EFFICIENT SOCIAL PERCEPTION IN ADULTS:

STUDIES ON VISUAL PERSPECTIVE-TAKING

AND VISUAL WORKING MEMORY

by

JEN JESSICA WANG

A thesis submitted to the

University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

School of Psychology

University of Birmingham

June 2011

ABSTRACT

Ten experiments examined the way that automatic processing of the visual perspectives and eye gaze of others affects adults' perception and encoding of the social world. I investigated the amount of flexibility that automatic visual perspective-computation accommodates. Experiments 1, 2, and 3 demonstrate that automatic visual perspective-computation shows some flexibility for enumerating and representing perspective contents. Experiments 4 and 5 further indicate that automatic visual perspective-taking allows selection of relevant perspective information. I also examined whether observing others' eye gaze affects adults' visual working memory encoding. Experiments 6, 7, and 8 indicate that agents' object-oriented gaze does not lead to more efficient encoding of agent and object information. Experiments 9 and 10 demonstrate that observing others' participant-oriented gaze disrupts visual working memory encoding. I argue that although adults have minimal conscious control over the activation of visual perspective-computation and processing of participant-oriented gaze, the efficient mindreading system shows some flexibility.

ACKNOWLEDGEMENTS

I would like to thank Professor Ian Apperly for his tireless encouragement, guidance, and support throughout the completion of this work.

I would like to acknowledge the helpful comments I received from Dr Dana Samson, Dr Christoph Teufel, Dr Andrew Surtees, and members of the Developmental Psychology Group and Social Cognition and Neuroscience Group at Birmingham.

I would like to thank everyone from Office 517, who created a very enjoyable working environment.

Thanks to my family, especially my parents, for being supportive in every possible way.

Finally, thanks to Dr Patrick Burns for his patience in proofreading of this thesis and in many other things.

CONTENTS

List of figures

List of tables

CHAPTER 1 INTRODUCTION

1.1 Overview

Successful social relationships require the consideration of others' feelings and points of view—for example, taking account of a friend's likes and dislikes when planning a dinner party menu. Holding others' minds in mind helps us to behave appropriately in social situations. An ability to make flexible decisions about our interpersonal behaviour is essential for maintaining successful relationships with those around us. Nevertheless, in navigating the social world one cannot solely rely on one's ability to make reasoned and careful considerations about others' minds. Rather, social interactions often require one to respond quickly to the behaviour of others, such as when dancing or playing ball games together. We spontaneously turn our heads to follow others' finger pointing, gaze directions, and other body movements. These rapid responses to social cues appear to occur without the aid of careful consideration or complex reasoning. However, it is the ability to respond efficiently to social cues that makes a wide range of dynamic social interactions possible.

The aim of the present thesis is to examine how the efficient processing of social cues affects the perceptual computation and encoding of the social world. In particular, I will focus on adults' computation of others' eye gaze. I shall first review evidence from the social cognition literature which indicates that efficient processing and flexible processing are in tension with one another. Social psychology studies provide compelling evidence for a two-system account (Bargh, 1994); one appears to be flexible but effortful, while the second is automatic and efficient. This account predicts that the efficient processes which support dynamic social interaction should be relatively inflexible.

The present thesis first investigates the operational limit of automatic visual perspective-taking. The second part of this thesis examines whether perceiving others' eye gaze might affect visual memory encoding. This thesis is informed by the mindreading literature. Mindreading, also known as 'Theory of mind', concerns one's ability to infer others' mental states, such as their beliefs, desires, intentions, and emotions. Although the majority of the mindreading literature has a developmental focus, I intend to demonstrate that bringing a social cognitive perspective to bear on the mindreading literature is beneficial.

In this first chapter I will review evidence from traditional mindreading studies, a large part of which indicates that mindreading is effortful. I will then review evidence from recent mindreading studies and other studies of social cognitive processing. This will highlight a distinct type of mindreading, which is automatic and efficient in character. Amongst the range of automatic social cognitive processes, I will be focusing on adults' social perception; that is, basic perceptual computation and the encoding of information modulated by its social contents.

1.2 Two Systems of Social Cognition

The notion of having separate systems of social cognition is not an unusual one (e.g., Chaiken & Trope, 1999). Various accounts share the same general features: one system operates under deliberate and conscious control, whereas the other is more stimulus-driven, operating automatically outside of conscious control, and requiring minimal effort (Bargh, 1994; Gilbert, 1998). For example, Devine (1989) demonstrated that there are two dissociable processes involved in stereotyping. Individuals who displayed either high-prejudice or low-prejudice were equally knowledgeable of stereotypic information. Moreover, the two groups of individuals were equally likely to interpret ambiguous behaviours using stereotypic associations when conscious control was precluded. The crucial distinction, however, was that individuals who appear to have little prejudice actively replaced stereotypic thoughts with counter-stereotypic thoughts, whereas high prejudice individuals did not do so. This study suggests that automatic stereotyping activates well-learnt associations without conscious control, whereas the controlled processes draw on personal belief, which is likely to be consciously applied. Evidence also indicates that resolving information that is inconsistent with stereotypic knowledge demands cognitive effort (Macrae, Bodenhausen, Schloerscheidt, & Milne, 1999). Macrae et al. demonstrated that the inconsistency resolution process was impaired when executive function was loaded by a secondary task. Importantly, the inconsistency resolution process remained intact when the secondary task only interfered with non-executive operations. This suggests that applying personal beliefs that are incongruent with stereotypes is an effortful process. Beyond stereotype studies, there is evidence from other types of operations (e.g., attitude change), which indicates that having two separate systems to cope with the coexisting but incompatible demands for efficiency and flexibility is representative of a broader set of phenomena.

In a similar vein, Apperly and Butterfill (2009) have argued for two distinct types of mindreading operations. One effortful and flexible system makes heavy demands on executive resources. It often depends on language and memory in supporting complex social reasoning (e.g., belief inferences). A second system in contrast is efficient yet inflexible. It makes minimal demands on executive resources, allowing one to make quick social inferences to cope with fast rolling daily interactions and communication. I will discuss evidence from the domains of flexible mindreading and efficient mindreading, which illustrates the distinct characteristics of these two systems.

1.3 Flexible Mindreading

1.3.1 *Development of Flexible Mindreading*

1.3.1.1 *Overview*

Traditionally, investigations of mindreading have examined children's ability to reason about others' mental states. Developmental studies largely focused on children's acquisition of concepts necessary for mindreading and their ability to reason about others' minds in an adult-like fashion. That is, the ability to correctly predict others' behaviours according to his/her beliefs (e.g., Wimmer & Perner, 1983).

Premack and Woodruff (1978) revealed that a chimpanzee recognised a problem that an actor faced and the actor's purpose to solve the problem. The chimpanzee was able to identify a photo that represented an appropriate solution to the actor's problem. Based on this observation, Premack and Woodruff concluded that this chimpanzee possessed the same mindreading abilities as a human being. Several philosophers (Bennett, 1978; Dennett, 1978; Pylyshyn, 1978) have criticised this conclusion. Bennett, Dennett, and Pylyshyn pointed out that Premack and Woodruff's methodology does not provide a definitive distinction between the chimpanzee's problem-solving abilities and their purported ability to infer others' mental states. There was no differentiation between the actor's perspective and that of the chimpanzee. Therefore, it was impossible to determine whether or not the chimpanzee truly represented others' mental states or derived the solution by coincidence. Wimmer and Perner (1983) adopted these philosophers' methodology to create a developmental task, the false belief task (or the unexpected transfer task). In this task, a protagonist (Maxi) in a story has a false belief about an object's location. In contrast, the child reading the story holds a true belief about the object's location, setting up a perspective difference between the protagonist and the child. The story proceeds as follows: Maxi and his mum had just arrived home from the shop. Maxi's mum had bought Maxi a chocolate bar in the shop, which Maxi put in the blue cupboard. While Maxi was out playing, Maxi's mum moved the chocolate bar from the blue cupboard to the green cupboard. The child is then asked where Maxi would look for the chocolate bar when he comes back into the house. Wimmer and Perner found that the majority of three- to four-year-olds answered that Maxi will look for the chocolate bar in the true location (the green cupboard), failing to take Maxi's false belief into account. By four to six years of age, around half of the children responded correctly but it was not until six to nine years of age that children consistently produced the correct answer. A large number of subsequent studies have explored the task by performing various methodological manipulations, for example, the false content task (Hogrefe, Wimmer, & Perner, 1986) and the appearance-reality task (Gopnik & Astington, 1988). A meta-analysis (Wellman, Cross, & Watson, 2001) was conducted with 178 studies of different versions of the false belief task. This analysis revealed a consistent developmental progression in the preschool years across cultures and task manipulations. Children do not acquire the ability to engage in adult-like mindreading until they are four years of age.

1.3.1.2 *The Role of Executive Function*

In order to pass false belief tasks, children need to be able to disengage from their own salient true beliefs and produce correct predictions about the protagonist's behaviour according to the protagonist's false belief. Furthermore, children also need to be able to follow and remember the object's location as it appears to the protagonist (Gordon & Olson, 1998; Hughes, 1998). Given these cognitive demands, it is unsurprising that the developmental trajectories of mindreading and executive function have considerable overlap with one another (Gerstadt, Hong, & Diamond, 1994; Zelazo, Frye, & Rapus, 1996). Both mindreading and executive functions elicit activation in the frontal lobes (e.g., Frith & Frith, 1999), and both abilities are impaired in autism (Ozonoff, Pennington, & Rogers, 1991; Hughes, Russell, & Robbins, 1994). Furthermore, individual difference studies neatly demonstrate the tight relationship between children's executive function and their ability to pass the false belief task (Carlson & Moses, 2001; Carlson, Moses, & Breton, 2002; Carlson, Moses, & Claxton, 2004). Carlson, Moses, and Breton have demonstrated that children's performances on conflict inhibition tasks (inhibiting a pre-potent response whilst acting to a conflicting response) are a strong predictor of false belief task performance over and above other factors, such as age, working memory, intelligence measures, and delay inhibition task, where children inhibit impulsive responses. Another study (Carlson et al., 2004) has found that inhibitory control predicts children's scores on a mindreading task after controlling for age, receptive vocabulary, and planning ability. These findings suggest that executive function is necessary for engaging in flexible mindreading.

1.3.1.3 *The Role of Language*

Research on typically developing children (e.g., Astington & Jenkins, 1999) as well as deaf children (e.g., Peterson & Slaughter, 2006) and autistic children (e.g., Happé, 1995), has demonstrated that individuals' language abilities correlate with their performance on false belief tasks. Some authors have argued that the ability to understand others' intentions is essential for communication (Grice, 1957; Sperber & Wilson, 1995). Without comprehension of a speaker's communicative intent, one cannot make the correct pragmatic interpretation from the sentences themselves. Others have stressed the importance of semantics (Gopnik & Meltzoff, 1997), including lexical knowledge for single words such as 'think' and 'know', as well as discourse semantics (Milligan, Astington, & Dack, 2007) for understanding collections of words. Still others (e.g., de Villiers & de Villiers, 2000) have argued that understanding syntactic structure provides the basis for decomposing sentences such as ‘Maxi thinks the chocolate is in the blue cupboard’ into main clause ‘Maxi thinks’ and complement clause ‘the chocolate is in the blue cupboard’. de Villiers and de Villiers have argued that in order to reason about another’s false belief, one has to recognise that the embedded complement clause could be entirely false, whist the sentence still remains true.

