Eddy's (1996b) study has been criticised for not truly capturing chimpanzee's natural tendency to beg for food. Kaminski et al. (2004) note that choosing to beg from one of two individuals is a complex task, requiring hundreds of training trials for chimpanzees to meaningfully participate in the experiment. Thus, they repeated Povinelli and Eddy's (1996b) paradigm with two crucial changes. Firstly, chimpanzees only had to beg from one human, where number of communicative signals indicated level of attentional understanding. Secondly, chimpanzees were not given training for this task to capitalise on their natural tendencies to beg for food. In their first experiment, they found chimpanzees begged significantly more when the experimenter's body was oriented to the front than the back, irrespective of whether the eyes were open (see Figure 1). In their second experiment chimpanzees also begged significantly more when the experimenter faced the front than the back. Crucially, if chimpanzees understand visual attention, then they should be sensitive to face orientation irrespective of body orientation. They found when the experimenter's body was oriented toward the front, chimpanzees begged significantly more when the head was also oriented toward the front. However, when the experimenter's body was oriented toward the back, chimpanzees begged equally whether the face was oriented to the front or back. Thus, chimpanzees seem to only be sensitive to face direction if the body is facing the front.

Figure 1

Stimuli and results from Kaminski et al.'s (2004) study.



Note. Left Experiment 1 Figure 2, Right Experiment 2 Figure 4.

The authors claim chimpanzees have a bivariate and hierarchical response to directional cues. In the case of their findings, if an agent is facing away then they do not have the ability to give food. Thus, begging to the agent whose face was to the front would not increase their chance of receiving food more than the agent whose face was to the back. However, sensitivity to face orientation when the agent is facing forward may influence receiving food as the agent is disposed to give it. As such, the authors claim chimpanzees are sensitive to being looked at, but only if body orientation indicates a high chance of goal success. One must take caution however to interpret the findings as chimpanzees understand seeing in terms of attention. They may instead reflect exploitation of behavioural regularities. For example, receival of food occurs more often if an agent's body is facing forward, and even more so if their head is too.

Competition studies have also been used to claim that chimpanzees understand seeing. Specifically, that they know whether a conspecific competitor's vision is being blocked by a barrier. For example, Hare et al. (2000; 2001) found that subordinate chimpanzees preferred to approach food that was occluded from a dominant chimpanzee rather than unoccluded food. Note, they did not prefer the hidden food in noncompetitive situations. Whilst this appears to demonstrate that chimpanzees know what another has seen, this does not distinguish between the dominant's gaze cues, general body direction, or various possible effects of the occluder. Thus, the minimum conclusion here is that chimpanzees are sensitive to whether another is facing a target object or not.

The picture that emerges from monkeys and apes is that they can follow gaze and can make use of body orientation to determine whether another is facing towards an object of interest. They however show little ability to make inferences from gaze cues.

Infant Gaze Following

Human new-borns prefer faces with direct gaze compared to averted gaze, as early as 2 days old (Farroni et al., 2002). Understanding direct gaze informs one about averted gaze: it indicates where or what someone is attending to (Shepherd, 2010), and that it could be potentially interesting or important. It also initiates social interaction through joint attention (Moore et al., 2014). Thus, gaze-following is fundamental for higher level social skills.

Infants' attention is directed by eye-gaze from an early age. For example, some studies have claimed infants to follow gaze as young as 3-months old (Scaife & Bruner, 1975; Hood et al., 1998). However, when accounting for looks to the other side and motion effects, it is evident children consistently follow gaze much later (Corkum & Moore, 1995; Brooks & Meltzoff, 2002; Moore & Corkum, 1998; Caron et al., 2002; Tang et al., 2024). Infants will follow gaze provided it is accompanied by a congruous head turn from 12-months. By 18-months, infants can reliably follow eye-gaze alone. This development occurs suddenly: around 15% of 15- to 16-month-olds successfully follow gaze alone, compared to 80% of 18- to 19-month-olds (Moore & Corkum, 1998). This suggests that infants are sensitive to eye-direction from 18-months but not before.

The spatial layout of infant, gazer, and target matters. Infants will follow head and eye direction as young as 6-months, providing the target was within their visual field, often focusing on the first object they see than what the gazer is looking at (Butterworth & Jarrett, 1991; D'Entremont 2000; Flom et al., 2004; Tang et al., 2024). This progresses to gaze-following towards targets in the periphery at 12-months, and targets behind the infant or occluders at 18-months (Butler et al., 2000; Caron et al., 2002; Butterworth & Jarrett, 1991; Flom et al., 2004; Deák et al., 2000; Tang et al., 2024). This finding is also true when the target is occluded from the infant's sight. Infants will look behind a barrier to see what the experimenter can see consistently at 18-months but not earlier (Moll & Tomasello, 2004). This suggests that 18-month-olds begin to represent line of sight as they demonstrate an understanding that gaze terminates at objects, or occluding barriers.

A common viewpoint is that gaze following is an important milestone in social cognitive development. For example, individual differences in gaze following are shown to be related to language acquisition (e.g. Carpenter et al., 1998; Morales et al., 2000; Brooks & Meltzoff, 2005), theory of mind (Charman et al., 2001) and emotion regulation (Morales et al., 2005). Taken together, gaze following may arise from a desire to perceive what others are looking at (Tomasello, 2001; Baron-Cohen, 1991; Bräuer et al., 2005) to fulfil the needs of higher-level socio-cognitive systems. This implies that infants who follow gaze go beyond simply using other's visual attention to locate interesting objects, to *thinking* about other's visual attention.

However, children do not have to think about the gazer's attention to be successful at gaze following. For example, infants may not follow the eyes of another because they believe there to be a communicative intent, but due to their natural tendency to attend to the eyes. De Bordes et al. (2013) tested this by assessing 20-month-olds looking behaviour when an agent made a gaze shift towards one of two targets. Crucially, before the agent made a gaze shift, infants' attention was drawn either to the eyes through mutual eye contact or flashing dots over them. It was predicted that if infants interpreted gaze shifts as communicative, then their first look would be to the correct object more often for mutual eye contact than flashing dots. However, they found infants equally followed gaze irrespective of how their attention was drawn to the eyes (mutual eye contact: 63%; flashing dots; 60%). In another condition where attention was drawn to the mouth instead of the eyes, infants' likelihood to look at the correct target significantly decreased. Together, these findings demonstrate infants' natural tendency to look at eyes promotes their gaze following behaviour, not their desire to interpret communicative intent.

Children also do not have to think about the gazer's attention to be successful at gaze following as they simply need to recognise the spatial compatibility of eye direction and potential areas of interest. Moore and Corkum (1994) propose infants may learn that following the gaze of others rewards them with an interesting sight more often than if they looked in the opposing direction. This reward reinforces gaze following behaviour. In support of this, Deák et al., (2014) analysed parent child interactions of 3- to 11-month-olds during a free play session. They found if the infant looked at their mother's face when she was looking at the object in her hand, their next look would be to that object. More importantly, they almost never looked at rewarding objects that were not congruent with the mother's gaze direction. Alongside computer simulations of reinforcement learning in gaze following (Jasso et al., 2012; Lewis et al., 2010), it is plausible infants are provided with adequate input to associate eye direction with potential areas of interest, without the need to attribute attentional states to agents.

Adult studies also show gaze following is driven by automatic attention orienting, with other spatial cues such as tongue and arrow direction eliciting similar cueing effects (Downing et al., 2004; Schmitz et al. 2024; Chacón-Candia et al., 2023). However, whilst adults appear to follow other directional cues equally, they do seem to process gaze as a specialised stimuli for social cognition. For example, Downing et al. (2004) presented participants with a spatial cueing task. In one condition, the eyes looked to the left or right and in another an extended tongue pointed to either side. They found that participants were cued to the location by the eyes and tongue equally. In a second experiment, they manipulated the predictability of target location such that it would appear in the uncued location four times more likely than the cued location. If cueing effects are merely spatial orienting, then both the eyes and tongue cues will be insensitive to probability manipulations. However, if the eyes are a special type of stimulus, then participants will be cued to the eye direction, but not to tongue direction. This is what they found. In the eyes conditions, participants were faster to respond to targets in the cued location than the more likely location, whereas in the tongue condition they responded faster to targets in the more likely location. Schmitz et al. (2024) used a similar method where arrow and gaze direction was 100% counter-predictive of target location and find the same pattern of results. This suggests that gaze direction is a special stimulus insensitive to top-down biases. Thus, by adulthood gaze following may involve more social processes than simple spatial compatibility effects.

Research has shown that children can follow others' head-and-eye turns from the age of about 12-months and eye turns along from 18- to 19-months. It is possible these abilities are present in some form much younger, particularly if the objects and gazer are simultaneously visible. However, for present purposes the question is simply whether and when children understand the significance of gaze.

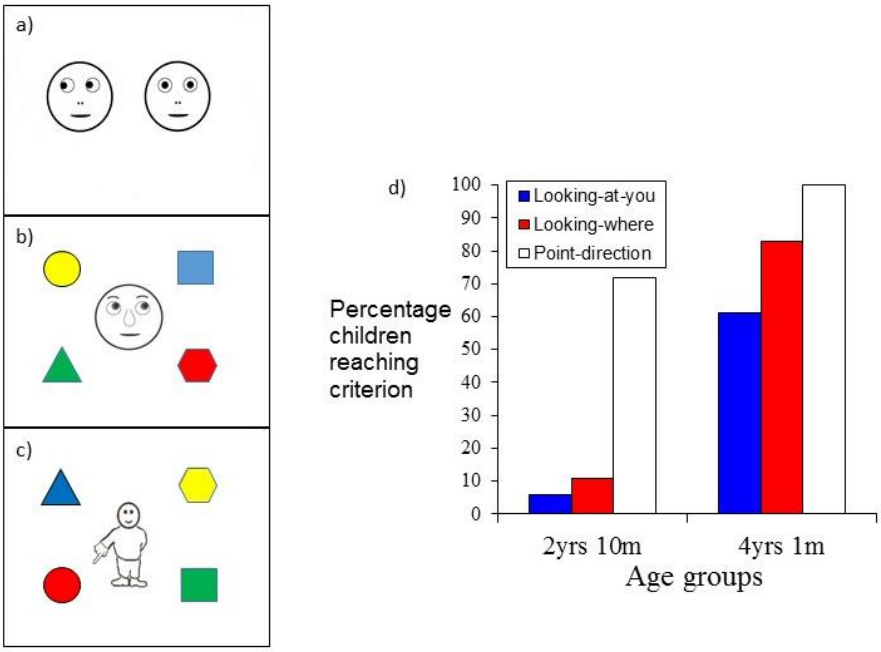
Development of Gaze-Judgement

Unlike non-human primates, one can often get useful data from preschool children simply by asking them questions. By following gaze, one's attention should be directed at the same location as the gazer, thus your view should inform about theirs. So, it would be fair to assume that a child that could follow gaze should be able to state the direction of gaze. However, this does not appear to be the case until 3- to 4-years of age (Doherty & Anderson, 1999).

Doherty and Anderson (1999) showed children images like those in Figure 2 Gaze direction was indicated by eye direction alone. They asked a) "Which one is looking at you?" or b) "Which one is Sam looking/pointing at?" Performance on the pointing control task (c) was good in younger children and perfect in older children. However, performance on the gaze tasks was poor among 2- to 3-year-olds, and not perfect in the 3- to 4-year-old group, as shown in the graph, d.

Figure 2

Example gaze stimuli (a, b, c) and children's performance (d) in Doherty & Anderson's (1999) gaze judgment study.



The authors ruled out trivial reasons for this finding. For example, when schematic faces were replaced with a real-life face, performance did not significantly differ (also true of photographs; Doherty et al., 2009). This suggests that 2- to 3-year-olds have difficulty judging gaze even when stimuli are highly salient. In their third experiment, they had a task where the head and eyes pointed toward the target. All children passed this task at near ceiling. This indicates that children had difficulty with eye-direction, and not comprehension of the task or question.

These were children who were expected to be able to follow gaze yet could not explicitly state another's eye-direction. The consensus view has been that the understanding underlying gaze following is conceptually rich. Butler et al. (2000) characterise this view nicely: following an adult's gaze indicates "that the infant understands … that there is a psychological and attentional relation between adult and target" (p. 360). Described in these terms, gaze following would effectively involve at least understanding attention as a representational mental state. That is that infants understand gaze as *referential*, where ‘all mental states are about or directed to some content (perceptible or representational) (Butler et al., 2000, p. 360). Such understanding should be sufficient for children to report where someone is looking, once they have the language to do so. The research described above demonstrates that they cannot.

A rich interpretation of infant gaze following could be reconciled with the findings if it is claimed that infant gaze understanding is implicit. In other words, children have knowledge about gaze direction in a form that is not available for verbal judgment. It could become available because of developments in the preschool period. Essentially, children are not learning new things about gaze direction. Instead, what they have already learned can be used more flexibly and in new ways. In contrast to this, Doherty et al. (2009) claimed that the new ability to make verbal judgments about eye direction at 3 years was entirely novel, not based on earlier gaze following abilities.

One way of distinguishing between the two possibilities is by examining the accuracy of preschool children’s gaze judgments. If children’s gaze judgments are based on a long-standing ability to determine where another is looking, then the difficulty is in using this knowledge verbally, not in accurately making the judgment. If the knowledge and ability is new, however, then like many new skills it is probable that initial accuracy is poor, and would improve thereafter (Miller et al., 2018).

Fine-grained judgement studies find support for the latter. Doherty et al. (2009) asked children to identify which of three rods the experimenter was looking at. The three rods were positioned to one side of the midline such that the experimenter could move her eyes, 10, 20 or 30 degrees to look at each rod. In another condition the rods were 15, 30 and 45 degrees to one side. They found 3-year-olds performance was at chance. Performance gradually improved over the 4- to 6-year-old period. Given that gaze discriminations are more accurate than 15-degree variations in adults, Vida and Maurer (2012a; 2012b) investigated when children reach adult-like performance. Children were told they needed to help an astronaut steer her spaceship towards the moon. If she was on course then she would be looking at the moon, but if she was off course then she would be looking to a space station either side of the moon. The children had to press a red button if the astronaut was looking at the red space station (right) or blue button if she was looking at the blue space station (left). The face would be looking towards either station at 1.6° increments. They found 6-year-olds could detect variations of around 1.94° gradually improving to adult-like detections of gaze variations (1.05°) at 14-years-old (Vida & Maurer, 2012a; 2012b). Thus, children become increasingly sensitive to eye-direction between 4- and 14-years-of age.

These results are not consistent with the idea that early implicit gaze understanding becomes available for explicit judgment or report. Instead, they claim that children's explicit gaze judgment ability is a new skill at 3 years. Initially their judgment is rather approximate, not sufficient to distinguish between differences of 15 degrees of visual angle. Over at least the following seven years, this skill becomes gradually better.

Another way of distinguishing between the two possibilities is by examining the differentiated use of luminance and geometric information from the eyes. Luminance information allows calculations of the ratio of light and dark regions of the eyes, whereas geometric information allows calculations of the sclera and iris edges. Doherty et al. (2015) proposed that young children have poor acuity and therefore will only have access to luminance cues, but older children would additionally have access to geometric cues. As such, gaze following should use luminance information only, whilst gaze judgements can additionally use geometric information. Doherty et al.'s (2015) findings support this claim: pre-school children follow the gaze of others when face stimuli retain luminance information but not when face stimuli retain geometrical information. However, they make judgements of gaze direction for face stimuli that retain either type of information. Thus, it is possible that the two abilities use different information in their processing. Evidence this dissociation persists into adulthood would be further support for two distinct abilities (see Chapter 4 for a discussion of the adult literature).

Thus, the evidence for human children demonstrates that younger children's gaze processing shows similar limitations to that of monkeys and chimpanzees: they can use others' gaze cues to orient their own attention to the object of the other's gaze but are not able to make judgments based on the gaze itself. Monkeys and chimpanzees seem never to overcome this limitation, whereas humans do. From about the age of 3-years, children become able to use the attentional relationship between observer and object to make judgements about gaze. The ability to do so appears to be a new skill as these judgments initially have low precision, which improves in a roughly linear fashion over the following few years. Gaze judgement additionally makes use of geometric information from the eyes which is not used for gaze following. I am not aware of any evidence that indicates continuity between the two abilities. This is the basis for the claim that there are two-systems for human gaze understanding (Doherty et al., 2009).

Relations Between Gaze Processing and Theory of Mind

With adults conceiving of attention in terms of representational mental states, children's understanding of representation is relevant to System 2 development. One such example is Theory of Mind. Whilst there is some debate, it appears to be a uniquely human skill (e.g. Premack & Woodruff, 1978; see Arre & Santos, 2021 for a review) that must be learnt and goes through protracted development (e.g. Wellman & Lui, 2004).

The term theory of mind was first coined by Premack and Woodruff (1978), who gave the definition "the individual imputes mental states to himself and others" (p. 515). Mental states include thoughts, beliefs and desires. The ability to attribute such mental states are important as they allow the prediction, explanation and manipulation of behaviour. The litmus test for this is the Maxi False Belief task (Wimmer & Perner, 1983). In short, children are told a story where Maxi places some chocolate into the green cupboard and then goes outside to play. Whilst he is absent, his mother takes the chocolate out of the green cupboard, uses some for her cake then puts in back in the blue cupboard and leaves to get some eggs. Maxi returns hungry and children are asked “where will Maxi look for the chocolate?”.

If children can attribute false beliefs, then they should respond with the green cupboard. This test has been replicated an astonishing number of times, with children consistently passing the task around 4-years-old (see Wellman et al., 2001 for a meta-analysis), around the same time as sophisticated gaze understanding develops.

This observation may come as no surprise. Gaze understanding is certainly a necessary part of proper understanding of mental states, since perception is the way knowledge, and beliefs are frequently formed. For example, in the case of Maxi, he forms a belief about the chocolate’s location by putting it there. His belief becomes outdated and therefore false by not seeing it moved elsewhere. As discussed above, most theorists would reason that gaze understanding is a precursor to theory of mind. In fact, Doherty and Anderson (1999) tested this. They compared children’s performance on False Belief and Looking Where tasks and found them to be equal difficulty. Performances were also associated ($r = 0.40$, $p < 0.05$) but this was no longer significant when controlling for age. This finding suggests that a concurrent development is more likely, or at least gaze judgement is a very immediate precursor.

The lack of association could be due to gaze judgements not requiring an understanding of misrepresentation, which the false belief task primarily tests. Arguably, any relationship would be with mental states that use visual attention as a core function of reasoning such as knowledge acquisition. To determine if someone knows the true state of affairs, children must understand that knowledge relates to experience. A simple way to view this is that seeing leads to knowing. It has been well established that children understand knowledge acquisition before they understand that beliefs can misrepresent (e.g. Wellman & Lui, 2004; Wellman et al., 2011; Peterson et al., 2012). Thus, gaze understanding and earlier forms of theory of mind may show stronger associations.

To my knowledge, only one study has investigated this. D'Entremont et al. (2012) presented 3- and 4-year-olds with a simplified gaze judgement task where they had to identify which of *two* objects an agent was looking at. Whilst 3-year-olds were better at this version than Doherty & Anderson's (1999) four object version, they did find a main effect of age where younger children performed significantly worse than older children. Alongside the gaze judgement task, children completed Wellman & Liu's (2004) theory of mind scale. They found successful gaze judgement was significantly easier than false belief understanding but was associated with earlier developing desire and knowledge understanding. From this, they claim that an understanding of desire, not belief, is necessary for understanding gaze direction.

Together, there are grounds to think understanding of mental states and gaze develop at roughly the same time. As children develop theory of mind, they become more interested in the cues that tell them about the contents of other's minds. One such cue is eye direction. It is effective at inferring the contents of other's minds as it plays a key role in knowledge and belief formation. Thus, children are motivated to attend to eye direction and the attentional relationship between observer and object. As such, rather than being a precursor to theory of mind, gaze understanding may be an integral part of its development.

Engagement. Fake It Until You Make It.

To be successful in the social world, children need to think about and respond to others behaviour on the basis of what others have attended to. Without the ability to make judgements of gaze or mental states, children have a problem that needs solving. Children may instead use a set of behavioural heuristics to fake it. For example, O'Neill (1996) found 2-year-olds were sensitive to whether their parents had attended to a particular event. In her study, 2-year-olds and an experimenter played with a toy, whilst the parent watched, read a magazine or waited outside the room. The experimenter then placed the toy on a high shelf and left the room. O'Neill found when requesting parents' help to retrieve the toy children communicate significantly more if the parent had not witnessed it being placed there compared to when they had. This has been said to indicate reasoning about knowledge states through perceptual understanding; the parent needed informing because they had not seen the object placed, so did not know where it was.

However, O'Neill argues that understanding of perception is not necessary, given that the concept seeing leads to knowing is emergent at 3- to 4-year-olds (O'Neill et al., 1992; Pratt & Bryant, 1990; Wimmer et al., 1988). Instead, she claims understanding general involvement – or *engagement* – in activities is sufficient. At the time of the child's request to retrieve the toy, the parent is fully engaged in the activity. The child's requests must therefore consider that their parents had been disengaged from the key event and needed to be updated. Thus, instead of tracking others' knowledge states children track their engagement.

Doherty (2011) formally defines engagement as "a relationship between a person and an object at its present location. The significance of the relationship is that another is likely to act on an object she is engaged with, and unlikely to act on an object she is not engaged with" (p. 312). Engagement persists over time until the object is moved from its current location initiating a disengagement event. There is an apparently equivalent concept "awareness relations" seen in monkeys (Martin & Santos, 2016). It is posited that monkeys compute the relations between agents and objectively true information – objects in their present location. As with engagement, the location of the object is considered a critical part of an awareness relation, which is "disrupted by *any manipulation* of the object's location when it is outside the agent's awareness" (Horschler et al., 2019, p. 73). Awareness relations are considered to be 'on' or ‘off’, so ‘disruption’ involves turning them off. Engagement has also either been established or not. Given the equivalence of these concepts, there is potentially a phylogenetically basic means of predicting the likelihood of an agent interacting with objects.

From experience, children infer that adults are more likely to engage with objects that they have previously had perceptual access to, typically established by broad general perceptual cues (eyes open, body posture, being there). In other words, an agent is potentially engaged with an object if it is front of them, and not occluded. These cues are easily monitored and indicate general attention over a period of time. However, Doherty (2009) notes that eye direction is transient in nature, and therefore not a good cue for engagement. For example, “a child’s mother may be fully engaged in a child’s pretend tea party but cast frequent glances towards the child’s younger sister asleep in the corner” (Doherty, 2009. pg. 123). Whilst the mother is attending to the younger sister *right now*, she remains engaged with the game. Thus, children are more likely to monitor cues that allow effective assessment of engagement, neglecting eye gaze as a specific cue.

There is some evidence that shows young children monitor broad cues. For example, 3-year-olds are more likely to monitor broad, stable cues and neglect eye gaze as their performance on tasks requiring judgement of eye direction is poor, but near ceiling on those with head direction (Doherty et al., 2009). Monitoring these cues may also explain children’s understanding of occlusion. Masangkay et al. (1974) showed children cards with a picture on either side. When placed upright, most 2-year-olds could say which picture they or the experimenter could see. Moll and Tomasello (2006) also found when observing an adult search for a desired object, 2-year-olds handed over the occluded object more often than chance. Thus, children seem to understand people cannot engage with things that are occluded.

Together, these findings suggest children do not have to understand how eye-direction or attentional implications maps onto the word *see*. They could be using a simple heuristic rule that adults are more likely to engage with things in front of people, that are not occluded. This heuristic is the foundation of *engagement*, with the development of gaze judgement marking the transition to an adult-like concept of *see*.

The transition from engagement to an adult-like understanding is particularly evident in a classic hiding task. Flavell et al. (1978) asked children to hide an object by either moving it behind an opaque screen, or by moving the screen in front of it. They found 2-and-a-half-year-olds could do the former, but not the latter. This result is surprising, as both tasks appear to require the same understanding – to occlude the object. The fact that the Move Screen task is difficult, suggests there are signature limits to this understanding.

McGuigan and Doherty (2002) replicated Flavell et al.'s (1978) findings and proposed when young children hear the word *see*, they interpret it as *engaged with*. However, *see* has a more specific meaning than general involvement. One can be engaged with things that one cannot see, such as through touch or anticipation of reappearance (e.g. a toy train reappearing from a tunnel). To succeed at the Move Object task, children do not need to distinguish between being engaged and seeing. At the beginning of the trial, the experimenter is not engaged with the object. Children understand that placing an object behind an occluder can prevent someone *becoming* engaged with it. This allows them to hide an object by placing it behind a screen.

On the other hand, the Move Screen task requires children to make a distinction between *being engaged with* and *seeing* the object. The trial begins with the experimenter already engaged with the object. Interposing the screen may stop the experimenter seeing the object but does not stop them being engaged with it. So, when asked to put the screen so that the experimenter cannot *see* (become engaged with) the object, children do not know how the screen could disrupt this engagement relation. Thus, they do not know where to place it. This explanation plausibly reflects children's social observations. They may have frequently seen others not react to objects that are occluded. Whereas if one object becomes occluded by another, people involved do not usually react as if the object is no longer there.

Understanding gaze direction may be a key development allowing children to distinguish between seeing and being engaged with. McGuigan and Doherty's (2002) findings are consistent with this claim. They demonstrate children's success on the Move Screen strongly associated with success on gaze judgement tasks after controlling for age. Thus, the two hiding tasks require different understanding of attentional relationships. Engagement is sufficient for the Move Object task, but sophisticated understanding of gaze is required for the Move Screen task.

In summary, individuals that do not have access to System 2 such as monkeys and infants still need to consider what others have attended to, to respond to their behaviour. Without the necessary concepts of seeing, they require a solution that allows them to fake it. Here it is proposed they use a concept of engagement which relies on a set of behavioural heuristics to determine another's general involvement with objects. Older children's sophisticated understanding of eye direction marks the transition from engagement-based strategies to predict behaviour to a mentalistic adult-like concept of *see*. This aligns with the proposal that understanding eye direction is an integral part of thinking about mental states.

The Problem of Infant Theory of Mind

The discussion so far presents a neat developmental timeline. Children consistently follow gaze from 18-months-old. When they develop a theory of mind around 4-years they become successful at making judgements of gaze direction. Until then, children use engagement-based strategies to predict behaviour.

A challenge for this account is that infants appear to demonstrate capacity to attribute mental states. For example, children are told a story about a mouse who places his cheese in a box and then goes to sleep. When he is asleep, the cheese is moved to a different box. After hearing the mouse express its desire to find the cheese, 3-year-olds look at the correct location (Clements & Perner, 1994; Garnham & Perner, 2001; Garnham & Ruffman, 2001). However, when they are then explicitly asked where the mouse will look for the cheese, the same children incorrectly say the location that it actually was. Southgate et al. (2007) claim that the verbal nature of the task limits children's ability. Specifically, they claim the verbal prompt "I wonder *where* he's going to look" is prematurely interpreted as referring to the actual location of the object. Thus, they adapt the task to be non-verbal and find children as young as 2-years can correctly look in anticipation for the protagonist's action. The youngest demonstration of false belief attribution is in 15-month-olds. Onishi and Baillargeon (2005) showed infants a sequence of events where an agent views an object placed in a box. The agent's view is then blocked whilst the object is moved to another box. Finally, the agent reaches into one of the boxes. They found infants looked longer when the agent reached into the box where the object actually was, than in the box the agent believed the object to be in. Thus, infants appear to be aware of an agent's false belief even if they cannot explicitly express it.

These findings ultimately changed the theoretical landscape of the theory of mind literature. Recently, however, there have been numerous failures to replicate (e.g. Grosse Wiesmann et al., 2017; Grosse Wiesmann et al., 2018; Kampis et al., 2021; Dörrenberg et al., 2018). Notably, Kulke and Rakoczy (2018) examined 65 published and un-published replication attempts and classified them as successful, partial, or non-replication. Overall, successful replications (26) were outnumbered by non-replications (17) and partial replications (22). To this end, there is currently an ongoing large-scale, *a priori* designed multi-lab study (Many Babies Project 2; Schuwerk & Kampis et al., accepted pending data collection) to test for the replicability of infant theory of mind measures.

With this in mind, the evidence thus far points towards children having a theory of mind around 4-years, but not before. However, how can infant looking behaviour that is consistent with belief understanding be explained? One view is that infants do understand belief, and that failures on explicit tasks are due to the complex demands of language and executive function (Onishi & Baillargeon, 2005; Leslie, 2005; Baillargeon et al., 2010). Opposing views suggest that infant success does not require belief attribution, but use of behavioural regularities such as “people look for things where they last saw them” (Perner & Ruffman, 2005, p.214; Ruffman & Perner, 2005, p. 462).

To resolve the two views, Apperly & Butterfill (2009) proposed infants use minimal theory of mind to correctly anticipate behaviour. That is, infants do not have to represent beliefs as beliefs. Instead, they can represent the relation between an agent and an object. To construct a minimal theory of mind, the agent must first *encounter* an object. This is defined as “a relation between the individual, an object and a location such that the relation obtains when the object is in the individual’s field” (Apperly & Butterfill, 2009, p. 33). In this sense, encountering is analogous to perception; the object needs to be within close proximity, in front of them and not occluded. If an agent has encountered an object, they can then form a *registration* towards it. This is “a relation much like encountering except that it continues to obtain even after an object is no longer in one’s field” (p. 34). This is analogous to belief; agents are likely to perform goal-direct actions towards objects in the location they register it in, much like they would towards the location they believe it to be.

Described in these terms, minimal theory of mind and explicit theory of mind rely on the same output: “Maxi last *saw/encountered* the chocolate in the green cupboard, thus *believes/registers* it is there.” Butterfill and Apperly (2013) claim that the distinction between the two is that explicit theory of mind works on propositional attitudes, in which situations can be described by different propositions. For example, if the chocolate in the drawer is Maxi’s birthday present, Maxi’s description of the chocolate’s location would differ if he knew this fact compared to if he did not. Registrations do not rely on propositions because they are simply relations between an agent and object, under no particular description. It does not matter whether Maxi describes the present as his or not, just that he has a relation with the present in the drawer. This distinction is evident by young children's difficulty understanding aspectuality, where success converges with that of explicit theory of mind (e.g., Fizke et al., 2017; Low & Watts, 2013; Low et al., 2014; Oktay-Gür et al., 2018; Rakoczy et al., 2015). However, this is beyond the scope of the current thesis and need not be pursued further.

What is relevant to the current thesis is the overlap between minimal theory of mind (Apperly & Butterfill, 2009) and engagement. Engagement is similar to registration but with a key difference the relation between the agent and object persists until *the object is moved*. As such, engagement only persists while the situation remains the same. When the object is moved in the absence of perception, the engagement relation ends. Thus, the engagement hypothesis makes different predictions for the False Belief task. That is Maxi has an engagement relation with the chocolate in the green cupboard. When the chocolate it moved the blue cupboard, the engagement relation ends. Without such a relation to refer to, children are left with no expectation of Maxi's actions. So, when they are asked to identify where Maxi will look, they favour the location that confers success despite it being the incorrect response. Thus, engagement may better explain 3-year-olds failures at explicit False Belief tasks than minimal theory of mind.

Engagement may also account for 2- to 3-year-olds apparent success on some implicit False Belief tasks. For example, in Southgate et al.'s (2007) anticipatory looking task, an agent sees a ball placed in a box, establishing an engagement relation with it in that location. Then the agent becomes distracted and turns away. At this point, the ball is removed from the scene entirely resulting in the engagement relation ending – a disengagement event. Doherty (2011) notes "aspects of the scene an agent is disengaged from remain salient" (p. 319). Given that the ball no longer exists in the situation, the only salient location is the location of disengagement. Therefore, children look in interest at the correct location without the need to attribute beliefs.

Ultimately, infant theory of mind studies do not pose a challenge to the notion that children use a concept of engagement for behaviour prediction in the preschool period. Whilst there have been three common explanations for infant theory of mind findings, they often do not resolve the conflicting outcomes across task variations or explicit failures. Here I propose young children's use of engagement account for this. Given engagement-based strategies may allow children to fake mental state understanding as well as gaze direction, it is further plausible that the development of these two abilities are intwined in development.

Conclusion

The aim of the current review was to discuss the literature on children's gaze understanding abilities and consider its association with theory of mind developments. In doing so, motivation for a two-system theory of gaze processing is evaluated. That being we have one system for locating the objects of others' attention, and one to think about the attentional relationship between observer and object. The latter is uniquely human and must be learnt, potentially developing as a result of children's increasing interest in others' mental states.

To summarize, the evidence shows that young children, monkeys, and chimpanzees can follow the gaze of another to orient their attention to objects of interest. However, they do not demonstrate the ability to make judgements or inferences based on gaze itself. From about the age of 3-years, children begin to understand the significance of gaze, with increasing proficiency until 14-years-old. This suggests gaze judgements are a new skill and not merely a progressional development from earlier abilities. As such, the abilities can be considered distinct, the basis of the claim that there are two-systems for human gaze understanding.

Gaze judgements develop at roughly the same time as understanding of mental states. It is plausible that understanding gaze direction is an integral part of theory of mind, given that the eyes are effective at inferring the contents of another's mind. However, until children develop either ability, they need a way to predict the behaviour of others based on what they have attended to. Thus, infants who do not have access to System 2 may use a concept of engagement (or awareness relations in non-human primates), which relies on a set of behavioural heuristics to determine another's general involvement with objects. It can plausibly explain signature limits in children's understanding of occlusion and conflicting outcomes in the theory of mind literature. As such, earlier conceptions of attention do not require mental state concepts. Once children develop a sophisticated understanding of eye direction, they transition from engagement-based strategies to predict behaviour to mentalistic concepts of seeing.

This review generates the following research questions for the current thesis:

1. Do young children predict behaviour using a concept of engagement?
2. Is engagement a way for young children to fake gaze and mental state understanding?
3. Is understanding of gaze direction integral to theory of mind development?
4. Does System 1 remain luminance based in adulthood?

Experiment 1 will primarily investigate research question 2. Using a cross-sectional sample, children will partake in a modified Move Screen task involving blindfolds to separately manipulate seeing and being engaged with. If the engagement hypothesis is supported, then the manipulation should see children's performance improve on the Move Screen task compared to the original. This experiment will also consider research question 3, by comparing performance on gaze judgement, theory of mind and Move Screen tasks. If children transition from engagement-based strategies to mentalistic ones to predict behaviour, then as their performance on gaze judgement and theory of mind tasks improves, so should their Move Screen performance.

Experiment 2 will primarily investigate research question 1. Using a longitudinal design, children will partake in gaze judgement and theory of mind tasks to determine the nature of their association. The aim is to discern whether the data reflect the traditional view of gaze judgement as a precursor to theory of mind, the currently proposed view of gaze judgement being integral to theory of mind development, or merely a coincidental contemporaneous association. This experiment will also investigate research question 3, by considering the cross-lagged effects between gaze judgement, theory of mind and Move Screen tasks. If children transition from engagement-based strategies to mentalistic ones to predict behaviour, then earlier performance on gaze judgement and theory of mind tasks should predict later Move Screen performance.

Finally, Experiments 3 and 4 will investigate research question 4. If there are in fact two distinct systems for human gaze understanding, then it is possible that they will remain distinct in adulthood. Specifically, adult System 1 should be identified by its sole use of luminance information like it is in childhood. Thus, these experiments aim to use attention orienting tasks to consider whether adult System 1 uses luminance information from the eyes more readily than geometric information.

Chapter 2

Testing the Engagement Hypothesis

The current chapter aims to investigate whether young children predict behaviour using a concept of engagement outlined in Chapter 1. As previously discussed, engagement may explain some of the signature limits to children's understanding of occlusion in a classic hiding task (Flavell et al., 1978; McGuigan & Doherty, 2002). That being in the Move Object version of the task, children understand that placing an object behind an occluder can prevent someone *becoming* engaged with it. This allows them to successfully hide an object by placing it behind a screen. However, the Move Screen version of the task begins with the other person already engaged with the unoccluded object. Interposing an occluder does not alter that. Thus, young children do not know how the screen could disrupt this engagement relation, so do not know where to place it.

The account has a straightforward test. In the Move Screen condition, if the experimenter has not experienced the object in its location, and is thus not yet engaged with it, the child should be better able to move the screen. One way to do this is to use a blindfold with a second experimenter placing the object after the first experimenter has been blindfolded. In this condition children should find it easier to correctly place the screen.

Research with blindfolds show that 12-month-olds will follow the head turns of blindfolded agents, but do not do so if they themselves have experienced being blindfolded (Meltzoff & Brooks, 2008; D'Entremont & Morgan, 2006). By 18 months infants do not need to have experienced the blindfold. Minimally this shows that infants do not follow the gaze of blindfolded people. Note that predicting the behaviour of people wearing blindfolds does not presuppose an understanding of seeing. More simply it may result from understanding that blindfolds prevent people becoming engaged with objects.

Another, less intuitive test derives from work with non-human primates. Monkeys have been posited to compute "awareness relations", relations between agents and objectively true information (Martin & Santos, 2016). The concept of awareness relations appears equivalent to that of understanding engagement. The 'objectively true information' concerns objects in their present location, as for the definition of engagement above. Location is also considered a critical part of an awareness relation, which is "disrupted by *any manipulation* of the object's location when it is outside the agent's awareness" (Horschler et al., 2019, p. 73). Awareness relations are considered to be 'on' or 'off', so 'disruption' involves turning them off. Engagement has also either been established or not.

To examine the importance of location, Horschler et al. (2019) examined monkeys' reactions to human behaviour after an object was very briefly moved from its location. Monkeys observed a human viewing an object enter one of two boxes, after which a screen blocked the human's view. The object was then either briefly moved out of the box mechanically along a small track then returned, or the box opened to reveal the object then closed. The screen was removed and the human either reached into the box with the object or the other box. If the object had not moved monkeys looked longer when the human reached into the empty box. Since the human had an awareness relation to the object in its location they could be expected to reach effectively. When they did not the monkeys behaved as if surprised. However, when the object had been briefly removed and replaced in the box, monkeys looked equally long regardless of which box the human reached into. They acted as if no relation existed. The conclusion was that the awareness relation had been disrupted by the removal of the object, so was effectively set to 'off'.

Given the equivalence of the two concepts, the same prediction can be made for engagement. Thus, in the present study there is a second condition in which the object is present in front of the experimenter before she is blindfolded. Once blindfolded the object is briefly removed and quietly replaced. Again, it is predicted that in this condition children should find it easier to correctly place the screen.

To control for other potential effects of a blindfold, a control condition is included in which the object is present in front of the experimenter before she is blindfolded but is not manipulated before the child is asked to move the screen. In this case the experimenter is already engaged with the object in its location, and nothing disrupts that. Moving the screen should be harder than in either experimental condition.

Understanding gaze direction may be a key development allowing children to distinguish between seeing and being engaged with. Specifically, McGuigan and Doherty (2002) compared performance on the Move Screen task with judgement of where a depicted viewer was looking. They found children succeeded at both tasks around the same time, and performances were correlated beyond common associations with age. Understanding occlusion and eye direction both plausibly involve projecting a line of sight from the eyes to an object. The Move Screen task additionally requires moving the screen to intersect that line of sight. Thus, the two hiding tasks require different understanding of attentional relationships. Engagement is sufficient for the Move Object task, but sophisticated understanding of gaze is required for the Move Screen task. The current study compares the two abilities.

Theory of mind is also relevant because engagement performs limited theory of mind functions. It allows behavioural prediction based on agents' involvement with events, approximately tracking knowledge of object location. However, it does not make proper distinctions between knowledge and perception required in the Move Screen task. Thus, it is predicted that success on this task marks the development of a more adult-like conception of knowledge, and thus the task will associate with theory of mind measures of knowledge understanding. The False Belief task is included as a well-validated measure of theory of mind development. However, this requires understanding of misrepresentation, and is expected to be harder than the Move Screen task.

Finally, the relation between theory of mind and gaze judgement is also of interest. It has frequently been speculated that understanding others' gaze is a precursor to theory of mind (Goméz, 1996; Gopnik et al., 1994). However, children appear to be able to judge eye-direction at a similar age to the ability to understand false belief. Doherty and Anderson (1999) found the two abilities were not different and were associated, although not significantly so once age had been taken into account. D'Entremont et al. (2012) found successful gaze judgement was significantly easier than false belief understanding but was associated with performance on a theory of mind battery including measures of understanding desire and knowledge access. It is predicted that gaze understanding will associate with understanding of knowledge.

Method

Participants

Participants were 172 children (93 girls) divided into three groups for analysis: 59 2-year-olds (*M* = 32 months, *SD* = 3 months, *range* = 24 – 37 months), 55 3-year-olds (*M* = 41 months, *SD* = 2 months, *range* = 38 – 44 months), and 58 4-year-olds (*M* = 48 months, *SD* = 3 months, *range* = 45 – 54 months) recruited from predominantly working-class preschools in Norwich, UK between 14th July 2022 and 2nd August 2023.

Inclusion criteria were informed parental consent, and child assent immediately prior to testing. No child met the exclusion criterion of a teacher- or parent-indicated special-needs diagnosis. The stopping criterion was all available children at contacted schools had been tested once planned sample size was achieved.

The experiment was carried out in accordance with ethical standards at University of East Anglia Research Ethics Committee (ETH2122-206, ETH2223-0400).

Statistical Analysis Plan

The experiment was originally planned as separate studies, so was preregistered in two parts (107365, 119091). However, circumstances allowed them to be run simultaneously and they differed only in the Engagement task. Thus, the original planned participant numbers were retained, but the data is analysed as a single study for greater statistical power.

For all planned analyses the G-Power 3.1.9.2 statistical tool (Faul et al., 2007) was used to estimate sample sizes required to achieve a power of 0.80 with two-tailed alpha set to 0.05. To estimate planned non-parametric analyses included Friedman's and Kruskall Wallis not covered by G-Power the parametric equivalent tests were used, increasing N by 15% (Lehmann, 2006).

Exactly half of children received the Blindfold task and half the Blindfold Remove and Replace task. Each subgroup took the same version of all other tasks. Subgroups were matched for age and drawn from overlapping clusters of local preschools; 69 children were drawn from schools that contributed to both subgroups and 103 from schools that contributed to only one or other subgroup. As a check on potential sample differences, performance on all tasks between subgroups were compared. There were no differences on Blindfold or Blindfold Remove and Replace tasks themselves, nor on any of the gaze judgement or theory of mind tasks (all *p*s > .124). However, the Blindfold Remove and Replace group performed better than the Blindfold group on the Move Screen (p =.006) and Blindfold Control (p = .05) tasks only. Thus, I cautiously analyse performance on the hiding tasks separately by subgroup using Friedman's test. Using a repeated measure within factors ANOVA as proxy, medium effect size of η^2 = .25, 1 group and 4 measurements, N = 28. Wilcoxon was used for post-hoc planned comparisons, medium effect size d = .50, N= 35. Thus, both subgroups were appropriately powered for this analysis.

Kruskal-Wallis was used for secondary planned analyses of age groups. For hiding task age effects this was analysed by subgroup, and for gaze judgement and theory of mind age effects the whole sample was analysed. Using one-way ANOVA plus 15% as proxy, medium effect size of $\eta^2 = .25$ and 3 groups, $N = 183$. Mann-Whitney was used for post-hoc planned comparisons, medium effect size of $d = .50$, $N = 134$. Thus, the hiding task subgroups and the whole sample were underpowered for this analysis.

Correlation analysis was used to analyse shared competencies between Move Screen, gaze judgement and theory of mind tasks, medium effect size $r = .30$, $N = 84$. Thus, the whole sample was appropriately powered for this analysis.

All non-parametric comparisons were checked with the equivalent parametric test (see Rasch & Guiard, 2004 for simulation-based evidence of robustness). The study sample size exceeds that of any study of its kind (see Wellman & Liu, 2004; McGuigan & Doherty, 2002).

Design

Tasks were blocked by type, with block presentation fully counterbalanced. For hiding tasks, 12 of 24 possible permutations were randomly chosen. Half of participants received the Blindfold condition and half the Blindfold Remove and Replace condition. Looking-where and Looking-at-you gaze judgement tasks were alternated beginning with a Looking-where task. The Knowledge Access task was presented before the False Belief task, following Wellman and Liu (2004).

Materials and Procedure

Testing was in a quiet, familiar location for around 15 minutes. The first author conducted the theory of mind tasks and the hiding tasks. A second female experimenter conducted the gaze judgement tasks and was the observer in the hiding tasks. The first author sat beside the child at a low table. The second experimenter sat on the floor opposite the child with her eyes level with the child's.

Hiding Tasks

Tasks were adapted from McGuigan and Doherty (2002). Materials were a free-standing cardboard screen (12 by 17cm), an elephant plush toy (8 x 9cm), and a blindfold (sunglasses with blacked out lenses) as shown in Figure 3. Blacked out sunglass were used in favour of traditional blindfolds to avoid potential ethical issues with blindfolding children in the practice trials. Additionally, this experiment was conducted soon after COVID-19 restrictions had fully lifted, meaning the apparatus needed to be wipeable between participants. Blacked out sunglasses were also used in favour of classical methods such as eye closure because from experience when asking children of this age range to close their eyes they will often peek and may therefore may not believe Experimenter 2 had fully closed their eyes.

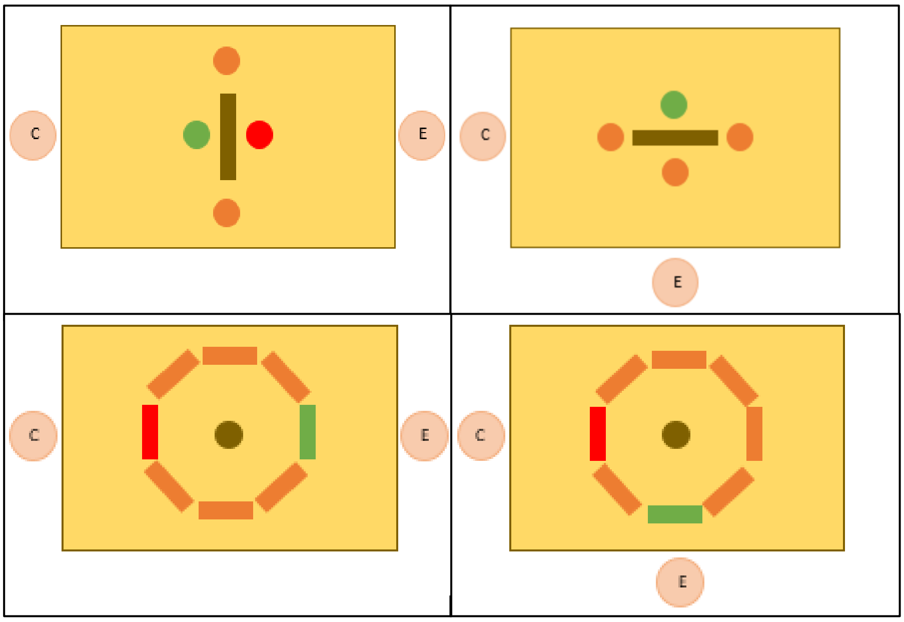
Each task had two trials. Experimenter 2 sat opposite (180°) or to the left or right (90°) of the child as shown in Figure 4, order counterbalanced. It is possible that if two 180° trials were presented, children may perceive the second trial as an indicator that their response for the first trial was incorrect. As such, children may artificially alter the response they would give. Thus, despite McGuigan and Doherty (2002) showing children find 90° trials harder than 180° trials, the current study retained one trial of each instead of two 180° trials.

Figure 3

Experiment 1 Hiding Task Stimuli

Figure 4

Hiding Task Set Up and Scoring Examples



Note. Green = Correct, Red = Egocentric, Orange = neutral, and any response not depicted = other. 180° Move Object (top left), 90° Move Object (top right), 180° Move Screen (bottom left) and 90° Move Screen (bottom right).

Move Object task. Experimenter 2 sat directly opposite the child (180°). The first experimenter placed the screen on the table, broadside to Experimenter 2, who fixed her gaze at the screen. Experimenter 1 handed the child the toy and said, "put Nellie somewhere on the table so [Experimenter 2] can't see her". The trial ended when the child had placed the object on the table and was no longer touching it.

The correct response was placing the object behind the screen out of view of Experimenter 2. An egocentric response was placing the object out of the child's own sight. This was only possible on 180° trials. A neutral response was placing the object near the screen, but in view of both Experimenter 2 and the child. Responses coded 'other' were making no attempt to occlude the object, for example, placing the object toward the extremities of the table.

Move Screen task. The Experimenter 1 placed the object in the centre of the table. Experimenter 2 fixed her gaze at the object. Experimenter 1 handed the screen to the child and said, “put this [pointing at the screen] somewhere on the table so [Experimenter 2] can’t see Nellie”. The trial ended when the child had placed the screen on the table and was no longer touching it.

The correct response was putting the screen between the object and Experimenter 2. An egocentric response was putting the screen between the object and themselves. This was possible on 180° and 90° trials. A neutral response was putting the screen upright near the object such that the object remained visible to both Experimenter 2 and the child. Responses coded ‘other’ were failing to occlude the object, for example, placing the screen flat or toward the extremities of the table.

Blindfold tasks.

Familiarization Phase. At the beginning of the block children were shown the blindfold and invited to wear it. The child was then asked, “what can you see?”. If the child did not respond or gave an answer of “dark” or similar, they were then prompted with “can you see anything?”. All children acknowledged they could not see anything. Experimenter 1 then told the child that “when [Experimenter 2] wears the blindfold, they won’t be able to see”.

The test question was asked at the end of each task, after Experimenter 2 had donned the glasses. The child was told in a whisper: “In a minute we will ask [Experimenter 2] to take her glasses off. Put this (pointing at the screen) somewhere on the table so that when [Experimenter 2] takes her glasses off, she can’t see Nellie”.

Once the child had placed the screen on the table and was no longer touching it, Experimenter 2 removed the blindfold. If Experimenter 2 could see Nellie she said, “there’s Nellie” pointing at the object. If not, she said, “where’s Nellie?”. Responses were scored using the same criteria as the Move Screen task.

Participants completed the Blindfold Control task and either the Blindfold or the Blindfold Remove and Replace:

Blindfold. Experimenter 2 fixed their gaze toward the centre of the table, put on the blindfold and exclaimed “I can’t see”. Experimenter 1 then quietly whispered to the child “shh, let’s put Nellie on the table”, quietly placing the object in the centre of the table. Experimenter 1 then handed the screen to the child and asked the test question.

Blindfold Remove and Replace. Experimenter 2 fixed their gaze toward the centre of the table and Experimenter 1 placed the object there. Experimenter 2 then put on the blindfold, exclaiming “I can’t see”. Experimenter 1 then quietly whispered to the child “shh, let’s take Nellie away and put her back”. Ensuring the child was watching, she removed the object from the table to position it by her side. It was immediately replaced in its original location, not on an alternative surface and did not leave Experimenter 1’s hand. Experimenter 1 then handed the screen to the child and asked the test question.

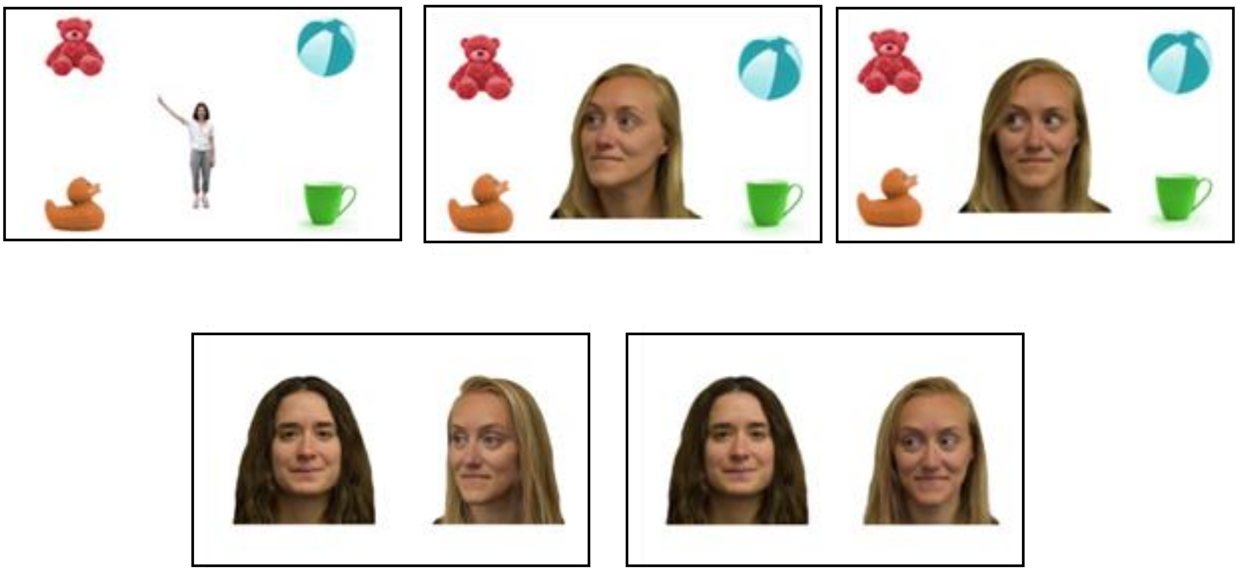
Blindfold Control task. Experimenter 2 fixed their gaze toward the centre of the table. Experimenter 1 placed the object in the centre of the table. Experimenter 2 put on the blindfold then exclaimed “I can’t see”. Experimenter 1 handed the screen to the child and asked the test question.

Gaze Judgement Tasks

There were two versions of each gaze task, one with viewer facing forward (Eyes only) and one where viewers’ heads pointed in the same direction as the eyes (Head turn). Viewers were two adult females. Example stimuli are shown in Figure 5. All images were on 21 by 29.7cm paper sheets. Children responded by pointing.

Figure 5

Experiment 1 Gaze Judgment Task Stimuli



Note. Point Where (top left), Looking Where Head (top middle), Looking Where Eyes (top right), Looking At You Head (bottom left), and Looking At You Eyes (bottom right).

Each task had four trials. Because the Looking Where tasks have four potential targets and Looking At You tasks have two potential targets the chance baseline differs. To facilitate comparison, criterion success for Looking Where tasks was at least three correct responses (p = .0508) and Looking At You tasks was four correct responses (p =.0625).

The Looking At You tasks showed two viewers, one looking forward and one with gaze averted 25°. All four combinations of left-right position of the two viewers and left-right direction of averted gaze were included. Children were asked "Which one is looking at you?".

The Looking Where tasks showed a viewer in the centre looking at one of four peripheral objects (mug, teddy bear, rubber duck and beach ball). Object colours (red, orange, yellow, green, blue, and purple) were adjusted using Gimp 1.10.30 to approximately equivalent hue. Object position and colour were varied across 24 possibilities to control for colour or object preferences and pseudo-randomly sampled into 6 sets so that the same pattern of object colour and position did not occur twice for a given participant. Children were asked “Which one is Emma looking at?”.

The Point Where task showed a woman in the centre of the image pointing at one of four objects, position and colour identical to those in the Looking Where tasks. The child was asked “Which one is Emma *pointing* at?”.

Theory of Mind Tasks

Knowledge Access. This task was adapted from Pratt and Bryant (1990). A short story was acted out with two Playpeople dolls (5cm), a penny and a box (6 x 8cm), as shown in Figure 6. Polly lifted the box and put it back down. Sam then opened the box, looked inside and closed the box. Children were asked in sequence:

Knowledge access test question: “Who knows what is in the box?”

Memory question: “Who looked in the box?”

Children passed the Knowledge Access task if they answered both questions correctly.

Figure 6

Experiment 1 Theory of Mind Task Stimuli

Note. Knowledge Access (left) and False Belief (right).

False Belief task. A short story was acted out with two Playpeople dolls (5cm), a ball, an opaque jar (6 x 4cm), and a box (5 x 5cm), as shown in Figure 6. Sally placed a ball in the box and left. Tony then moved the ball to the jar. Sally returned and children were asked the following questions:

False belief test question: "Where will she look first for her ball?"

Reality question: "Where is the ball really?"

Memory question: "Where did Sally put the ball in the beginning?

Children passed the False Belief task if they answered all three questions correctly.

Results

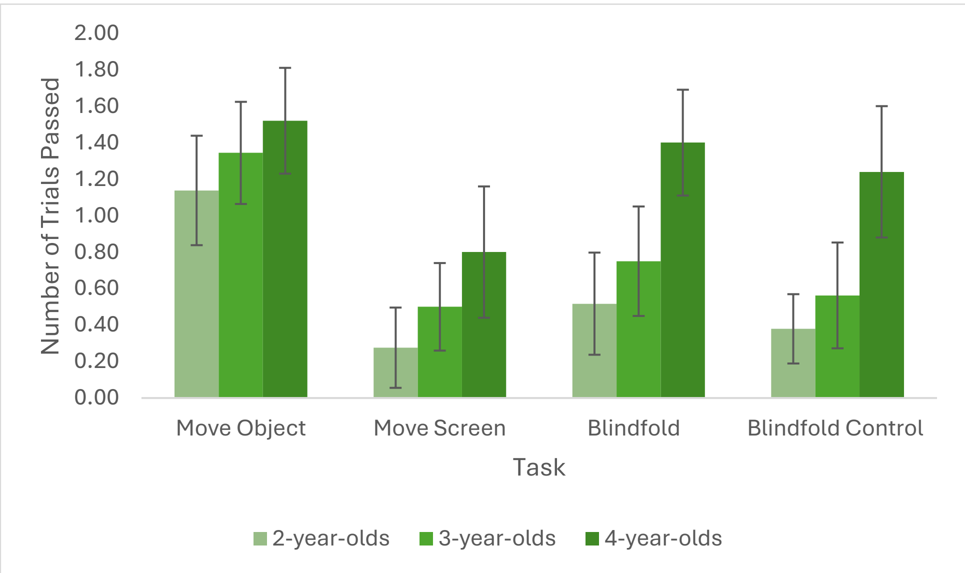
Age Effects

Hiding Tasks Blindfold Subgroup

Figure 7 shows performance on each hiding task improved with age (Move Screen: Kruskal-Wallis $H(2) = 6.57$, $p = .037$; Blindfold: $H(2) = 15.86$, $p < .001$; Blindfold Control: $H(2) = 14.30$, $p < .001$). Blindfold and Blindfold Control task performances improved significantly between the 3-year-old and 4-year-old group (respectively; $p = .004$, $p = .005$). Move Screen task performance improved between the 2-year-old and 4-year-old groups ($p = .013$) but not between the 3-year-old and 4-year-old group ($p = .202$).

Figure 7

Experiment 1 Performance on Hiding Tasks By Age, Blindfold Subgroup



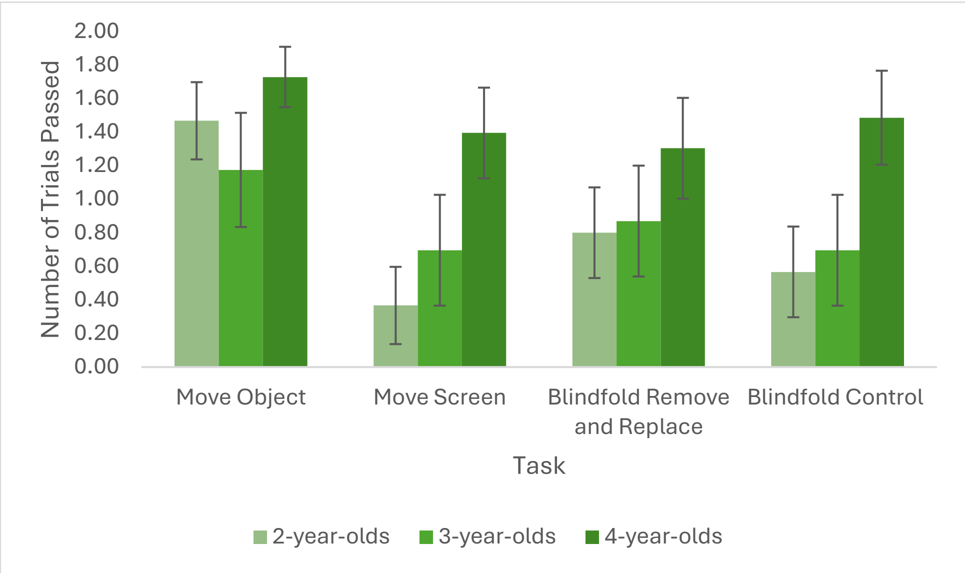
Note. Error bars represent 95% CI.

Hiding Tasks Blindfold Remove and Replace Subgroup

Figure 8 shows performance on each hiding task improved with age (Move Screen: Kruskal-Wallis $H(2) = 24.76$, $p < .001$; Blindfold Remove and Replace: $H(2) = 20.11$, $p < .001$; Blindfold Control: $H(2) = 7.19$, $p = .028$). All tasks showed significant improvement between the 3- 4-year-old groups (all *ps* < .046).

Figure 8

Experiment 1 Performance on Hiding Tasks By Age, Remove and Replace Subgroup



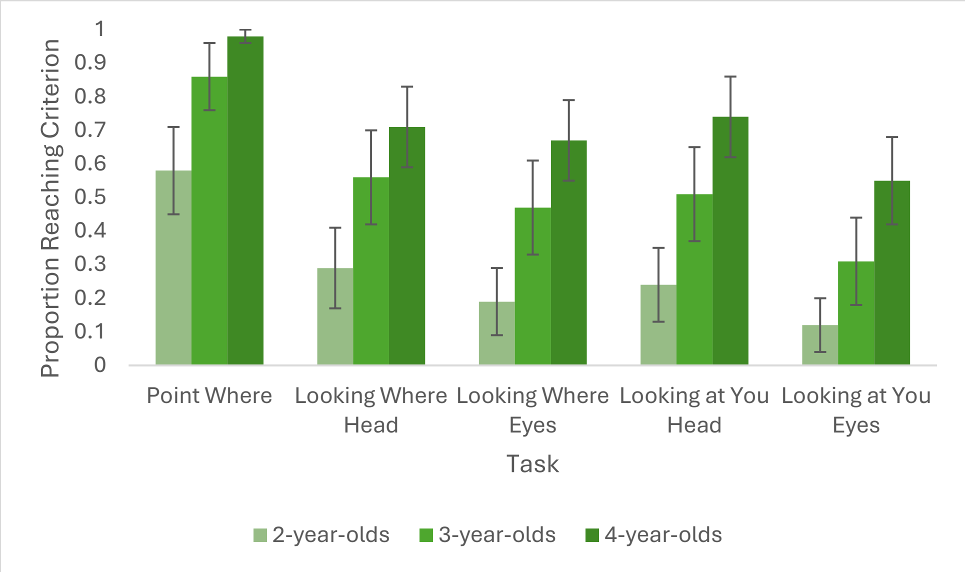
Note. Error bars represent 95% CI.

Gaze Judgement Tasks

Figure 9 and Table 1 show performance on the Point Where task approached ceiling in 3-year-old and 4-year-old children. For all tasks there was a main effect of age (all *ps* <.001). There were significant improvements between all age groups (all *ps* < .033) except for the Looking Where Head which did not differ between the 3-year-old and 4-year-old group (p = .115).

Figure 9

Experiment 1 Performance on Gaze Judgment Tasks



Note. Error bars represent 95% CI.

Table 1

Experiment 1 Mean (SD) Proportion Gaze Judgement Task Performance by Age Group

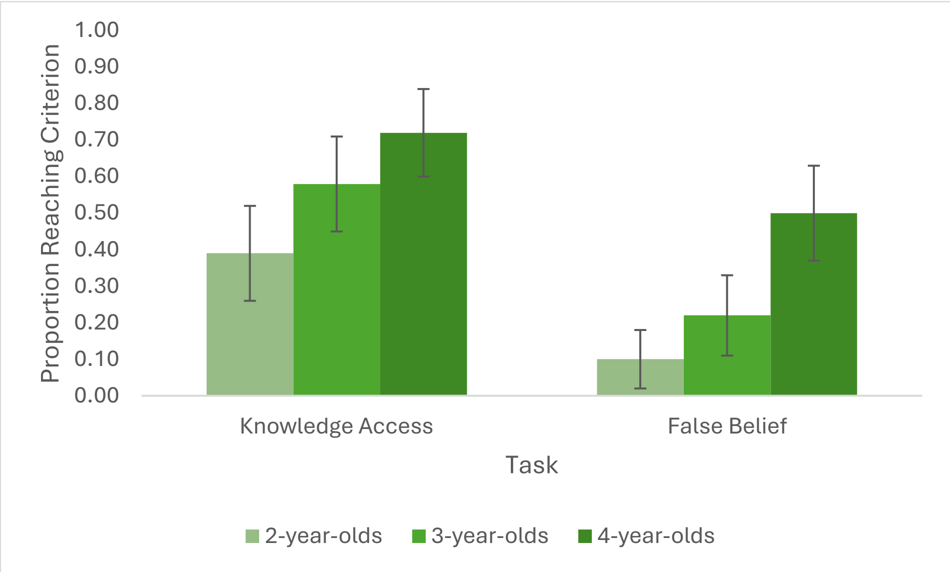
Variable	**2-year-olds**	**3-year-olds**	**4-year-olds**	**Total**
Point Where	.68(.36)	.89(.26)	.97(.12)	.85(.29)
Looking Where Head	.49(.27)	.69(.30)	.78(.24)	.65(.30)
Looking Where Eyes	.41(.28)	.60(.33)	.77(.26)	.59(.33)
Looking At You Head	.66(.25)	.78(.27)	.87(.28)	.77(.28)
Looking At You Eyes	.59(.24)	.68(.26)	.76(.32)	.67(.29)
Overall Looking Where	.45(.23)	.64(.27)	.77(.22)	.62(.27)
Overall Looking At You	.62(.18)	.73(.22)	.81(.27)	.72(.24)
Overall Gaze Judgment	.54(.17)	.69(.20)	.79(.19)	.67(.21)

Theory of Mind Tasks

Figure 10 and Table 2 show performance on both tasks improved with age (both *ps* ≤ .001), with significant improvements on Knowledge Access between 2- and 3-year-olds (Mann-Whitney $U = 1311.00$, $p = .041$), and False Belief between 3-year-old and 4-year-old participants (Mann-Whitney $U = 1145.5$, $p = .002$).

Figure 10

Experiment 1 Performance on Theory of Mind Tasks



Note. Error bars represent 95% CI.

Table 2

Experiment 1 Mean (SD) Theory of Mind Task Performance by Age Group

Variable	**2-year-olds**	**3-year-olds**	**4-year-olds**	**Total**
Knowledge Access (Max. 1)	.39(.49)	.58(.50)	.72(.45)	.56(.50)
False Belief (Max. 1)	.10(.31)	.22(.42)	.50(.50)	.27(.45)

Task Effects

Hiding Tasks

Blindfold Subgroup. Figure 7 shows children performed well on the Move Object task. Performance on the Classic Move Screen task was lowest, followed by the Blindfold Control task, then Blindfold.