A meta-analysis (Milligan et al., 2007) has revealed that children’s performance on mindreading tasks is associated with their performance on a range of language tests (general language, receptive vocabulary, semantics, syntax, and memory for complements). There is an ongoing debate about the precise roles of each linguistic component in the development of full-blown mindreading. However, there is little doubt that the development of children’s language is closely related to the development of mindreading.

1.3.1.4 *Summary of Developmental Studies*

In summary, studies that have focused on the development of mindreading revealed that there is a tight relationship between children’s social reasoning ability and the development of executive function along with different aspects of language. Therefore, a certain degree of executive function and language is necessary for the development of adult-like mindreading. However, without data from adults, it will remain unclear whether or not executive function and language are necessary for the online operation of mature mindreading. In the next section, I will review evidence from healthy adults as well as adults with acquired brain injury, which implicate the role of executive function and language in mature mindreading.

1.3.2 *Adults' Flexible Mindreading*

1.3.2.1 *Evidence from Brain-injured Patients*

Studies of patients with acquired brain injury are often informative in the identification of functions and mental operations associated with the injured brain regions. For the purposes of examining the role of executive function and language in flexible mindreading, evidence from brain-injured patients provides unique insights that studies with neurologically intact adults do not offer.

It is important to first of all identify the principal brain regions involved in mature mindreading. Brain imaging studies achieve this by comparing brain activation during belief reasoning tasks to physical reasoning tasks. Healthy adults show activation in the medial prefrontal cortex (mPFC) and temporal parietal junction (TPJ) when successfully attributing mental states (Gallagher et al., 2000; Fletcher et al., 1995; Saxe & Kanwisher, 2003). One hypothesis is that the frontal lobes are implicated in the holding of separate perspectives (e.g., Gallagher & Frith, 2003). Another hypothesis is that the frontal lobes are involved in resisting interference from one's own perspective (Ruby & Decety, 2003). There has been some debate about whether the TPJ is involved in mental state reasoning *per se* (e.g., Saxe & Kanwisher, 2003) or merely in the processing of lower level social stimuli, such as human movements (Allison, Puce, & McCarthy, 2000; Frith & Frith, 1999). Apperly, Samson, Chiavarino, and Humphreys (2004) examined 12 brain-damaged patients' performance on a reduced incidental task demand version of the false belief task as well as tasks on executive function and language. Their results implicated the left TPJ in false belief reasoning. All three patients with lesions in this brain region showed impairment on the false belief trials but not on working memory control trials within the same task. Interestingly, another four patients with frontal lesions showed impairment on both false belief trials and working memory control trials. This suggests that the errors they made on false belief trials may have been due to impairments in executive function[1].

1.3.2.1.1 *The Role of Executive Function*

Happé, Malhi, and Checkley (2001) conducted a study of patient P.B., who had brain damage of the orbito-frontal region. This patient showed impairments on mindreading tasks, which required attribution of others' thoughts, feelings, and intentions (story task and cartoon task). P.B. also exhibited impairment on a number of frontal lobe tasks (inhibition, set shifting, and generativity) and memory tasks (recognition memory task and verbal recall task). A link between the orbito-frontal cortex and flexible mindreading has also been demonstrated by Stone, Baron-Cohen, and Knight (1998). These authors revealed that patients with bilateral orbito-frontal lesions showed similar level of performance on social reasoning tasks to that of individuals with Asperger's syndrome. These patients passed tasks that involved representing the protagonist's belief (first-order belief reasoning) and representing the protagonist's belief about another protagonist's belief (second-order belief reasoning). However, they struggled with faux pas tasks, which involve identifying others' naïve mental states about something they should not have done but do not realise.

Furthermore, Samson, Apperly, Kathirgamanathan, and Humphreys (2005) studied a stroke patient W.B.A. with a lesion of the right frontal region. This region overlaps with the area associated with the ability to infer others' metal states (Vogeley et al., 2001). W.B.A. was tested on both a low inhibitory control and a high inhibitory control version of a non-verbal false belief task. The difference between the low inhibition and high inhibition conditions was achieved by manipulating the saliency of

[1] It is worth noting, however, some studies do point out a role for mPFC in mentalizing (e.g., Frith & Frith, 2006; Gallagher & Frith, 2003).

W.B.A.'s perspective. In the high inhibition condition, W.B.A acquired a true belief about an object's location as the trial sequence unfolded; hence he held a discrepant belief to the actor, who had a false belief. In the low inhibition condition, W.B.A. did not find out about the object's true location until after the actor indicated her belief. The actor did this at the end of the trial sequence by pointing to the false location (simultaneously revealing object's true location, i.e., the location not pointed to). Therefore, W.B.A. did not hold a contradictory belief to that of the actor prior to his own response. Samson et al. found that W.B.A.'s performance was only impaired in the high inhibition condition, where he committed egocentric errors. In contrast, in the low inhibition condition, W.B.A. consistently gave the correct responses. Furthermore, W.B.A.'s impaired performance on the control memory trials revealed that his brain lesion led to impaired executive abilities. This may have account for his impaired performance on social reasoning tasks when the incidental task demand was high. However, once the incidental task demand was reduced, Samson et al. showed that W.B.A.'s ability to infer the actor's belief was intact. Posing a potential counterexample, Bird, Castelli, Malik, Frith, and Husain (2004) studied patient G.T., who showed mostly intact mindreading ability despite the severe executive dysfunction. However, G.T. did not show impaired inhibitory control, which has been previously demonstrated to be closely related to children's false belief task performance (Carlson et al., 2004). Studies with brain-injured patients suggest that executive function, especially inhibitory control, plays an essential role in flexible mindreading. It appears to be particularly important for inhibiting one's own salient belief (Samson et al.).

1.3.2.1.2 *The Role of Language*

Varley and Siegal (2000) studied patient S.A., who had a lesion of the left temporal lobe that led to severe impairments in grammatical understanding. In spite of this, S.A. showed intact belief reasoning on unexpected content tasks. Interestingly, S.A.'s performance on the Wisconsin card-sorting task was also intact, demonstrating preserved executive function. A further investigation (Varley, Siegal, & Want, 2001) revealed S.A.'s normal performance on a modified picture mindreading task. In the same study, Varley et al. also examined patient M.R., who was severely aphasic and had impaired executive function as measured by the Wisconsin card sorting task. Nevertheless, M.R. also performed flawlessly on belief reasoning task in both true belief and false belief conditions. Apperly, Samson, Carroll, Hussain, and Humphreys (2006) studied patient P.H., who showed severe impairment on grammatical tests and executive tests. However, P.H. exhibited near perfect performance on non-verbal first-order and non-verbal second-order belief reasoning tasks. P.H. also demonstrated intact comprehension of semantics associated with mindreading. This dissociation between grammatical abilities and mindreading abilities indicates that the former is not necessary for the latter.

1.3.2.1.3 *Summary of Findings from Brain-Injured Patients*

The role of executive function and language in mature mindreading is in sharp contrast to the role they play in the development of mindreading. Although language might provide the initial structure for reasoning about others' minds, study of aphasic patients suggests that once mindreading abilities have been acquired, one does not have to rely on intact grammar to engage in flexible mindreading. Executive function, on the other hand, has been demonstrated to be essential for most social reasoning tasks that require inhibiting one's own salient belief. Various accounts have been offered in explanation of the role that executive function plays in the development of mindreading (e.g., Carlson & Moses, 2001; Frye, Zelazo, & Palfai, 1995). I will not discuss these accounts further here. Instead, in the following section, I will point out instances where neurologically intact adults show imperfect performance on mindreading tasks.

1.3.2.2 *Evidence from Neurologically Intact Adults*

1.3.2.2.1 *The Involvement of Executive Function*

Consistent with the evidence gathered from brain-injured patients, studies with neurologically intact adults also point towards a role for inhibition in mindreading. Bull, Phillips, and Conway (2008) put participants under a range of secondary executive function tasks while reasoning on two different mindreading tasks. One of the mindreading tasks was ‘Reading the Mind in the Eyes’ (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). In this task participants are required to match thoughts and feelings to arrays of eye stimuli. The second mindreading task was a story-based one (Happé, 1994), where participants answered questions about a protagonist’s mental states. Both types of mindreading tasks were matched with non-mindreading control tasks. The secondary executive function tasks included an inhibition task, a switching task, and an updating task. Bull et al. revealed that when participants performed a secondary executive function task along with a mindreading task, inhibition produced more dual tasking cost than switching and updating on the mindreading eyes task, but not on the control eyes task. Nonetheless, participants’ performance on both the mindreading story-based tasks and the control story-based tasks were found to be disrupted when they were loaded with all three executive function components. This suggests that attributing mental states to the eye stimuli requires inhibitory control, whereas the story-based tasks make high demands on a range executive function components.

In a similar vein, German and Hehman (2006) showed that reduced executive function may account for the compromised mindreading abilities observed in older adults. The authors systematically manipulated the complexity of belief-desire reasoning task by pairing true belief with approach desire, true belief with avoid desire, false belief with approach desire, and false belief with avoid desire. Evidence indicates that false belief problems are harder than true belief problems (e.g., Apperly, Warren, Andrews, Grant, & Todd, 2011). Furthermore, in a typical false belief task, the protagonist has the desire to approach the object. Children pass these false belief tasks at four-years of age. However, when the protagonist has the desire to avoid the object, children do not produce correct responses until six-years of age (Friedman & Leslie, 2004a). German and Hehman demonstrate that performance costs with increased task difficulty were greater in the older adults than in the younger adults. Furthermore, the authors' regression analysis showed that adults' processing speed and inhibitory control accounted for a significant proportion of variance in the performance. These two studies suggest that executive function, especially inhibition, plays a critical role in healthy adults' flexible mindreading.

1.3.2.2.2 *Flexible Mindreading in Typical Adults- Biases and Demands*

Despite the fact that most adults find false belief tasks trivially easy, it does not follow that adults infer others' mental states effortlessly. German and Hehman (2006) demonstrated that young healthy adults experience greater difficulty reasoning about others' false belief and avoid desire compared to their true belief and approach desire. Furthermore, when making predictions about the location where an ignorant protagonist seeks an object, adults tend to choose the location with a favourable outcome. In other words, rather than choosing at chance, their judgement is influenced by their own privileged knowledge (Birch & Bloom, 2007; Friedman & Leslie, 2004b). This result suggests that individuals are biased by extra information that is only accessible to themselves, this so called ‘curse of knowledge’ (Birch & Bloom, 2007; Mitchell, Robinson, Issacs, & Nye, 1996).

A similar type of bias has been shown in an online communication game, in which adults failed to account for another person’s discrepant perspective or knowledge (Apperly et al., 2010; Keysar, Lin, & Barr, 2003). In this particular game, adult participants are required to move objects around a 4 x 4 grid. Participants sit opposite a director. The grid is placed between the director and the participants. Some of the slots in the grid are occluded from the director’s point of view. This ensures that participants can see all the objects in all the slots whereas the director could only see some of the objects. The director gives specific instructions to the participants to move the objects around. For example: ‘move the small ball one slot down’. In the experimental condition, the director’s commands on occasion involve an ambiguous object (e.g., ‘mouse’ could be a computer mouse or a real mouse). In order to identify the correct referent, one must take the director’s perspective. The director’s command may also specify an object using an adjective that is relative to one’s perspective. For example, the director may make reference to a ‘large ball’ when three balls, (large, medium, and small) are visible to the participants, but only two of the balls (medium and small) are visible to the director. Therefore the ‘large ball’ in the director’s perspective is actually the medium ball in the participants’ perspective. In the baseline condition, the director’s commands only refer to objects that are unambiguous from both the director and the participant’s point of views. Keysar et al. found that participants looked much longer at objects occluded from the director’s perspective in the experimental condition than in the baseline condition. Moreover, a high proportion of participants also attempted to reach for the occluded objects.

By adapting Keysar et al.'s task into a computerised task, Apperly et al. (2010) demonstrate that adults' difficulty with the task is unlikely to be caused the incidental cognitive demand of the task. This study examined two main cognitive components of the task: switching between perspectives and discounting occluded slots. In the first set of experiments, participants were required to follow two directors; one shares participants' informed perspective, the other only has partial visual access to the grid. On trials where instructions came from the same director as the preceding trial, participants were not required to switch between perspectives. On trials where instructions came from a different director, participants had to switch to the new director's perspective in order to identify the correct referent. Findings reveal that participants performed identically on switch trials and non-switch trials. This suggests that the demand to switch between participants' own perspective and the director's perspective is unlikely to account for adults' biased performance. In a further experiment, Apperly et al. showed that participants performed more accurately when following instructions with a simple additional rule to discount objects in dark background slots compared to when following instructions of a visible director. This suggests that inferring others' perspective can be demanding even for adults who have the prerequisite cognitive ability. Another study using the same paradigm has shown that participants' judgements are even more biased towards their own perspective when given limited time to respond (Epley, Keysar, Van Boven, & Gilovich, 2004). This suggests that making inferences about others' perspectives is a demanding process. Qureshi (2009) carried out a test of the communication game, as well as a variety of executive function tests (*go/ no-go task* with letters and pictures, where participants responded stimuli from one category but not those from the other category; *stop-signal task*, in which the majority of the trials required participants respond to one of two stimuli apart from the stop trials, which included a tone closely following a stimulus to acts as a stop-signal; *cued recall task*, where participants recalled a word from one of two word lists whilst overcoming interference generated by a word from the list not being recalled; *Simon task*, in which participants responded to the left/ right orientation of an arrow whilst ignoring the left/ right spatial position of the arrow; *shape matching task*, where participants judged whether two coloured stimuli had matching shapes, whilst a third colour stimuli matching one of the two target colour stimuli were also present on half of the trials). Qureshi showed that the best-fit structural equation model includes a direct relationship between individuals' ability to inhibit responses in a go/ no-go task and error rate in the communication game. This suggests that one's ability to use information from a director's perspective is accounted by one's ability to hold in mind and select between responses. Studies described above provide consistent evidence that perspective differences are difficult to overcome and tap executive resources in healthy adults.

Furthermore, evidence suggests that merely holding another person's false beliefs in mind without making any inferences is also cognitively demanding. Apperly, Back, Samson, and France (2008) presented participants with two sentences; one made reference to a protagonist's belief about the colour of an object, and the other referred to the object's actual colour. A picture probe was presented after the two sentences. One part of the picture depicts two boxes containing two objects, one on a table, and one on a chair. The other part of the picture had a schematic figure that illustrated either the protagonist's belief about a particular object's colour or the actual colour of an object. Sometimes the protagonist held a false belief, in which case the two sentences provided conflicting information (e.g., Sentence 1: 'he thinks the object on the table is <u>yellow</u>'. Sentence 2: 'really, the object on the table is <u>red</u>').

Sometimes, the two sentences were unrelated to each other (e.g., Sentence 1: ‘he thinks the object on the table is red’. Sentence 2: ‘really, the object on the chair is yellow’). Participants judged whether the picture probe accurately represented the situation described in the sentences. A combined measurement of error rate and response times suggests that when the protagonist held a false belief rather than an unrelated belief, it was more difficult for participants to make judgements about both the protagonist’s belief and the reality. This was the case regardless of the amount of time given to participants to encode the information.

Apperly and colleagues have demonstrated that adults do not automatically make inferences about others’ beliefs (Apperly, Riggs, Simpson, Chiavarioni, & Samson, 2006; Back & Apperly, 2010). Participants were presented with a non-verbal video version of the false-belief task, in which an object was moved from one hidden location to another when an actor was either present or absent. Therefore the actor either held a true belief or a false belief about the object’s location. Participants were presented with either a belief probe or a reality probe. The belief probe describes the actor’s belief about the object’s location; the reality probe describes the actual location of the object. Participants were slower to respond to the belief probe than the reality probe when no explicit instruction was given to encode the actor’s belief. However, when participants were given explicit instructions to track the actor’s belief, they were equally quickly to respond to the belief probe and the reality probe. This suggests that adults do not automatically ascribe others’ mental states regardless of it being a false belief (Apperly et al., 2006) or a true belief (Back & Apperly, 2010). Interestingly, Cohen and German (2009) provided evidence to show that adult participants encode an actor’s belief without overt instructions to do so. However, this information was only maintained for a very short period of time. When there was a three-second delay between the onset of the actor's belief and the probe to recall the actor's belief, this information could be easily retrieved. However, when the delay was as long as 23 seconds, participants suffered additional processing cost for recovering the information about the actor's belief. These studies suggest that although adults can encode belief information automatically, this information also decays quickly unless overt instruction was given to maintain it.

1.3.2.2.3 *Summary of Typical Adults' Flexible Mindreading*

These findings suggest that even mature mindreading abilities produce biases and require effort when making inferences about others' perspectives. More specifically, executive resources are still necessary for engaging in flexible mindreading (i.e., overcoming one's own salient perspective, holding in mind others' false beliefs, and making inferences about others' true and false beliefs).

However, taking this as an accurate picture of adults' ability to mindread raises important questions about how online social interaction is supported. Observations of day-to-day social interaction suggest that we have a less flexible but highly efficient alternative mean of supporting such interactions. In the following section, I will review evidence that suggests that efficient mindreading may operate independently of flexible mindreading.

1.4 Efficient Mindreading

Recent studies have demonstrated that infants as young as 15 months of age can pass certain types of mindreading-like tasks. Given that infants at this age have limited language abilities and immature executive function suggests that neither language nor executive function is 'underwriting' behaviours on certain mindreading-esque tasks. Furthermore there is evidence from adults to demonstrate automatised processes as diverse as stereotyping, visual perspective-taking, gaze cueing, and task-sharing. I will review these processes in turn in later sections.

1.4.1 *Infants' Mindreading*

Until recent years, developmental studies have largely focused on declarative inferences in mindreading tasks amongst children three to five years of age. However, new evidence focusing on behaviour indices suggests that infants have some understanding of others' minds (e.g., Onishi & Baillargeon, 2005; Southgate, Senju, & Csibra, 2007; Surian, Caldi, & Sperber, 2007). Onishi and Baillargeon demonstrated that at 15 months of age, infants correctly anticipate an actor's behaviour based on her beliefs, both true beliefs and false beliefs. Onishi and Baillargeon tested this using a violation of expectation paradigm, which measure infants' surprise at an event. Infants were familiarised through repeated presentation with a set sequence of events: an actor hid a toy in one of two boxes and then subsequently reached for the box where she hid the toy. During the test phase, infants were presented with sequences where the actor either produced an action that was consistent with her knowledge about the toy's location or an action that violated her knowledge about the toy's location. When infants saw the actor performing an action that violated her knowledge, they looked longer at the event, indicating that infants found this novel relative to the event to which they had been habituated. This finding demonstrates that infants are sensitive to others' belief-like mental states and their corresponding actions. It has also been shown that 13-month-old infants are also sensitive to the relationship between agents' action-goals and perceptual knowledge about objects (Surian et al., 2007). Furthermore, by 25 months of age, infants are able to accurately anticipate others' behaviours according to their false beliefs (Southgate et al., 2007).

Intriguingly, these infancy studies employed non-verbal tasks that closely resembled sequences from false belief tasks. However, they placed very little demand on executive function and language. Therefore, one argument against the operation of a separate system for efficient mindreading is that infants may be capable of engaging in adult-like social reasoning provided that the incidental task demand of language and executive function were reduced (e.g., Onishi & Baillargeon, 2005). However, the distinction between flexible mindreading and efficient mindreading is not exhausted by the differential demands the systems place on language and executive resources. As described earlier, there may be further distinctions amongst the fundamental characteristics featured in the flexible and efficient systems. I will review further evidence to show that efficient mindreading (as well as for the neighbouring social cognitive processes, such as stereotyping and gaze attentional cueing) is not subject to conscious control and is often stimulus-driven. Therefore it is unlikely that efficient mindreading and flexible mindreading belong to the same system.

1.4.2 *Adults' Efficient Mindreading*

In this section, I will illustrate three types of operations in the social cognition literature that are similar to efficient mindreading and have already been shown to be automatic. Automaticity has been defined as processes that are stimulus-driven, cannot be interrupted once triggered, operate effortlessly, and run in parallel with the non-automatic processes (Payne & Bishara, 2009). In this thesis, the term 'automatic' will be used in a broader sense to include processes that might feature key characteristics of automaticity without necessarily meeting all the defining criteria.

1.4.2.1 *Automatic Stereotyping*

A number of studies have demonstrated that person perception operates automatically and unconsciously (Perdue & Gurtman, 1990; Winter & Uleman, 1984).

Evidence suggests that stereotypes may be activated by the mere presence of features that are associated with the stereotyped group (Devine, 1989). Various phenomena, such as biased impression formation (Pratto & Bargh, 1991), distorted memory (Predue & Gurtman, 1990; Winter & Uleman, 1984), and speeded or delayed responses (Kawakami, Young, & Dovidio, 2002) have been taken as evidence for automatic stereotyping. Furthermore, individuals were more susceptible to stereotype bias when they could only allocate little attention to person perception (Pratto & Bargh, 1991), suggesting that stereotyping indeed operates with minimal attentional resources. Finally, studies have demonstrated reduced or modulated involvement of stereotyping in person perception when individuals undergo extensive training (Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000), intention implantation (Mendoza, Gollwitzer, & Amodio, 2010), or encounter someone they are very familiar with (Quinn, Mason, & Macrae, 2010). However, these studies nevertheless observed some level of stereotyping, suggesting that the stereotypic activation operates independently of these modulatory factors. In sum, it is clear that the activation of stereotype is stimulus-driven, demands little attentional resources, and is subject to minimal conscious control.

1.4.2.2 *Shared Action Representation*

Sebanz, Knoblich, and Prinz (2003) studied action representation in a variation of the go/ no-go task, in which the relevant colour responses were either paired with a spatially compatible or incompatible cue. A computer screen displayed a hand stimulus, which either pointed to the left or the right. The index finger of the hand had a ring coloured red or green. Participants responded to one ring colour by pressing a response button on their right, the other ring colour by pressing a button on their left. The irrelevant stimuli was a finger-pointing stimulus that either pointed to the right or left. Participants responded slower to colours when the spatial cues were incompatible with the location of the colour responses, and faster when the spatial cues were compatible with the location of the colour responses. Interestingly, when two participants shared a task, each responsible for responding to one of the two colour responses, the spatial compatibility effect was equal to that of having only one participant covering both colour responses. This was only the case when another person shared the task with the participants but not when the participants performed half of the task by themselves. Furthermore, Sebanz et al. showed that participants appeared to form shared action representation even when there was no visual or audio feedback from their partner during the task. Moreover, the presence of another person who did not actually share the task was not by itself sufficient for the formation of shared action representation. A good strategy of the task is to completely ignore the other person's role, and to solely concentrate on one's own task. Therefore, the very fact that participants appear to represent the entire task even when it is not beneficial suggests that this process is stimulus-driven and lacks flexible control.