Performance on the Move Object task was significantly better than all Move Screen tasks (all ps < .001). Friedman's test indicated there was a main effect of Move Screen task $\chi2(2) = 15.66$, p <.001. Blindfold Control (*M* = .35, *SD* =.41) and Blindfold task (*M* =.43, *SD* =.42) performances were significantly better than the Classic Move Screen task performance (*M* = .26, *SD* = .37) (respectively: *p* = .013, *p* <.001), and Blindfold performance was significantly better than Blindfold Control (*p* = .023).

Move screen tasks were significantly associated (all *p*s < .001) and remained so after controlling for age (all *p*s <.001).

Errors on the Move Object task were predominantly neutral then egocentric, and rarely other. In contrast, errors on all Move Screen tasks were mostly egocentric, then equally neutral or other (see Table 3).

Table 3

Experiment 1 Proportion Errors Blindfold Subgroup

Task	**Egocentric**	**Neutral**	**Other**	**Total Errors**
Move Object	.29	.67	.04	55
Move Screen	.57	.24	.19	129
Blindfold	.61	.22	.17	98
Control	.66	.19	.15	113

Blindfold Remove and Replace Subgroup. Figure 8 shows children performed well on the Move Object task. Performance on the Classic Move Screen task was lowest, followed by the Blindfold Control task, then Blindfold Remove and Replace.

Performance on the Move Object task was significantly better than all other Move Screen tasks (all *p*s $< .001$). Friedman's test indicated there was not a main effect of Move Screen task $\chi2(2) = 3.30$, p =.192. Blindfold Remove and Replace performance (*M* = .51, *SD* =.40) was better than Classic Move Screen (*M* = .42, *SD* =.42) but this fell short of conventional significance (*p* = .051). Blindfold Control performance (*M* = .48, *SD* =.43 was intermediate between the other two tasks and did not differ significantly (Blindfold Remove and Replace, *p* = .582; Classic Move Screen, *p* =.189).

Move screen tasks were significantly associated (all *p*s $< .001$) and remained so after controlling for age (all *p*s $<.001$).

Errors on the Move Object task were predominantly neutral then egocentric, and rarely other. In contrast, errors on all Move Screen tasks were mostly egocentric then neutral, and rarely other (see Table 4).

Table 4

Experiment 1 Proportion Errors Remove and Replace Subgroup

Task	**Egocentric**	**Neutral**	**Other**	**Total Errors**
Move Object	.23	.72	.05	44
Move Screen	.75	.19	.06	99
Remove and Replace	.58	.34	.08	85
Control	.60	.35	.05	91

Gaze Judgement Tasks

To gain a stable overall view of gaze judgement ability I combined the raw scores of head-and-eye and eyes-only versions of each task into an overall Looking Where and an overall

Looking At You variable. These were also combined for an overall measure of judgement ability. For descriptives, see Table 1.

To account for differences in chance baseline, tasks were categorised as pass or fail using the criteria described in the Methods (pg. 51). Considering the gaze judgement tasks, Friedman's test indicated there was a main effect of task $\chi 2(3) = 24.49$, $p < .001$. Performance on the Looking Where task was significantly better when head direction (*M* =.52, *SD* = .50) was a cue compared to eyes (*M* =.44, *SD* = .50) (*p* = .047) and for the Looking At You task (Head: *M* =.49, *SD* = .50; Eyes: *M* =.33, *SD* = .47) (*p* < .001). When only eyes were a cue, Looking Where performance was better than Looking At You (*p* = .008). These did not differ when head direction was also a cue (*p* = .611).

All gaze judgement tasks were significantly associated (see Table 5) and remained so after controlling for age.

Table 5

Experiment 1 Correlations Gaze Judgement Tasks Using Criterion Scores. Partial Correlations Controlling for Age are Shown Below the Diagonal.

Variable	**Looking Where Head**	**Looking Where Eyes**	**Looking At You Head**	**Looking At You Eyes**
Age (Months)	.37***	.42***	.46***	.41***
Looking Where Head	-	.51***	.28***	.30***
Looking Where Eyes	.42***	-	.39***	.33***
Looking At You Head	.14 †	.24**	-	.55***
Looking At You Eyes	.18*	.19*	.45***	-

Note. Correlations are Spearman's Rho. †$p<.10$, *$p<.05$, **$p<.01$, ***$p<.001$, 2-tailed.

If participants extracted some directional information in the Point and Looking Where tasks, errors would be more likely to be on the same side as the target. This was the case for all tasks, (p = 1/3, all *p*s <.002): Point Where task, 50% (53 out of 106) errors; Looking Where Eyes task, 66% (158 out of 241 errors); Looking Where Head task, 52% (147 out of 281 errors). Proportion errors to the correct side increased with age and for the gaze judgement tasks was above chance for all age groups.

Table 6

Experiment 1 Proportion Errors to the Correct Side on Gaze Judgement Tasks by Age Group.

Variable	**2-year-olds**	**3-year-olds**	**4-year-olds**
Point Where	.52**	.38	.71 †
Looking Where Head	.54***	.70***	.87***
Looking Where Eyes	.47***	.56***	.60***

Note. One-Sample t-test *p=1/3* . †*p*<.10, **p*<.05, ***p*<.01, ****p*<.001, 2-tailed.

Theory of Mind Tasks

Performance on the theory of mind control questions was relatively poor, potentially reflecting the high number of 2-year-olds in the sample. Thirty-six (20.1%) children failed the Knowledge Access memory question, 10 of whom passed the test question. On the False Belief task 58 (33.7%) children failed the memory question, 5 (2.9%) failed the reality question and 7 (4.1%) children failed both questions; of these 70 children 11 correctly answered the test question. Previous studies differ in leniently requiring success on the belief question only or strictly on all three questions. Analyses conducted using both criteria did not substantially differ. The analysis is reported using the strict criteria.

Performance was higher for the Knowledge Access task than the False Belief task, (M_1 = .56, M_2 = .27; McNemar, p <.001, 59-9). Task performances were moderately associated (r_s =.302, p<.001; controlling for age $r_{partial}$ =.204, p=.008).

Comparison of Tasks

Performances were classified as pass or fail for comparison. For the Move Screen task, McGuigan and Doherty (2002) use passing criterion of 3 out 4. However, the current study only has two trials. Thus, to apply a similar level of leniency as previous studies passing criterion is defined as 1 out of 2. The Looking Where tasks have 4 potential targets and Looking At You tasks have 2 potential targets, so the chance baselines differ. To allow direct comparison, the passing criterion for Looking Where tasks was set as 4 out of 8 (p = .114) and for Looking At You tasks least 6 out of 8 (p = .145).

False Belief performance was harder than all other tasks (all ps <.001). Classic Move Screen task and Knowledge Access were of equal difficulty (all ps >.081). Classic Move Screen task was significantly harder than gaze judgement tasks (all ps <.05).

Table 7

Experiment 1 Binomial Differences of Move Screen, Gaze Judgement and Theory of Mind Tasks.

Variable	**Looking Where**	**Looking At You**	**Knowledge Access**	**False Belief**
Move Screen	14 – *52****	19 – *35**	30 – *46* †	51 – *17****
Looking Where	-	37 – *15***	44 – *22***	78 – *6****
Looking At You		-	35 – *35*	61 – *11****
Knowledge Access			-	59 – *9****
False Belief				-

Note. The first number in each pair represents the number of children passing the row task and failing the column task; the second number represents the number of children failing the row task and passing the column task. †p<.10, *p<.05, **p<.01, ***p<.001, 2-tailed.

Due to subgroup differences, I cautiously analyse correlations on the experimental hiding tasks separately by subgroup. For both subgroups, the experimental task performance significantly associated with Move Screen, gaze judgement and Knowledge Access task performance, but not with False Belief task performance. After controlling for age, only associations with Move Screen and Overall Gaze Judgement Scores remained stable. The experimental tasks are not analysed further.

Table 8

Experiment 1 Correlations Experimental Hiding Tasks Using Raw Scores.

Variable	**Move Screen**	**Looking Where**	**Looking At You**	**Overall Gaze Judgement**	**Knowledge Access**	**False Belief**
Correlation						
Blindfold	.521***	.375***	.275*	.398***	.223*	.155
Remove and Replace	.553***	.335**	.237*	.323**	.275*	.211†
Age Partial Correlation						
Blindfold	.458***	.206†	.120	.222*	.140	.078
Remove and Replace	.480***	.204†	.114	.194†	-.082	-.051

Note. Correlations are Spearman's Rho. †*p*<.10, **p*<.05, ***p*<.01, ****p*<.001, 2-tailed.

Performances on Move Screen, gaze judgement and theory of mind tasks were significantly associated with each other and age, as shown in Table 9. After controlling for age, Knowledge Access and False Belief performances were associated with each other but not with other tasks. Both Classic Move Screen task performances was associated with Overall Gaze Judgement.

Table 9

Experiment 1 Correlations All Tasks Using Raw Scores. Partial Correlations Controlling for Age are Shown Below the Diagonal.

Variable	**Move Screen**	**Looking Where**	**Looking At You**	**Overall Gaze Judgement**	**Knowledge Access**	**False Belief**
Age (Months)	.431***	.530***	.432***	.540***	.302***	.416***
Move Screen	-	.337***	.369***	.397***	.165*	.262***
Looking Where	.141†	-	.434***	.848***	.204**	.300***
Looking At You	.224**	.268***	-	.834***	.184*	.274***
Overall Gaze Judgement	.217**	.787***	.792***	-	.219**	.315***
Knowledge Access	.040	.055	.062	.070	-	.302***
False Belief	.101	.103	.115	.118	.204**	-

Note. Correlations are Spearman's Rho. †p<.10, *p<.05, **p<.01, ***p<.001, 2-tailed.

Regression analysis was used further to examine associations between task performances. Hierarchical linear regression examined whether Overall Gaze Judgement performance explained the variance in Move Screen performance beyond that of age, as shown in Table 10. Variation in Move Screen performance was best explained by age and Overall Gaze Judgement performance, with age contributing significantly more to the explained variance than Overall Gaze Judgement.

Table 10

Experiment 1 Summary of Hierarchical Regression Analysis for Variables Predicting Move Screen Performance.

Variable	***F***	R^2	ΔR^2	***B***	***SE***	***t***
Step 1	41.978***	.198	.198**			
Age Months				.445	.008	6.479***
Step 2	25.210***	.230	.032**			
Age Months				.329	.009	4.093***
Overall Gaze Judgement				.212	.019	2.640**

Note. $*p < .05$, $**p < .01$, $***p < .001$.

Discussion

The results replicate those of Flavell et al. (1978) and Mcguigan and Doherty (2002). Children of all ages performed well on the Move Object task relative to their Move Screen performance. Thus, whilst both tasks appear to require understanding of occlusion, children's difficulty with the Move Screen task suggests there are signature limits to this understanding. The general hypothesis concerns a potential explanation for this finding - engagement. McGuigan and Doherty (2002) propose that when young children hear the word *see*, they interpret it as *engaged with*. The Move Object task is easy because at the start of the trial, the experimenter is not engaged with the object. However, at the start of the Move Screen task the experimenter is engaged with the object, leaving children unsure how interposing the screen stops disrupts engagement. In short, young children cannot make a distinction between *being engaged with* and *seeing* the object. This hypothesis was tested in two ways.

Firstly, the engagement hypothesis proposes that young children cannot distinguish between *seeing* and *engagement*. As such, it predicts that if the experimenter is not engaged with the object when the child places the screen, then the task should be easy. The results of the Blindfold condition support this. In the Blindfold condition the agent donned a blindfold before the object was placed. Thus, the trial begins with the agent not engaged with the object. Performance on this condition was better than on the classic task in which there is no blindfold and relative to a control task in which the blindfold is donned after the object is placed (success: blindfold, 43%; classic 25%, control, 35%). This suggests young children understand occlusion in terms of spatial and behavioural relations and not in terms of attention or mental states.

Secondly, the engagement hypothesis proposes that engagement relations persist over time until the object is moved from its current location. Horschler et al. (2019) found that manipulations as simple as removing the object and replacing it immediately back, result in monkeys behaving as though no relation existed. Thus, it was predicted that children would behave similarly to this type of manipulation. The results of the Remove and Replace condition do not support this. At the start of the Remove and Replace trial, the experimenter is engaged with the object. Once they donned the blindfold, the object was then removed and immediately replaced. Thus, the trial begins with the agent engaged with the object, but the manipulation of the object results in the agent becoming disengaged from the object at the point of response. Performance on this condition was better than on the standard task but the difference fell short of significance (p = .051), nor was it significantly better than on the control task, which was intermediate between the Remove and Replace and the classic task (success: Remove and Replace, 51%; classic, 42%; control, 48%).

One interpretation of this finding is that spatial manipulations need to be substantial to disrupt engagement relations. Even though removing the object creates a *disengagement* event, traces of the engagement relation may not perish until the object is placed in an entirely new location. As O'Neill (1996) posits, the location of the disengagement event remains salient. Thus, returning the object to the original location may restore the relation creating the opportunity to cancel out the disengagement event. In the current study, the object is not placed in a new location at any point during the remove and replace phase – it remains suspended in space, never leaving the experimenter's hand. Thus, this kind of spatial manipulation is not strong enough to overcome persisting engagement traces resulting in children acting as though the experimenter remained engaged with the object.

This may also explain why the monkeys in Horschler et al.'s (2019) study act as though this type of manipulation turns off awareness relations. In their task, the ball is not suspended in space. Instead, it moves along an edge. So, when it moves towards the distractor box, this central location is perceived as 'new'. This should be sufficient to diminish any awareness traces at the original location, allowing the awareness relation to be permanently switched off. Meaning when the ball returns to the original location, there is no active awareness relation for the monkeys to expect behaviour towards. This small, but important detail may be the key to the incongruence of the current findings. As such, the definition of engagement should be revised to "a relationship between a person and an object, which persists until it is moved to a *new location*".

Together, these findings suggest that children can have improved success on the Move Screen task when seeing and engagement are separately manipulated. Thus, young children are likely using a concept of engagement to predict the behaviour of others. However, children's performance on the control condition demonstrates that simply adding a blindfold is enough to confer greater success on the task. At the start of the trial, the experimenter is engaged with the object. After donning the blindfold no manipulations occur, meaning the experimenter remains engaged with the object. Thus, placing the screen should be equally as difficult as the original Move Screen task. As such, children's improvement in performance on this condition was unexpected. Despite the improvement, success was still lower than that of the experimental conditions. It is plausible that the blindfold makes the eyes highly salient, encouraging children to pay greater attention to them as a more informative resource compared to the broad cues that they would typically monitor.

Another potential explanation for this finding may be that the task requires inhibitory control. Specifically, the task asks children to disrupt line of sight from the eyes. Inhibiting the presence of the eyes would make it easier to focus on the line of sight itself and therefore successfully disrupt it. This may be harder to achieve in the Move Screen task when the eyes are open compared to when they are covered by a blindfold. As such, any condition that uses a blindfold, including the control condition, would confer improvement in performance. However, this is not a full explanation of the improvement in performance as the experimental conditions were significantly easier than the control condition.

Relatedly, inhibitory control may also explain why any variation of the Move Screen task is harder than the Move Object task. In all the Move Screen tasks, children must refer to the eyes and the object to determine line of sight. This makes it hard to inhibit the presence of the eyes and therefore distracts from the line of sight itself. On the other hand, in the Move Object task the line of sight is already disrupted by the screen. This means to place the object outside of line of sight, children must refer to the screen only. As such, children do not need to inhibit the presence of the eyes. They must simply refer to them after object placement to confirm it is not within the line of sight. As inhibitory control develops in the preschool period (Carlson, 2005), it is likely as a child's inhibitory control improves, they are better able to inhibit the presence of the eyes in the Move Screen task, conferring greater success at the task.

It is important to note that performance on neither the control nor experimental conditions improves to that of the Move Object task. This suggests that to successfully interpose the screen, understanding of line of sight alone may not be sufficient. A potential explanation may be the differing spatial reasoning needed for the two tasks. The Move Object task involves placing the object *behind* the screen, whereas the Move Screen task involves placing the screen *between* the experimenter and object. In Simms and Gentner's (2019) work on children's midpoint encoding, they note that the spatial relation *between* involves two referent points whereas *behind* only involves a singular reference, meaning it is a potentially more challenging relation. Thus, the Move Screen task difficulty may reflect the complexity of the spatial relation.

The additional complexity of *between* is particularly evident in children's spatial coding strategies in search tasks. In such tasks, an object is hidden one of many locations in reference to one or more landmarks. When spatial relations involve a single landmark, children can locate objects as young as one-years-old. For example, in Bushnell et al.'s (1995) study one-year-olds correctly searched for a toy under a distinctive cushion, whilst two-year-olds can do this even if the distinctive cushion is an indirect landmark. However, Spetch and Parent (2006; Simms & Gentner, 2008) showed that when there are two landmarks children's use of spatial relations go through a longer development. In their task, children were presented with a row of 15 boxes in which they had to find hidden stickers. At the start of each trial, the experimenter selected an array of five boxes, placing teddy bears on boxes 1 and 5 as landmarks for the sticker's location which was placed inside box 3. Children were told that the bears would help them to find the sticker. The children were however not told that the sticker was always located in the centre box. They found that 3-year-olds first look location was nearly always adjacent to a teddy bear but was not systematically between or outside the two landmarks. By four-years, children were systematically searching between the two landmarks. This suggests that children develop an understanding of the spatial relation *between* later than relations that rely on a single landmark such as *behind*. Whilst the above studies refer to search strategies rather than object placement, they allude to spatial understanding being a potential barrier to children's success at the Move Screen task beyond line of sight.

Additional hypotheses were that performance on the classic Move Screen task, gaze judgement performance and measures of theory of mind ability would all be associated. Performance on the gaze judgement tasks replicated those of Doherty and colleagues (1999; 2009). All children performed well on the point where task, approaching ceiling by 3-years-old. However, 2- and 3-year-olds had difficulty making judgements of eye direction and saw rapid improvement across the age range. Judgements of where someone was looking were easier than judging which of two faces was looking at you. Children found it easier to make judgements when eye direction was accompanied by a congruent head turn. However, unlike Doherty and Anderson's (1999) findings, children find it difficult to make judgements even with an accompanied head turn. Thus, children's difficulty making judgements of gaze direction may be more generalised to other gaze related cues than previously thought.

As hypothesised, overall gaze judgement ability correlated with Move Screen performance beyond common associations with age, which was confirmed by regression analysis. When considering each gaze judgement task individually, they had disparate associations. Specifically, the Looking Where task fell short of conventional significance when controlling for age. This does not replicate McGuigan and Doherty (2002) who found a strong and specific relation between the Looking Where and Move Screen performance. On the other hand, the more difficult Looking At You task remained significant. This is the first study to investigate the association with the Looking At You task, so further investigation is needed to understand the nature of this association. Nevertheless, given the strong association between the two gaze judgement tasks, the general finding prevails that gaze judgement ability and the

Move Screen task shares a common understanding of line of sight, “that what someone is attending to is determined by the direction of his or her eyes” (McGuigan & Doherty, 2002, p. 423). The finding that the Move Screen task is more difficult than gaze judgement ability suggests it requires an additional understanding of how to intersect line of sight. Thus, gaze judgement ability may be integral to success on the Move Screen task, rather than merely a contemporaneous development. Together, these findings suggest that engagement may allow children to respond to others’ behaviour based on what they have attended to, without the necessary concepts of seeing. Once they develop a concept of line of sight, children can map eye-direction or attention onto the word *see* resembling an adult-like concept of seeing.

Contrary to predictions, knowledge understanding did not associate with Move Screen performance after controlling for age. This is surprising as the two tasks were equivalent difficulty, with children typically passing the Move Screen task around the same time they understand seeing leads to knowing. Thus, it would be expected that once children conceive of attention as a mental state, they would be able to make the distinction between knowledge and perception required in the Move Screen task. The current findings suggest this is not the case. In particular, the False Belief task a well-validated measure of theory of mind development, also did not associate with Move Screen performance further suggesting line of sight understanding does not require mental state attribution. The False Belief task additionally requires understanding of misrepresentation, resulting in it being a harder task than both Knowledge Access and Move Screen. Thus, success on the Move Screen task may not mark the transition from engagement-based to mentalistic-based strategies to predict behaviour.

Finally, the relation between theory of mind and gaze judgement was of interest. In the current study, neither theory of mind measure associated with gaze judgement ability beyond common associations with age. This replicates Doherty and Anderson’s (1999) findings, who speculated that the shift from engagement to sophisticated understanding of gaze may be motivated by children's increasing interest in the mind. However, both their findings and those of the current study suggest that a concurrent development is more likely than a precursor relationship. At the very least, gaze judgement is likely to be an immediate precursor as children find these measures significantly easier than either theory of mind measure. However, the current results do not provide strong evidence that gaze understanding is a precursor to theory of mind as often claimed (Gopnik et al., 1994; Goméz, 1996).

The lack of association with the False Belief task is unsurprising as it additionally tests misrepresentation. However, lack of association with Knowledge Access is surprising. Gaze understanding is a necessary part of mental states, since this is the way knowledge is formed. This is implied by Knowledge Access testing for understanding that seeing leads to knowing. In this case, visual attention is a core function of reasoning. D'Entremont et al. (2012) investigated this by comparing children's performance on a gaze judgement task with Wellman and Liu's (2004) theory of mind scale. They found gaze judgement associated with Knowledge Access but not false belief. Thus, it was expected in the current study that gaze understanding would show associations with this earlier form of theory of mind. This was not the case. Together, these findings can be interpreted that children's shift from engagement to sophisticated understanding of gaze may not be motivated by children's increasing interest in the mind. Instead, children may become interested in the mind once they understand that people are psychologically connected to objects through their eyes.

In summary, the current study finds support for the engagement hypothesis. When individually manipulating seeing and engagement during the Move Screen task, children's performance improves. It is plausible that when young children hear the word see, they interpret it as engaged with. Additionally, the current study shows engagement relations may only be disrupted after substantial spatial manipulations, such as to a new location. Thus, I propose a small but important revision to its definition, that engagement persists until the object is moved *to a new location*. Finally, the current study presents evidence that children are likely mapping eye-direction onto the word see once they develop a concept of line of sight. However, this understanding does not share commonality with mental state attribution. Understanding gaze direction may be a contemporaneous development, or at the very least an immediate precursor. Thus, sophisticated gaze understanding may not mark the transition from engagement-based to mentalistic-based strategies to predict behaviour.

Chapter 3

A Longitudinal Study of Gaze Understanding Development

The current thesis considers the claim that we have two systems for gaze processing. That being we have one system for locating the objects of others' attention, and one to think about the attentional relationship between observer and object. The latter is uniquely human and must be learnt, potentially developing as a result of children's increasing interest in others' mental states. Thus far I have presented partial evidence for this claim. Namely that gaze understanding goes through protracted development in the pre-school period. However, it remains unclear if theory of mind motivates System 2 development.

Experiment 1 replicates findings that gaze judgements develop at roughly the same time as understanding of mental states, but that the abilities do not associate when controlling for age (Doherty & Anderson, 1999). This would suggest that gaze judgement and theory of mind develop concurrently but are not specifically related to one another. However, whilst the two abilities may not share a common understanding, it does not rule out the possibility that one motivates the development of the other. Thus, the primary aim of the current experiment is to longitudinally replicate the developmental trajectory of the two abilities and explore whether there is a precursor relationship between them. Specifically, the current experiment is interested in whether the data reflects the claims of the two-systems theory that theory of mind ability predicts later gaze judgement (Doherty, 2006; Doherty et al., 2009), or the consensus view that gaze judgement ability predicts later theory of mind (Goméz, 1996; Gopnik et al., 1994). Alternatively, the two abilities may have merely coincidental contemporaneous development.

What is apparent thus far is that until System 2 develops, children predict the behaviour of others based on what they have attended to. Once children can make judgements of gaze direction, they transition to an adult-like concept of seeing that "what someone is attending to is determined by the direction of his or her eyes" (McGuigan & Doherty, 2002, p. 423).

However, this understanding is not associated with mental state attribution. Thus, the secondary aim of the current experiment is to examine the transition to mentalistic-based strategies to predict behaviour, and whether this is marked by a sophisticated understanding of gaze. As per the findings of Experiment 1, it is predicted that at each timepoint Move Screen performance will associate with gaze judgement ability, but not with theory of mind measures Knowledge Access or False Belief. Additionally, if engagement is a way for children to fake understanding of gaze or mental states, then both abilities should predict later Move Screen performance.

The current experiment additionally controls for age related developments, inhibitory control and vocabulary. Inhibitory control is of interest as it is shown to develop across a similar time period as theory of mind and gaze understanding (Carlson, 2005). Whilst it is currently unknown if inhibitory control may be related to gaze understanding, it has been shown to be a requirement of theory of mind tasks, longitudinally predicting its performance (Carlson et al., 2002; Carlson et al., 2015; Hughes & Ensor, 2007; Marcovitch et al., 2015). As discussed in Chapter 2, it is also speculated that the Move Screen task may have an inhibitory component. Language development is also shown to be a significant predictor of theory of mind development (e.g. Astington & Jenkins, 1999; Farrar & Maag, 2022; Schick et al., 2007), and given the verbal nature of all tasks, language should be accounted for.

Method

Participants

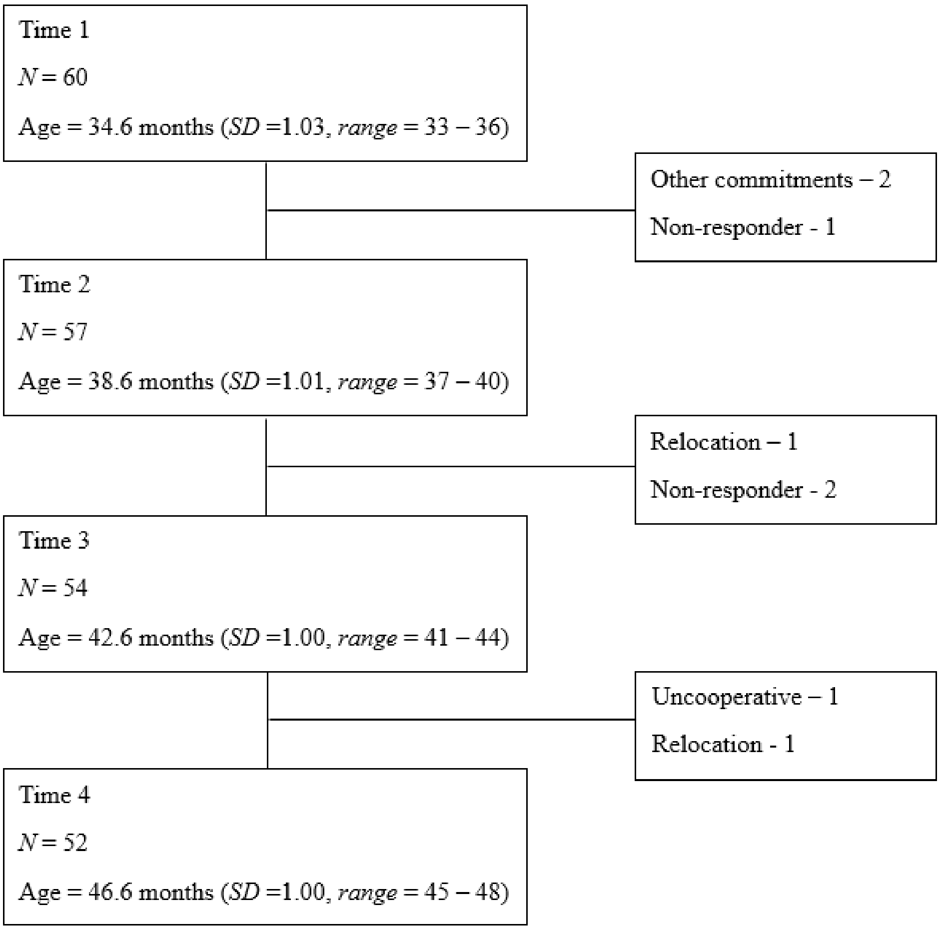
Participants were 60 children (36 girls) recruited from the University of East Anglia Child Participant Panel (see Figure 11 for timepoint descriptives). Participants were tested at four timepoints spaced four months apart. To compensate their time, children received a small gift, and parents received a £5 Love2Shop voucher at each visit.

Inclusion criteria were informed parental consent, and child assent immediately prior to testing. No child met the exclusion criterion of parent-indicated special-needs diagnosis.

Once a participant withdrew from a timepoint, they were then withdrawn from any remaining time points. The stopping criterion was once planned sample size was achieved at Time 1.

Figure 11

Experiment 2 Timepoint Descriptives and Drop-out Reasons



The experiment was carried out in accordance with ethical standards at University of East Anglia Research Ethics Committee (ETH2122-1672).

Statistical Analysis Plan

This study was conducted in line with the pre-registration (108902). Since, the researcher's statistical scope for longitudinal analysis has expanded and the statistical analysis plan below has updated to reflect this. For all planned analyses the G-Power 3.1.9.2 statistical tool (Faul et al., 2007) was used to estimate sample sizes required to achieve a power of 0.80 with two-tailed alpha set to 0.05.

Within task comparisons will be tested in multiple ways. The primary analysis is a two-way ANOVA. This compares performance differences between sub-tasks (e.g. Move Screen vs. Move Object) and timepoints. Planned comparisons require a sample size of 34. Correlations are used to see commonalities between sub-tasks, requiring a sample size of 84.

Whilst the ANOVA analysis can give an indication of how gaze judgement, Move Screen and theory of mind ability performance changes over time, they treat each timepoint as independent from one another. This means they provide discrete comparisons rather than continuous change. However, the timepoints are dependent on one another in longitudinal datasets. Thus, I also report latent growth curve models as they model change as a continuous process. These models also consider that each child varies in their baseline performance and rate of development. For example, a child who scores low at Time 1 may improve quicker than a child that scores high at Time 1.

Within timepoint comparisons are tested with correlations and McNemar tests. These require sample sizes of 84 and 20 respectively. Between timepoint comparisons were originally planned to be tested with a Random Intercept Cross Lagged Panel Model. However, the complexity of the model resulted in it not converging, meaning that it could not find a solution that best fit the data. Convergence is important for reliable and interpretable results. Instead, correlations and linear regressions are used requiring a sample size of 84 and 55 respectively.

Except for correlational analyses, the sample was suitably powered at Time 1 and remained so at Time 4 following a 10% attrition in sample size. To retain as much data as possible, listwise deletion was used for every analysis where there was missing data. Thus, all participants were retained in the dataset irrespective of whether they participated in all four timepoints or not. This was also true of any participants that did not complete all tasks within a timepoint.

Design

Tasks were blocked by type, and arranged in four task orders such that a child will complete each task order across the four timepoints (see Table 11). To reduce any potential carryover effects, all task types had superficially different versions that were equally balanced within and across timepoints as outlined below.

Table 11

Experiment 2 Task Orders

Task	**Task Order 1**	**Task Order 2**	**Task Order 3**	**Task Order 4**
Task 1	Gaze Cueing	Gaze Cueing	Gaze Cueing	Gaze Cueing
Task 2	Gaze Judgement	Gaze Judgement	Gaze Judgement	Gaze Judgement
Task 3	Bear-Dragon	Bear-Dragon	Hiding	Hiding
Task 4	Hiding	Theory of Mind	Theory of Mind	Bear-Dragon
Task 5	Theory of Mind	Hiding	Bear-Dragon	Theory of Mind

One of the initial aims of the study was to explore how System 1 may change as System 2 develops. Previous research has shown that children's gaze following uses different information from the eyes than adult gaze following (Doherty et al., 2015, see Chapter 4 for further discussion). It is possible that as System 2 develops, System 1 becomes more flexible in its processing allowing adult gaze following to recruit different information. It was hoped that by replicating Doherty et al.'s (2015) gaze cueing task longitudinally, this hypothesis could be investigated. However, the task lacked sensitivity as no gaze following effects were found at any timepoint. Thus, I do not explore these findings within the current thesis.

Gaze judgement tasks were always presented second as this was the main variable of interest. Looking Where and Looking At You gaze judgement tasks were alternated beginning with a Looking Where task. For two timepoints, participants received a schematic version of the task, and for two timepoints a photograph version of the task.

The remaining three tasks were pseudo randomly assigned. For hiding tasks, there were 2 permutations either starting with a Move Object task or a Move Screen task. Participants received each permutation at two of the four time points. The Knowledge Access task was presented before the False Belief task, following Wellman and Liu (2004). There were two versions of each task using different names and objects, and participants received each version at two of the four timepoints. Bear-Dragon task had four versions. The trials were randomly assigned 2 orders and started with either a Bear trial first or a Dragon trial. Each participant received each version once across the four timepoints.

Whilst the child completed the above tasks, parents completed the Developmental Vocabulary Assessment for Parents.