1.4.2.3 *Automatic Gaze Attentional Cueing*

A number of studies show that participants respond slower in a simple target detection and identification task when uninformative gaze directions were incompatible with target locations (e.g., Friesen & Kingstone, 1998; Langton & Bruce, 1999). This demonstrates that participants automatically orient their attention towards agents' gaze directions, despite the fact that the gaze information is completely irrelevant and disruptive to the current task. In a similar vein, evidence also shows that when participants viewed a natural scene consisting of an agent and a number of objects, the objects closer to the agent's uninformative gaze directions are detected more quickly (Langton, O'Donnell, Riby, & Ballantyne, 2006). This suggests that participants' attention is drawn to actors' gaze directions, hence objects located nearby are subject to reduced levels of change blindness. Furthermore, participants' preference ratings for objects are higher when the object is looked at by an uninformative face than when it is not (Bayliss, Paul, Cannon, & Tipper, 2006). The preference ratings for gazed-upon objects are modulated by both positive and negative emotional expressions of the uninformative face. Importantly, modulation of preference ratings by emotional expressions does not occur when gaze is directed away from the object (Bayliss, Frischen, Fenske, & Tipper, 2007). These studies consistently demonstrate that participants automatically allocate attention to others' gaze direction even when gaze information is not relevant to the current task. This suggests that adults' processing of gaze direction is likely to be stimulus-driven and outside of conscious control.

1.4.2.4 *Automatic Visual Perspective-Computation*

Evidence of automatic visual perspective-computation has been provided by Samson, Apperly, Braithwaite, Andrews, and Bodley Scott (2010). In this study, participants were cued to judge either their own perspective or an avatar's perspective at the start of each trial. This was followed by a number probe, which either correctly or incorrectly describes a number of dots within the cued perspective. These probes were followed by a display picture that contains an avatar facing a number of dots, with either some more dots or nothing behind him. Participants either shared the same visual perspective with that of the avatar or held a different perspective from that of the avatar (for more detailed methodology, see Section 2.2). Samson et al. found that making speeded judgements about the avatar's perspectives is more difficult for the participants when they held a different perspective from the avatar rather than when they held the same perspective. This indicates that participants' own perspective generated interference in their simple judgements about the avatar's perspective. This finding is in line with the social reasoning tasks reviewed in Section 1.3.2.2.2 (e.g., Keysar et al., 2003), which showed that adults' own salient perspectives often affect their judgements about others' perspectives, and that overcoming such interference demands cognitive resources.

Additionally, when participants judged their own perspectives, they suffered interference from the avatar's discrepant perspective. The interference is manifested in a higher processing cost. This is likely to be due to automatic computation of the avatar's perspective. Interestingly, the same pattern of results was found even when participants were never required to make explicit judgements about the avatar's perspective. These findings suggest that the implicit computation of the avatar's perspectives is likely to occur prior to participants' responding to confirm their own perspectives. Moreover, participants' automatic taking of the avatar's perspective cannot be inhibited even when given obvious opportunities to do so (i.e., in an experiment where they only ever judged their own perspective).

Qureshi, Apperly, and Samson (2010) further examined these effects by introducing a secondary task that draws on executive function resources. The rationale was that if the computation for another's perspective is effortful and demands cognitive resources, then once those resources have been taken up by a secondary task, the computation of another's perspective should be interrupted. Their findings did not lend support to the view. Rather, participants suffered more interference from the avatar's discrepant perspective when judging their own perspective. There was also an exaggerated interference from participants' own perspective when they made judgements about the avatar's perspective. This suggests that the computation of others’ perspectives is indeed automatic and cognitively effortless. Nonetheless, managing and selecting computed perspectives does demand executive resources.

1.4.2.5 *Summary of Adults’ Efficient Mindreading*

Studies reviewed in this section provide consistent evidence that efficient mindreading characterises of distinct features that are not present in flexible mindreading. Firstly, adults’ automatic visual perspective-computation reveals that processing of others’ mental states requires only minimal executive control. Moreover, the infant work suggests that this capacity may be early-developing. The common features of these efficient mindreading processes are consistent with those of the prominent two-system accounts in the social cognition literature. These operations have been shown to activate automatically, be stimulus- or context- driven, and appear to be immune from conscious control. Moreover, the literature suggests that these features are absent in flexible mindreading, which instead appears to be much more controllable and allows one to adopt different responses according varying social contexts.

1.5 Focus of Current Experimental Work

The current experimental work is heavily influenced by the work on efficient mindreading reviewed above. The main aim of this thesis is to examine how automatic processing of others’ visual perspectives and eye gaze could affect our perceptual computation and encoding of the social world.

Chapters 2, 3, and 4 provide discussion and examination on the scope of automatic visual perspective-computation. Using Samson et al.’s (2010) visual perspective-taking paradigm, I examined the effect of further processing load on automatic visual perspective-computation. Chapters 5, 6, and 7 discuss and examine the possibility that observing others’ eye gaze might affect individuals’ visual working memory encoding. This was accomplished by manipulating agents' gaze towards objects.

CHAPTER 2 INTRODUCTION TO AUTOMATIC LEVEL-1 VISUAL PERSPECTIVE-TAKING

2.1 *Level-1 Visual Perspective-Taking*

Visual perspective-taking concerns one's ability to predict the visual scene as it appears from another agent's point of view. This operation can be divided into two levels according to the aspects of information required. Level-1 visual perspective-taking concerns *what* an agent has visual access to, this could be done by tracing the agent's line of sight (Michelon & Zacks, 2006). Level-2 visual perspective-taking concerns *how* an object appears to an agent from different viewing points (Flavell, Everett, Croft, & Flavell, 1981). Some authors have argued that individuals make such predictions by rotating their egocentric reference frames (Michelon & Zacks, 2006) or by employing a visual-matching strategy for low angles (Kessler & Thomson, 2010).

Level-1 visual perspective-taking ability emerges early in development. Sodian, Thoermer, and Metz (2007) demonstrated that 14-month-olds are sensitive to others' visual perspectives. Infants were familiarised to a sequence where the actor announced that her goal was to look for her toy. Subsequently, she reached for one of the two particular toys that were present. During the test phase, the actor either reached for her old-goal toy or reached for the other toy. Crucially, the authors found that infants only showed longer looking time when the actor could see her old-goal toy but still reached for the other toy. When the actor could not see her old-goal toy and reached for the only toy visually available to her, infants did not look longer at the event. This suggests that 14-month-old infants are influenced by an actor's visual perspectives when observing her goal-oriented actions.

By 24 months of age, infants can act upon a request whilst taking the speaker's visual access into consideration (Moll & Tomasello, 2006). In this study, a speaker requests that the infant give her an object. Twenty-four-month-old infants responded to the request by giving the speaker the object that could be seen by both herself and the speaker, rather than a second object, which could only be seen by the infants. O'Neill (1996) showed that at 24 months of age, infants are also sensitive to mothers' past visual experiences. They produced more helpful gestures when directing the mothers' attention to a wanted object when the mothers did not witness where the objects were placed. This suggests that despite immature executive function, infants are nevertheless sensitive to discrepancies between their own visual perspectives and those of others.

2.2 *Adults' Automatic Computation of Level-1 Visual Perspective*

As briefly described in Chapter 1, the work of Samson et al. (2010) indicates that adults automatically compute both their own and the avatar's visual perspectives even when it is unnecessary for the task and disadvantageous to do so. In this study, participants were presented with display pictures containing a room, an avatar, and a number of dots on the walls of a room. The avatar faces one wall and has his back to another wall. Dots appearing on the wall in front of the avatar are within both the participants' and the avatar's visual fields. However, dots which appear on the wall behind the avatar are only within the participants' visual field; they are not within the avatar's visual field. On consistent trials, all the dots appear in front of the avatar. On inconsistent trials, one or more dots appear behind the avatar. On any given trial, participants judged either the avatar's perspective or their own perspective. At the beginning of each trial, they were cued with the perspective they should judge. This was followed by a number probe, which reported the number of dots visible to either the avatar or the participants. Finally, they saw the display picture revealing the perspective contents. Participants had to indicate whether or not the perspective cue and the number probe matched the display picture. If participants were automatically computing the perspective of the avatar, then one would expect this to interfere with their own perspective judgements when they held a different perspective to the avatar. Samson et al. found when participants made judgements about their own perspective, they suffered additional processing cost in the inconsistent condition compared to the consistent condition. This showed that participants automatically computed the perspective of the avatar (so called altercentric intrusion). Likewise, when participants made judgements about the avatar's perspective, they also suffered additional processing cost in the inconsistent condition compared to the consistent condition (so called egocentric intrusion).

These findings demonstrate that adults automatically compute others' Level-1 visual perspectives even when it is unnecessary and disadvantageous for task performance. Moreover, Samson et al. (2010) found altercentric intrusion even when participants solely judged their own perspectives throughout the entire experiment. As described in Chapter 1, Qureshi et al. (2010) showed that participants' computation of their own and the avatar's visual perspectives were not interrupted by a concurrent executive task. This indicates that visual perspective-computations are not heavily reliant on executive resources. Taken together, these studies suggest that computations of one's own and others' Level-1 visual perspectives is an automatic, efficient, and effortless process.

2.3 *Rationale for Chapter 3 and Chapter 4*

In Chapters 3 and 4, I will examine the scope of the effects demonstrated by Samson et al. (2010). In Chapter 3, I will first investigate the amount of computational capacity of automatic visual perspective-taking by implementing varied enumeration load. Enumerative operations can be divided into two types. The first is limited to small number sets and requires little effort. The second concerns large numbers but demands more effort (e.g., Trick & Pylyshyn, 1994). One hypothesis which I shall test is that automatic visual perspective-computation operates solely within the scope of efficient enumeration, and not when the processing demand for enumeration is high. In Chapter 4, I will examine whether automatic visual perspective-taking allows participants to selectively process visual information. I will investigate the extent that relevant and irrelevant information is respectively selected and ignored whilst participants are automatically computing visual perspectives.

These investigations advance our understanding of the scope of automatic visual perspective-computation. They have both theoretical and practical importance. Theoretically, operations of efficient mindreading should also be characterised by limited cognitive flexibility (Apperly & Butterfill, 2009). Evidence from the literature suggests that efficient mindreading is early-developing and independent of executive resources. However, the limitation of these operations has as yet been demonstrated. The aim of Chapters 3 and 4 is to address this issue directly by examining the limitations as well as flexibilities within efficient mindreading. From a pragmatic point of view, it is important to understand whether or not automatic visual perspective-computation operates with the flexibility to select relevant information. In a social world, one often encounters situations where one is required to extract relevant visual information in order to make veridical inferences about the minds of others. In Chapter 4, I will experimentally investigate whether efficient mindreading allows for the selective processing of visual information.

CHAPTER 3 AUTOMATIC LEVEL-1 VISUAL PERSPECTIVE-TAKING: BEYOND THE LIMITS OF SUBITIZATION

3.1 Introduction

As described in Chapter 2, Samson et al. (2010) demonstrated that adults automatically compute an avatar's Level-1 visual perspective even when it is disadvantageous for task performance. Greater processing costs were observed when participants and avatar held different visual perspectives than when participants and avatar held the same visual perspective. In Samson et al.'s study, the enumerative cost for computing perspective contents was designed to be constant. This was achieved by ensuring the number of dots displayed never exceeded three. The enumeration literature shows that adults can enumerate sets of four or fewer items rapidly and effortlessly. In contrast, the enumeration of items beyond four is slower and effortful. The efficient visual discrimination of small number sets is known as 'subitizing' (e.g., Kaufman, Lord, Reese, & Volkmann, 1949; Trick & Pylyshyn, 1994). Participants' response times for one to four items reveal a small increase of between 40 to 100 ms processing time per additional item, whereas their response time for enumerating number of items beyond four elicits 250 to 350 ms additional processing time per item. This suggests that counting large numbers demands greater processing cost compared to subitizing small numbers.