Materials and Procedure

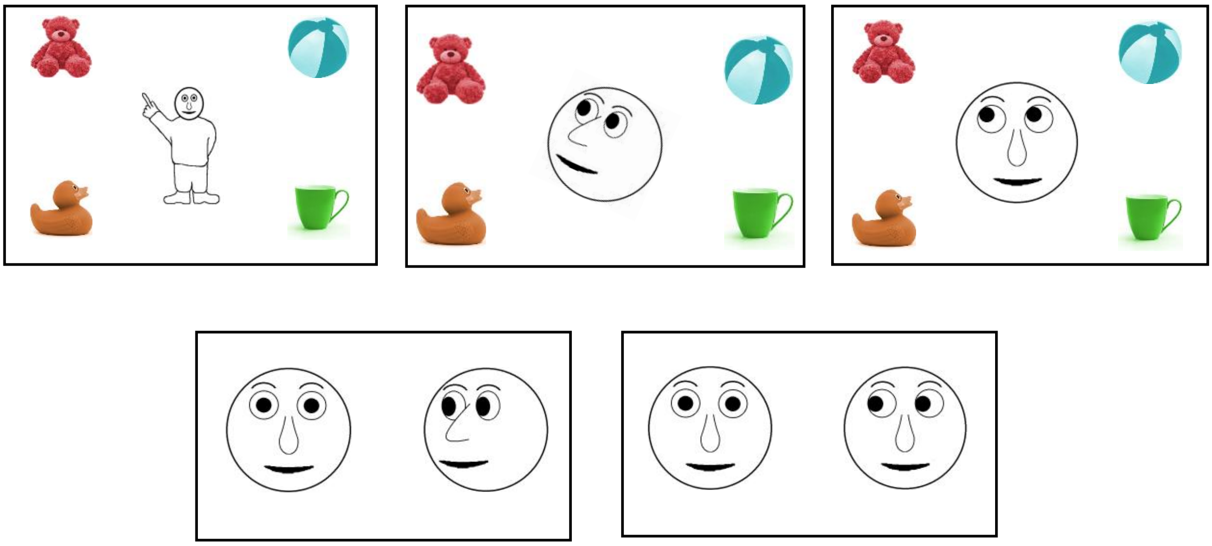
Testing was in a quiet location either on the university campus or at the child's home for around 20 minutes. The experimenter sat at a low table on the floor opposite the child with her eyes level with the child's.

Gaze Judgement Tasks

These tasks were identical to those in Experiment 1, except for the schematic faces (see Figure 12) for superficial variation between timepoints. Doherty and Anderson (1999) find children's performance on gaze judgement tasks using schematic faces does not differ to photographed faces, and therefore this is an appropriate manipulation to reduce any potential carryover effects without influencing the natural change in performance seen with age.

Figure 12

Experiment 2 Gaze Judgment Task Stimuli Schematic Versions



Note. Point Where (top left), Looking Where Head (top middle), Looking Where Eyes (top right), Looking At You Head (bottom left), and Looking At You Eyes (bottom right).

Hiding Tasks

Tasks were identical to the Move Object and Move Screen tasks in Experiment 1. As neither blindfold task was presented here, only a single experimenter was required. The current study also differed as each task had four trials, where pass criterion was a correct response on at least 3 trials. The experimenter sat opposite (180°), then to the left (90°), right (90°) and finally opposite (180°) the child.

Figure 13

Experiment 2 Hiding Task Stimuli

Move Object task. The experimenter sat directly opposite the child (180°) and placed the screen broadside on the table and fixed her gaze at the screen. The experimenter then handed the child the toy and said, "put Nellie somewhere on the table so I can't see her". The trial ended when the child had placed the object on the table and was no longer touching it. Trials were scored as per Experiment 1.

Move screen task. The Experimenter placed the object in the centre of the table and fixed her gaze at the object. The experimenter handed the screen to the child and said, "put this [pointing at the screen] somewhere on the table so I can't see Nellie". The trial ended when the child had placed the screen on the table and was no longer touching it. Trials were scored as per Experiment 1.

Theory of Mind Tasks

Both Knowledge Access and False Belief tasks were identical to Experiment 1. In the current study, both tasks had superficially different versions where the protagonist names and objects were replaced with alternatives (see Figure 14).

Figure 14

Experiment 2 Theory of Mind Task Stimuli.

Note. Knowledge Access (left) and False Belief (right). Top is version 1 and bottom is version 2.

Bear-Dragon Task

Before the task began, the experimenter verified that children could follow the direction of the 10 target instructions, such as ‘touch your tummy’. Once they were able to demonstrate they could complete all 10 instructions they were then introduced to each puppet.

For the bear children were told:

“This is Mr Bear and he is our friend (smiling and waving Mr Bear’s hand). So, when the bear asks us to do something, we do what he says (Experimenter and Mr Bear nod their heads)”.

For the dragon children were told:

"This is Mr Dragon and he is very naughty (gasp), we don't listen to him. So, when the dragon asks us to do something, we don't do it (Experimenter and Mr Dragon shake their heads)".

Children were then given a practice trial for each puppet. Moving the bear's mouth and talking in a high-pitched voice, the experimenter said, "touch your nose". Moving the dragon's mouth and talking in a low, gruff voice the experimenter said, "touch your tummy". If children followed the bear but ignored the dragon they passed the practice. Children were given up to four practice trials. On the fifth, parents were be asked to hold the children's hands down on the table to 'help' them. Children were then reminded of the rules.

Following the practice trials, children completed 10 test trials. They were half bear and half dragon, presented in alternating fashion. No assistance is given at these trials. For Dragon trials, a full command movement would score 0, a partial command movement would score 1, a wrong movement would score 2 and no movement would score 3. These were reversed for Bear trials. Children could score between 0 and 15 for each trial type. Only the Dragon Trials are analysed as children must supress their instinct to follow the instructions, whereas the Bear Trials simply measure a child's ability to follow the instructions.

Developmental Vocabulary Assessment for Parents

Parents completed the Developmental Vocabulary Assessment for Parents (DVAP; Libertus et al., 2015), to measure children's vocabulary levels. This consists of the first 8 practice items and first 204 words from Form A of the PPVT-4 (Dunn & Dunn, 2007). Parents had to indicate whether a child understood each word or not. This gave an overall score for the number of words the child understands. Raw scores are used as there are no age normed data available.

When designing the study, there was no standardised vocabulary measure that could be used across the whole age range, nor was quick and simple to administer alongside the other tasks. Thus, whilst the DVAP is based on American vocabulary tests it was considered the most appropriate measure to use for the current study. In addition, the DVAP has been validated on American word items (Libertus et al., 2015) so it was decided that it was an appropriate compromise to not swap out the items for those in the BPVS-3 (Dunn & Dunn, 2009).

Results

Within Task Comparisons

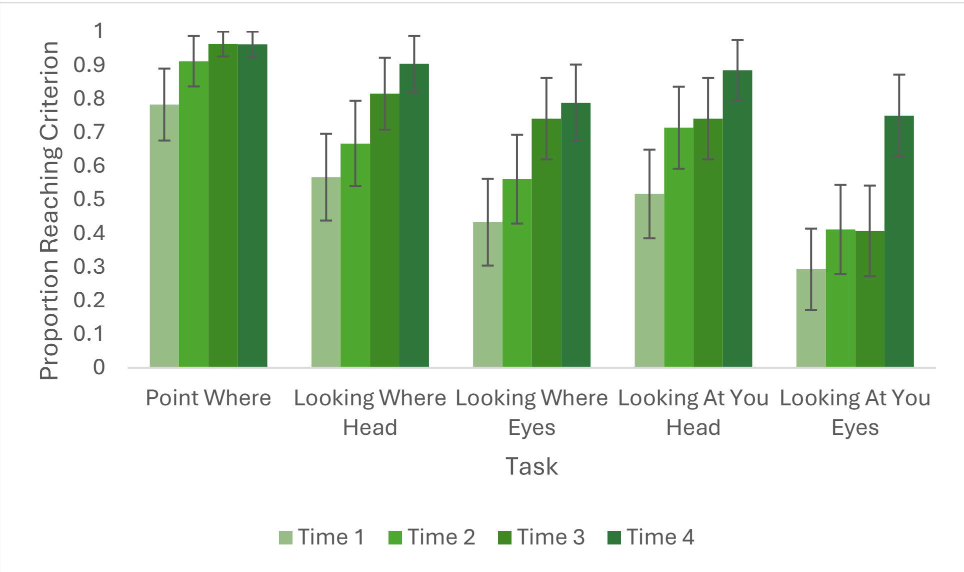
Gaze Judgement

Preliminary analyses show that the only difference between schematic and photo task versions was for the Looking Where Head task at Time 2 ($M1$ = .83, $M2$ = .50; Mann Whitney U = 273, p = .002). All other tests of difference were non-significant (all ps > .100). Thus, the scores for both versions of the tasks are combined for the remaining analyses.

Figure 15 and Table 12 show performance on the Point Where task approached ceiling across all timepoints. All other tasks improved over time, with Head tasks approaching ceiling, and Eyes tasks reaching ~75% pass rate by Time 4.

Figure 15

Experiment 2 Performance on Gaze Judgment Tasks.



Note. Error bars represent 95% CI.

Table 12

Experiment 2 Mean (SD) and Number of Missing Datapoints of All Tasks by Timepoint.

Task	**Time 1**		**Time 2**		**Time 3**		**Time 4**	
	Mean (*SD*)	Missing	Mean (*SD*)	Missing	Mean (*SD*)	Missing	Mean (*SD*)	Missing
Gaze Judgement								
Point Where	3.28(*1.12*)	0	3.77(*.60*)	0	3.87(*.44*)	0	3.87(*.60*)	0
Looking Where Head	2.63(*1.21*)	0	3.07(*1.05*)	0	3.48(*.84*)	0	3.62(*.80*)	0
Looking Where Eyes	2.23(*1.38*)	0	2.67(*1.35*)	0	3.20(*1.17*)	0	3.42(*.96*)	0
Looking At You Head	3.14(*1.00*)	2	3.45(*1.01*)	1	3.46(*1.06*)	0	3.71(*.92*)	0
Looking At You Eyes	2.59(*1.11*)	2	2.89(*1.12*)	1	2.83(*1.19*)	0	3.54(*.90*)	0
Hiding Task								
Move Object	2.53(*1.34*)	1	2.89(*1.56*)	1	3.35(*.89*)	0	3.71(*.85*)	0
Move Screen	.66(*.91*)	2	1.16(*1.19*)	1	1.52(*1.50*)	0	2.83(*1.37*)	0
Theory of Mind								
Knowledge Access	.37(*.49*)	1	.54(*.50*)	1	.72(*.45*)	0	.65(*.48*)	0
False Belief	.23(*.43*)	0	.32(*.47*)	0	.52(*.50*)	0	.62(*.49*)	0
Inhibitory Control								
Dragon Task	3.79(*5.54*)	8	5.18(*6.48*)	2	7.52(*6.81*)	2	10.90(*6.43*)	2
Vocabulary								
Understands	104(*39.9*)	1	109(*27.9*)	0	121(*26.9*)	2	132(*27.1*)	1

To account for differences in chance baseline, tasks were categorised as pass or fail using the criteria described in the Methods. A repeated measures ANOVA indicated there was a main effect of task F (4) = 32.57, $p < .001$, $\eta_p^2 = .399$. Performance on the Point Where task was significantly better than all other tasks (all ps < .001). Performance on the Looking Where task was significantly better when head direction was a cue compared to eyes ($p = .002$) and for the Looking At You task ($p < .001$). When only eyes were a cue, Looking Where performance was better than Looking At You ($p < .001$). These did not differ when head direction was also a cue ($p = .322$). There was also a main effect of timepoint F (3) = 18.59, $p < .001$, $\eta_p^2 = .275$. Performance significantly improved between Time 1 and Time 2 ($p = .019$), and Time 3 and Time 4 ($p < .001$), but not between Time 2 and Time 3 ($p = .145$). Finally, there was a significant interaction between timepoint and task F (12) = 1.80, $p = .045$, $\eta_p^2 = .035$.

Planned comparisons show that improvements for the Point Where and Looking Where Head tasks were not significant between consecutive timepoints (all ps > .182). Significant improvements in performance occurred between Time 2 and Time 3 ($p = .038$) for the Looking Where Eyes task, between Time 1 and Time 2 ($p = .038$), and Time 3 and Time 4 ($p = .019$) for the Looking At You Head task, and between Time 3 and Time 4 for the Looking At You Eyes task ($p < .001$).

Consistent correlations across timepoints are between the Looking Where tasks and between the Looking At You tasks (see Table 13). The Point Where task correlated with all other gaze judgement tasks. As the Point Where task is largely an attention check, performance is near ceiling at all timepoints and is significantly easier than all gaze tasks, this will not be analysed further. The remaining analyses will consider the Eyes tasks only to provide a specific picture of the developing understanding of eye direction alone.

Table 13

Experiment 2 Correlations Gaze Judgement Tasks Using Criterion Scores By Timepoint.

Time 1	**Looking Where Head**	**Looking Where Eyes**	**Looking At You Head**	**Looking At You Eyes**
Point Where	.520***	.297*	.273*	.346**
Looking Where Head	-	.629***	.065	.178
Looking Where Eyes		-	.041	.204
Looking At You Head			-	.410**
Time 2	**Looking Where Head**	**Looking Where Eyes**	**Looking At You Head**	**Looking At You Eyes**
Point Where	.044	.351**	.356**	.007
Looking Where Head	-	.350**	.157	.031
Looking Where Eyes		-	.227	.239
Looking At You Head			-	.367**
Time 3	**Looking Where Head**	**Looking Where Eyes**	**Looking At You Head**	**Looking At You Eyes**
Point Where	.159	.331*	.331*	-.037**
Looking Where Head	-	.479***	.153	.104
Looking Where Eyes		-	.421**	.147
Looking At You Head			-	.405**
Time 4	**Looking Where Head**	**Looking Where Eyes**	**Looking At You Head**	**Looking At You Eyes**
Point Where	.274*	.386**	.241	.346*
Looking Where Head	-	.310*	-.118	-.038
Looking Where Eyes		-	.108	.353*
Looking At You Head			-	.487***

Note. Correlations are Spearman's Rho. $^{\dagger}p<.10$, $*p<.05$, $**p<.01$, $***p<.001$, 2-tailed.

Hiding Tasks

Figure 16 and Table 12 show performance on both tasks improved over time, with Move Object approaching ceiling, and Move Screen reaching 63% pass rate by Time 4.

Figure 16

Experiment 2 Performance on Hiding Tasks.

Note. Error bars represent 95% CI.

Tasks were categorised as pass or fail where a pass was at least 3 out of 4 trials correct. A repeated measures ANOVA indicated there was a main effect of task F (1) = 132.51, $p <$ .001, η_p^2 = .730. Performance on the Move Object task was significantly better than the Move Screen task ($p <$ .001). There was also a main effect of timepoint F (3) = 34.06, $p <$ .001, η_p^2 = .410. Performance significantly improved between Time 2 and Time 3 ($p <$ .001), and Time 3 and Time 4 ($p <$ .001), but not between Time 1 and Time 2 (p = .308). Finally, there was not a significant interaction between timepoint and task F (3) = 2.13, p =.099, η_p^2 = .042.

Planned comparisons show that improvements for both tasks occurred between Time 2 and Time 3 (Move Object: p = .002; Move Screen: p =.028), and Time 3 and Time 4 (Move Object: p = .032; Move Screen: $p <$ .001). The hiding tasks were significantly associated at Time 2 (r_s (54) = .266, p =.047) and Time 4 (r_s (50) = .326, p = .018), but not at Time 1 (r_s (56)

= .170, p = .201) or Time 3 (r_s (52) = .118, p = .396). The Move Object task largely serves as an attention check here, and thus the remaining analysis will only include Move Screen performance.

Errors on the Move Screen task were predominantly egocentric, then neutral, and rarely other. This pattern was consistent across all timepoints (see Table 14).

Table 14

Experiment 2 Proportion Move Screen Errors By Timepoint

Task	**Egocentric**	**Neutral**	**Other**	**Total Errors**
Time 1	.61	.29	.10	194
Time 2	.67	.32	.01	159
Time 3	.71	.29	.00	130
Time 4	.54	.46	.00	57

Theory of Mind

Figure 17 and Table 12 show performance on both tasks improved over time. Neither task approached ceiling by Time 4, reaching ~63% pass rate.

Figure 17

Experiment 2 Performance on Theory of Mind Tasks.

Note. Error bars represent 95% CI.

A repeated measures ANOVA indicated there was a main effect of task F (1) = 9.598, p = .003, η_p^2 = .161. Performance on the Knowledge Access task was significantly better than the False Belief task (p = .003). There was also a main effect of timepoint F (3) = 12.594, $p <$ .001, η_p^2 = .201. Performance significantly improved between Time 2 and Time 3 (p = .002), but not between Time 1 and Time 2 (p = .096), or Time 3 and Time 4 (p = .627). Finally, there was not a significant interaction between timepoint and task F (3) = .868, p =.459, η_p^2 = .017.

Planned comparisons show that improvements for both tasks occurred between Time 2 and Time 3 (Knowledge Access: p =.015; False Belief: p =.028). The theory of mind tasks were significantly associated at Time 3 (r_s (52) = .313, p =.021) and Time 4 (r_s (50) = .339, p = .014), but not at Time 1 (r_s (57) = -.018, p = .891) or Time 2 (r_s (54) = .027, p = .841).

Bear-Dragon

Table 12 show performance the Dragon task improved over time reaching 73% by Time 4.

A repeated measures ANOVA indicated there was a main effect of timepoint *F* (3) = 14.6, $p < .001$, $\eta_p^2 = .250$. Performance significantly improved between Time 2 and Time 3 ($p = .003$), and Time 3 and Time 4 ($p = .004$), but not between Time 1 and Time 2 ($p = .380$). Finally, there was a significant interaction between timepoint and task *F* (3) = 7.82, $p <.001$, $\eta_p^2 = .151$.

DVAP

Table 12 shows the number of words a child understands improved with time, reaching 65% of the total word by Time 4.

A repeated measures ANOVA indicated there was a main effect of timepoint *F* (3) = 33.6, $p < .001$, $\eta_p^2 = .406$. Performance significantly improved between Time 2 and Time 3 ($p < .001$), and Time 3 and Time 4 ($p <.001$), but not between Time 1 and Time 2 ($p = .128$).

Latent Growth Curve Models

The following latent growth curve models were conducted on the pass/fail data for each of the target tasks. As such, the models were run using an estimator called Weighted Least Squares Mean and Variance adjusted (WLSMV) as it is good at handling binomial data (Li, 2016). However, this estimator did not allow the model to converge for the False Belief task, meaning it could not find a solution. To overcome this, estimator Maximum Likelihood with Robust Standard Errors (MLR) was used instead (Li, 2016). Quadratic models which show how changes may speed up or slow down over time were also tested. These did not converge for any of our measures, thus linear models are reported.

To test whether the models fit the data well, $\chi 2$ test is commonly used where a non-significant result indicates a good fit. However, the current study has a small sample size so the test may say there is a good fit simply because it lacks power (Satorra & Saris, 1985). Thus, additional measures of model fit are used: RMSEA < .08, CFI >= .90 and TLI >= .95 (Hu & Bentler, 1999). The models for the Looking Where Eyes, Knowledge Access and False Belief tasks fit the data well as all measures of model fit were above the acceptable cutoffs. However, the models for Looking At You and Move Screen tasks did not fit the data well as the CFI and TLI values are lower than acceptable. This could be a result of the small sample size and data not missing at random. The theoretical value of the data exploration warrants cautious interpretation of the current models.

Table 15

Experiment 2 Latent Growth Curve Models for Gaze Judgement, Move Screen and Theory of Mind Tasks.

Model fit	**Looking Where Eyes**	**Looking At You Eyes**	**Move Screen**	**Knowledge Access**	**False Belief**
Chi-Square	1.214	9.754	6.520	5.451	3.959
CFI	1.00	.606	.559	.956	1.00
TLI	1.239	.528	.471	.947	1.11
RMSEA	0.000	.139	0.079	0.042	.000
Intercepts					
Intercept	.488***	.288***	.022	.414***	.214***
Slope	.027***	.031***	.046***	.025**	.033***
Covariance					
Int-slope	-.001	.000	.000	-.012*	.002
Variance					
Intercept	.057	.057	-.006	.115*	.030
Slope	.000	-.000	.001	.002**	-.000
Time 1	.197***	.165***	.045	.128***	.147***
Time 2	.191***	.200***	.118***	.202***	.179***
Time 3	.121***	.209***	.186***	.151***	.172***
Time 4	.074	.198***	.141***	.094	.177***

The models show similar intercepts and slopes for the Looking Where Eyes and Knowledge Access task, and Looking At You and False Belief task demonstrating they have concurrent trajectories. However, the Move Screen task performance is not similar to either gaze judgement or theory of mind ability as it is near floor at Time 1 and development is more rapid. Children who performed better on the Knowledge Access task at Time 1 tended to show less improvement over time (β = -.012, p = .036). For all other tasks, children's rate of development is not related to their performance at Time 1, meaning that if a child fails a task at Time 1, their performance improves at the same rate as a child who passes the task at Time 1. The variance of the intercept for Knowledge Access is significant, suggesting that there is meaningful between-person variability at Time 1. This is not significant for all other tasks.

Within Timepoint Comparisons

Performances were classified as pass or fail for comparison. As in Experiment 1 passing criterion for Looking Where tasks was set as 3 out of 4 and for Looking At You tasks least 4 out of 4. For the Move Screen task, the passing criterion was 3 out of 4.

The Move Screen task was the most difficult task throughout Time 1, Time 2 and Time 3. By Time 2, gaze judgement tasks were easier than theory of mind tasks, and by Time 3 theory of mind tasks were of roughly equal difficulty. At Time 4, False Belief was the most difficult task, with all other tasks equivalent. Thus, there is a general trend for passing gaze judgement tasks before theory of mind tasks. Passing the Move Screen task appears only possible after children can pass gaze judgement and theory of mind tasks equally.

Table 16

Experiment 2 Binomial Differences of Move Screen, Gaze Judgement and Theory of Mind Tasks By Timepoint.

Time 1	**Looking Where Eyes**	**Looking At You Eyes**	**Knowledge Access**	**False Belief**
Move Screen	2 – *26****	2 – *17****	2 – *20****	1 – *13****
Looking Where Eyes	-	15 – *7* †	16 – *12*	19 – *7**
Looking At You Eyes		-	8 – *13*	13 – *9*
Knowledge Access			-	17 – *9*
False Belief				-
Time 2	**Looking Where Eyes**	**Looking At You Eyes**	**Knowledge Access**	**False Belief**
Move Screen	3 – *26****	3 – *17***	7 – *17**	6 – *10*
Looking Where Eyes	-	15 – *7* †	19 – *7**	24 – *5****
Looking At You Eyes		-	15 – *11*	18 – *8**
Knowledge Access			-	17 – *9*
False Belief				-
Time 3	**Looking Where Eyes**	**Looking At You Eyes**	**Knowledge Access**	**False Belief**
Move Screen	4 – *27****	12 – *17*	4 – *26****	9 – *20**
Looking Where Eyes	-	22 – *4****	9 – *8*	17 – *5**
Looking At You Eyes		-	7 – *24***	9 – *15*
Knowledge Access			-	15 – *4**
False Belief				-
Time 4	**Looking Where Eyes**	**Looking At You Eyes**	**Knowledge Access**	**False Belief**
Move Screen	6 – *14*†	6 – *12*	9 – *10*	7 – *6*
Looking Where Eyes	-	7 – *5*	13 – *6*	15 – *6**
Looking At You Eyes		-	12 – *7*	14 – *7*
Knowledge Access			-	9 – *21**
False Belief				-

Note. The first number in each pair represents the number of children passing the row task and failing the column task; the second number represents the number of children failing the row task and passing the column task. †p<.10, *p<.05, **p<.01, ***p<.001, 2-tailed.

Correlation analysis (see Table 17) showed associations between the Looking Where Eyes and Looking At You Eyes task across all four timepoints. Theory of mind tasks were associated at Time 3 and Time 4, and Move Screen associated with Looking At You Eyes and both theory of mind tasks at Time 4.

Previous literature shows inhibitory control is related to theory of mind performance (Carlson et al., 2002; Carlson et al., 2015; Hughes & Ensor, 2007; Marcovitch et al., 2015), and potentially contributes to Move Screen performance. Here the Dragon task associates with Knowledge Access at Time 4, and Move Screen at Time 2 and Time 4. Additionally, associations with language were of interest given the verbal nature of all tasks and existing literature on the influence of language development on theory of mind success (e.g. Astington & Jenkins, 1999; Farrar & Maag, 2022; Schick et al., 2007). Here vocabulary size associates with theory of mind performance at Time 2 to 4, and Move Screen at Time 3. As such, inhibitory control and language should be controlled for.

When controlling for inhibitory control and language, Move Screen associations with Looking At You Eyes and Knowledge Access were no longer significant. All other associations persisted. Thus, the lack of between task associations within timepoints suggest gaze judgement, theory of mind and Engagement do not share contemporaneous competencies.

Table 17

Experiment 2 Correlations All Tasks Using Raw Scores By Timepoint. Partial Correlations are Shown Below the Diagonal Controlling for Dragon Task and DVAP.

Time 1	**Move Screen**	**Looking Where Eyes**	**Looking At You Eyes**	**Knowledge Access**	**False Belief**	**Dragon Task**	**DVAP**
Move Screen	-	-.069	.095	-.03	.095	.053	-.101
Looking Where Eyes	-.109	-	.290*	.022	.110	-.063	-.012
Looking At You Eyes	.052	.311*	-	.223†	-.052	-.005	-.061
Knowledge Access	-.022	.041	.226	-	-.018	.029	-.152
False Belief	.122	.025	-.039	.024	-	.153	.213
Time 2	**Move Screen**	**Looking Where Eyes**	**Looking At You Eyes**	**Knowledge Access**	**False Belief**	**Dragon Task**	**DVAP**
Move Screen	-	.022	.197	.065	-.124	-.308*	.146
Looking Where Eyes	.024	-	.263*	.029	.022	.026	.125
Looking At You Eyes	.179	.266 †	-	-.057	.076	-.089	.120
Knowledge Access	.108	.014	-.092	-	.027	.005	.006
False Belief	-.172	-.002	.065	.040	-	.120	.271*
Time 3	**Move Screen**	**Looking Where Eyes**	**Looking At You Eyes**	**Knowledge Access**	**False Belief**	**Dragon Task**	**DVAP**
Move Screen	-	.096	-.066	.223	.175	.200	.305*
Looking Where Eyes	.134	-	.199	.180	.238 †	.058	.059
Looking At You Eyes	.079	.180	-	-1.31	.140	.101	-.227
Knowledge Access	.079	.184	.010	-	.313*	.079	.306*
False Belief	.082	.226	.256†	.232	-	.109	.236 †
Time 4	**Move Screen**	**Looking Where Eyes**	**Looking At You Eyes**	**Knowledge Access**	**False Belief**	**Dragon Task**	**DVAP**
Move Screen	-	.045	.296*	.308*	.438**	.420**	.213
Looking Where Eyes	-.044	-	.342*	-.012	.021	.067	-.017
Looking At You Eyes	.223	.332*	-	.128	.059	.098	-.028
Knowledge Access	.137	-.051	.001	-	.339*	.291*	.281*
False Belief	.356*	.006	.095	.323*	-	.253†	.330*

Note. Correlations are Spearman's Rho. †p<.10, *p<.05, **p<.01, ***p<.001, 2-tailed.

With children's failures being mostly egocentric, it is possible that inhibitory control may allow children to overcome this egocentric bias to correctly place the screen. To explore this, the association between the Dragon task and egocentric responses was analysed. Egocentric responses on the Move Screen task were associated with Dragon Task performance at Time 4 only, where the higher the inhibitory control the fewer egocentric responses are given (r_s (48) = -.514, $p < .001$). At all other timepoints, this association was not significant (all *p*s > .078).

Between Timepoint Comparisons

To gain a stable overall view of gaze judgement ability for the following analyses, Looking Where Eyes and Looking At You Eyes raw scores were combined (see Table 18).

Table 18

Experiment 2 Mean (SD) Overall Gaze Judgement Ability Indicated by the Total Number of Eyes-Only Task Trials Passed.

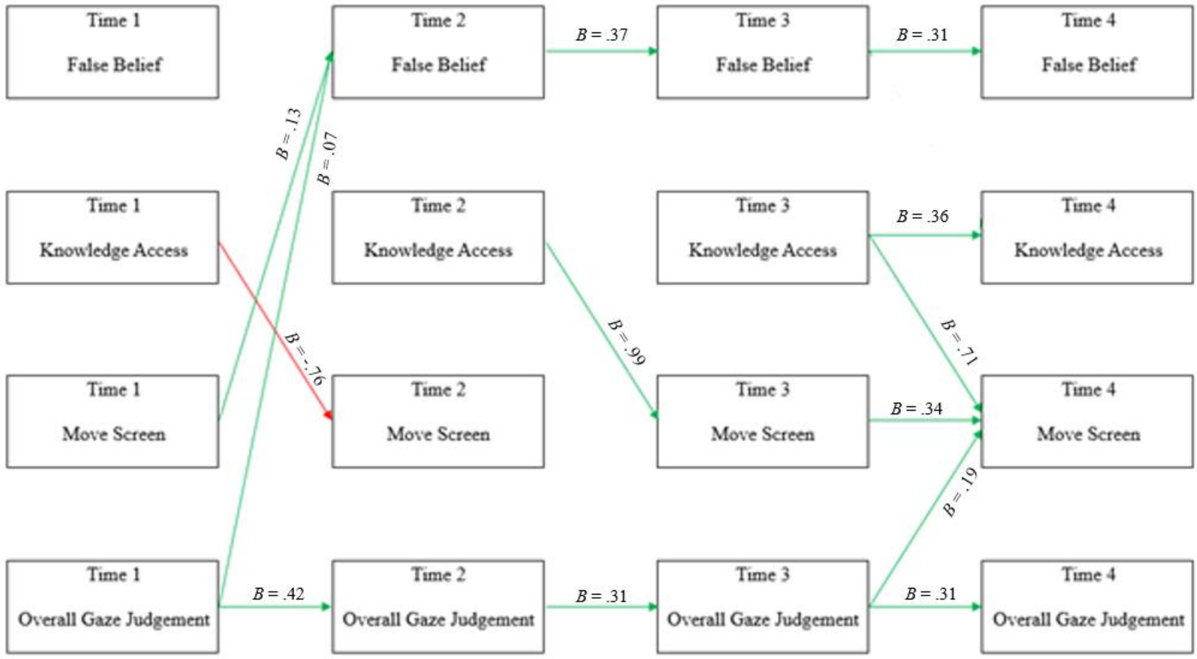
Overall Gaze Judgement Ability	**Mean**	**SD**	**Missing**
Time 1	4.79	2.01	2
Time 2	5.54	*1.93*	1
Time 3	6.04	*1.83*	0
Time 4	6.96	*1.51*	0

Below the general pattern emerges that earlier Overall Gaze Judgement predicts later ability across all timepoints. This is also true of False Belief performance between Time 2 and Time 4, and Move Screen and Knowledge Access between Time 3 and Time 4. Between tasks, earlier Knowledge Access performance predicts later Move Screen performance. Between Time 1 and Time 2 this is negative, whilst between Time 2 and Time 4 this is positive. Move Screen and Overall Gaze Judgement ability are significant predictors of False Belief between Time 1 and Time 2 only. Overall Gaze Judgement ability was a significant predictor of Move

Screen between Time 3 and Time 4 only. See Figure 18 for a visual summary of the cross-lagged predictions.

Figure 18

Experiment 2 Summary of Cross-lagged Regression Analyses for All Tasks.



Note. Red arrows refer to negative predictors. Green arrows refer to positive predictors.

Table 19

Experiment 2 Summary of Backward Regression Analysis for Cross-lagged Prediction.

Variable	*F*	R^2	*B*	*SE*	*t*
Time 2 Move Screen	5.13*	.09			
Time 1 Knowledge Access			-.76	.34	-2.27*
Time 2 Overall Gaze Judgement	12.42***	.20			
Time 1 Gaze Judgement			.42	.12	3.56***
Time 2 False Belief	4.18*	.14			
Time 1 Move Screen			.13	.07	2.02*
Time 1 Overall Gaze Judgement			.07	.03	2.19*
Time 3 Move Screen	4.47*	.16			
Time 2 Overall Gaze Judgement			-.11	.065	$-1.84^{\dagger}$
Time 2 Knowledge Access			.99	.39	2.54*
Time 3 Overall Gaze Judgement	5.96*	.11			
Time 2 Overall Gaze Judgement			.31	.13	2.44*
Time 3 False Belief	5.47**	.19			
Time 2 Knowledge Access			.24	.13	$1.81^{\dagger}$
Time 2 False Belief			.37	.14	2.66**
Time 4 Move Screen	10.53***	.40			
Time 3 Move Screen			.34	.10	3.26**
Time 3 Overall Gaze Judgement			.19	.06	3.43**
Time 3 Knowledge Access			.71	.34	2.08*
Time 4 Overall Gaze Judgement	7.80**	.14			
Time 3 Overall Gaze Judgement			.31	.11	2.79**
Time 4 Knowledge Access	6.52*	.12			
Time 3 Knowledge Access			.36	.14	2.55*
Time 4 False Belief	5.56*	.10			
Time 3 False Belief			.31	.13	2.36*

Note. $^{\dagger}p<.10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

Discussion

The primary aim of the current experiment was to investigate longitudinally the developmental trajectory of gaze understanding, and consider its precursor relationship with theory of mind. The results show that gaze judgement does go through a protracted development, which is poor at 2-years-old and rapidly improves until 4-years. Contrary to the claims of the two-system theory, gaze understanding is not integral to the development of theory of mind. Instead, the data suggests it is an immediate precursor to the onset of its development. The secondary aim was to examine whether a sophisticated understanding of gaze marks the transition to mentalistic-based strategies to predict behaviour. The lack of association between the three abilities suggests that they may not share a common understanding. However, earlier understanding of gaze at Time 3 and knowledge acquisition at all timepoints predicts later Move Screen performance. Thus, an understanding of knowledge formation may allow children to distinguish between engagement and perception, motivating the transition to mentalistic-based strategies to predict behaviour.

Performance on the gaze judgement tasks replicated those of Doherty and colleagues (1999; 2009) and Experiment 1. Performance on the point where task was near ceiling at all timepoints, demonstrating that children could respond to the task demands even at a young age. However, at late 2-years-old children had difficulty making judgments of eye direction, identifying where someone was looking around 49% of the time and which of two faces were looking at them 29% of the time. Their performance improves to around 75% by the time they turn 4-years-old, with notable improvements between Time 1 and Time 2, and Time 3 and Time 4. Thus, the finding that children's understanding eye direction goes through protracted development is robust. The current findings add value to the literature, as they demonstrate no meaningful between-child variance in initial performance, or rate of development. Specifically, children who perform worst at Time 1 did not have a faster rate of development than those who perform best at Time 1. Thus, the developmental trajectory of gaze understanding is likely to be stable across children.

The children in the current study demonstrate a typical trajectory for theory of mind development (Wellman & Liu, 2004). Both Knowledge Access and False Belief tasks are poor at 2-years-old, showing significant improvements across the timepoints. Neither task had near ceiling performance at Time 4. However, this could reflect the fact that Time 4 age was bordering 4-years-old and did not capture consistent success that is seen in older 4-year-olds. As expected, Knowledge Access was easier than False Belief and showed greater between child variance for initial performance and rate of development. Specifically, children who performed better at Knowledge Access at Time 1 had a slower rate of development than children who performed worse.

The hiding task results replicate those of the literature (Flavell et al., 1978; McGuigan & Doherty, 2002) and Experiment 1. Children performed better on the Move Object task than the Move Screen task at all timepoints. Performance on the Move Screen task started at near floor, with success reaching around 63% by Time 4, with notable improvements occurring between Time 2 and Time 3, and Time 3 and Time 4. The current findings add value to the literature, as they demonstrate no meaningful between child variance in initial performance, or rate of development. Thus, overcoming the signature limits of children's understanding of occlusion goes through protracted development, which is stable between children.

Interestingly, Move Screen performance was strongly correlated with inhibitory control at timepoints 2 and 4. However, this was negative at Time 2 and positive at Time 4. The inconsistency of this association begs the question if this is a robust finding. In any case that this may be robust, inhibitory control may confer success on the Move Screen task by allowing children to overcome egocentric biases. Across all timepoints, children's incorrect responses on the Move Screen task were egocentric more often than not. At the earlier timepoints, children are more likely using a concept of engagement in the Move Screen task. So, when they are asked to put the screen so that the experimenter cannot see the object, children do not know how placing it will disrupt the engagement relation. What they do know, is how placing the screen can stop themselves from seeing the object. Thus, egocentric placement becomes the default response. Around Time 4, children are likely transitioning towards use of line of sight concepts, where a persisting barrier to success on the task may be the default egocentric response. Without the necessary inhibitory control, children are likely to have difficulty suppressing this salient response (Rothbart & Posner, 1985). This is supported by the finding that inhibitory control and egocentric responses are significantly associated at Time 4 only, such that a child with greater inhibitory control made fewer egocentric responses. Thus, children who are better able to inhibit the default egocentric response may be able to better navigate the demands of the Move Screen task. This may also contribute to the lack of near ceiling performance on the three Blindfold variations of this task in Experiment 1.

In line with the previous literature (Doherty & Anderson, 1999; McGuigan & Doherty, 2002) and Experiment 1's findings, the trajectories of all three abilities go through a similar developmental time course. However, gaze understanding develops slightly sooner than the other two abilities. This is most apparent when comparing the within timepoint effects, where the Move Screen and theory of mind tasks were consistently harder than gaze judgement tasks. Despite the concurrent development, there is a lack of association between all three abilities at timepoints 1-3. Time 4 is the exception where Move Screen associates with both abilities, however only associations with False Belief persist when controlling for inhibitory control and language. First, one can conclude that understanding of eye direction does not result from children attributing mental states as neither task involving line of sight associates with theory of mind. Second, the lack of association between the Move Screen and gaze judgement tasks is surprising, as both tasks by their nature require understanding of line of sight, thus their performance should reflect this shared understanding. However, it may be that understanding such as spatial reasoning (see Chapter 2 for a discussion) contributes to Move Screen success more than that of line of sight.

The lack of association between gaze judgement and theory of mind ability, begs the question of whether there would be a precursor relationship between them. To start, the predictions of the two-system theory are tested. The two-system theory claims that children become motivated to understand eye direction because they are interested in the mind (Doherty, 2006; Doherty et al., 2009). As such, earlier theory of mind ability would predict later gaze understanding. Across all four timepoints, earlier theory of mind ability did not predict later gaze judgement performance. Earlier gaze judgement performance was the only significant predictor. This is a clear finding that does not support the claims of the two-system theory.

Alternatively, the predictions of the consensus view are of interest. That being that gaze understanding is a precursor to theory of mind development, as visual perception is how knowledge and beliefs are formed (Goméz, 1996; Gopnik et al., 1994). Given the relative ease of gaze judgement in comparison to theory of mind in the current findings, there is credence for this hypothesis. If the hypothesis is supported, then earlier understanding of eye direction would predict later theory of mind ability. The current findings partially support this. Specifically, gaze understanding and Move Screen performance at Time 1 is a significant predictor of False Belief performance at Time 2. Thereafter, earlier False Belief performance is the only significant predictor of later False Belief performance. There are no earlier predictors of Knowledge Access, with the exception of Time 4 which is predicted by Time 3 Knowledge Access. Thus, understanding of line of sight may be an immediate precursor to understanding of beliefs. Thereafter, the abilities develop concurrently but do not share a common understanding.

This precursor relation however seems implausibly direct. Specifically, it is well documented that children acquire Knowledge Access before False Belief (e.g. Wellman & Liu, 2004; Wellman et al., 2011; Peterson et al., 2012), meaning any precursor relationship would be expected to be towards Knowledge Access first. Additionally, understanding visual attention is a core function of reasoning in Knowledge Access task. To determine if someone knows the true state of affairs, children must understand that seeing leads to knowing. Thus, understanding that people are psychologically connected to objects through their eyes should show stronger and earlier predictions with Knowledge Access than False Belief. It is likely that this was not reflected in the data as Time 1 may have been measured too late in development given that performance on neither the Knowledge Access nor gaze judgement tasks was near-floor. As such, it cannot be strongly concluded that gaze understanding is an immediate precursor to theory of mind, but that the data lends partial support to the consensus viewpoint (Goméz, 1996; Gopnik et al., 1994).

The current experiment was also interested in whether children use the concept of engagement to fake understanding of gaze or mental states. It was predicted that if this was the case, then a sophisticated understanding of gaze would mark the transition to mentalistic-based strategies to predict behaviour. Specifically, earlier gaze understanding and theory of mind ability would predict later Move Screen performance. Firstly, earlier gaze understanding predicted later Move Screen performance between Time 3 and Time 4 only. Secondly, earlier understanding of knowledge predicts later Move Screen performance. This could be interpreted that conceiving of attention as a mental state may provide children with the necessary tools to distinguish between knowledge and perception required to succeed at the Move Screen task. It is not surprising that earlier False Belief performance only predicted Move Screen at Time 4, as the False Belief task additionally requires understanding of misrepresentation which develops later than that of knowledge acquisition. Together, the results suggest as children develop an understanding of knowledge formation, they move towards mentalistic-based strategies. This transition is likely marked by a sophisticated understanding of gaze, allowing children to conceive of attention in terms of mental states.

To conclude, the results of the current experiment partially replicate previous findings. Gaze understanding goes through protracted development in the preschool period, simultaneously with theory of mind. Despite the contemporaneous development, neither ability share a common understanding. The results allude to gaze understanding being an immediate precursor to theory of mind, however measurements at earlier ages are required to better explore this relation. Contrary to previous findings, Move Screen performance did not associate with gaze understanding. Instead, other factors such as spatial reasoning and inhibitory control may be larger contributing factors to success on the Move Screen task. Never-the-less, gaze understanding and knowledge acquisition may provide children with the tool necessary to distinguish between seeing and engagement, motivating the transition to mentalistic-based strategies to predict behaviour.

Chapter 4

Investigating if System 1 Remains Distinct in Adulthood

Research has shown whilst infants can follow gaze from their second year (Corkum & Moore 1995), they cannot make judgements about gaze direction until the preschool period (Doherty & Anderson 1999). These abilities appear to be distinct to one another not only in development, but in function (Doherty et al., 2009), suggesting gaze understanding involves two systems. As with other two-system theories (Kahneman, 2011) if there are two systems for processing gaze, then both systems are likely to be available in adulthood. With adults having a high degree of accuracy for discriminating gaze direction (Vida & Maurer, 2012a; 2012b), one way to dissociate the systems could be through their use of different information from the eyes. Doherty et al. (2015) presented evidence from child data suggesting gaze-following is luminance-based, using the ratio of light to dark areas of the eye, whereas gaze-judgement is also geometric-based, calculating the spatial location of the pupil in the eye region from edges. However, they also find that adult gaze following appears to rely on both luminance and geometric cues (Doherty et al., 2015).

The initial aim of the current chapter was to investigate whether gaze following is consistent with a luminance-based System 1 in childhood, and at what point System 1 incorporates geometric information in its processing as seen in adulthood. The hypothesis was that if System 1 becomes more flexible in its use of geometric cues, then gaze following would gradually incorporate them over the adolescent period. If however geometric cues are available for use immediately after they are acquired, then they would be incorporated into gaze following suddenly at 4- to 5-years.

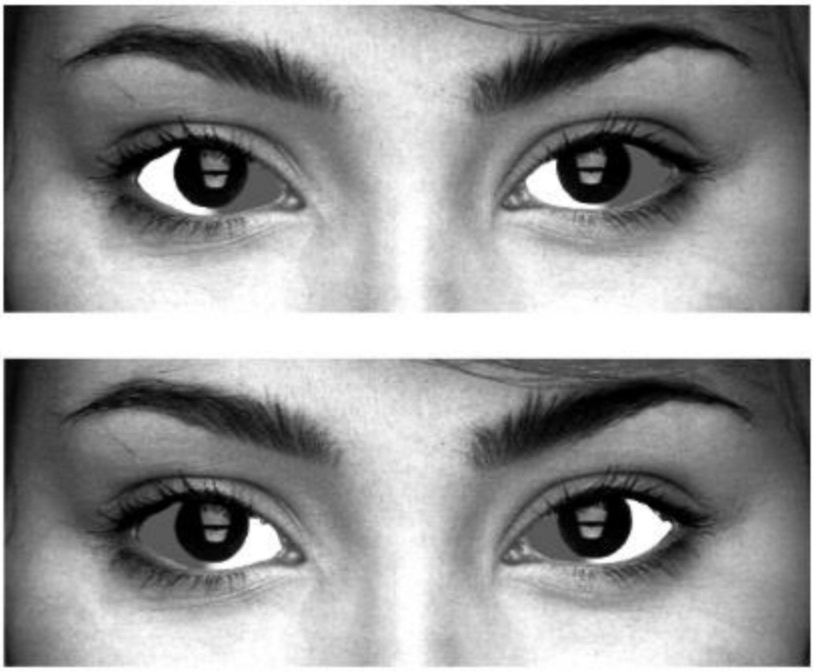
One hundred and twenty-eight children aged 2- to 11-years-old (M = 6.45, SD = 2.30) were recruited in a direct replication of Doherty et al.'s (2015) gaze cueing task. However, the task lacked sensitivity as no gaze effects were found. Specifically, target identification was not more accurate when they appeared in the gazed at location than when they appeared in the not gazed at location. Typically, younger children chose one of the two targets to respond with on every trial, whilst typically older children correctly identified the target on every trial. This resulted in children having floor or ceiling performance for cued and non-cued targets meaning no gaze cueing effects could be detected. As such, the hypothesis could not be tested. Reaction time data may have overcome the lack of sensitivity. Given the target age this would have required eye tracking data which was not possible to collect with time and funding restraints on this project. Thus, the new aim of the current chapter is to investigate whether adult gaze following is consistent with a luminance-based System 1, or if it additionally processes geometric cues. This chapter also considers alternative measures of automatic attention orienting to assess robustness of the finding.

There are two types of visual information that we obtain from the eyes. Traditional accounts emphasise computing the edges of the eye and sclera to make a geometrical calculation of a gaze vector (see Jenkins & Langton, 2003 for a discussion). The white sclera in humans makes such edge extraction easier than in non-human primates. This requires some precision, potentially affording a slow, calculated judgement of where the eyes are directed. Thus, using edge information would be consistent with the properties of System 2. However, the sclera also offers another perhaps more direct cue: the ratio of luminance given by the white on each side of the dark iris. This allows an approximate measure of the rotation of the eye, from which gaze direction can be estimated. This approximation suggests a much more efficient, coarser processing of eye direction through luminance cues, which would be consistent with the properties of System 1. Thus, it is possible the two systems could be dissociated by the visual information they use.

The two types of cues can be separately disrupted (Ando, 2002; Jenkins & Langton 2003) and seem to make separate contributions to adult judgment. Ando's bloodshot illusion is a particularly striking demonstration for luminance information. Ando (2002) presented participants with faces that had their eyes directed forward. Without manipulating the edges of the eye or the iris, the scleral region to the left or the right of the iris was darkened (see Figure 19). A rod was attached at the top of the monitor, centred over the nose of the face. At the end of the rod was a vertical stick with a white mark on the bottom. The rod could be rotated over a range of 180°, allowing the participant to indicate when the white mark was in the line of sight of the eyes. The angle of rotation from the midline was used to indicate the degree of gaze perception. Participants perceived an average shift in gaze direction of 16 degrees from luminance information alone. Regarding geometrical information, Jenkins and Langton (2003) demonstrated that inverting the eye region severely disrupted processing. This type of manipulation does not alter the luminance ratio of the eyes but does alter the appeared geometry of the eye region. Specifically, an iris that is on the left hand side of the eye region in an upright face, will appear to be in the right hand side in an inverted face. This appeared geometrical change is sufficient to alter the perception of gaze direction.

Figure 19

Apparent Shift of Gaze Due to Luminance (Ando, 2002; Figure 1).



However, young children may only have access to luminance cues. The visual acuity of a 4-month-old is around 40 times lower than that of an adult (de Heering et al., 2008). To put this into perspective, an infant this old would see eyes as clearly as if they were 150 feet away. Visual acuity remains low throughout toddlerhood with norm acuity scores for tests such as the Cardiff acuity test, being 20/63 for 18- to 24-month-olds, reaching 20/20 by the time a child turns 3-years-old (see Figure 20) (Adoh & Woodhouse, 1994; Deves et al., 1996). Thus, infants and toddlers would be unable to effectively resolve the edges of the eyes, limiting them to the use of luminance cues. As such, it is possible luminance cues alone are used for gaze-following.

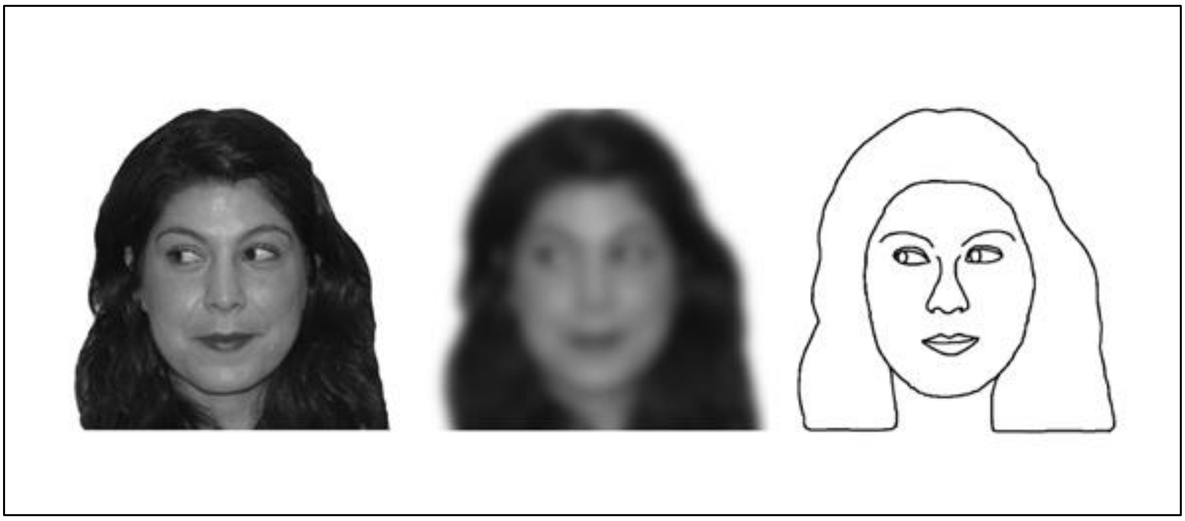
Figure 20

Visual Acuity Simulation of a 4-month-old (left), 18-month-old (middle), and 3-year-old (right). *SIGHT-SIM™*

Doherty et al. (2015) tested this hypothesis in a gaze cueing. Children were first shown an image of a woman facing forward, with eyes directed to either the left or right. Shortly after an animation character appeared very briefly, either to the side looked to or the opposite side. Children were asked simply to identify the character. If children were cued by the woman's gaze, they should be more likely to make a correct identification when the character appeared in the gazed-to location. To test the effect of psychophysical cues, they manipulated the faces in three ways, as shown in Figure 21. One set of photographs were filtered to remove the edges of eyes and face but left most of the luminance information intact. Another set were of the same face, but with only the lines intact. This was created by the low-tech method of tracing round the edges in Photoshop. They used these stimuli, plus the original greyscale photographs. They found that 2- and 3-year-old participants were cued by the greyscale photographs and the blurred photographs but were not cued by the line drawings. This could be interpreted as System 1 being luminance-based.

Figure 21

Normal, Blur, and Line Drawn Stimuli Used by Doherty et al. (2015). The Actual Stimuli Were Greyscale as Shown.



On their own these results do not support a two-systems claim. A single system might begin sensitive primarily to luminance information, and over time incorporate geometric information. Doherty et al. (2015) addressed this in a second experiment. They used the same gaze judgement task as Doherty and Anderson (1999), using the three manipulated images. Here, the two-systems theory predicts that children who could correctly judge gaze-direction should do so for both luminance and geometric cues. They found 2-year-olds were relatively good when the face was a photograph, getting just over half of trials correct. However, they were at or just above chance for the line drawing and blurred faces, respectively. In contrast, 3- and 4-year-old children were better at all three types of face, with a slight superiority for the photographs. Importantly, each age group was equally good at the blur and line images, and better still when they were combined. For explicit judgment each cue was about equally effective.

It is key to note here is that psychophysical cues used were not dependent on age, but on what children were asked to do. In the first experiment gaze processing was implicit: they did not ask or tell children anything about the face. Children were simply trying to spot the animation characters. Under these conditions, children were only cued by luminance cues. In the second experiment, gaze processing was explicit: children were asked what the lady was looking at. Under these conditions, children used both luminance and geometric cues, roughly equally and additively. Doherty et al (2015) conclude this as evidence that there are distinct systems for gaze-following and explicit gaze-judgements.

Doherty et al. (2015) also considered whether this dissociation persists into adulthood. They replicated their gaze-cueing study with adults and unlike the 3-year-olds adults were equally cued by all three stimuli. This suggests that gaze cueing involves different stimuli in adults and children: young children use only the luminance information in gaze, but adults are cued by both luminance and geometric information, either when combined or separately. This finding suggests that children's dissociation of luminance cues does not necessarily characterise System 1. Instead, it may be that as children's use of geometric cues for gaze-judgement increases, their use becomes more flexible. Therefore, a corresponding increase in use of geometric cues for gaze following would be seen. This is consistent with the findings described above.

A limitation of Doherty et al.'s (2015) study is that they only measure response accuracy. Accuracy scores simply demonstrate whether the participant was cued by either geometric or luminance cues. They do not however, indicate whether there is a difference in how these cues contribute to the processing of eye direction. One way this can be measured is reaction time. If one cue is more important than the other, then it would be most efficient for gaze following. As such, it would result in faster reaction times compared to the other cue. Support for this supposition comes from Munsters et al. (2016). They conducted a gaze-cueing task measuring manual and saccadic reaction times. Their face stimuli were passed through a Gaussian filter to create three images: low spatial frequency (i.e., luminance), unfiltered, and high spatial frequency (i.e., geometric). They found an overall cueing effect where cued targets elicited faster manual and saccadic reaction times than for uncued targets. The magnitude of this cueing effect did not differ between spatial frequencies, meaning participants were equally cued by luminance or geometric information. However, they did find a main effect of filtering where manual and saccadic reaction times were generally faster for luminance cues than geometric. This suggests that whilst both types of cues are used to follow gaze, luminance cues are processed faster than geometric. As such, adult System 1 may use both cues to achieve the same end goal but is primarily luminance-based in its processing.

If System 1 serves the function of using another's gaze to locate objects of attention, then similar patterns for luminance and geometric cues should be seen in other measures of attention orienting. One such measure is Samson et al.'s (2010) dot perspective task. In this task participants are shown a room with red dots on the walls. Inside the room there is also an avatar with their whole body facing to the left or right wall. On half the trials, the number of dots seen by the participant and avatar were the same (consistent), and the other half they were different (inconsistent). The participant's job was to identify if they themself could see a target number of dots. For example, if the target number was '2', the participant had to say whether they could see 2 dots in the room. They found reaction times were slower to respond when the perspectives were different compared to when the perspectives were the same. This finding has been replicated many times (e.g. Conway et al., 2017; Michael et al., 2018; Ramsey et al., 2013; Vestner et al., 2022).

There is large debate about what these findings suggest. The first explanation offered by the original authors (e.g. Samson et al., 2010; Qureshi et al., 2010; Schurz et al., 2015; Furlanetto et al., 2016) is that we spontaneously compute the other person's perspective even when it is irrelevant, which interferes with the processing of our own. The opposing explanation is the submentalising account, claiming that the interference is simply a result of our attention being automatically oriented to the same location as that of the avatar (e.g. Santiesteban et al., 2014; Conway et al., 2017; Heyes, 2014; Langton, 2018; Vestner et al., 2022). Various methods have been used to untangle the two, including brain imaging, use of barriers and non-social agents with no success of reaching a consensus. Holland et al.'s (2021) meta-analysis has come close to a consensus by suggesting that a large portion of the variance in dot perspective task score can be explained by attention orienting.

More convincingly, breaking down the dot perspective task to its constituent parts allows analogies with Doherty et al.'s (2015) gaze cueing task. First, the stimuli themselves have the same key features: an agent whose gaze is directed to either the left or right hand side of the screen and targets of interest that are either in the same or opposite direction as the gaze. The positioning of these features within a blank space (Doherty et al., 2015) or within a room (Samson et al., 2010) is arguably irrelevant. Second, both tasks ask participants to identify what they can see. Whether it involves identifying one of two cartoon characters (Doherty et al., 2015) or counting dots (Samson et al., 2010) is irrelevant. Irrespective of whether participants are spontaneously computing the avatar's perspective, slower responses when it differs is more likely due to participant's attention being directed to the wrong place, which takes time to correct.

Together, there is good reason to conclude that attention orienting explains the dot perspective task and can thus be an effective measure of System 1. As such, it would be expected that task performance would match that of gaze cueing tasks. That being, for both luminance and geometric information participants will be faster to respond when the avatar and the participant see the same number of dots compared to when they see a different number of dots. However, responses will be generally faster for faces retaining luminance cues compared to geometric.

In summary, we obtain two types of visual information from the eyes for gaze processing. The first is relative luminance of the iris and sclera, which is efficient to compute and available from birth. The second is the computation of the edges of the eye and sclera, which is more calculated to compute, and potentially not available until 3-years-old once visual acuity is sufficiently developed. Initially cued gaze relies on luminance cues, incorporating geometric cues into adulthood. However, luminance cues may remain important for efficient processing of cued gaze. The aim of the following studies it to test this prediction by considering the speed at which the two cues are processed for two measures of System 1 behaviour, Doherty et al.'s (2015) gaze cueing and Samson et al.'s (2010) dot perspective tasks.

Experiment 3

Experiment 3 will be a replication of Doherty et al.'s (2015) gaze cueing task, with the additional measure of reaction time. Previous research suggests that adult use of luminance and geometric cues does not dissociate for accuracy, but through their processing speed (Munsters et al., 2016). Thus, Doherty et al.'s (2015) finding that adult gaze following responds to luminance and geometric cues equally does not account for this fact. As such, it is hypothesised that there will be a main effect of cueing where participants will be more accurate and faster to respond to targets in the looked at location than the not looked at location. It is hypothesised that there will be a main effect of psychophysical cue where accuracy and reaction time will be higher and faster for luminance cues than geometric. Finally, there will not be an interaction between cueing and psychophysical cue for accuracy or reaction time.

Method

Participants

Forty-Five undergraduate students (35 Female, 8 Male, 2 prefer not to say) were recruited through SONA and received 1 SONA credit for their participation. Participants had a mean age of 20.8 years (*SD* = 4.18, *Range* = 18-38).

Inclusion criterion was informed consent prior to testing. The pre-registration (107757) specified no exclusion criteria. However, due to multiple conditions having skewed data (see Appendix A) 4 participants were identified as outliers and removed from the sample. Stopping criterion was 45 as outlined in the statistical analysis plan. Participants could withdraw their consent by closing the browser or responding 'I wish to withdraw when asked at the debrief. However, no participants chose to do so. The new sample included 41 participants with a mean age of 20.5 years (*SD* = 3.53, *Range* = 18-38) (3 Female, 6 Male, 2 prefer not to say).

The School of Psychology Ethics Committee granted this study ethical approval (ETH2122-2206).

Statistical Power and Analytical Plan

A 2x3 within-subject ANOVA is planned for analysis. To achieve a power of .80 with two-tailed alpha set to .05, the G-Power 3.1.9.2 statistical tool (Faul et al., 2007) indicates medium effect size $f = .25$, $N = 36$ (see Appendix C). Thus, I aimed to recruit 45 participants as this would exceed the required sample size for sufficient power and account for any withdrawals of consent.

Despite the data in 2 conditions being non-normally distributed (see Appendix A) data will be analysed using parametric tests of repeated measures ANOVA and post-hoc tests, as per the pre-registration. This is because the sample size is larger than 30, which according to central limit theorem the data distribution should approximate to normal (Field, 2003). It is also argued that parametric tests are robust to non-normal data (Rasch & Guiard, 2004).

Design

A 2x3 within-subject design was used. The independent variable cueing was split into two levels: cued and non-cued. The second independent variable, psychophysical cue, had three levels: normal, luminance and geometric. Response accuracy and speed were the dependent measures. Each psychophysical cue was presented in a single block, with block presentation fully counterbalanced across participants.

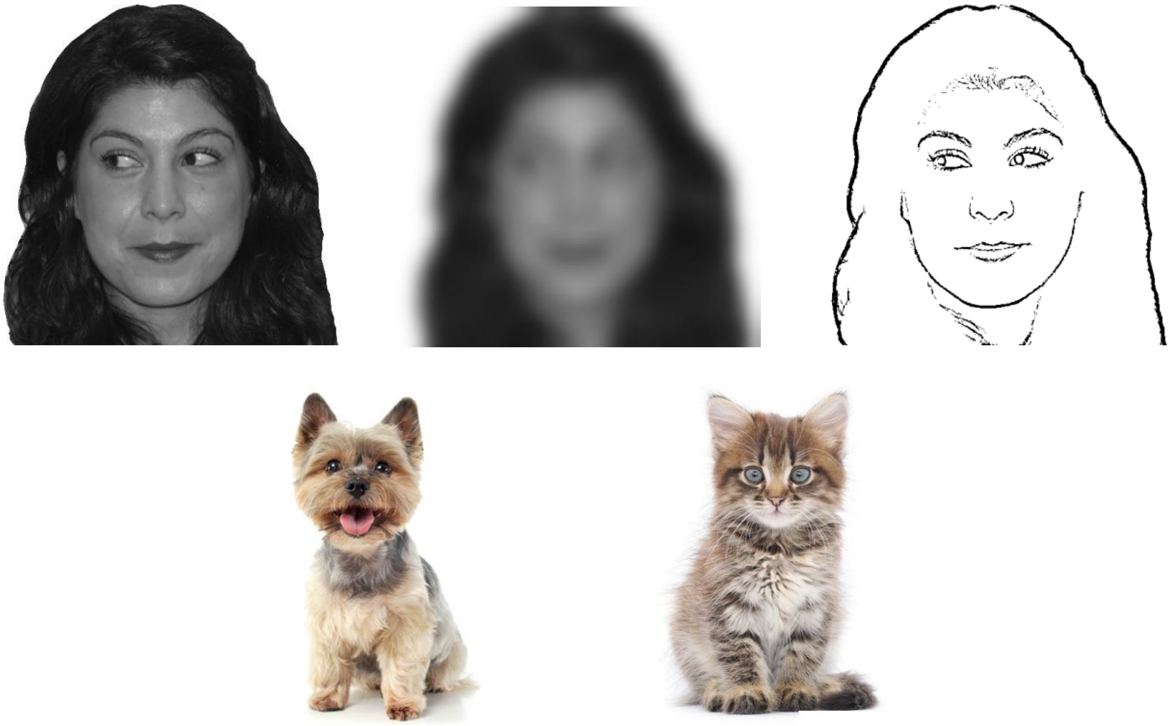
Materials

The study was conducted online using Gorilla. The set up was not standardised for each participant, however participants were instructed to complete the task on a desktop device only. They had a screen size range of 1280 x 720 to 3440 x 1440 and web browsers included Microsoft Edge, Google Chrome, Safari and 1 Unknown.

Face stimuli were adapted from Doherty et al. (2015). Using their unfiltered photographs of a Caucasian woman looking left, right and centre (539 x 489 pixels), new luminance and geometric images were produced (see Figure 22). Images were decomposed into six non-overlapping frequency bands using Laplcaian Pyramid filter in Matlab. The bottom two subbands were used for the luminance condition. For the geometric condition, the top two subbands were converted to a binary black and white image using Matlab's im2bw function, set at 0.82 level. The face stimuli were presented centre screen with target objects 'dog' or 'cat' scaled to 310 x 430 pixels, positioned either side of the face. The dog and cat stimuli were specifically selected as they were highly similar in terms of shape and colour. This meant participants had to attend to the correct location to detect the correct target, rather than identifying distinct features in their periphery.

Figure 22

Experiment 3 Face and Target Stimuli Examples.



Note. Top Left, a normal face (normal); Top Centre, a blurred face (luminance); Top Right, a line-drawn face (geometric); Bottom Left, dog target; Bottom Right, cat target.

Stimuli were presented in three counterbalanced blocks, one for each psychophysical cue. Each block consisted of 32 trials; half had the face looking in the direction of the target (cued), and half away from it (non-cued). Participants had to respond with a keyboard press which target object they saw. Accuracy was calculated by number of correct responses. Speed was calculated by taking the mean of response time from only trials with correct responses.

Procedure

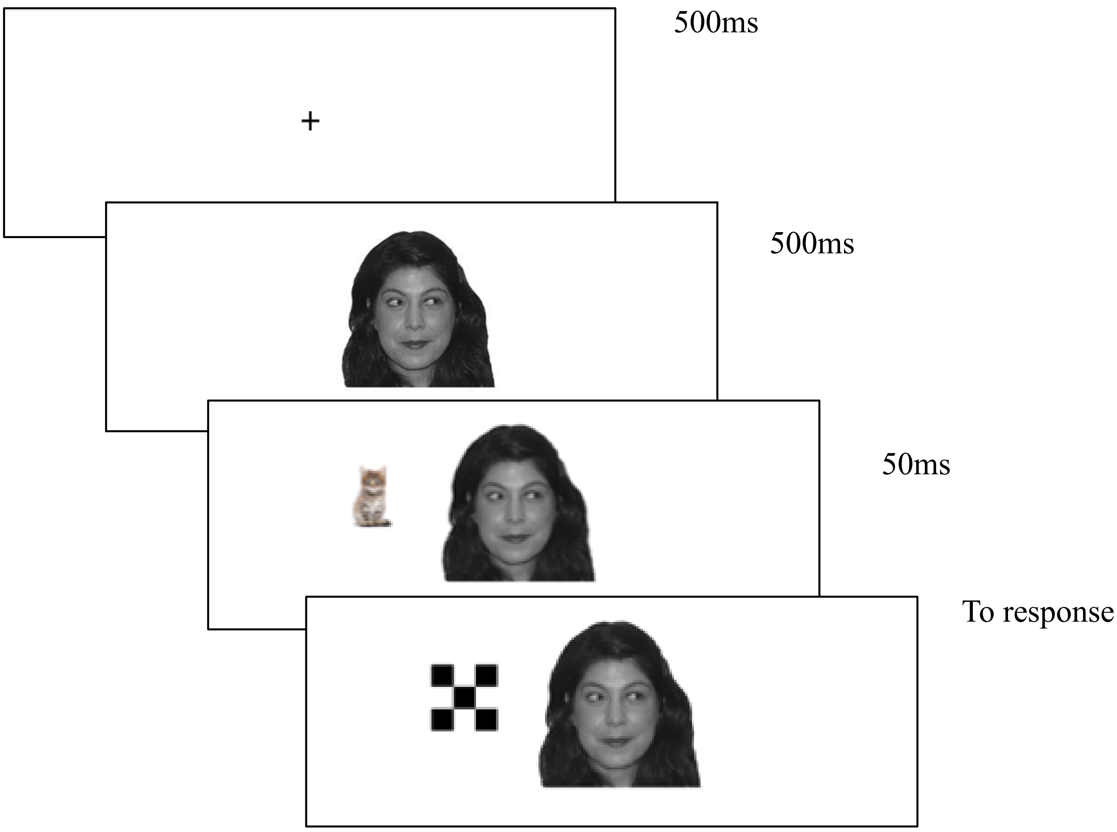
To start, participants were presented an information sheet and consent form, followed by a demographics survey for their age, gender and ethnicity.