In the current chapter, I report three experiments that examine the limits of automatic visual perspective-computation. I do so by further developing Samson et al.'s (2010) visual perspective-taking paradigm. Experiment 1 examined automatic visual perspective-computations under increased enumerative demands. Experiments 2 and 3 further investigated the precision of participants' perspective-computation and their representation for the avatar's perspective content under varied enumerative load.

3.2 Experiment 1

In the current experiment, a set of larger perspective contents was employed to produce higher enumerative demands on participants' perspective-computation. This allows examination of the limits of automatic visual perspective-computation. If automatic visual perspective-computation only operates within the boundaries of subitization, then once participants must effortfully count the perspective contents, they should cease to automatically compute perspectives. However, if automatic visual perspective-computation is equipped for enumeration beyond the boundaries of subitization, then participants should automatically compute both perspectives even when the perspective contents had to be counted.

3.2.1 *Method*

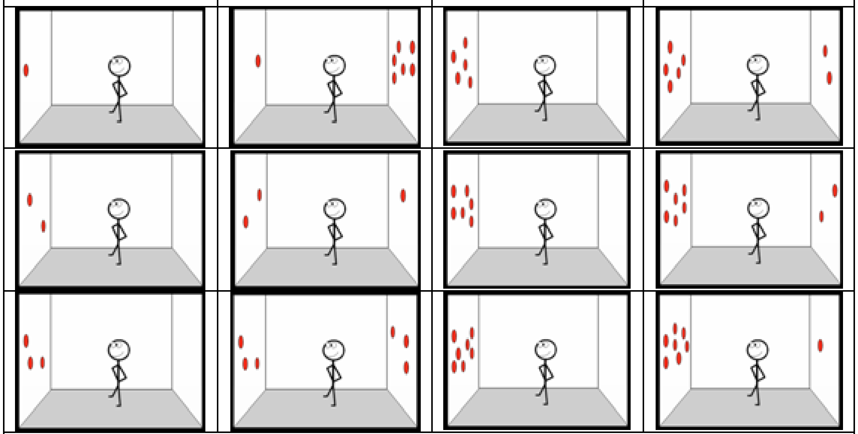
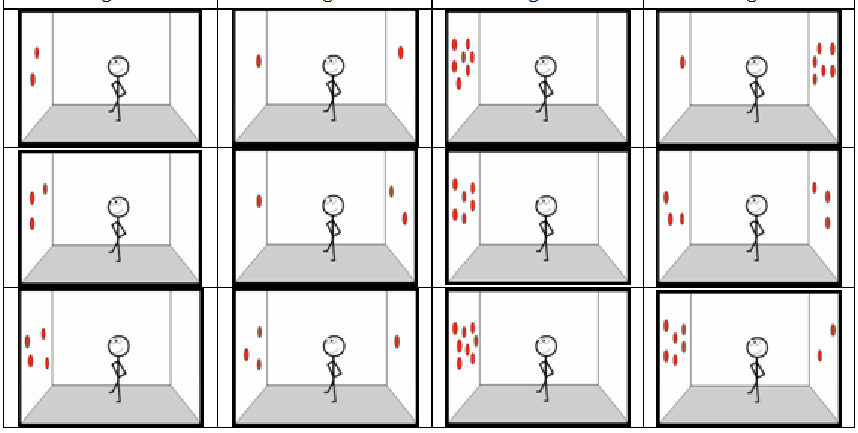
3.2.1.1 *Participants*

Twenty students (15 female, average age 20.5 years, age range 18 to 28) from the University of Birmingham participated in this study in return for study credits. Participants who failed to perform above chance in either the self or the other condition were excluded from analysis. One participant's data were replaced due to not performing significantly above chance in the other condition.

3.2.1.2 *Design and Procedure*

A 2 x 2 x 2 within subject design was constructed (see Table 3.1) with congruency (congruent, incongruent), perspective (other, self), and number size (large-number, small-number; small-number referred to one to four dots, which are numbers that could be subitized; large-number referred to five to eight dots, which are likely to be counted) as factors. Each condition occurred equally frequent throughout the experiment.

Table 3.1 Examples of displays in Experiment 1: examples were taken from each condition to show possible combinations of dots across the two walls. Each set of dots had further three types of pattern to avoid participants applying strategies induced by pattern recognition (Wolters, van Kempen, & Wijlhuizen, 1987).

Other			
Small-number		Large-number	
Congruent	Incongruent	Congruent	Incongruent
			
Self			
Small-number		Large-number	
Congruent	Incongruent	Congruent	Incongruent
			

Each trial sequence began with a perspective cue, either ‘He’ or ‘You’, presented on a screen for 750 ms. When ‘He’ was presented, participants made explicit judgements about the number of dots the avatar saw; when ‘You’ appeared, participants judged the number of dots they saw on the screen. This was followed by a number probe representing the perspective contents. The number probe remained on the screen for 750 ms. At the end of the trial sequence, participants saw a display picture depicting an avatar standing in the centre of a room with various numbers of red dots on the walls. The dots could be on either the participants' left or right hand side, or on both sides of the walls. The total number of the dots ranged from one to eight (see Figure 3.1).

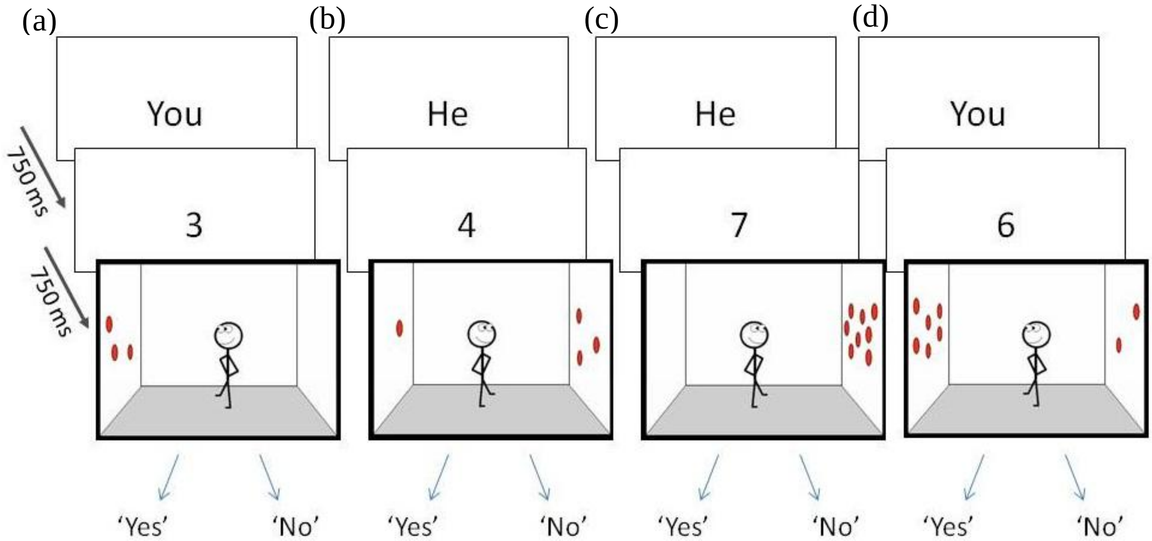


Figure 3.1 Examples of trials from Experiment 1: (a) small-number congruent condition, (b) small-number incongruent condition, (c) large-number congruent condition, (d) large-number incongruent condition. Experiments 2 to 5 employed identical trial sequence to Experiment 1, without necessarily including all the above conditions.

Participants were instructed to judge whether the perspective cues and number probes correctly described the final display pictures. Participants made yes/no judgements by clicking on a computer mouse, pressing the left button for 'yes' responses, the right button for 'no' responses. All participants were instructed to respond as quickly and accurately as possible. All trial types were mixed within blocks in a pseudo-random fashion so that participants never encountered three continuous trials from the same condition. There were 202 trials in total: one practice block of 10 trials, with four testing blocks of 48 trials. Four running versions of the experiment were generated by rotating four testing blocks to prevent order effects (Cozby, 2009). The experiment was presented using E-Prime (Schneider, Eschman, & Zuccolotto, 2002a; 2002b).

3.2.1.3 *Predictions*

Notwithstanding the inclusion of a large-number condition in the experiment, it is predicted that participants will implicitly process their own and the avatar's *small-number* perspectives as demonstrated in Samson et al. (2010). That is, effects of both egocentric intrusion and altercentric intrusion in the small-number condition are predicted. Furthermore, there are two possibilities with regards to the limits of automatic visual perspective-computation. One is that automatic visual perspective-computation may comprise the capacity for both fast and slow enumerations. If this is correct, then participants should automatically process their own and the avatar's perspectives regardless of the size of the perspective contents. That is, they should show effects of egocentric intrusion and altercentric intrusion in both the small-number and the large-number conditions. In contrast, automatic perspective-computation may lack cognitive flexibility for enumerating numbers beyond subitization. If this is correct, once the enumerative load on perspective contents exceeds the boundaries of subitizing, then participants should show no automatic computation of perspectives. In that case, effects of egocentric intrusion and altercentric intrusion in the large-number conditions should not be observed.

3.2.2 *Results*

In the current experiment, only the 'yes' response trials were included for analysis. This was because the number probes for the 'no' response trials in the incongruent condition always corresponded to the perspective content that was not to be judged on that trial. For the congruent condition, the number probes on ‘no’ response trials was by definition a number that did not correspond to the number of dots on the wall. Since the number probes in the two conditions were not designed to match each other, the comparisons and interpretations would provide little information. Analysis of response time data only included the trials to which participants responded correctly, 3.82% of the data were removed due to erroneous responses. Data points that were more than two standard deviations away from the overall mean were removed, 4.80% of the correctly responded data were removed due to slow response time.

The current design ensured that participants saw identical displays for trials in which they judged their own perspective and trials in which they judged the avatar’s perspective. All displays were restricted to contain a maximum of eight dots in total. As a consequence, in the incongruent condition there were more available combinations of dots when the avatar held a small perspective content compared with a large perspective content. For example, Figure 3.2A shows that when the avatar sees two dots, the wall behind him could display any number of dots between one and six, generating six different combinations. When the avatar sees a large number, for example, in Figure 3.2B the avatar sees six dots, the wall behind him could only display one or two dots, generating two combinations.

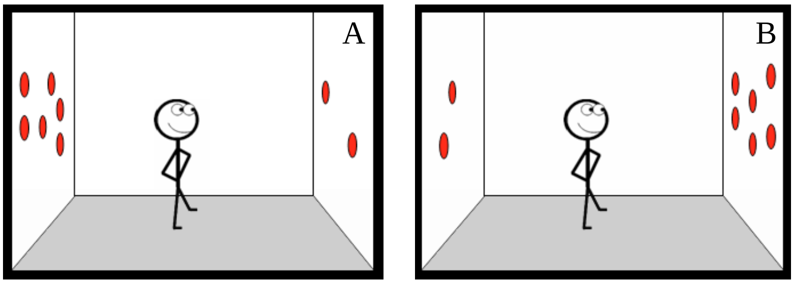


Figure 3.2 Demonstration of the number of available combinations for the avatar's small and large number perspectives.