Next participants were instructed "On the screen you will see a face. An animal will then appear, and you need to respond as quickly as possible what it is." They then had to press the D key on their keyboard if they saw a 'dog' and the C key for a 'cat'.

Each block started with six practice trials, where the face always looked straight ahead, then continued onto the 32 experimental trials. All trials followed the same sequence (see Figure 23). A fixation cross was shown in the centre of the screen for 500ms, followed by the face stimuli for a further 500ms. The target object was then presented for 50ms (100ms in practice trials) then immediately covered by a pattern mask. The face and pattern mask remained on screen until the participant made a response. Upon completion, participants were debriefed.

Figure 23

Experiment 3 Example Trial Event Sequence



Results

Data Cleaning

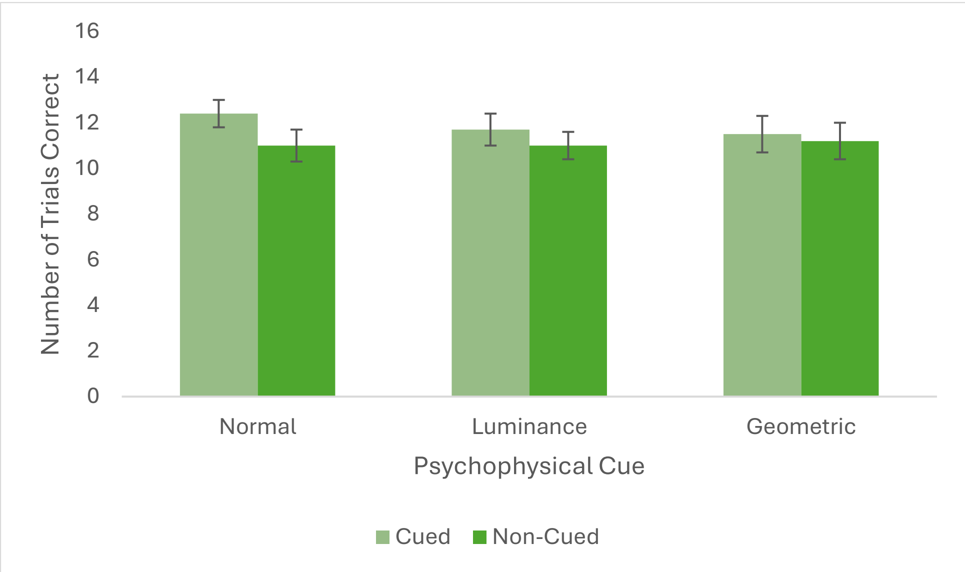
Preliminary analysis of normality indicated that for reaction time, 5 of the 6 conditions had Zskew and Zkurtosis scores that deviated from the criterion of ±2 (see Appendix A). The largest exception was 11.86 and 33.53 respectively, motivating the exploration of outliers within the speed dependent variable. Initially, outlier trials were removed within participant per condition, where a trial was outside ±3 SDs of the participant mean. Between participant outliers were then identified, where a participant's mean was outside ±3 SDs of the sample mean. Across all conditions, four participants were identified as outliers and removed from the dataset.

Preliminary Analysis

Figures 24 and 25 suggest there is little difference in accuracy or speed across psychophysical cues, with cued targets near identical to non-cued targets.

Figure 24

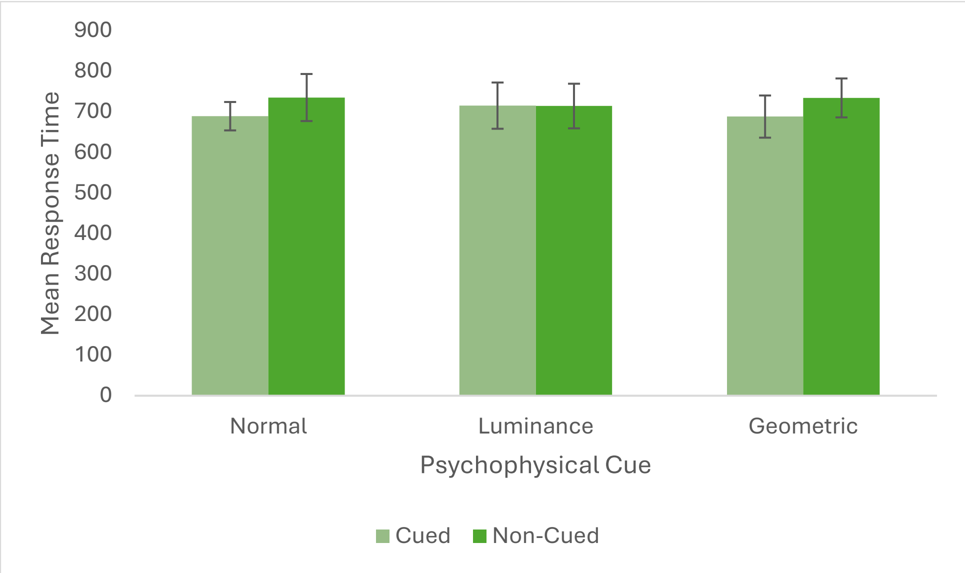
Experiment 3 Number of Correct Responses By Psychophysical Cue and Cueing Type.



Note. Error bars represent 95% CI.

Figure 25

Experiment 3 Mean Response Time By Psychophysical Cue and Cueing Type.



Note. Error bars represent 95% CI.

Accuracy

A 2x3 within-subject ANOVA found a main effect of cueing type F (1, 40) = 12.141, p =.001, η_p^2=.233, where accuracy was higher for cued targets than non-cued targets. No main effect for psychophysical cue was found F (2, 80) = .795, p =.455, η_p^2=.019 (for descriptive statistics, see Table 20). There was no interaction between cueing type and psychophysical cue F (2, 80) = 1.858, p =.163, η_p^2=.044. This demonstrates accuracy was better for cued locations, but did not differ between psychophysical cues. The magnitude of the cueing effect was also not influenced by the psychophysical cue.

Table 20

Experiment 3 Descriptive Statistics Number of Trials Correct

	Normal	**Luminance**	**Geometric**	**Total**
Cued	12.4(*1.97)*	11.7(2.29*)*	11.5(*2.54)*	11.9(*2.30)*
Non-Cued	11(*2.42)*	11(*2.21)*	11.2(*2.33)*	11.1(*2.32)*
Total	11.7(*2.20)*	11.4(*2.25)*	11.3(*2.44)*	11.5(*2.30)*

Response Time

A 2x3 within-subject ANOVA found a main effect of cueing type F (1, 40) = 7.251, p =.010, η_p^2=.153, where response time were faster for cued targets than non-cued targets. No main effect for psychophysical cue was found F (2, 80) = .009, p =.991, η_p^2=.000 (for descriptive statistics, see Table 21). There was no interaction between cueing type and psychophysical cue F (2, 80) = 1.320, p =.273, η_p^2=.032. This demonstrates response times were faster for cued locations, but did not differ between psychophysical cues. The magnitude of the cueing effect was also not influenced by the psychophysical cue.

Table 21

Experiment 3 Descriptive Statistics Mean Response Time

	Normal	**Luminance**	**Geometric**	**Total**
Cued	689(*110)*	715(*178)*	688(*165)*	697(*151)*
Non-Cued	735(*183)*	714(*172)*	734(*154)*	727(*170)*
Total	712(*147)*	714*(175)*	711*(160)*	712(*161)*

Discussion

The aim of Experiment 3 was to investigate the contribution of luminance and geometric cues to adult gaze following behaviour through a replication of Doherty et al.'s (2015) study. The current results replicate those found by Doherty et al. (2015). Participants made more correct target identifications when the target appeared in the gazed at location than the non-gazed at location. This adds strength to their claim that adult gaze following effects can be found with accuracy measures contrary to previous research (Prinzmetal et al., 2008; Stevens et al. 2008). In line with Doherty et al. (2015) the results show participants were equally good at identifying targets whether the face retained luminance cues or geometric cues. Though it did not reach significance, there was a slight benefit of the cues in combination. Thus, adult System 1 likely relies on both luminance and geometric cues.

However, it is possible that the outcome of gaze following in adults incorporates geometric cues, but the processing demands may remain distinct. Previous literature has suggested that the processing of psychophysical cues for gaze following can be dissociated through reaction time measures (Munsters et al. 2016). Thus, Experiment 3 built upon Doherty et al.'s (2015) study by including reaction time measures. As hypothesised, there was a main effect of cueing where participants response times were faster to targets in the gazed at location than the non-gazed at location. Contrary to the hypothesis, reaction times did not differ across the three face stimuli. There was also no interaction of gaze cueing and psychophysical information. These findings suggest luminance and geometric cues are equally important for the processing of gaze following in adults. As speculated by Doherty et al. (2015), this indicates that the use of geometric cues in gaze following is a developmental progression rather than functional.

In summary, Experiment 3 replicated the findings of Doherty et al. (2015) but found no dissociation of psychophysical cues in reaction time as hypothesised.

Experiment 4

Experiment 4 will investigate whether the use of luminance and geometric cues translates to another potential measure of adult System 1, Samson et al.'s (2010) dot perspective task. I will only be measuring self-perspective trials, as explicit interpretation of others visual perspective is not a target behaviour for System 1. This experiment will be presented in two parts.

Experiment 4a

The stimuli for the dot perspective task will be adapted to replace the avatar with the face stimuli from Experiment 3. Importantly, the only directional cue obtained from these stimuli are the eyes, unlike the original task which had the full body. Whilst there is evidence to suggest congruent head and eyes directional cues are just as effective as full body (Gardner et al. 2018), Part a will primarily confirm whether the self-consistency interference is just as effective for eyes only cues.

A further consideration is that the eyes in these stimuli are not directed 90 degrees towards the wall like a full body avatar would be. This could give the impression that the face is not looking at the discs on the wall and disrupt and consistency effects. Thus, the face stimuli will be placed behind a window rather than inside the room. This will give the illusion of depth and therefore the impression of looking at the discs on the wall. Additionally, other directional cues such as the body and head will be directed at the same visual angle as the eyes. Thus, Part a will also demonstrate whether the self-consistency effect can be found when the avatar is not in the room or directed at 90 degrees.

It is hypothesised that there will be a main effect of consistency, where participants will be more accurate and faster to respond when the avatar and their viewpoint is consistent than when it is inconsistent. Additionally, there will be no main effect of directional cue, where accuracy and response times will not differ between body, head and eyes, or eyes only cue. Finally, it is hypothesised there will be no interaction such that the magnitude of the self-consistency effect for accuracy or reaction time will not differ between directional cues.

Method

Participants

Participants were 45 participants (39 Female, 5 Male, 1 Other) with a mean age of 19.9 years (*SD* = 2.59, *Range* = 18-36). Participants were recruited through SONA and received 1 SONA credit for their participation.

Inclusion criterion was informed consent prior to testing. The pre-registration (110531) specified no exclusion criteria. However, due to multiple conditions having skewed data (see Appendix A) 7 participants were identified as outliers and removed from the sample. Stopping criteria was when the target sample size of 45 was reached, as outlined in the Statistical Analysis Plan. Participants could withdraw their consent by closing the browser or responding 'I wish to withdraw when asked at the debrief. However, no participants chose to do so. The new sample included 38 participants with a mean age of 20.0 years (*SD* = 2.79, *Range* = 18-36) (32 Female, 5 Male, 1 Other).

The School of Psychology Ethics Committee granted this study ethical approval (ETH2122-2206).

Statistical Power and Analytical Plan

For both Experiment 4a and 4b, a 2x3 within-subject ANOVA was planned for analysis. Using the G-Power 3.1.9.2 statistical tool (Faul et al., 2007), a sample size of 36 will achieve a power of .80 with a two-tailed alpha set to .05, medium effect size f = .25. The indicated sample size also exceeds that of Samson et al.'s (2010; N =16) original study. I therefore aimed to recruit 45 participants per experiment to account for any participant withdrawal.

The data is not normally distributed (see Appendix A), but as per the pre-registration, parametric tests are used for analysis (Field, 2003; Rasch & Guiard, 2004).

Design

A 2x3 multifactorial design was used to investigate two within-subject independent variables: consistency and directional cue. Consistency was categorised as two levels: consistent and inconsistent. Directional cue was three level: body, head, and eyes. The dependent variables were accuracy of response (0-12 per condition), and speed of response (ms). Each directional cue was presented in a single block, with blocks fully counterbalanced across participants.

Materials

The study was conducted online using Gorilla. Participant set up was not standardised, apart from limiting study completion to desktop devices. However, information regarding screen size and browser was collected (*Range:* 1126 x 701 to 1920 x 1080; *Browser*: Google Chrome, Safari & Microsoft Edge).

Stimuli presented were of a room, showing the left, back and right wall. On one, two or no walls were red discs (138 x 64 pixels) (see Figure 26). In the centre of the back wall was a window, of which an agent was looking from the outside in, with only their upper body visible (753 x 753 pixels). The agent would have either their entire body, head and eyes, or eyes only directed at a 25 degree angle to the left or right wall to look at the red discs (see Figure 27). This is because when a model is asked to keep their head straight but move their eyes to the left or right by 25 degrees, participants can detect changes in head direction (Doherty & Anderson, 2001). So, I wanted to ensure there was a detectable directional cue, without confounding the conditions. I placed the agent behind the window to create the illusion of depth, and of the agent looking directly at the discs.

Figure 26

Experiment 4a Dot Perspective Task Target Presentation Combinations

	0	**1**	**2**	**3**
0				
1				
2				
3				

Note. Columns represent self-perspective and rows represent other-perspective.

Figure 27

Experiment 4a Directional Cue Examples.

Note. Left Body; Middle, Head; Right, Eyes.

Trials were presented over three counterbalanced blocks, one for each directional cue. Each block consisted of 48 trials. Twenty Four trials were 'consistent', where the number of targets visible to the participant and the agent were the same. The other 24 were 'inconsistent', where the number differed. Participants had to respond with a keyboard press if the number of discs they could see matched a target number at the start of the trial. Half the trials required a 'no' response and half a 'yes' response. The 'no' trials were included as fillers to prevent participants automatically answering yes to every question. The 'no' trials were excluded from analysis because 'no' trials corresponded to neither perspective. Participants also saw four 'filler' trials per block, where no dots were presented. Accuracy was calculated by number of correct responses. Speed was calculated by taking the mean of response time from only trials with correct responses.

Procedure

Following Samson et al.'s (2010) Experiment 3 procedure, participants had to judge how many objects they could see. They were given the following instructions:

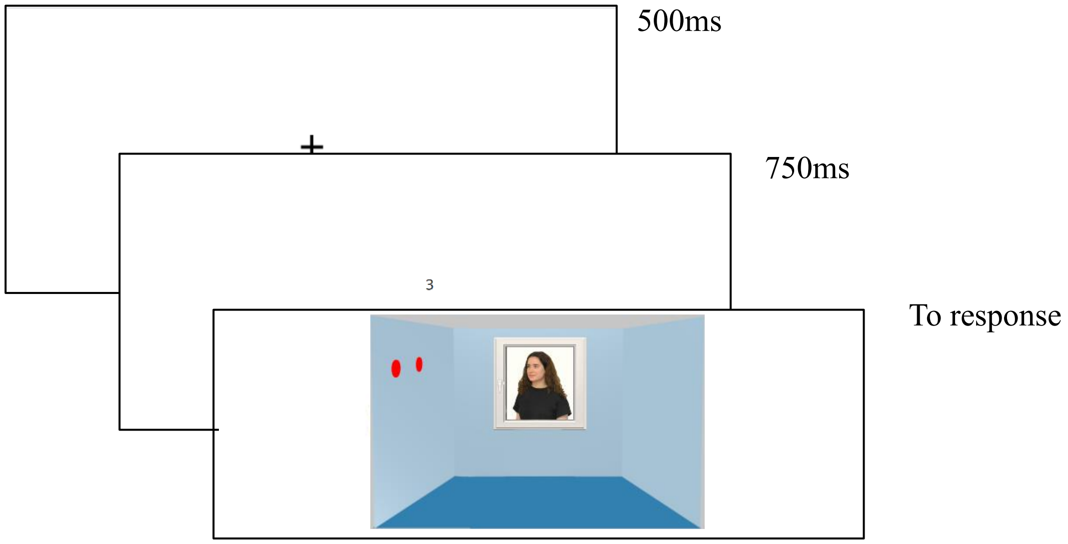
"On the screen you will see a number from 0 to 3. This will be followed by a picture of a person in a room with dots to their left and right. In this task you have to judge if you can see the same number of dots as the number specified. Press the Y key for yes and the N key for no.

Respond as quickly as possible. (Practice Block only: You will be given feedback on your answers.)"

Each block started with thirteen practice trials and received feedback (correct/incorrect), then continued onto the 48 experimental trials. All trials followed the same sequence (see Figure 28). After a fixation cross was displayed for 500ms, a number from 0 to 3 was then displayed for 750ms, followed by the agent and room until a response was given. For an example trial event sequence, see Figure. Upon completion, participants were debriefed.

Figure 28

Experiment 4a Example Trial Event Sequence



Results

Data Cleaning

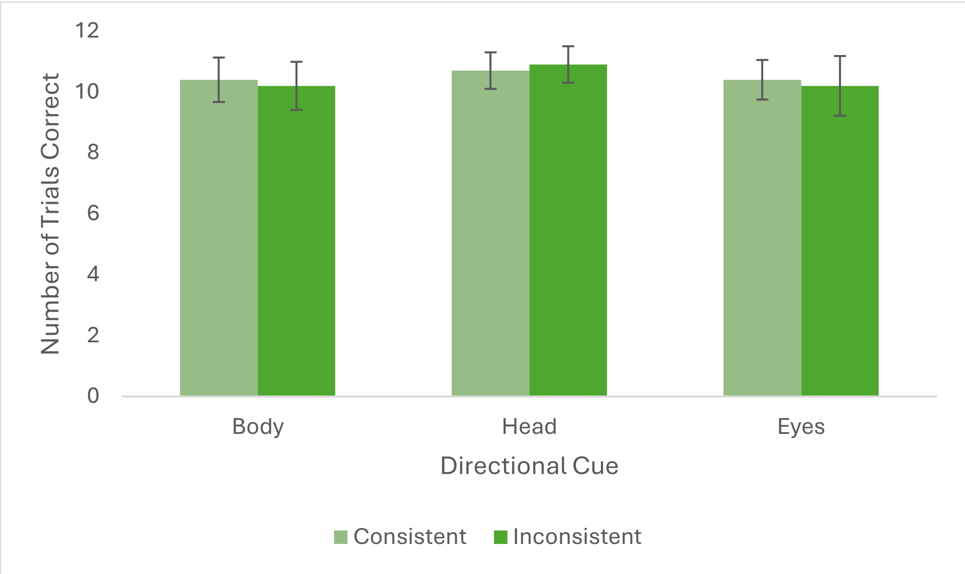
Preliminary analysis of normality indicated that all conditions had Zskew and Zkurtosis scores that deviated from the criterion of ±2 (see Appendix A). The largest exception was for response time, motivating the exploration of outliers within the speed dependent variable. Data cleaning methods matched that of Experiment 3. Across all conditions, seven participants were identified as outliers and removed from the dataset.

Preliminary Analysis

Figures 29 and 30 suggest there is little difference in accuracy or speed across directional cues. Participants responses to consistent trials is slightly faster than to inconsistent trials, but no clear differences in accuracy.

Figure 29

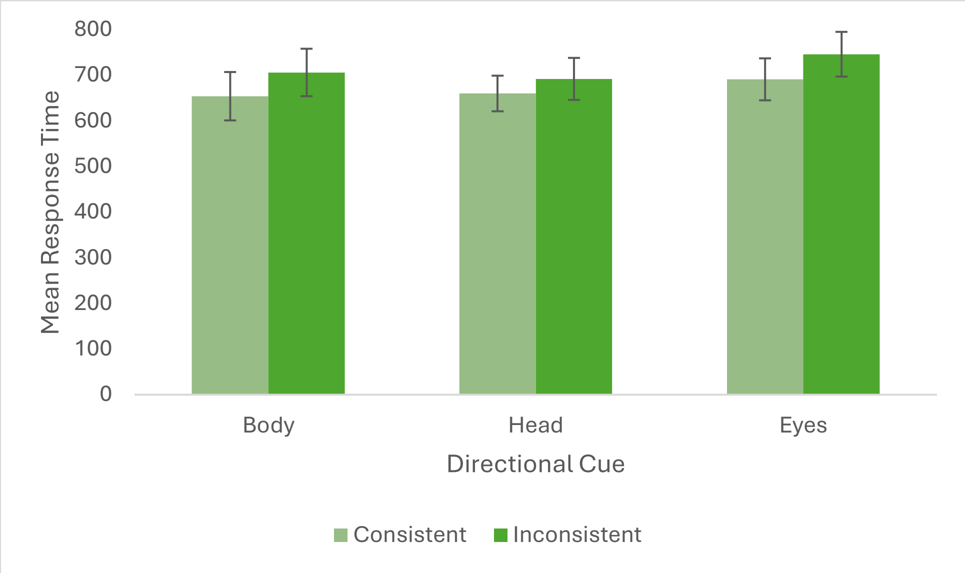
Experiment 4a Number of Correct Responses By Consistency and Directional Cue.



Note. Error bars represent 95% CI.

Figure 30

Experiment 4a Mean Response Time By Consistency and Directional Cue.



Note. Error bars represent 95% CI.

Accuracy

A 2x3 within-subject ANOVA found no main effect of consistency F (1, 37) = 0.305, p =.584, η_p^2=.008, or directional cue F (2, 74) = 2.088, p =.131, η_p^2=.053 (for descriptive statistics, see Table 22). No interaction between consistency or directional cue was found F (2, 74) = 1.512, p =.227, η_p^2=.039. This demonstrates accuracy did not differ when the agent and participants view was the same or different, or when directional cue became less salient.

Table 22

Experiment 4a Descriptive Statistics Number of Trials Correct

	Body	**Head**	**Eyes**	**Total**
Consistent	10.4(2.27*)*	10.7(1.93*)*	10.4(*2.13)*	10.5(*2.11)*
Inconsistent	10.2(*2.43)*	10.9(1.92*)*	10.2(3.01*)*	10.4(*2.45)*
Total	10.3(*2.35)*	10.8(1.93*)*	10.3(*2.57)*	11.5(*2.28)*

Response Time

A 2x3 within-subject ANOVA found a main effect of consistency F (1, 37) = 31.792, p < .001, η_p^2=.462, where response times were faster for consistent trials than inconsistent trials. No main effect for directional cue was found F (2, 74) = 2.487, p =.090, η_p^2=.063 (for descriptive statistics, see Table 23). There was no interaction between cueing type and psychophysical cue F (2, 74) = 0.775, p =.464, η_p^2=.021. This demonstrates participants responded faster when their view matched the agent than when it differed. The magnitude of this effect was consistent across all directional cues, with all cues eliciting equivalent response times.

Table 23

Experiment 4a Descriptive Statistics Mean Response Time

	Body	**Head**	**Eyes**	**Total**
Consistent	654(*160)*	660(*118)*	691(140*)*	668(*139)*
Inconsistent	706(*160)*	692(*140)*	746(*149)*	715(*150)*
Total	680(*160)*	*676(129)*	719*(145)*	712(*145)*

Discussion

The aim of Experiment 4a was to test the feasibility of using stimuli that contained directional cues from the eyes only within the dot perspective task. Previous research had found the self-consistency effect to be equivalent for avatars with congruent head and eye directional cues as full body cues (Gardner et al., 2018). However, the eyes are much less salient cue than the head, so the current study intended to confirm whether the self-consistency effect persists when they are the only cue. The results show no effect of consistency or directional cue on the accuracy of response. However, as hypothesised participants were faster to respond when their viewpoint was consistent with the avatar than when it was inconsistent. Participant response time did not differ when the directional cue was the full body, head and eyes, or eyes only, and this did not interact with the self-consistency effect.

Together this suggests that despite being less salient than other directional cues, the self-consistency effect can be found with eyes only, and it is just as strong as full body cues. Additionally, this also demonstrates the other stimuli adaptions required for eyes only avatars did not impact the findings. Thus, any differences in the self-consistency effect when using the face stimuli from Experiment 3 can be attributed to the dissociation of psychophysical cues and not other stimuli changes.

Experiment 4b

Experiment 4b will investigate if there is dissociation of psychophysical cues within the dot perspective task. A strength of using the dot perspective task is that it controls for the potential that System 1 operates at a faster timescale than System 2. In the gaze cueing task, the target appears at 500ms. This is potentially long enough for conscious processing and therefore recruitment of the more calculated geometric cues. Thus, luminance and geometric processing speed may not dissociate because both have been processed at the point the target appears. In the dot perspective task however, the target appears at 0ms. If luminance cues are more efficient than geometric cues, then the target is likely to be located before geometric cues have begun to be processed. In this sense, the dot perspective task should allow the two cues to be processed on their own timescale, better demonstrating any differences in speed.

It is hypothesised that there will be a main effect of consistency, where participants will be more accurate and faster to respond when the avatar and their viewpoint is consistent than when it is inconsistent. It is also hypothesised that there will be a main effect of psychophysical cue where accuracy and reaction time will be higher and faster for luminance cues than geometric. Additionally, there will not be an interaction between cueing and psychophysical cue for accuracy or reaction time.

Method

Participants

Participants were 102 participants (87 Female, 12 Male, 2 Other, 1 Prefer not to say) with a mean age of 19.2 years (*SD* = 1.23, *Range* = 18-26). Participants were recruited through SONA and received 1 SONA credit for their participation.

Inclusion criterion was informed consent prior to testing. The pre-registration (110531) specified no exclusion criteria. However, due to multiple conditions having skewed data (see results for details) 10 participants were identified as outliers and removed from the sample. Pre-registered stopping criteria was when the target sample size of 45 was reached, however circumstances allowed for more data to be collected. Data collection ceased after the study had been made available for 10 days. Participants could withdraw their consent by closing the browser or responding ‘I wish to withdraw when asked at the debrief. Two participants chose to do withdraw. One participant was excluded for misunderstanding the task requirements. The new sample included 91 participants with a mean age of 19.2 years (*SD* = 1.25, *Range* = 18-26) (79 Female, 9 Male, 2 Other, 1 Prefer not to say).

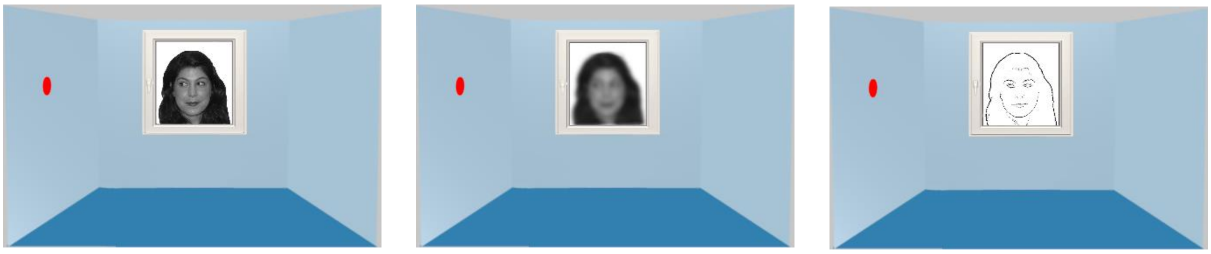
The School of Psychology Ethics Committee granted this study ethical approval (ETH2122-2206).

Design

The same design as Experiment 4a was used. However, the directional cue independent variable was replaced with psychophysical cue, consisting of three levels: normal, luminance and geometric.

Figure 31

Experiment 4b Dot Perspective Task Psychophysical Examples.



Note. Left, Normal; Middle, Luminance; Right, Geometric.

Materials and Procedure

The study was conducted online using Gorilla. Participant set up was not standardised, apart from limiting study completion to desktop devices. However, information regarding screen size and browser was collected (*Range:* 1080 x 810 to 1920 x 1080; *Browser*: Google Chrome, Safari, Opera & Microsoft Edge).

Stimuli, test blocks and trial presentation were identical to Experiment 4a with the following exception: the face stimuli from Experiment 3 were used in place of the agent.

Results

Data Cleaning

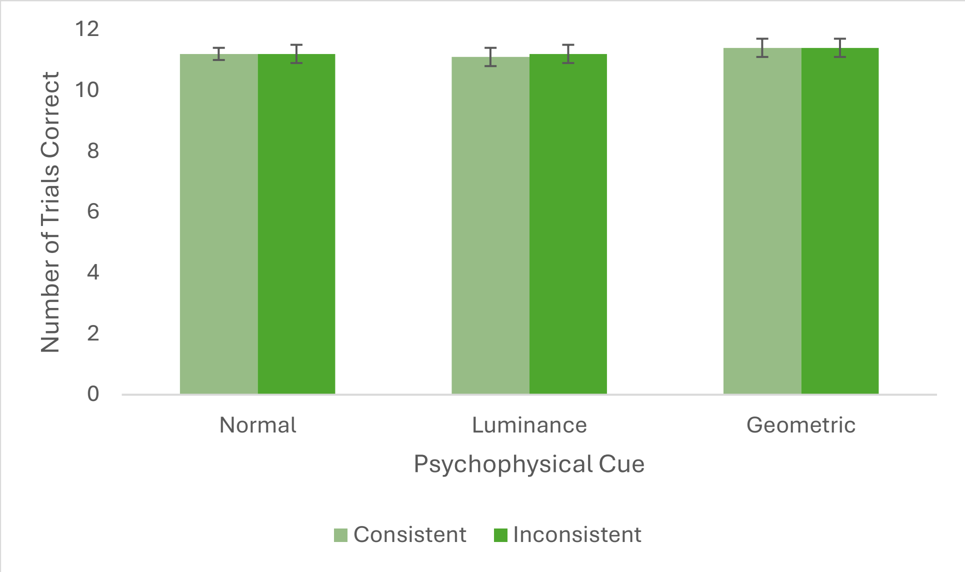
Preliminary analysis of normality indicated that all conditions had Zskew and Zkurtosis scores that deviated from the criterion of ±2 (see Appendix A). The largest exception was for response time, motivating the exploration of outliers within the speed dependent variable. Data cleaning methods matched that of Experiment 3. Across all conditions, ten participants were identified as outliers and removed from the dataset.

Preliminary Analysis

Figures 32 and 33 suggest accuracy did not differ between consistent and inconsistent trials and was greater for geometric faces than normal or luminance. Response times were faster to consistent trials than inconsistent trials, but did not differ across psychophysical cues.

Figure 32

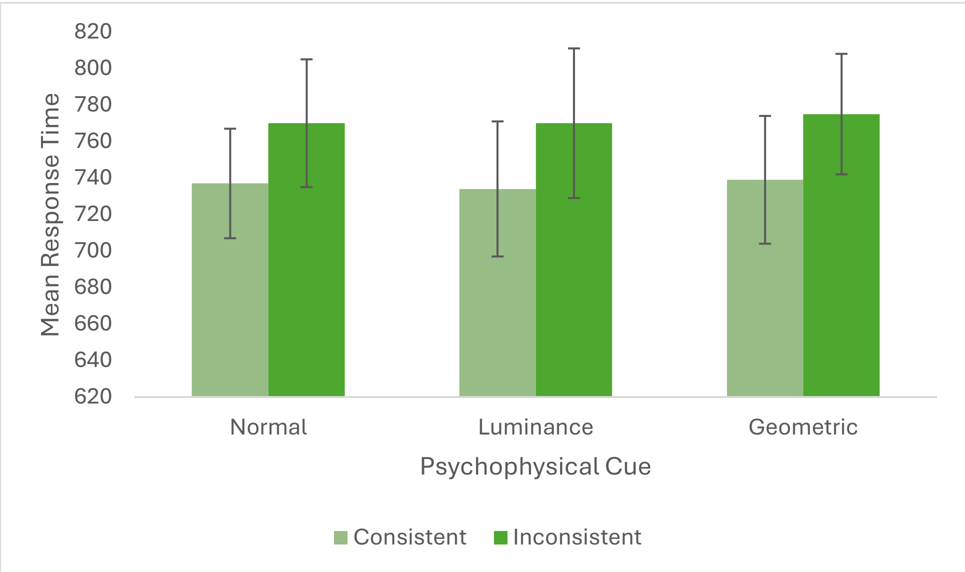
Experiment 4b Number of Correct Responses By Consistency and Psychophysical Cue.



Note. Error bars represent 95% CI.

Figure 33

Experiment 4b Mean Response Time By Consistency and Psychophysical Cue.



Note. Error bars represent 95% CI.

Accuracy

A 2x3 within-subject ANOVA found no main effect of consistency F (1, 90) = 0.143, p =.706, η_p^2=.002, or psychophysical cue F (2, 180) = 2.102, p =.125, η_p^2=.023 (for descriptive statistics, see Table 24). No interaction between consistency or psychophysical cue was found F (2, 180) = 0.451, p =.638, η_p^2=.005. This demonstrates accuracy did not differ when the avatar and participant's view was the same or different, or with psychophysical cue.

Table 24

Experiment 4b Descriptive Statistics Number of Trials Correct

	Normal	**Luminance**	**Geometric**	**Total**
Consistent	11.2(1.18)	11.1(1.41)	11.4(1.22)	11.2(1.27)
Inconsistent	11.2(1.54)	11.2(1.36)	11.4(1.15)	11.3(1.35)
Total	11.2(1.36)	11.2(1.39)	11.4(1.19)	11.3(1.31)

Response Time

A 2x3 within-subject ANOVA found a main effect of consistency F (1, 90) = 31.636, p < .001, η_p^2=.260, where response times were faster for consistent trials than inconsistent trials. No main effect for psychophysical cue was found F (2, 180) = 0.082, p =.921, η_p^2=.001 (for descriptive statistics, see Table 25). There was no interaction between cueing type and psychophysical cue F (2, 180) = 0.057, p =.945, η_p^2=.001. This demonstrates participants responded faster when their view matched the avatar than when it differed. The magnitude of this effect was consistent across all psychophysical cues, with all cues eliciting equivalent response times.