For the same reason described above, there were more available combinations of dots when the participants' perspective contents were larger. For example, when participants see eight dots, the avatar could see any number of dots between one and seven. When participants see three dots, the avatar could only see one or two dots. However, this uneven range of numerosities was only the case for the incongruent conditions. Since the congruent conditions only display dots in front of the avatar, these conditions were not subject to the same limitation as the incongruent conditions. Therefore, the congruent conditions contained an even range of numerosities. To ensure any differences observed between the congruent and incongruent conditions were purely a consequence of the perspective-computation, and not enumeration differences, each condition was weighted before the analysis. The weighting procedure was applied so that the range of numerosities in the incongruent condition matched with that of the congruent condition[2]. That is, all set sizes of perspective contents contributed equally to the overall processing cost. To achieve this, the mean processing cost for each set size of perspective contents was calculated then averaged with the other set sizes from the same condition.

[2] A counter design would be to allow the set sizes contained in the congruent condition to match the uneven range of numerosities in the incongruent condition. This design was employed in Experiment 3 to demonstrate replication of the current results without applying the weighting procedure.

Analysis was conducted using processing costs, which was calculated by dividing each participant's response time by the proportion of trials to which they responded correctly in each condition. This score provided a concise summary of both speed and accuracy (for separate condition means of response time and error rate, see Appendix A). The self condition and other condition were analysed separately, as the current investigation included no *a priori* hypothesis regarding differences in the judgements for these two perspectives *per se*.

3.2.2.1 *Other Condition*

A 2 x 2 repeated-measure analysis of variance (ANOVA) was conducted with congruency (congruent & incongruent) and number size (small-number & large-number) as the main contrasts. There were main effects of congruency, $F(1, 19) = 38.18$, $p < .001$, $\eta_p^2 = .668$, which shows that the incongruent conditions required greater processing cost compared to the congruent conditions (*M*s = 1586.23 and 825.95, respectively), and number size, $F(1, 19) = 145.74$, $p < .001$, $\eta_p^2 = .885$, which confirms that large-numbers demand greater processing cost to compute than small-numbers (*M*s = 1287.02 and 1125.16, respectively). There was no interaction between congruency and number size, $F(1, 19) = 2.68$, $p = .118$, $\eta_p^2 = .124$. Planned comparisons revealed a significant effect of egocentric intrusion in the small-number condition, $t(19) = 9.02$, $p < .001$ (other-small-number-congruent = 720.48, other-small-number-incongruent = 931.42), and a significant effect of egocentric intrusion in the large-number condition, $t(19) = 2.20$, $p = .040$ (other-large-number-congruent = 1529.84, other-large-number-incongruent = 1642.62; see Figure 3.3). Additionally, congruency effects were calculated by subtracting the processing cost in the congruent conditions from the matching incongruent conditions. Comparison between the small-number and the large-number conditions showed no difference between the congruency effect observed in these conditions, $t(19) = 1.64$, p = .118 (other-large-number = 112.79, other-small-number = 210.95). This suggests that egocentrism is unlikely to be a function of enumerative demand.

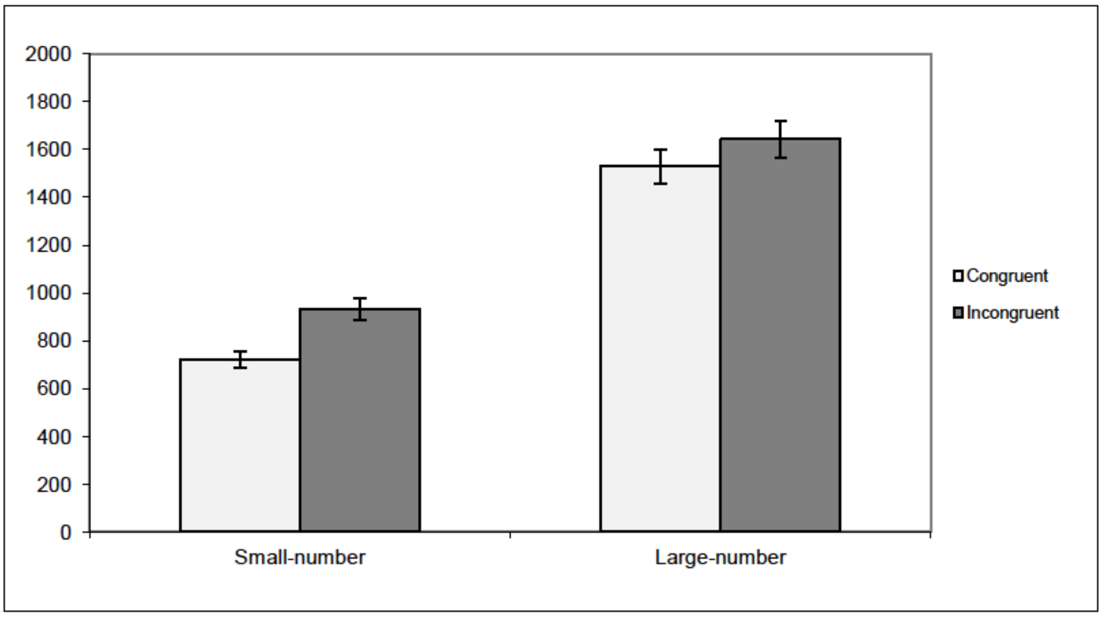


Figure 3.3 Processing cost for Experiment 1 'yes' response from the other condition. Error bars represent standard errors.

3.2.2.2 *Self Condition*

A 2 x 2 repeated-measures ANOVA was carried out with congruency (congruent & incongruent) and number size (small-number & large-number) as the main comparisons. There were main effects of congruency, $F(1, 19) = 20.40$, $p < .001$, $\eta_p^2 = .518$, which shows that the incongruent conditions required greater processing cost compared to the congruent conditions (*M*s = 1699.48 and 858.63, respectively), and number size, $F(1, 19) = 309.55$, $p < .001$, $\eta_p^2 = .942$, which confirms that large numbers demand greater processing cost to compute than small numbers (*M*s = 1365.02 and 1193.08, respectively). There was no significant interaction between congruency and number size, $F(1, 19) = 0.02$, $p = .903$, $\eta_p^2 = .001$. Planned comparisons showed a significant effect of small-number altercentric intrusion, $t(19) = 2.52$, $p = .021$ (self-small-number-congruent = 769.86, self-small-number-incongruent = 947.39), accompanied by a significant effect of large-number altercentric intrusion, $t(19) = 3.68$, $p = .002$ (self-large-number-congruent = 1616.31, self-large-number-incongruent = 1782.65; see Figure 3.4). Additionally, congruency effects were calculated by subtracting the processing cost in the congruent conditions from the matching incongruent conditions. Comparison between the small-number and the large-number conditions showed no difference between the congruency effect observed in these conditions, $t(19) = 0.12$, p = .903 (self-large-number = 166.35, self-small-number = 177.54). This suggests that altercentric intrusion is unlikely to be a function of enumerative demand.

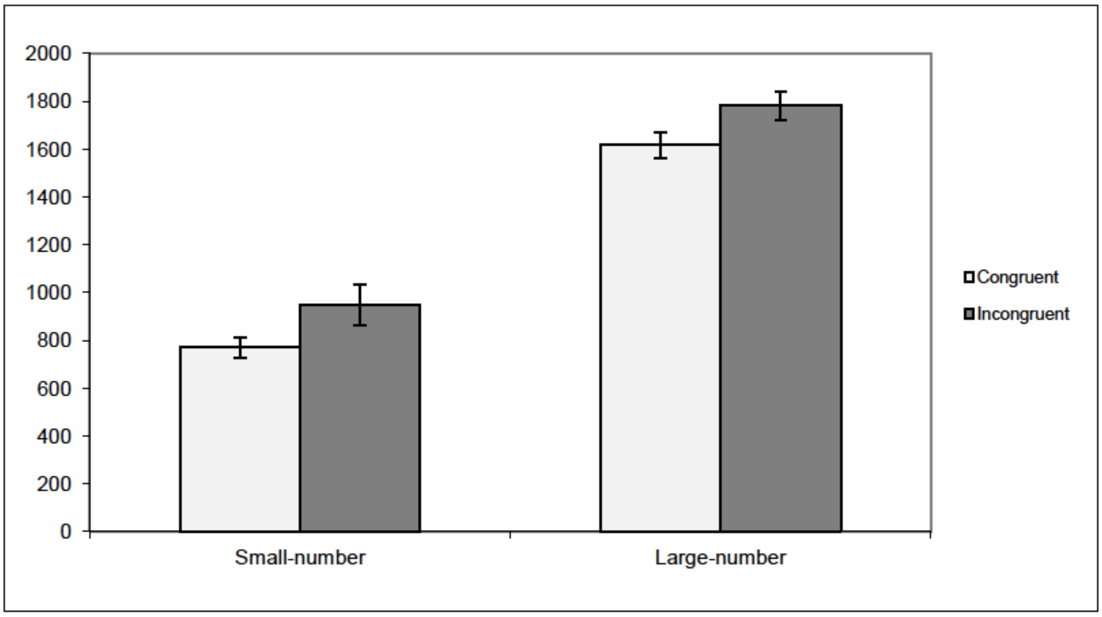


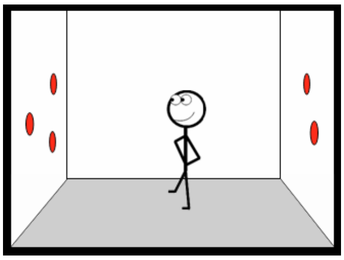
Figure 3.4 Processing cost for Experiment 1 from the self condition. Error bars represent standard errors.

3.2.3 *Discussion*

Results from the small-number condition replicated Samson et al.'s (2010) findings, demonstrating effects of egocentric intrusion and altercentric intrusion. When participants made explicit judgements about the avatar's perspective, they showed greater processing cost in the incongruent condition compared to the congruent condition. This indicated that participants' own discrepant perspective interfered with their judgements about the avatar's perspective. When participants made explicit judgements about their own perspective, a larger processing cost was also observed in the incongruent condition than in the congruent condition. This suggests that participants implicitly compute an avatar's perspective; therefore, when the avatar holds a discrepant perspective, it causes interference with participants' judgements about their own perspective. The effects of egocentric intrusion and altercentric intrusion indicate that participants automatically process both their own perspective and an avatar's perspective even when it is disadvantageous for task performance. Interestingly, effects of both egocentric intrusion and altercentric intrusion were also found in the large-number condition. This finding indicates that participants do not only automatically process both perspectives when the perspective contents are subitizable, but also when the perspective contents require greater enumerative processing. The current findings suggest that automatic Level-1 visual perspective-computation is not restricted to operate within the limits of subitization (Kaufman et al., 1949; Trick & Pylyshyn, 1994).

The current experiment and that of Samson et al. (2010) indicate that participants automatically compute an avatar's visual perspective even when it is unnecessary to do so and disadvantageous to task performance. In order for participants to produce a correct 'yes' response on trials about their own perspective, they have to sum up the two subsets of dots in the display; one being solely the avatar's perspective content. Present evidence suggests that participants' responses are influenced by the avatar's discrepant perspective contents. However, it is not clear whether participants represent the subset seen by the avatar any differently from the subset not seen by the avatar. Furthermore, there has been no evidence to indicate whether the interference participants suffer comes from a precise number that corresponds to the avatar's perspective or just an impression that the avatar sees a different number of dots. There are three types of possible representation of the avatar's perspective. Firstly, a *precise enumeration plus binding*: the participants suffer interference from a precisely represented number of dots that is the avatar's perspective content (not necessarily in the form of an actual numerical figure, but an exact number of dots). For example, when there are three dots in front of the avatar with two dots behind him, this hypothesis predicts that the effect of altercentric intrusion reflects the competition between three dots and five dots (see Table 3.2). Secondly, a *precise enumeration of own perspective*: although the avatar's perspective content is always a subset of participants' own perspective, they do not specifically retain concrete numeral information about the avatar's perspective content. On this hypothesis, altercentric intrusion comes in the form of an impression that the avatar sees fewer dots than participants themselves. Lastly, it could be that participants undergo a *precise enumeration without binding* that does not distinguish between the subset seen by the avatar and the subset not seen by the avatar: in this case, participants represent both subsets of dots that are part of their own perspective as precise numbers, and are equally likely to experience interference from either of these subsets when judging their own perspective. To distinguish between these hypotheses, investigations in Experiments 2 and 3 will focus on participants' implicit computation and representation of the avatar's perspective whilst they make explicit judgements about their own perspectives.