Table 25

Experiment 4b Descriptive Statistics Mean Response Time

	Normal	**Luminance**	**Geometric**	**Total**
Consistent	737(*148)*	734(*177)*	739(*167)*	737(*164)*
Inconsistent	770(*164)*	770(*197)*	775(*161)*	772(*174)*
Total	753(*156)*	*752(187)*	*757(164)*	755(*169)*

Discussion

The aim of Experiment 4b was to investigate whether the two-systems predictions for psychophysical cues translate to other potential measures of System 1 behaviour. The results partially replicated Samson et al.'s (2010) findings. Participants were faster to respond when their viewpoint was consistent with the avatar than when it was inconsistent. However, unlike Samson who found accuracy was better when the avatar's viewpoint was the same as the participant, the current study found no difference.

Contrary to the hypothesis, accuracy and response time did not differ across psychophysical cues and no interaction was found with consistency. The lack of difference for reaction time is particularly notable as targets in the dot perspective task are presented at 0ms, which would have allowed a greater distinction between the timescale that the two cues are processed at. This finding further supports the notion that adult gaze following incorporates both luminance and geometric cues. However, it does not suggest that either cue is more important for processing cued gaze. This finding is also consistent with Experiment 3, demonstrating the use of psychophysical cues for automatically attention orienting is consistent across multiple tasks.

In summary, Experiment 4b replicated the findings of Samson et al. (2010) but found no dissociation of psychophysical cues as hypothesised. This pattern of findings aligns with that of gaze cueing tasks, further supporting the dot perspective task as a suitable measure of adult System 1.

General Discussion

It has been proposed there are two systems for processing gaze that can be distinguished by the information they use from the eyes. The aim of the current study is to test the plausibility of this claim by investigating what information adult System 1 uses. Specifically, System 1 is claimed to be luminance-based serving the function to locate objects of attention through automatic attention orienting. The main findings are as follows. Experiment 3 shows that either luminance or geometric information results in gaze cueing. However, the speed at which we respond to gaze cues does not differ between luminance and geometric information. As such, gaze following uses either luminance or geometric cues in both outcome and processing. Experiment 4 demonstrated that the self-consistency effect can be elicited through eye direction cues in isolation. However, as in Experiment 3, the speed of processing does not differ if the faces retain luminance cues compared to geometric cues.

Doherty et al. (2009; 2015), outlined the two-system theory for gaze understanding. System 1 serves the function to *direct* attention, by following gaze to objects of attention. This is available early in development and automatic. System 2 serves the function to *understand* others' attention, by allowing thoughts about the attentional relationship between observer and object. It has protracted development and computationally demanding. One hypothesis is that the two systems remain distinct throughout the lifespan, with System 1 being luminance-based and System 2 incorporating geometrical information. However, the findings of the current chapter suggest System 1 may not remain distinct in adulthood, at least in terms of the psychophysical cues it uses.

A novel consideration of the current study was whether adult System 1 may remain distinct in the contribution of the psychophysical cues how it processes cued gaze. Specifically, evidence has suggested that faces that remain luminance cues cue gaze faster than those that retain geometric cues (Munsters et al., 2016). This more efficient processing of luminance cues has also been seen in attention related ERPs for both adults and children (de Jong et al., 2008; Mares et al., 2018). In contrast, the results of the current chapter find not speed differences between the two cues. A potential consideration for contrasting results is that previous research had the illusion of eye movement, with eyes facing forward initially then shifting to the left or right. It has been shown that motion processing is primarily determined by luminance information, and less reliant on geometric (Edwards & Badcock, 1996; Geisler et al., 2007). Thus, it is possible in the above studies the faces that retain luminance cues may have a processing advantage over those with geometric cues. In the current tasks, there was no illusion of motion, thus neither cue had an advantage over another. Therefore, the use of luminance and geometric cues for purely processing gaze direction may in fact be equivalent. As such, adult System 1 may use both cues to achieve the same end goal and are processed with equal importance.

These findings have two potential implications for the two-system theory. Either gaze processing is best explained by a single system, or that the systems are only distinct in childhood, and through experience System 1 becomes more flexible in its processing. Doherty et al. (2015) suggest this contradiction could be untangled by assessing how System 1 changes over the lifespan. For example, the gradual development of fine-grained gaze judgements in System 2 between 4- to 14-years-old (Doherty et al., 2009; Vida & Maurer, 2012) is pivotal to the two-systems claim. Thus, it would be appropriate to consider whether the use of geometric cues in System 1 is as gradual process or not. Doherty et al. (2015) explain that a single system account would predict the use of geometric cues to be developmental and not functional. As such, once geometric cues are available, they will see a gradual increase in their use for *both* gaze following and gaze judgements. On the contrary, a two-system account would suggest the use of geometric cues would gradually increase for gaze judgements alone. Once gaze judgement reaches an adult-like ability, geometric cues can then be applied to System 1. With the findings from the current study replicating that of Doherty et al. (2015), this proposal warrants investigation.

In conclusion, the current study aimed to provide support for a two-systems theory of gaze understanding by investigating whether System 1 remains distinct into adulthood. The present findings show adult attention orienting relies on both luminance and geometric information from the eyes. This finding is robust across dependent measures of accuracy, response time or task type. As such, there is no clear evidence to conclude that adult System 1 remains distinct in adulthood. Whilst the results of the current study are more plausibly explained by a single system account, future studies should consider the 4- to 14-year-old period to investigate how geometric cues become incorporated in adult gaze following.

Chapter 5

General Discussion

The aim of the current thesis was to explore the development of gaze understanding, if young children use a simpler concept of attention prior to this development, and whether the two-system theory can be supported. Traditionally, it was thought that once infants could follow the gaze of others' they are able to think about visual attention (Butler et al., 2000; Corkum & Moore, 1998). This milestone has been suggested to be a precursor to understanding belief (Goméz, 1996). However, many 3-year-olds are unable to judge where someone is looking from their eye-direction alone (Doherty & Anderson 1999; Doherty et al., 2009). This suggests that whilst young children can direct their attention toward the objects of attention, they do not understand the attentional relation between the observer and object. Evidence from comparative (e.g. Anderson & Mitchell, 1999; Povinelli & Eddy, 1996a; 1996b) and infant (e.g. Corkum & Moore, 1995; Brooks & Meltzoff, 2002; Doherty et al., 2009) research suggests that gaze following, and gaze judgement are two distinct abilities and not directly related developmentally. This is the basis for the claim that humans have two systems to process gaze (Doherty, 2006). It is thought that System 2 development is motivated by children's increasing interest in the mind and is an integral part of theory of mind development. Prior to this, children use a simpler concept of attention that does not require any mental state concepts.

Do young children predict behaviour using a concept of engagement?

Children need to think about and respond to others' behaviour on the basis of what others have attended to. Without an understanding of how eye-direction maps onto the word see, children may instead rely on a simple behavioural heuristic that adults are more likely to engage with things in front of them and are not occluded. This is the foundation of engagement. This concept has been used to explain the signature limits of children's understanding of occlusion in Flavell et al.'s (1978) classic hiding task, suggesting that children interpret the word *see* as *engaged with*. Experiment 1 primarily investigated this claim.

I first replicated the finding that preschool children can hide an object by placing it behind a screen but have difficulty moving the screen in front of it (Flavell et al., 1978; McGuigan & Doherty, 2002). The engagement account posits that the Move Object task is easy because children understand that placing the object behind the screen can prevent the experimenter becoming engaged with it. However, in the Move Screen task the experimenter is already engaged with the object, so children do not know how interposing the screen can disrupt that.

The most straightforward test of this was to create a scenario where the experimenter was not engaged with the object. To do so, I modified the Move Screen task so that the experimenter was blindfolded at the start of the experiment, after which a second experimenter then placed the object on the table. This meant they had not experienced the object in its location and had thus not established an engagement relation with it. As predicted, performance on this task improved by around 72% compared to the classic task. This would suggest that children understand seeing in terms of spatial and behavioural relations, not resembling an adult-like concept of attention.

A second test of this claim was to consider the conditions under which engagement relations can be disrupted. Specifically, whether engagement relations persist until the object is moved from its current location (Doherty, 2011). When testing an equivalent concept of awareness relations, Horschler et al. (2019) found that simply removing the object and replacing it immediately back outside of the experimenter's perception resulted in monkeys behaving as though no relation existed. Thus, it was predicted that engagement would also be sensitive to this type of manipulation. To test this, I modified the Move Screen task to mirror Horschler et al.'s (2019) study. At the start of the trial, the experimenter had visual access to the object so was engaged with it. The experimenter then donned a blindfold, after which a second experimenter removed and immediately replaced the object. It was predicted that children would treat the experimenter as disengaged from the object, making it easier to correctly place the screen. The results of Experiment 1 show this type of manipulation confers an improvement in performance of around 21%, but this fell short of conventional significance.

It could be interpreted that spatial manipulations are important for engagement relations to persist, but that these need to be more substantial than simply removing and replacing it. Returning the object in its original location may cancel out the disengagement event, meaning children continue to represent the experimenter as engaged with the object. As such, I proposed a revision to the definition of engagement to be "a relationship between a person and an object, which persists until it is moved to a *new location*". Future research should empirically test this revision with a version of the Move Screen task where the object is placed in a new location before being returned to its original location. According to the new definition, this additional step should be sufficient to deteriorate traces of engagement relations, making it easier for children to succeed. The new location could be varied in terms of distance, surface and purpose to consider if this effect is graded or on-off.

Both modifications of the Move Screen task were compared with a control task, to check that the effects were not simply due to the addition of a blindfold. In this condition, the experimenter started the trial engaged with the object. They then donned the blindfold, but no manipulations occurred. This meant that the experimenter remained engaged with the object, and thus performance was predicted to be equivalent to that of the classic task. On the contrary, the control task did result in children's performance improving, but not as much as that of the experimental manipulations. It is possible that the blindfold encourages children to go beyond monitoring broad perceptual cues to pay greater attention to the eye region as a more specific cue for perceptual experience. Whilst this can give children a better insight into what the experimenter can see, it is not sufficient to overcome the conceptual distinction between seeing and engagement required in the Move Screen task.

It is important to note that use of engagement may not be a complete explanation of the Move Screen task difficulty. Neither the control nor experimental conditions in Experiment 1 improves to the level of the Move Object task, meaning the task may require understanding beyond that of line of sight. As previously discussed, the Move Object task involves the spatial relation *behind* whereas the Move Screen task involves *between*. Between is a plausibly more challenging relation than behind as it refers to two points of reference (Simms & Gentner, 2019) and appears to be available for use around 4-years-old (Spetch & Parent, 2006; Simms & Gentner, 2008). This is around the same time that children consistently pass the Move Screen task. Thus, children may have the ability to block line of sight, but not in situations which require spatial concepts they haven't properly developed.

The findings of Experiment 2 indicate inhibitory control to be another contributing factor of the Move Screen task. It was found that inhibitory control was negatively associated with Move Screen performance at Time 2, and positively associated at Time 4. Despite this inconsistency, further exploration of the data showed that at Time 4 increased inhibitory control was associated with a reduction in egocentric responses. This could be interpreted as inhibitory control conferring success on the Move Screen task, by allowing children to overcome egocentric biases. Young children using engagement do not know how placing the screen disrupts the engagement relation, but they do know how to place the screen to they themselves cannot see the object. This is plausibly a salient, default response that without the necessary inhibitory control children have difficulty overcoming even when they have a developed concept of line of sight. The current findings only allow for this to be speculation, and future research should consider the role of inhibitory control may have.

Together, children's improved success on the Move Screen task when separately manipulating seeing and engagement provides support for engagement being a potential explanation for the signature limits of their understanding of occlusion. However, this understanding may be additionally limited by spatial reasoning and inhibitory control which should be explored in future research.

Is engagement a way for young children to fake gaze and mental state understanding?

Without the ability to conceive of attention in terms of mental states or seeing, children need an alternative measure. The evidence from the thesis thus far would suggest that engagement is a good solution. If this is the case, then a sophisticated understanding of gaze and others' minds would mark the transition to an adult-like concept of attention, "what someone is attending to is determined by the direction of his or her eyes" (McGuigan & Doherty, 2002, p. 423). To test this claim, Experiments 1 and 2 investigated the association and cross-lagged effects between Move Screen performance, and gaze judgement and theory of mind ability.

Experiment 1 found a strong, significant relation between overall gaze judgement ability and the Move Screen task. However, in Experiment 2 there was only an association at Time 4 specifically with the Looking At You task which does not persist beyond common associations with inhibitory control and language (see Experiment 2 for a discussion of what this may mean). This pattern of results is interesting for two reasons. Firstly, it makes a novel contribution to the literature as to my knowledge no other study compares the performance of the Looking At You and the Move Screen task. Secondly, the lack of association with the Looking Where task is incongruent with McGuigan and Doherty (2002) findings. Whilst their findings may not be replicable, there is certainly some relevance of gaze judgement for the Move Screen task. Regardless of what the gaze is directed at, both gaze judgement tasks plausibly involve projecting a line of sight from the eyes to an object. This understanding is fundamental to the Move Screen task as it requires children to intersect that line of sight, which is most evident between late 3- and 4-years. It could be interpreted that a sophisticated understanding of gaze allows children map eye direction onto the word *see*, marking the transition to an adult-like conception of attention.

When comparing Move Screen and theory of mind ability, Experiment 1 found no associations beyond age related developments. However, Experiment 2 both Knowledge Access and False Belief associated with Move Screen performance at Time 4. As above, when controlling for inhibitory control only associations with False Belief persisted. This would suggest that the two abilities do not share a common understanding of mental representation, but children are going through some conceptual change around 4-years that impacts both abilities.

Of greater interest, Experiment 2 found that earlier Knowledge Access performance predicted later Move Screen success across all timepoints. The Move Screen task requires children to make distinctions between knowledge and perception. Engagement does not make this distinction, whereas an adult-like conception of knowledge does. The current findings could be interpreted that representing seeing as a cause for knowing provides children the ability to differentiate the two, lending to their later success on the Move Screen task.

In summary, children who do not have access to System 2 need an alternative way to respond to others' attention. Here it is suggested they use a concept of engagement which relies on a set of behavioural heuristics to determine another's general involvement with objects. The findings of this thesis would suggest that young children use engagement. As they develop an understanding of knowledge formation, they transition to using mentalistic-based strategies. This transition likely requires flexible use of line of sight concepts, to be able to map eye direction onto the word see for an adult-like conception of attention.

What does the developmental trajectory of gaze understanding look like?

One aim of the current thesis was to investigate the development of children's gaze understanding. Doherty and colleagues (1999; 2009) found that preschool children have difficulty making judgements of eye-direction. This is a surprising finding given that it is well established that children can follow gaze from around 18-months old (Corkum & Moore, 1995; Brooks & Meltzoff, 2002; Moore & Corkum, 1998; Caron et al., 2002; Tang et al., 2024). Thus, using a cross-sectional design in Experiment 1 and a longitudinal design in Experiment 2 I used Doherty and Anderson's (1999) Looking Where and Looking At You tasks to measure this development.

In both experiments, performance on the gaze judgement tasks replicated those of Doherty and colleagues (1999; 2009). Performance on the point where task was near ceiling at all ages, demonstrating that even young children could respond to the task demands. Two- and 3-year-olds had difficulty making judgements of eye-direction and saw rapid improvement by the time they turn 4-years. Experiment 2 demonstrates that this development has no meaningful between child variance, such children have similar rates of development irrespective of their starting point. Thus, it appears that children's protracted development of gaze understanding is a robust finding and is likely to be stable across children.

However, the studies of the current thesis do not capture the whole developmental trajectory. Namely, children are passing the Looking Where task around 41% to 49% and the Looking At You task between 29% and 59% of the time at 2-years-old, both improving to around 75% at 4-years-old. This means that at least 50% of their development is unaccounted for. In terms of the two-systems theory, this lends itself to contradictory support for the basic claim that gaze judgement is a new skill arising around 3-years-old. The basis of this claim is that new skills take practice, and thus initial accuracy would be poor and improve thereafter. In support of this, the findings of Experiments 1 and 2 show gaze judgement ability improves incrementally over time. This is in line with previous research that found fine-grained gaze judgements gradually improve between 3- and 14-years-old (Doherty et al., 2009; Vida & Maurer, 2012a; 2012b).

This gradual improvement in gaze judgement performance may reflect how children learn to represent gaze as a straight line. In particular, Experiment 1 finds that children failures to make correct gaze judgements are more often than not in the correct direction. Children are therefore extracting some directional information from the eyes just not with great accuracy. As such, young children may represent sight in a cone like manner similar to car headlights shining diffused light over a large space. This is plausible given that engagement monitors broad cues, so it is likely that children view eye direction in a broad sense too. Through experience, children realize their cone must get narrower to be successful at making gaze judgements. Eventually, their representation of sight becomes analogous to that of a laser beam, which would allow them to recruit line of sight concepts. Future research should therefore consider how children spatially represent gaze, and how this interacts with their conceptual understanding.

These findings can be interpreted as children gradually acquire the ability to make judgements of gaze, and that it is not existent in some implicit form that becomes suddenly available when children have the necessary language. However, the findings of the current thesis implies that this development starts earlier than late 2-years. Without this data one can only make speculations. However, it may be that gaze judgement is directly related to gaze following ability, with children using what they have already learned in more flexible and accurate ways.

In line with previous research (Doherty & Anderson, 1999), the findings of both experiments show that judgements of where someone is looking are easier than judging which of two faces is looking at you. Given that the tasks are strongly correlated, it is likely that they both share the same core ability but for different functions. Most notably, the Looking Where task refers to objects, whereas the Looking At You task refers to the child themselves. The Looking At You task may thus be harder because it relies on more complex social skills such as introspection and intrapersonal skills.

A novel contribution of Experiments 1 and 2 concerns judgement of eye-direction with an accompanied head turn. Previous literature found performance on trials with an accompanied head turn was near ceiling (Doherty & Anderson, 1999). This was interpreted as eye-direction being a special stimulus that children had specific difficulty with. However, the findings from the current thesis show that whilst children find it easier to make judgements with an accompanied head turn, this was still a difficult to process. This is surprising as head direction is a large, salient cue that does not rely on complex calculations such as iris-sclera comparisons. However, human eye turns have a corresponding head turn around 40% of the time (Kobayashi & Koshima, 2001), meaning it is a prominent cue for gaze direction. As such, children's difficulty making judgements of gaze direction may not be specific to the eyes, but more generally towards all cues for gaze direction.

Together, the findings of Experiments 1 and 2 indicate that children develop the ability to judge eye direction in the preschool period. This understanding appears to apply generally to cues of gaze and potentially has multiple functions. Gaze judgement is likely to be a novel ability, that requires practice over many years. However, without the necessary infancy data it cannot be concluded with certainty that this is indicative of two-systems for gaze processing as previously claimed. Future research should consider extending the age that children are tested on gaze judgement tasks into infancy and comparing it to their gaze following ability. If this line of research does find the two abilities are directly related, then a single system account may better reflect human gaze processing.

Is understanding of gaze direction integral to theory of mind development?

Understanding the significance of eye direction appears to be a uniquely human ability. For example, non-human primates can follow gaze cues but show little ability to make inferences from them (e.g. Call et al., 2000; Anderson & Mitchell, 1999; Povinelli & Eddy, 1996a; 1996b; Anderson et al., 1996). Thus, the two-systems theory posits that humans are motivated to understand the significance of gaze as a result of our higher social skills (Doherty et al., 2009). Specifically, as children develop a theory of mind, they become more interested in the cues that tell them about the contents of others' minds. This is an intuitive prediction as perception is the way that knowledge and beliefs are frequently formed. As such, gaze understanding may be an integral part of theory of mind development. Experiments 1 and 2 tested this hypothesis by investigating concurrent and precursor relationships of gaze judgement and theory of mind ability.

Children were tested on gaze judgement, Knowledge Access and False Belief tasks across the ages of 2- to 4-years. In line with previous literature, the findings of both experiments show the two abilities are similar in their developmental trajectory (Doherty & Anderson, 1999; McGuigan & Doherty. 2002). Both abilities go through rapid development in the preschool period, with children demonstrating good performance by 4-years. The latent growth curve models from Experiment 2 demonstrate the most compelling evidence, as the intercept and slope values of the gaze judgement tasks are similar to that of the theory of mind tasks. However, it is important to highlight that in both experiments gaze judgement tasks were significantly easier than theory of mind tasks at all ages. This would suggest that gaze understanding develops slightly earlier than theory of mind.

Despite the concurrent development, there was a lack of association between the two abilities beyond common associations with age (Experiment 1) or at the same timepoint (Experiment 2). Specifically, the lack of association with the False Belief task replicates the previous literature (Doherty & Anderson, 1999; D'Entremont et al., 2012). This is unsurprising as the False Belief task goes beyond children using perception to determine if someone holds a belief or if it has become outdated, additionally testing misrepresentation. Knowledge Access is plausibly a purer test of whether children conceive of attention as a mental state, as perception is a core function of its reasoning. In line with this, D'Entremont et al. (2012) found that knowledge understanding associated with gaze understanding, motivating the prediction that there would be an association in the current thesis. Thus, the lack of association in the current thesis was unexpected.

Previous findings may reflect the simplicity of the tasks used. D'Entremont et al. (2012) asked children to identify which of *two* objects an agent was looking at. This only requires children to distinguish between gaze to the left or right but does not necessarily mean they judge the agent to be looking at that object. The findings of Experiment 1 would support this claim. When making an incorrect response, children are selecting the correct side more often than they would by chance. This suggests that children do interpret some directional information from the eyes, but do not have sufficient understanding to determine the exact object of attention. Thus, D'Entremont et al.'s (2012) findings may reflect the simpler attention directing System 1, than that of the gaze judgement-based System 2. Together, it could be concluded that understanding of eye direction does not require children to attribute mental states.

Even though the two abilities may not share a common understanding, one may still motivate the other to develop. Specifically, the two-system theory would claim that theory of mind ability should predict later gaze judgement ability. I tested this in Experiment 2, where I was interested in the cross-lagged effects of the two abilities. Across all timepoints, the only predictor of later gaze judgement performance was itself. Neither theory of mind task performance predicted later gaze judgement. Time 1 gaze judgement did however predict Time

2 False Belief performance. Thereafter, earlier False Belief performance was the only significant predictor of later False Belief performance. This is a clear finding that could be interpreted as gaze understanding being an immediate precursor to theory of mind.

This precursor relation seems implausibly direct. This is because gaze judgement was not an earlier predictor of Knowledge Access, which children acquire before False Belief (e.g. Wellman & Liu, 2004; Wellman et al., 2011; Peterson et al., 2012). Neither gaze understanding nor Knowledge Access was measured at an age where there was near-floor performance. This means that there may be a precursor relation that occurs much earlier than the current data set allows to be seen. As such, it cannot be strongly concluded that gaze understanding is an immediate precursor to theory of mind, but the data available lends partial support for this. Thus, it could be interpreted that gaze understanding allows children to connect perception with experience, to form representations of others' beliefs.

Together, the findings of the current thesis do not support the claims of the two-system theory. Whilst the two abilities develop concurrently, gaze understanding is easier than theory of mind and does not appear to require the ability to attribute mental states. It is unlikely to be motivated by children's increasing interest in others' mind as it is not predicted by earlier theory of mind performance. Instead, the findings of the current thesis partially support the consensus view that gaze understanding is a precursor to theory of mind development since perception is the way that knowledge and beliefs are frequently formed (Goméz, 1996; Gopnik et al., 1994). More research is needed with younger children to consolidate this claim in terms of knowledge formation.

Does System 1 remain luminance based in adulthood?

Thus far, the current thesis has concerned children's developing System 2. It has been suggested that System 2 is distinct from System 1 in childhood (Doherty et al., 2009), but it is unclear if they remain distinct in adulthood. To explore this, Doherty et al. (2015) showed that the two systems in childhood could be distinguished by the visual information they use from the eyes. Specifically, System 1 uses the relative luminance of the iris and sclera, whereas System 2 incorporates geometrical calculations. In adulthood however, they found System 1 uses both types of information suggesting the systems may not remain distinct through the lifespan. Doherty et al. (2015) only considered which cues are being used for effective System 1 behaviour but did not consider if either cue is more important for System 1 processing. For example, adult System 1 may use both luminance and geometric cues, but may rely primarily on luminance cues in its processing. Evidence from reaction time data suggests this is likely (Munsters et al., 2016). Thus, the aim of Experiments 3 and 4 was to investigate whether adult System 1 may use luminance cues more efficiently than geometric cues, and whether this finding is robust across other measures of System 1.

To test this, in Experiment 3 I replicated Doherty et al.'s (2015) gaze cueing task additionally recording measurements of reaction time. Firstly, participants made more correct and faster target identifications, when the target appeared in the gazed at location than the non-gazed at location. However, pertinent to the hypothesis being tested, gaze cueing did not differ between luminance and geometric cues for either accuracy or reaction time. These findings suggest that luminance and geometric cues are equally important for System 1 processing.

Experiment 4 aimed to see if this pattern of results is seen with other measures of adult System 1. Here I adapted Samson et al.'s (2010) dot perspective task, replacing the avatar with faces retaining luminance or geometric cues. This task was chosen as it draws many similarities with gaze cueing tasks and Holland et al.'s (2021) meta-analysis found that a large portion of the task effects can be explained by attention orienting. This task additionally has the advantage of controlling for the potential that System 1 operates at a faster timescale than System 2. Specifically, in Experiment 3 targets are not presented until 500ms, which is potentially long enough for the recruitment of the more calculated geometric cues. In the dot perspective task however, the target is presented at 0ms. If luminance cues are more important for System 1 processing, then reaction times are likely to be faster as the target should be located before geometric cues have begun to be processed. As such, it was thought the dot perspective task would better demonstrate any differences in processing efficiency than Experiment 3. On the contrary, the findings of Experiment 4 were identical to Experiment 1. As such, Chapter 4 presents strong evidence that adult System 1 uses luminance and geometric cues in both outcome and processing, suggesting it does not remain distinct in adulthood.

These findings present a query for the two-system theory: is gaze processing best explained by a single system, or does System 1 become more flexible in its processing due to System 2 developments? Future work should consider this, by investigating how System 1 change over the lifespan. For example, Doherty et al. (2015) suggest that a gradual increase in the use of geometric cues for both gaze following and gaze judgements would be indication of a single system, whereas a gradual increase for gaze judgement followed by sudden use for gaze following would be indication of two systems.

In summary, Experiments 3 and 4 aimed to investigate whether System 1 remains distinct in adulthood. Using two measures of attention orienting it was found that System 1 relies on both luminance and geometric information from the eyes in both outcome and processing. Thus, System 1 does not remain distinct in adulthood. The results do not provide strong support for either a two-system or single system account, suggesting future research should consider how geometric cues become incorporated into System 1.

What can be concluded about the two-system theory?

Central to the thesis is the two-systems theory of gaze processing (Doherty, 2006). This posits that we respond to gaze direction in two ways. System 1 uses other’s gaze cues to locate object of attention, whilst System 2 uses the same cues but to allow thoughts about the attentional relationship between observer and object. Evidence that humans are able to follow gaze from infancy (e.g. Corkum & Moore, 1995; Brooks & Meltzoff, 2002), but gradually develop the ability to make gaze judgements in the preschool period (Doherty & Anderson, 1999) and beyond (Doherty et al. 2009; Vida & Maurer, 2012a; 2012b) is the basis that these systems are distinct and not directly related in development. The theory additionally hypothesizes that System 2 development is motivated by children's increasing interest in the mind. Thus, to find support for the theory the findings of the current thesis should meet the following basic claims:

- *Gaze judgements develop over a protracted period, gradually increasing in accuracy.*

The findings of Experiments 1 and 2 are in support of this claim – children do develop the ability to judge eye direction over a protracted period with gradually increasing accuracy.

- *Gaze judgements reflect a conceptual development in both seeing and mental state understanding.*

The findings of Experiments 1 and 2 find partial support for this claim. Before children have access to System 2, they understand seeing in terms of engagement. To be able to use effectively use eye direction and therefore access System 2, children need to instead conceive of seeing in terms of line of sight. The current thesis presents no evidence that System 2 develops as a result of children's increasing theory of mind ability. However, System 2 may become prioritized over engagement to monitor others' attention due to an increasing understanding of knowledge formation.

- *Gaze systems likely remain distinct throughout the lifespan.*

The findings of Experiments 3 and 4 do not support this claim – adult System 1 incorporates both luminance and geometric information for both its outcome and processing.

From this, I would conclude that children have two systems for processing gaze. However, I would argue that System 2 does not develop as a result of children's increasing interest in others' mind, but instead due to changes in their spatial representation of gaze.

Specifically, young children possibly represent gaze as a cone and need to learn to represent it in lines to be able to recruit line of sight concepts necessary for System 2 functions. However, it cannot be strongly concluded that adults have two systems for gaze processing. Adult System 1 may become more flexible in its use as a result of System 2 developments, or adults may simply not require two systems once they have a sophisticated understanding of gaze. As such, more research is needed across the adolescent period, to determine how the two systems come to co-exist in adulthood.

References

Adoh, T. O., & Woodhouse, J. M. (1994). The Cardiff acuity test used for measuring visual acuity development in toddlers. *Vision research*, *34*(4), 555-560. https://doi.org/10.1016/0042-6989(94)90168-6

Anderson, J. R., & Mitchell, R. W. (1999). Macaques but not lemurs co-orient visually with humans. Folia Primatologica, 70, 17-22. https://doi.org/10.1159/000021670

Anderson, J. R., Montant, M., & Schmitt, D. (1996). Rhesus monkeys fail to use gaze direction as an experimenter-given cue in an object-choice task. *Behavioural Processes*, *37*(1), 47-55. https://doi.org/10.1016/0376-6357(95)00074-7

Ando, S. (2002). Luminance-induced shift in the apparent direction of gaze. *Perception, 31*(6), 657-674. https://doi.org/10.1068/p3332

Andrew, R.J. (1963). The origins and evolution of calls and facial expressions of the primates. *Behaviour, 20*, 1-109. https://doi.org/10.1163/156853963X00220

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, *116*(4), 953. https://doi.org/10.1037/a0016923

Arre, A. M. & Santos, L. R. (2021). Mentalizing in non-human primates. In *The Neural Basis of Mentalizing* (pp.131-147). Cham: Spinger International Publishing

Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental psychology*, *35*(5), 1311.

Baillargeon, R., Scott, R. M. & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences, 14*(3), 110-118. https://doi.org/10.1016/j.tics.2009.12.006

Baron-Cohen, S. (1991). Precursors to a theory of mind: Understanding attention in others. *Natural theories of mind: Evolution, development and simulation of everyday mindreading*, *1*(233-251), 1.

Bräuer, J., Call, J., & Tomasello, M. (2005). All great ape species follow gaze to distant locations and around barriers. *Journal of Comparative Psychology, 119*(2), 145–154. https://doi.org/10.1037/0735-7036.119.2.145

Brooks, R., & Meltzoff, A. (2002). The importance of eyes: How infants interpret adult looking behavior. *Developmental Psychology, 38*, 701–711. https://doi.org/10.1037/0012-1649.38.6.958

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental science*, *8*(6), 535-543. https://doi.org/10.1111/j.1467-7687.2005.00445.x

Bushnell, E. W., McKenzie, B. E., Lawrence, D. A., & Connell, S. (1995). The spatial coding strategies of one-year-old infants in a locomotor search task. *Child development*, *66*(4), 937-958. https://doi.org/10.1111/j.1467-8624.1995.tb00914.x

Butler, S., Caron, A., & Brooks, R. (2000). Infant understanding of the referential nature of looking. *Journal of Cognition and Development, 1,* 359–377.flave https://doi.org/10.1207/S15327647JCD0104_01

Butterfill, S. A., & Apperly, I. A. (2013). How to construct a minimal theory of mind. *Mind & Language*, *28*(5), 606-637. https://doi.org/10.1111/mila.12036

Butterworth, G., & Jarrett, N. (1991). What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology, 9*(1), 55–72. https://doi.org/10.1111/j.2044-835X.1991.tb00862.x

Call, J., Agnetta, B., & Tomasello, M. (2000). Cues that chimpanzees do and do not use to find hidden objects. *Animal Cognition, 3*, 23-34. https://doi.org/10.1007/s100710050047

Carlin. J. D., & Calder, A. J. (2013). The neural basis of eye gaze processing. *Current Opinion in Neurobiology, 23*(3), 450-455. *https://doi.org/10.1016/j.conb.2012.11.014*

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. In *Measurement of executive function in early childhood* (pp. 595-616). Psychology Press.

Carlson, S. M., Claxton, L. J., & Moses, L. J. (2015). The relation between executive function and theory of mind is more than skin deep. *Journal of Cognition and Development*, *16*(1), 186-197. https://doi.org/10.1080/15248372.2013.824883

Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development: An International Journal of Research and Practice*, *11*(2), 73-92. https://doi.org/10.1002/icd.298

Caron, A. J., Kiel, E., Dayton, M., & Butler, S. (2002). Comprehension of the referential intent of looking and pointing between 12 and 15 months. Journal of Cognition and Development, 3, 445–464. https://doi.org/10.1080/15248372.2002.9669677

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the society for research in child development*, i-174. https://doi.org/10.2307/1166214

Chacón-Candia, J. A., Román-Caballero, R., Aranda-Martín, B., Casagrande, M., Lupiáñez, J., & Marotta, A. (2023). Are there quantitative differences between eye-gaze and arrow cues? A meta-analytic answer to the debate and a call for qualitative differences. *Neuroscience & Biobehavioral Reviews*, *144*, 104993. https://doi.org/10.1016/j.neubiorev.2022.104993

Charman, T., Swettenham, J., Baron-Cohen, S., Cox, A., Baird, G., & Drew, A. (1997). Infants with autism: an investigation of empathy, pretend play, joint attention, and imitation. *Developmental psychology*, *33*(5), 781. https://doi.org/10.1037/0012-1649.33.5.781

Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, *9*(4), 377-395. https://doi.org/10.1016/0885-2014(94)90012-4

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), 454-465. https://doi.org/10.1037/xhp0000319

Corkum, V. L., & Moore, C. (1995). Development of joint visual attention in infants. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 61-83). Hillsdale, NJ: Erlbaum.

Corkum, V., & Moore, C. (1998). The origins of joint visual attention in infants. Developmental Psychology, 34(1), 28. https://doi.org/10.1037/0012-1649.34.1.28

Coss, R. G. (1970). The perceptual aspects of eye-spot patterns and their relevance to gaze behaviour. *Behaviour Studies in Psychiatry (Hutt, C. & Hutt, SJ, eds). Pergamon Press, London*, 121-147.

D'Entremont, B., Seamans, E. & Boudreau, E. (2012). The relationship between children's gaze reporting and theory of mind. *Journal of Cognition and Development, 13*(4), 505-523. https://doi.org/10.1080/15248372.2011.602653

de Bordes, P. F., Cox, R.F.A., Hasselman, F., & Cillessen, A. H. N. (2013). Toddlers' gaze following through attention modulation: Intention is in the eye of the beholder. *Journal of Experimental Child Psychology, 116*(2). 443-452. https://doi.org/10.1016/j.jecp.2012.09.008

de Heering, A., Turati, C., Rossion, B., Bulf, H., Goffaux, V., & Simion, F. (2008). Newborns' face recognition is based on spatial frequencies below 0.5 cycles per degree. *Cognition*, *106*(1), 444–454. https://doi.org/10.1016/j.cognition.2006.12.012

de Jong, M. C., van Engeland, H., & Kember, C. (2008). Attentional effects of gaze shifts are influenced by emotion and spatial frequency, but not in Autism. *Journal of the American*

Academy of Child & Adolescent Psychiatry. 47(4), 443-454. https://doi.org/10.1097/CHI.0b013e31816429a6

Deák, G. O., Flom, R. A., & Pick, A. D. (2000). Effects of gesture and target on 12- and 18-month-olds' joint visual attention to objects in front of or behind them. *Developmental Psychology, 36*, 511–523. https://doi.org/10.1037/0012-1649.36.4.511

Deak, G. O., Krasno, A. M., Triesch, J., Lewis, J., & Sepeta, L. (2014). Watch the hands: Infants can learn to follow gaze by seeing adults manipulate objects. *Developmental science*, *17*(2), 270-281. https://doi.org/10.1111/desc.12122

D'Entremont, B. (2000). A perceptual-attentional explanation of gaze-following in 3- and 6-month-olds. *Developmental Science, 3*, 302–311. https://doi.org/10.1111/1467-7687.00124

D'Entremont, B., & Morgan, R. (2006). Experience with visual barriers and its effects on subsequent gaze-following in 12-to 13-month-olds. *British Journal of Developmental Psychology*, *24*(3), 465-475. https://doi.org/10.1348/026151005X51248

D'Entremont, B., Hains, S. M. J., & Muir, D.W. (1997). A demonstration of gaze following in 3- to 6-month-olds. *Infant Behavior and Development*, *20*(4), 569-572. https://doi.org/10.1016/S0163-6383(97)90048-5

Deves, S., Williams, C., Parker, J., Harvey, I., Sparrow, J. M., & Harrad, R. A. (1996). Visual acuity testing in children up to age 3: Normal ranges and comparisons between tests, from the"ALSPAC"study. In *Investigative Ophthalmology And visual science* (Vol. 37, No. 3, pp. 3335-3335). 227 East Washington Sq, Philadelphia, Pa 19106: Lippincott-Raven Publ.

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

Doherty, M. (2011). A two-systems theory of social cognition. *Perception, causation, and objectivity*, 305-323.

Doherty, M. J. (2006). The development of mentalistic gaze understanding. *Infant and Child Development: An International Journal of Research and Practice*, *15*(2), 179-186. doi.org/10.1002/icd.434

Doherty, M. J., & Anderson, J. R. (1999). A new look at gaze: preschool children's understanding of eye-direction. *Cognitive Development*, *14*(4), 549-571. https://doi.org/10.1016/S0885-2014(99)00019-2

Doherty, M. J., & Anderson, J. R. (2001). People don't keep their heads still when looking to one side, and other people can tell. *Perception*, *30*(6), 765-767. https://doi.org/10.1068/p2998

Doherty, M. J., Anderson, J. R., & Howieson, L. (2009). The rapid development of explicit gaze judgment ability at 3 years. *Journal of Experimental Child Psychology*, *104*(3), 296-312. https://doi.org/10.1016/j.jecp.2009.06.004

Doherty, M. J., McIntyre, A. H., & Langton, S. R. H. (2015). Developmentally distinct gaze processing systems: Luminance versus geometric cues. Cognition, 137, 72-80. https://doi.org/10.1016/j.cognition.2015.01.001

Dörrenberg, S., Rakoczy, H. & Liszkowski, U. (2018). How (not) to measure infant Theory of Mind: Testing the replicability and validity of four non-verbal measures. *Cognitive Development, 46*, 12-30. https://doi.org/10.1016/j.cogdev.2018.01.001

Downing, P. E., Dodds, C. M., & Bray, D. (2004). Why does the gaze of others direct visual attention? *Visual Cognition, 11*(1), 71–79. https://doi.org/10.1080/13506280344000220

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test--Fourth Edition (PPVT-4)* [Database record]. APA PsycTests. https://doi.org/10.1037/t15144-000

Dunn, L. M., & Dunn, D. M. (2009). *The British picture vocabulary scale*. GL Assessment Limited.

Edwards, M., & Badcock, D. R. (1996). Global-motion perception: Interaction of chromatic and luminance signals. *Vision Research*, *36*(16), 2423-2431. https://doi.org/10.1016/0042-6989(95)00304-5

Emery, N. J. (2000). The eyes have it: The neuroethology, function and evolution of social gaze. Neuroscience & Biobehavioral Reviews, 24, 581-604. https://doi.org/10.1016/S0149-7634(00)00025-7

Emery, N.J., Lorincz, E.N., Perrett, D.I., Oram, M.W., & Baker, C.I. (1997). Gaze following and joint attention in rhesus monkeys (Macaca mulatto). *Journal of Comparative Psychology*, *111*(3), 286-293. https://doi.org/10.1037/0735-7036.111.3.286

Farrar, M. J., & Maag, L. (2002). Early language development and the emergence of a theory of mind. *First language*, *22*(2), 197-213. https://doi.org/10.1177/014272370202206504

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. Proceedings of the National Academy of Sciences of the United States of America, 99(14), 9602-9605. https://doi.org/10.1073/pnas.152159999

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175-191. https://doi.org/10.3758/BF03193146

Fizke, E., Butterfill, S., van de Loo, L., Reindl, E., & Rakoczy, H. (2017). Are there signature limits in early theory of mind?. *Journal of Experimental Child Psychology*, *162*, 209-224. https://doi.org/10.1016/j.jecp.2017.05.005

Flavell, J. H., Shipstead, S. G., & Croft, K. (1978). Young children's knowledge about visual perception: Hiding objects from others. *Child Development, 49*(4), 1208-1211 https://doi.org/10.2307/1128761

Flom, R., Deák, G. O., Phill, C. G., & Pick, A. D. (2004). Nine-month-olds' shared visual attention as a function of gesture and object location. *Infant Behavior and Development, 27*(2), 181–194. https://doi.org/10.1016/j.infbeh.2003.09.007

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), 158–163. https://doi.org/10.1037/xhp0000138

Gao, X., & Maurer, D. (2011). A comparison of spatial frequency tuning for the recognition of facial identity and facial expressions in adults and children. *Vision Research, 51*(5), 508-519. https://doi.org/10.1016/j.visres.2011.01.011

Gardner, M. R., Bileviciute, A. P., & Edmonds, C. J. (2018). Implicit mentalising during level-1 visual perspective taking indicated by dissociation with attention orienting. *Vision, 2*(1), 3. https://doi.org/10.3390/vision2010003

Garnham, W. A., & Perner, J. (2001). Actions really do speak louder than words—But only implicitly: Young children's understanding of false belief in action. *British Journal of Developmental Psychology*, *19*(3), 413-432. https://doi.org/10.1348/026151001166182

Garnham, W. A., & Ruffman, T. (2001). Doesn't see, doesn't know: is anticipatory looking really related to understanding or belief? *Developmental Science*, *4*(1), 94-100. https://doi.org/10.1111/1467-7687.00153

Geisler, W. S., Albrecht, D. G., & Crane. A. M. (2007). Responses of neurons in primary visual cortex to transient changes in local contrast and luminance. *Journal of Neuroscience, 27*(19), 5063-5067. https://doi.org/10.1523/JNEUROSCI.0835-07.2007

Gibson, J. J., & Pick, A. D. (1963). Perception of another person's looking behaviour. *American Journal of Psychology, 78*, 386-394. https://doi.org/10.2307/1419779

Goméz, J.C. (1996). Non-human primate theories of (non-human primate) minds: some issues concerning the origins of mind-reading. In: P. Carruthers and P. Smith (Eds.) *Theories of theories of mind* (pp. 330-343). Cambridge, Cambridge University Press.

Gopnik, A., Slaughter, V., & Meltzoff, A. (1994). Changing your views: How understanding visual perception can lead to a new theory of mind. In Lewis, C. & Mitchell, P. (Eds.) *Children's early understanding of mind*. Hove: Lawrence Erlbaum Associates, pp. 157–181.

Grosse Wiesmann, C., Friederici, A. D., Disla, D., Steinbeis, N., & Singer, T. (2018). Longitudinal evidence for 4-year-olds' but not 2-and 3-year-olds' false belief-related action anticipation. *Cognitive Development*, *46*, 58-68. https://doi.org/10.1016/j.cogdev.2017.08.007

Grosse Wiesmann, C., Friederici, A. D., Singer, T., & Steinbeis, N. (2017). Implicit and explicit false belief development in preschool children. *Developmental Science*, *20*(5), e12445. https://doi.org/10.1111/desc.12445

Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? Animal Behaviour, 61, 139–151. https://doi.org/10.1006/anbe.2000.1518

Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. Animal Behaviour, 59,771–785. https://doi.org/10.1006/anbe.1999.1377

Hauser, M. D., Glynn, D., & Wood, J. (2007). Rhesus monkeys correctly read the goal-relevant gestures of a human agent. *Proceedings of the Royal Society B, 274*, 1913-1918. *https://doi.org/10.1098/rspb.2007.0586*

Heyes, C. (2014). Submentalizing: I am not really reading your mind. *Perspectives on Psychological Science*, *9*(2), 131-143. https://doi.org/10.1177/1745691613518076

Holland, C., Shin, S. M., & Phillips, J. (2021). Do you see what I see? A meta-analysis of the Dot Perspective Task. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 43, No. 43).

Hood, B.M., Willen, J.D., & Driver, J. (1998). Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science*, *9*(2), 131-134. https://doi.org/10.1111/1467-9280.00024

Horschler, D. J., Santos, L. R., & MacLean, E. L. (2019). Do non-human primates really represent others' ignorance? A test of the awareness relations hypothesis. *Cognition*, *190*, 72-80. https://doi.org/10.1016/j.cognition.2019.04.012

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, *6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Hughes, C., & Ensor, R. (2007). Executive function and theory of mind: Predictive relations from ages 2 to 4. *Developmental psychology*, *43*(6), 1447. https://doi.org/10.1037/0012-1649.43.6.1447

Itakura, S., & Tanaka, M. (1998). Use of experimenter-given cues during object-choice tasks by chimpanzees (Pan troglodytes), an orangutan (Pongo pygmaeus), and human infants (Homo sapiens). *Journal of Comparative Psychology*, *112*(2), 119. https://doi.org/10.1037/0735-7036.112.2.119

Jasso, H., Triesch, J., Deák, G., & Lewis, J. M. (2012). A unified account of gaze following. *IEEE Transactions on Autonomous Mental Development*, *4*(4), 257-272. https://doi.org/10.1109/TAMD.2012.2208640

Jenkins, J., & Langton, S. R. H. (2003). Configural processing in the perception of eye-gaze-direction. *Perception, 32*(10), 1181-1188. https://doi.org/10.1068/p3398

Jones. W., & Klin, A. (2013). Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism. *Nature, 504*, 427-431. *https://doi.org/10.1038/nature12715*

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kaminiski, J., Call, J., & Tomasello, M. (2004). Body orientation and face orientation: Two factors controlling apes' begging behaviour from humans. *Animal Cognition, 7*, 216-223. https://doi.org/10.1007/s10071-004-0214-2

Kampis, D., Karman, P., Csibra, G., Southgate, V., & Hernik, M. (2021). A two-lab direct replication attempt of Southgate, Senju and Csibra (2007). *Royal Society open science*, *8*(8), 210190. https://doi.org/10.1098/rsos.210190

Kobayashi, H., & Hashiya, K. (2011). The gaze that grooms: Contribution of social factors to the evolution of primate eye morphology. *Evolution and Human Behavior, 32*(3), 157-165. https://doi.org/10.1016/j.evolhumbehav.2010.08.003

Kobayashi, H., & Kohshima, S. (1997). Unique morphology of the human eye. Nature, 387, 767-768. https://doi.org/10.1038/42842

Kobayashi, H., & Kohshima, S. (2001). Unique morphology of the human eye and its adaptive meaning: comparative studies on external morphology of the primate eye. *Journal of Human Evolution*, *40*(5), 419-453. https://doi.org/10.1006/jhev.2001.0468

Kulke, L., & Rakoczy, H. (2018). Implicit Theory of Mind–An overview of current replications and non-replications. *Data in brief*, *16*, 101-104. https://doi.org/10.1016/j.dib.2017.11.016

Langton, S. R. (2018). I don't see it your way: The dot perspective task does not gauge spontaneous perspective taking. *Vision*, *2*(1), 6. **https://doi.org/10.3390/vision2010006**

Lehmann, E. L., & D'Abrera, H. J. (2006). *Nonparametrics: statistical methods based on ranks* (Vol. 464). New York: Springer.

Leslie, A. M. (2005). Developmental parallels in understanding minds and bodies. *Trends in Cognitive Sciences*, *9*(10), 459-462. https://doi.org/10.1016/j.tics.2005.08.002

Lewis, J., Jasso, H., Triesch, J., & Deák, G. O. (2010). Building a model of infant social interaction.

Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior research methods*, *48*, 936-949. https://doi.org/**10.3758/s13428-015-0619-7**

Libertus, M. E., Odic, D., Feigenson, L., & Halberda, J. (2015). A Developmental Vocabulary Assessment for Parents (DVAP): Validating parental report of vocabulary size in 2-to 7-year-old children. *Journal of Cognition and Development*, *16*(3), 442-454. https://doi.org/10.1080/15248372.2013.835312

Low, J., & Watts, J. (2013). Attributing false beliefs about object identity reveals a signature blind spot in humans' efficient mind-reading system. *Psychological Science*, *24*(3), 305-311. https://doi.org/10.1177/0956797612451469

Low, J., Apperly, I. A., Butterfill, S. A., & Rakoczy, H. (2016). Cognitive Architecture of Belief Reasoning in Children and Adults: A Primer on the Two-Systems Account. *Child Development Perspectives*, *10*(3), 184-189. https://doi.org/10.1111/cdep.12183

Low, J., Drummond, W., Walmsley, A., & Wang, B. (2014). Representing how rabbits quack and competitors act: Limits on preschoolers' efficient ability to track perspective. *Child development*, *85*(4), 1519-1534. https://doi.org/10.1111/cdev.12224

Marcovitch, S., O'Brien, M., Calkins, S. D., Leerkes, E. M., Weaver, J. M., & Levine, D. W. (2015). A longitudinal assessment of the relation between executive function and theory of mind at 3, 4, and 5 years. *Cognitive development*, *33*, 40-55. https://doi.org/10.1016/j.cogdev.2014.07.001

Mares, I., Smith, M. L., Johnson, M. H., & Senju, A. (2018). Revealing the neural time-course of direct gaze processing via spatial frequency manipulation of faces.

Martin, A., & Santos, L. R. (2016). What cognitive representations support primate theory of mind? *Trends in cognitive sciences*, *20*(5), 375-382. https://doi.org/10.1016/j.tics.2016.03.005

McGuigan, N., & Doherty, M. J. (2002). The relation between hiding skill and judgment of eye direction in preschool children. *Developmental Psychology*, *38*(3), 418. https://doi.org/10.1037/0012-1649.38.3.418

Mearing, A. S., Burkart, J. M., Dunn, J., Street, S. E., & Koops, K. (2022). The evolutionary drivers of primate scleral coloration. *Scientific Reports*, *12*(1), 14119. https://doi.org/10.1038/s41598-022-18275-9

Meltzoff, A. N., & Brooks, R. (2008). Self-experience as a mechanism for learning about others: a training study in social cognition. *Developmental psychology*, *44*(5), 1257. https://doi.org.uk/10.1037/a0012888

Michael, J., Wolf, T., Letesson, C., Butterfill, S., Skewes, J., & Hohwy, J. (2018). Seeing it both ways: Using a double-cuing task to investigate the role of spatial cuing in Level-1 visual perspective-taking. *Journal of Experimental Psychology: Human Perception and Performance, 44*(5), 693–702. https://doi.org/10.1037/xhp0000486

Miller, S. D., Chow, D., Wampold, B. E., Hubble. M. A., Del Re, A. C., Maeschalck, C., & Bargmann, S. (2018). To be or not to be (an expert)? Revisiting the role of deliberate practice in improving performance. *High Ability Studies, 31*(1), 5-15. https://doi.org/10.1080/13598139.2018.1519410

Moll, H., & Tomasello, M. (2004). 12- and 18-month-old infant follow gaze to spaces behind barriers. *Developmental Science, 7*, 1–9. https://doi.org/10.1111/j.1467-7687.2004.00315.x

Moll, H., & Tomasello, M. (2006). Level 1 perspective-taking at 24 months of age. *British Journal of Developmental Psychology*, *24*(3), 603-613. https://doi.org/10.1348/026151005X55370

Moore, C., & Corkum, V. (1994). Social understanding at the end of the first year of life. *Developmental Review*, *14*(4), 349-372. https://doi.org/10.1006/drev.1994.1014

Moore, C., & Corkum, V. (1998). Infant gaze-following based on eye-direction. British Journal of Developmental Psychology, 16, 495–503. https://doi.org/10.1111/j.2044-835X.1998.tb00767.x

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

Morales, M., Mundy, P., Crowson, M., Neal, A. R., & Delgado, C. (2005). Individual differences in infant attention skills, joint attention, and emotion regulation behaviour. *International Journal of Behavioral Development*, *29*(3), 259-263. https://doi.org/10.1080/01650250444000432

Morales, M., Mundy, P., Delgado, C. E., Yale, M., Messinger, D., Neal, R., & Schwartz, H. K. (2000). Responding to joint attention across the 6-through 24-month age period and early language acquisition. *Journal of applied developmental psychology*, *21*(3), 283-298. https://doi.org/10.1016/S0193-3973(99)00040-4

Munsters, N. M., van den Boomen, C. Hooge, I. T. C., & Kemner, C. (2016). The role of global and local visual information during the gaze-cued orienting of attention. *PLoS ONE, 11*(8), e0160405. https://doi.org/10.1371/journal.pone.0160405

Muroyama, Y. (1991). Mutual reciprocity of grooming in female Japanese macaques (Macaca Fuscata). *Behaviour, 119*(3-4), 161-170. https://doi.org/10.1163/156853991X00427

Okamoto-Barth, S., Call, J., & Tomasello, M. (2007). Great apes' understanding of other individuals' line of sight, *Psychological Science, 18*(5), 462-468. https://doi.org/10.1111/j.1467-9280.2007.01922.x

Oktay-Gür, N., Schulz, A., & Rakoczy, H. (2018). Children exhibit different performance patterns in explicit and implicit theory of mind tasks. *Cognition*, *173*, 60-74. https://doi.org/10.1016/j.cognition.2018.01.001

O'Neill, D. K. (1996). Two-year-old children's sensitivity to a parent's knowledge state when making requests. *Child development*, *67*(2), 659-677. https://doi.org/10.1111/j.1467-8624.1996.tb01758.x

O'Neill, D. K., Astington, J. W., & Flavell, J. H. (1992). Young children's understanding of the role that sensory experiences play in knowledge acquisition. *Child Development*, *63*(2), 474-490. https://doi.org/10.1111/j.1467-8624.1992.tb01641.x

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? S*cience*, *308*(5719), 255-258. https://doi.org/10.1126/science.1107621

Perner, J., & Ruffman, T. (2005). Infants' insight into the mind: How deep? *Science*, *308*(5719), 214-216. https://doi.org/10.1126/science.1111656

Peterson, C. C., Wellman, H. M. & Slaughter, V. (2012). The mind behind the message: Advancing theory of mind scales for typically developing children, and those with deafness, autism, or asperger syndrome. *Child Development, 83*(2), 469-485. https://doi.org/10.1111/j.1467-8624.2011.01728.x

Povinelli, D. J., & Eddy, T. J. (1996a). Chimpanzees: joint visual attention. *Psychological Science*, *7*(3), 129-135. https://doi.org/10.1111/j.1467-9280.1996.tb00345.x

Povinelli, D.J., & Eddy, T.J. (1996b). What young chimpanzees know about seeing. *Monographs of the Society for Research in Child Development*, *61*(3).

Povinelli, D.J., Bierschwale, D.T., & Čech, C.G. (1999). Comprehension of seeing as a referential act in young children, but not juvenile chimpanzees. *British Journal of Developmental Psychology*, *17*(1), 37–60. https://doi.org/10.1348/026151099165140

Pratt, C., & Bryant, P. (1990). Young children understand that looking leads to knowing (so long as they are looking into a single barrel). *Child development*, *61*(4), 973-982. https://doi.org/10.1111/j.1467-8624.1990.tb02835.x

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515-526. https://doi.org/10.1017/S0140525X00076512

Prinzmetal, W., Leonhardt, J., & Garrett, R. (2008). Does gaze direction affect accuracy?. *Visual Cognition*, *16*(5), 567-584. https://doi.org/10.1080/13506280801981341

Qureshi, A. W., Apperly, I. A., & Samson, D. (2010). Executive function is necessary for perspective selection, not Level-1 visual perspective calculation: Evidence from a dual-task study of adults. *Cognition*, *117*(2), 230-236. https://doi.org/10.1016/j.cognition.2010.08.003

Rakoczy, H., Bergfeld, D., Schwarz, I., & Fizke, E. (2015). Explicit theory of mind is even more unified than previously assumed: Belief ascription and understanding aspectuality emerge together in development. *Child Development*, *86*(2), 486-502. https://doi.org/10.1111/cdev.12311

Ramsey, R., Hansen, P., Apperly, I., & Samson, D. (2013). Seeing it my way or your way: Frontoparietal brain areas sustain viewpoint independent perspective selection processes. *Journal of Cognitive Neuroscience, 25*(5), 670-684. https://doi.org/10.1162/jocn_a_00345

Rasch, D., & Guiard, V. (2004). The robustness of parametric statistical methods. *Psychology Science*, *46*, 175-208.

Rothbart, M. K., & Posner, M. I. (1985). Temperament and the development of self-regulation. In *The neuropsychology of individual differences: A developmental perspective* (pp. 93-123). Boston, MA: Springer US.

Ruffman, T., & Perner, J. (2005). Do infants really understand false belief? *Cognitive Development*, *9*, 377-395. https://doi.org/10.1016/j.tics.2005.08.001

Ruiz, A., Gómez, J.C., Roeder, J.J., & Byrne, R. W. (2008). Gaze following and gaze priming in lemurs. *Animal Cognition, 12*, 427-434. https://doi.org/10.1007/s10071-008-0202-z

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(5), 1255-1266. https://10.1037/a0018729

Santiesteban, I., Catmur, C., Hopkins, S. C., Bird, G., & Heyes, C. (2014). Avatars and arrows: Implicit mentalizing or domain-general processing? *Journal of Experimental Psychology: Human Perception and Performance, 40*(3), 929–937. https://doi.org/10.1037/a0035175

Satorra, A., & Saris, W. E. (1985). Power of the likelihood ratio test in covariance structure analysis. *Psychometrika, 50*(1), 83–90. https://doi.org/10.1007/BF02294150

Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature, 253*, 265–266. https://doi.org/10.1038/253265a0

Schick, B., De Villiers, P., De Villiers, J., & Hoffmeister, R. (2007). Language and theory of mind: A study of deaf children. *Child development*, *78*(2), 376-396. https://doi.org/10.1111/j.1467-8624.2007.01004.x

Schmitz, I., Strauss, H., Reinel, L., & Einhäuser, W. (2024). Attentional cueing: Gaze is harder to override than arrows. *Plos one*, *19*(3), e0301136. https://doi.org/10.1371/journal.pone.0301136

Schurz, M., Kronbichler, M., Weissengruber, S., Surtees, A., Samson, D., & Perner, J. (2015). Clarifying the role of theory of mind areas during visual perspective taking: Issues of spontaneity and domain-specificity. *NeuroImage*, *117*, 386-396. https://doi.org/10.1016/j.neuroimage.2015.04.031

Schuwerk, T., Kampis, D., Baillargeon, R., Biro, S., Bohn, M., Byers-Heinlein, K., … Rakoczy, H. (2021). Action anticipation based on an agent's epistemic state in toddlers and adults. https://doi.org/10.31234/osf.io/x4jbm

Shepherd, S. V. (2010). Following gaze: Gaze-following behavior as a window into social cognition. *Frontiers in Integrative Neuroscience, 4*, 5. https://doi.org/10.3389/fnint.2010.00005

Shepherd, S. V., & Platt, M. L. (2007). Spontaneous social orienting and gaze following in ringtailed lemures (Lemur catta). *Animal Cognition, 11*, 13-20. https://doi.org/10.1007/s10071-007-0083-6

Simms, N. K., & Gentner, D. (2019). Finding the middle: Spatial language and spatial reasoning. *Cognitive Development*, *50*, 177-194. https://doi.org/10.1016/j.cogdev.2019.04.002

Simms, N., & Gentner, D. (2008). Spatial language and landmark use: can 3-, 4-, and 5-year-olds find the middle?. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 30, No. 30).

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, *18*(7), 587-592. https://doi.org/10.1111/j.1467-9280.2007.01944.x

Spetch, M. L., & Parent, M. B. (2006). Age and sex differences in children's spatial search strategies. *Psychonomic Bulletin & Review*, *13*(5), 807-812. https://doi.org/10.3758/BF03194001

Stevens, S. A., West, G. L., Al-Aidroos, N., Weger, U. W., & Pratt, J. (2008). Testing whether gaze cues and arrow cues produce reflexive or volitional shifts of attention. *Psychonomic Bulletin & Review*, *15*(6), 1148-1153. https://doi.org/10.3758/pbr.15.6.1148

Tang, Y., Gonzalez, M. R., & Deák, G. O. (2024). The slow emergence of gaze-and point-following: A longitudinal study of infants from 4 to 12 months. *Developmental Science*, *27*(3), e13457. https://doi.org/10.1111/desc.13457

Tomasello, M. (2001). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Tomkins, S. S. (1963). *Affect, imagery, consciousness: The negative affects*. Springer Publishing Company.

Vestner, T., Balsys, E., Over, H., & Cook, R. (2022). The self-consistency effect seen on the dot perspective task is a product of domain-general attention cueing, not automatic perspective taking. *Cognition, 224*, 105056. https://doi.org/10.1016/j.cognition.2022.105056

Vida, M. D., & Maurer, D. (2012a). Gradual improvement in fine-grained sensitivity to triadic gaze after 6 years of age. *Journal of Experimental Child Psychology, 111*(2), 299-318. https://doi.org/10.1016/j.jecp.2011.08.009

Vida, M. D., & Maurer, D. (2012b). Fine-grained sensitivity to vertical differences in triadic gaze is slow to develop. *Journal of Vision, 12*, 634. https://doi.org/10.1167/12.9.634

Vida, M. D., & Maurer, D. (2015). A comparison of spatial frequency tuning for judgements of eye gaze and facial identity. *Vision Research, 112,* 45-54. https://doi.org/10.1016/j.visres.2015.04.018

Wellman, H. M. & Lui, D. (2004). Scaling of theory of mind tasks. *Child Development, 75*, 523-541. https://doi.org/10.1111/j.1467-8624.2004.00691.x

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. https://doi.org/10.1111/1467-8624.00304

Wellman, H. M., Fang, F. & Peterson, C. C. (2011). Sequential progressions in a theory of mind scale: Longitudinal perspectives. *Child Development, 82*(3), 780-792. https://doi.org/10.1111/j.1467-8624.2011.01583.x

Wimmer, H. & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103-128. https://doi.org/10.1016/0010-0277(83)90004-5

Wimmer, H., Hogrefe, G. J., & Perner, J. (1988). Children's understanding of informational access as source of knowledge. *Child development*, 386-396. https://doi.org/10.2307/1130318

Yorzinski, J. L., Thorstenson, C. A., & Nguyen, T. P. (2021). Sclera and iris color interact to influence gaze perception. *Frontiers in Psychology, 12*, 632616. https://doi.org/10.3389/fpsyg.2021.632616

Appendices

Appendix A

Table 26

Experiment 3 Zskew and Zkurtosis Full Sample

Dependent Variable	Trial Type	Cue Type	Zskew	Zurtosis
Number Correct	Cued	Normal	-0.79	0.09
		Luminance	-1.61	0.36
		Geometric	-2.24	0.69
	Non-cued	Normal	-0.95	-1.20
		Luminance	-0.33	-1.26
		Geometric	-1.84	0.45
Average Reaction Time	Cued	Normal	0.79	-0.20
		Luminance	1.67	0.62
		Geometric	1.98	-0.31
	Non-cued	Normal	5.08	9.48
		Luminance	11.86	33.53
		Geometric	5.00	5.25

Table 27

Experiment 3 Zskew and Zkurtosis After Removal of Outliers

Dependent Variable	Trial Type	Cue Type	Zskew	Zurtosis
Number Correct	Cued	Normal	-0.39	-0.20
		Luminance	-1.76	0.69
		Geometric	-2.23	1.16
	Non-cued	Normal	-0.77	-1.13
		Luminance	-0.21	-1.40
		Geometric	-1.76	0.27
Average Reaction Time	Cued	Normal	0.62	-0.72
		Luminance	-0.20	1.52
		Geometric	2.07	0.69
	Non-cued	Normal	0.47	-1.04
		Luminance	2.93	2.09
		Geometric	2.03	0.47

Table 28

Experiment 4a Zskew and Zkurtosis Full Sample

Dependent Variable	Consistency	Directional Cue	Zskew	Zurtosis
Number Correct	Consistent	Body	-5.14	3.06
		Head	-5.88	6.12
		Eyes	-4.75	2.89
	Inconsistent	Body	-3.98	1.13
		Head	-7.85	11.24
		Eyes	-4.92	2.37
Average Reaction Time	Consistent	Body	2.97	2.76
		Head	3.05	3.22
		Eyes	2.04	0.25
	Inconsistent	Body	12.46	36.69
		Head	1.30	2.32
		Eyes	1.21	-0.60

Table 29

Experiment 4a Zskew and Zkurtosis After Removal of Outliers

Dependent Variable	Consistency	Directional Cue	Zskew	Zurtosis
Number Correct	Consistent	Body	-4.93	3.21
		Head	-5.51	6.05
		Eyes	-4.15	1.95
	Inconsistent	Body	-3.94	1.41
		Head	-7.81	12.49
		Eyes	-4.52	2.19
Average Reaction Time	Consistent	Body	0.16	-0.47
		Head	0.40	-0.52
		Eyes	2.90	2.28
	Inconsistent	Body	-0.42	0.64
		Head	-1.52	3.61
		Eyes	1.90	0.16

Table 30

Experiment 4b Zskew and Zkurtosis Full Sample

Dependent Variable	Consistency	Psychophysical Cue	Zskew	Zurtosis
Number Correct	Consistent	Normal	-8.23	7.34
		Luminance	-9.14	-9.56
		Geometric	-12.18	20.04
	Inconsistent	Normal	-9.38	9.65
		Luminance	-9.75	11.02
		Geometric	-12.22	18.13
Response Time	Consistent	Normal	27.53	114.97
		Luminance	7.12	9.02
		Geometric	12.18	39.09
	Inconsistent	Normal	8.72	13.20
		Luminance	4.98	4.51
		Geometric	6.58	11.31

Table 31

Experiment 4b Zskew and Zkurtosis After Removal of Outliers

Dependent Variable	Consistency	Psychophysical Cue	Zskew	Zurtosis
Number Correct	Consistent	Normal	-7.11	6.08
		Luminance	-11.46	20.00
		Geometric	-12.65	28.80
	Inconsistent	Normal	-13.00	25.80
		Luminance	-11.26	18.98
		Geometric	-13.56	29.60
Response Time	Consistent	Normal	3.67	3.56
		Luminance	4.31	3.56
		Geometric	3.92	3.00
	Inconsistent	Normal	3.91	2.48
		Luminance	4.90	4.34
		Geometric	3.89	2.08