Table 3.2 Three types of possible representation of the avatar's perspective

<table>
<tr><td></td><td>Participants' Perspective</td><td>Avatar's Perspective</td></tr>
<tr><td>Precise Enumeration + Binding</td><td rowspan="3">3 dots + 2 dots = 5 dots</td><td>3 dots</td></tr>
<tr><td>Precise Enumeration of Own Perspective</td><td>Something that is not 5 dots</td></tr>
<tr><td>Precise Enumeration without Binding</td><td>Either 3 dots or 2 dots</td></tr>
</table>

3.3 Experiment 2

Previous visual perspective-taking studies (Samson et al., 2010; Experiment 1 of the current chapter) only analysed 'yes' responses as a measure of perspective judgements. In these studies, the number probes on 'no' response trials in self conditions always correspond to the avatar's perspective contents. As discussed in Section 3.2.3, the avatar's perspective content is always one of the two subsets that make up the participants' perspective content. If automatic visual perspective-taking involves an unique representation of the avatar's perspective, the number probe that corresponds to the avatar's perspective should be more difficult to reject compared to number probes that do not correspond to the avatar's perspective. Nonetheless, there has been no direct comparison of the processing cost required to reject number probes that does and does not correspond to the avatar's perspective content.

The current experiment examines participants' representation of the avatar's perspective content when making explicit judgements about their own perspective. In particular, the current experiment aims to test three different interpretations of what participants represent of an avatar's perspective. Participants may bind subset of dots that corresponds to the avatar's perspective content to the avatar, in which case they should find it difficult to reject a number probe that directly corresponds to the avatar's perspective content. Alternatively, participants may hold precise representations of both of the subsets computed for their own perspective content without binding the avatar to his perspective content, in which case both types of number probes should be equally difficult to reject. And finally, participants may not have a precise representation of either subset of dots, in which case the number probes that correspond to either subsets should not be more difficult to reject compared to a number that does not correspond to either subsets. The present experiment focuses on the representation of perspective contents that are subitizable. This ensures that differences in the processing cost of subitizable and non-subitizable sets are eliminated. Potential discrepancies between participants' representations for perspective contents that can and cannot be subitized will be investigated in Experiment 3.

3.3.1 *Method*

3.3.1.1 *Participants*

Sixteen students (15 female, mean age 20.13 years, age range 18 to 33) from the University of Birmingham participated in this experiment in return for study credits. All 16 participants performed above chance for both types of perspective judgements, therefore data from all participants were included in the analysis.

3.3.1.2 *Design and Procedure*

A 3 x 2 within-participant design was respectively constructed for the 'no' response trials in the self and the other conditions. The design included probe-type (for the self condition, an *avatar-probe*, a number probe that corresponds to the avatar's perspective content; for the other condition, a *self-probe*, a number that corresponds to the participants' perspective content; for both the self and the other conditions *wall-probe*, a number probe that corresponds to the number of dots on the wall behind the avatar; and *novel-probe*, a number probe that was not either subsets of dots) and congruency (congruent, incongruent) as factors[3]. The three probe-types appear in equal numbers of trials for both the self and the other conditions. The 'yes' response trials comprised a 2 x 2 design with perspective (other, self) and congruency (congruent, incongruent) as factors. There were equal numbers of self and other trials, congruent and incongruent trials, and 'yes' response and 'no' response trials. A total of 202 trials were presented comprising one practice block of 10 trials followed by four blocks of 48 trials. The rest of the procedure was identical to that of Experiment 1.

3.3.1.3 *Predictions*

If participants represent the avatar's perspective content as a precise number, then they should find it difficult to reject a number that corresponds to the avatar's perspective content. Therefore an effect of altercentric intrusion on self judgements should be observed exclusively in the avatar-probe condition but not in other probe-types. If participants do not represent the avatar's perspective content as a precise number, then they would make no distinctions between the three types of number probes. In this case, effects of altercentric intrusion should be observed in all three probe-types. However, if participants spontaneously represent both subsets of dots attended to without distinguishing between the subset that is and is not the avatar's

[3] Since the 'no' response trials in the congruent condition always contained number probes that do not correspond to either perspective contents or either subset of dots, it was not possible to divide the congruent condition into the three probe-type conditions. In order to ensure any differences observed between the congruent and incongruent conditions was a result of perspective representation, and not enumeration, the congruent trials were allocated to one of the three probe-type conditions that contains matching number probes.

perspective, then both the wall-probe and the avatar-probe conditions should show effects of altercentric intrusion.

3.3.2 *Results*

3.3.2.1 *Basic Effects*

Data from ‘yes’ response trials and ‘no’ response trials were processed respectively, in the same way as in Experiment 1. ‘Yes’ response trials due to erroneous responses (4.03% of the data) and correctly responded trials slower than two standard deviations from mean response time (4.70% of the data) were eliminated from the data set. ‘No’ response trials due to erroneous responses (4.96% of the data) and correctly responded trials slower than two standard deviations from mean response time (4.77% of the data) were eliminated from the data set. A general analysis was first conducted to examine the same phenomena of egocentric intrusion and altercentric intrusion observed in Samson et al. (2010) and in Experiment 1. The self and the other condition were analysed separately for reasons described in Experiment 1. For the ‘yes’ responses, two pairs of t-tests were conducted to examine the effects of egocentric intrusion and altercentric intrusion, respectively. A significant effect of egocentric intrusion was found, $t(15) = 5.72$, $p < .001$ (other-congruent = 534.81, other-incongruent = 626.76), as well as a significant effect of altercentric intrusion, $t(15) = 2.57$, $p = .021$ (self-congruent = 611.13, self-incongruent = 682.12; see Figure 3.5).

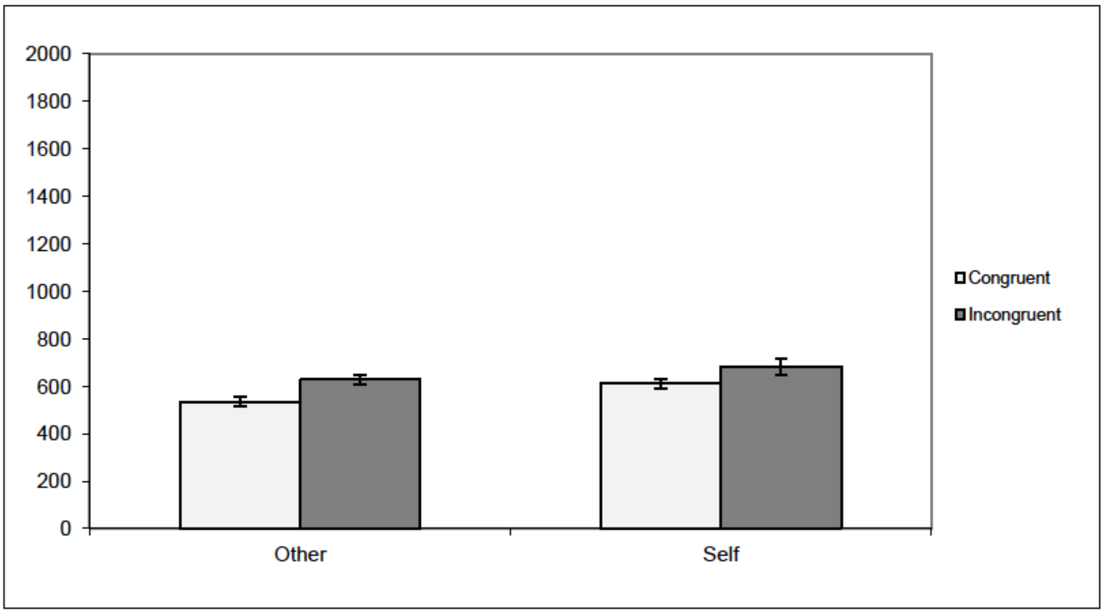


Figure 3.5 Processing cost for Experiment 2 'yes' response trials. Error bars represent standard errors.

A preliminary analysis was conducted on the 'no' response trials to investigate the phenomena of egocentric intrusion and altercentric intrusion before extending to test the three probe-types. A significant effect of egocentric intrusion was found, $t(15) = 5.50$, $p < .001$ (other-congruent = 559.60, other-incongruent = 624.89), along with a significant effect of altercentric intrusion, $t(15) = 2.31$, $p = .036$ (self-congruent = 605.81, self-incongruent = 650.26; see Figure 3.6).

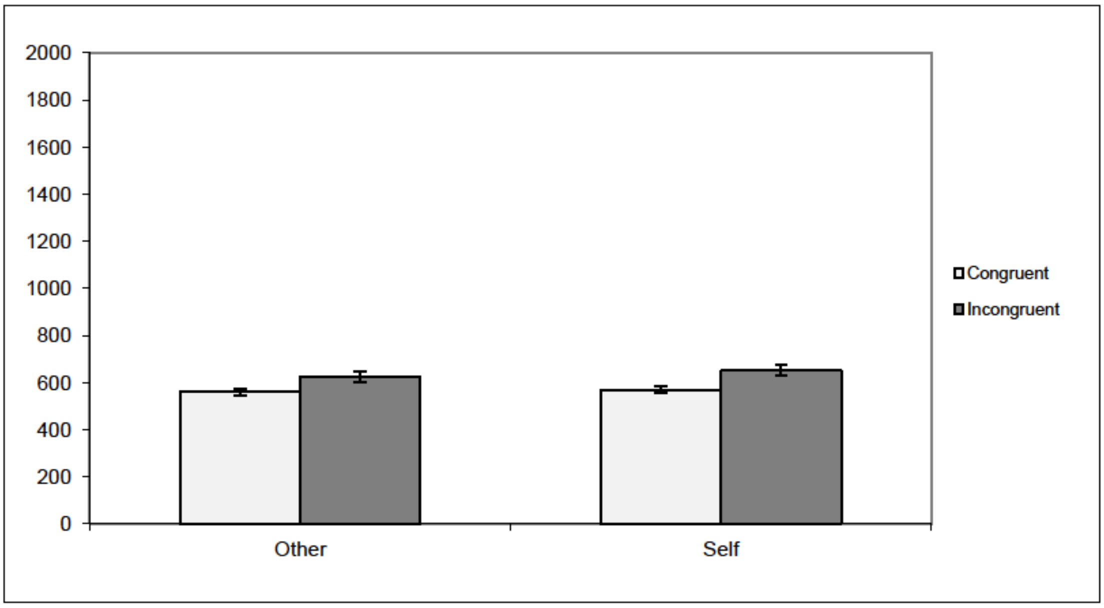


Figure 3.6 Processing cost for Experiment 2 'no' response trials. Error bars represent standard errors.

3.3.2.2 *Altercentric Intrusion Probe Type*

A 3 x 2 repeated-measures ANOVA was carried out with probe-type (avatar-probe, wall-probe, & novel-probe) and congruency (congruent & incongruent) in the 'no' response trials from the self condition. A significant main effect of congruency was found, $F(1, 15) = 6.14$, $p = .026$, $\eta_p^2 = .290$, which shows that the incongruent conditions required greater processing cost compared to the congruent conditions (*M*s = 650.49 and 596.00, respectively), with a significant main effect of probe-type, $F(1, 15) = 28.49$, $p < .001$, $\eta_p^2 = .655$ ($M_{avatar} = 659.82$, $M_{wall} = 543.90$, $M_{novel} = 666.02$). There was a significant interaction effect between the two factors, $F(1, 15) = 6.39$, $p < .001$, $\eta_p^2 = .299$; see Figure 3.7. Planned comparisons revealed a significant effect of altercentric intrusion in the avatar-probe condition, $t(15) = 2.89$, $p = .011$ (avatar-probe-congruent = 593.02, avatar-probe-incongruent = 726.62), but not in the wall-probe condition, $t(15) = 0.02$, $p = .982$ (wall-probe-congruent = 544.14, wall-probe-incongruent = 543.66) or in the novel-probe condition, $t(15) = 1.54$, $p = .144$ (novel-probe-congruent = 650.85, novel-probe-incongruent = 681.19).

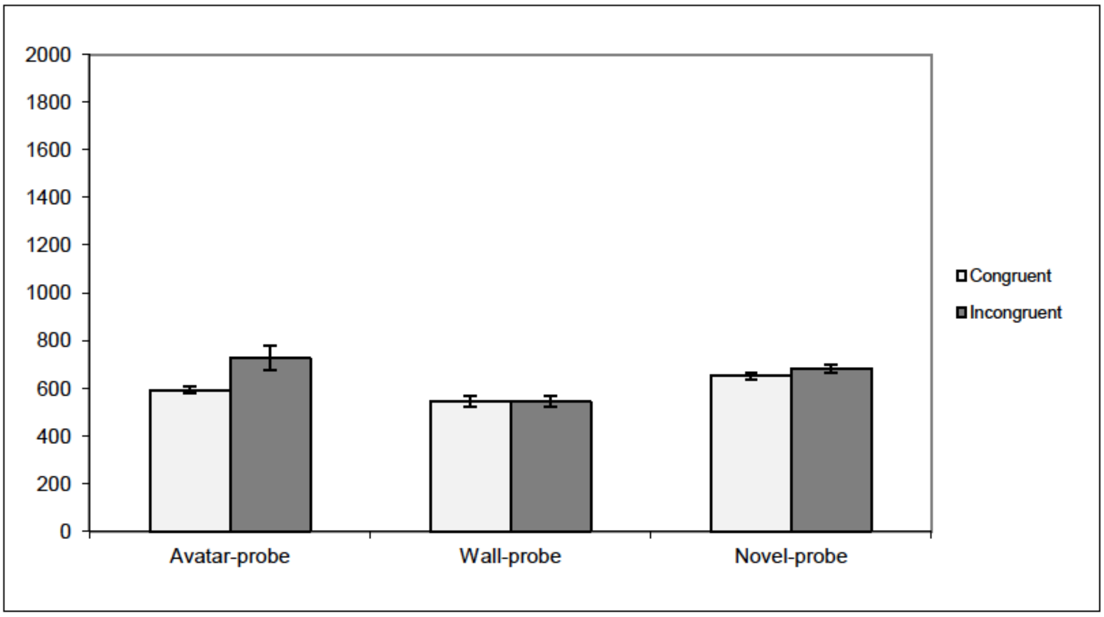


Figure 3.7 Processing cost for Experiment 2 'no' response from the self condition probe-types. Error bars represent standard errors.

3.3.3 *Discussion*

The current experiment demonstrated effects of egocentric intrusion and altercentric intrusion when participants made 'yes' responses, replicating findings from Experiment 1 and Samson et al. (2010). These effects were also found when participants made 'no' responses to reject an incorrect display. Significantly, the effect of altercentric intrusion was only observed when participants rejected number probes that corresponded to the avatar's perspective contents. When participants rejected number probes that corresponded to the content on the wall behind the avatar or to a novel number, there was no effect of altercentric intrusion. This suggests that participants' implicit processing of the avatar's perspective involves binding the avatar's perspective content with the avatar.

The findings of Experiments 1 and 2 suggest that automatic visual perspective-computation operates with some flexibility for automatic computation of large number perspective contents, and binding the avatar's small number perspective contents to the avatar. Nonetheless, given the high processing cost for enumerating large number perspective contents, it is difficult to predict whether participants also represent the avatar's large number perspective content as a precise number of dots. It is possible that the demand of enumeration leaves little flexibility for participants to precisely represent the avatar's perspective content. In Experiment 3, therefore, the upper enumerative limits of automatic visual perspective-computation will be investigated.

3.4 Experiment 3

The present investigation focuses on participants' computation and representation of the avatar's perspective content when both their own and the avatar's perspectives contain large numbers of dots. Since the results of Experiment 2 rules out the possibility that participants spontaneously represent both small number subsets of dots, this possibility will not be examined further with larger sets of dots in the current experiment. Instead, the current investigation focuses on a potential binding between the avatar and his large number perspective contents.

3.4.1 *Method*

3.4.1.1 *Participants*

Twenty students (13 female, mean age 24.55 years, age range 18 to 37) from the University of Birmingham participated in this study in return for study credits. One participant's data were replaced due to a significantly below chance performance in the self condition.

3.4.1.2 *Design and Procedure*

A 2 x 2 x 2 within-participant design was employed for the 'no' response trials, with probe-type (for the self condition: *avatar-probe*, *novel-probe*; for the other condition: *self-probe*, *novel-probe*), congruency (congruent, incongruent), and number size (large-number, small-number) as the factors. The *wall-probe* trials were included as filler trials in order to match the design for Experiment 2[4]. The 'yes' response trials were constructed with a 2 x 2 x 2 design with perspective (other, self), congruency (congruent, incongruent), and number size (large-number, small-number) as factors. The avatar-probe condition contained number probes that corresponded to the avatar's perspective contents. The self-probe contained number probes that corresponded to the participants' perspective contents. The novel-probe condition contained the same collection of number probes to the avatar-probe condition, with the numbers reallocated so that they did not correspond to the avatar's perspective, the participants' perspective, or the content on the wall[5]. This design also ensured that differences observed between the congruent and incongruent conditions were results of perspective-computation without applying the weighting procedure from Experiment 1. This was achieved by assigning each congruent trial to one of the three probe-type conditions, and reallocated the number probes from the corresponding incongruent conditions to the congruent conditions so that they contained identical number probes. A total number of 202 trials were presented in pseudo-random order

[4] The wall-probe trials were treated as filler trials for two reasons. Firstly, as previously described, the possibility of spontaneous representation for both subsets of dots has already been excluded in Experiment 2. Secondly, the large-number condition in the current experiment only comprised trials where both participants' and the avatar's perspective contents were large-numbers, which meant that the wall-probe condition only contained small numbers. Since the wall condition does not address novel theoretical issues and contains inevitably distinct enumeration load, these trials were excluded from the analysis.

[5] All the number probes in the novel-probe condition were taken directly from the avatar-probe condition apart from one number probe, where a seven instead of six had to be presented, as six corresponded to the avatar's perspective content in that trial. This adaptation was not considered to confound with the results as a larger number was now included. Therefore this could only increase the likelihood of observing an effect of altercentric intrusion in the novel-probe condition, which does not favour the current predictions.

across one practice block and four testing blocks. There were an equal numbers of avatar-probe/ self-probe, wall-probe, and novel-probe trials; congruent and incongruent trials; large-number and small-number trials; self and other trials; 'yes' response and 'no' response trials.

3.4.1.3 *Predictions*

As demonstrated in Experiment 1, effects of egocentric intrusion and altercentric intrusion in the 'yes' response small-number and large-number conditions should be observed. Furthermore, as demonstrated in Experiment 2, the 'no' response small-number trials should also exhibit effects of egocentric intrusion and altercentric intrusion, along with an exclusive effect of altercentric intrusion in the avatar-probe condition. For the 'no' response large-number condition, there were no specific predictions regarding the effect of egocentric intrusion. Given that the saliency of one's own perspective often interferes with one's judgements about others' perspective, one might expect to find an effect of egocentric intrusion. However, it is not entirely clear whether this should be the case when participants reject mismatched probes and displays. As discussed earlier, the current theoretical interest lies with participants' implicit computation and representation of the avatar's perspective when making explicit judgements about their own perspective. Three possible outcomes can be identified. The first is that if participants bind the avatar's large-number perspective content with the avatar, then an effect of altercentric intrusion should be observed exclusively in the avatar-probe condition. Participants should find it more difficult to reject a number probe that corresponds to the avatar's perspective. Alternatively, if participants do not bind the avatar with his perspective content and only represent the avatar's large-number perspective content as an impression that the avatar sees differently from themselves, then the effects of altercentric intrusion should be found in both the avatar-probe condition and the novel-probe condition. That is, participants would not distinguish between a subset of dots that corresponds to the avatar's perspective and a number that does not correspond to either subset of dots in the display. Finally, there has been no previous demonstration of any effects of altercentric intrusion when participants make 'no' responses to large-number perspectives. Since there is little justification for assuming participants' 'no' responses to large-number perspective contents undergoes identical computation to that of small-number perspective contents, the possibility of not observing an effect of altercentric intrusion in 'no' response large-number perspectives remains open.

3.4.2 *Results*

3.4.2.1 *Basic Effects*

3.4.2.1.1 *Other Condition*

Data from 'yes' response trials and 'no' response trials were processed respectively, in the same way as in Experiment 2. 'Yes' responses trials due to erroneous responses (4.22% of the data) and correctly responded trials slower than two standard deviations from mean response time (4.92% of the data) were eliminated from the data set. 'No' response trials due to erroneous responses (7.24% of the data) and correctly responded trials slower than two standard deviations from mean response time (2.94% of the data) were eliminated from the data set. For the 'yes' response trials, a 2 x 2 repeated-measures ANOVA was conducted with congruency (congruent & incongruent) and number size (small-number & large-number) as factors. There were significant main effects of congruency, $F(1, 19) = 30.45$, $p < .001$, $\eta_p^2 = .616$, which shows that the incongruent conditions required greater processing cost compared to the congruent conditions (*M*s = 1084.85 and 942.86, respectively), and number size, $F(1, 19) = 95.87$, $p < .001$, $\eta_p^2 = .835$, which confirms that large-numbers demand greater processing cost to compute than small-numbers (*M*s = 1331.94 and 695.77, respectively). There was a significant interaction effect between the two factors, $F(1, 19) = 4.64$, $p = .044$, $\eta_p^2 = .196$; see Figure 3.8. Planned comparisons showed a significant effect of small-number egocentric intrusion, $t(19) = 3.14$, $p = .005$ (other-small-number-congruent = 650.43, other-small-number-incongruent = 741.10), and a significant effect of large-number egocentric intrusion, $t(19) = 4.80$, $p < .001$ (other-large-number-congruent = 1235.28, other-large-number-incongruent = 1428.60). Additional comparison between the small-number and the large-number conditions revealed greater egocentric intrusion in the large-number condition compared to the small-number condition, $t(19) = 2.16$, p = .044 (other-large-number = 193.31, other-small-number = 90.67). This suggests that when participants make 'yes' responses, egocentrism could be a function of enumerative demand.

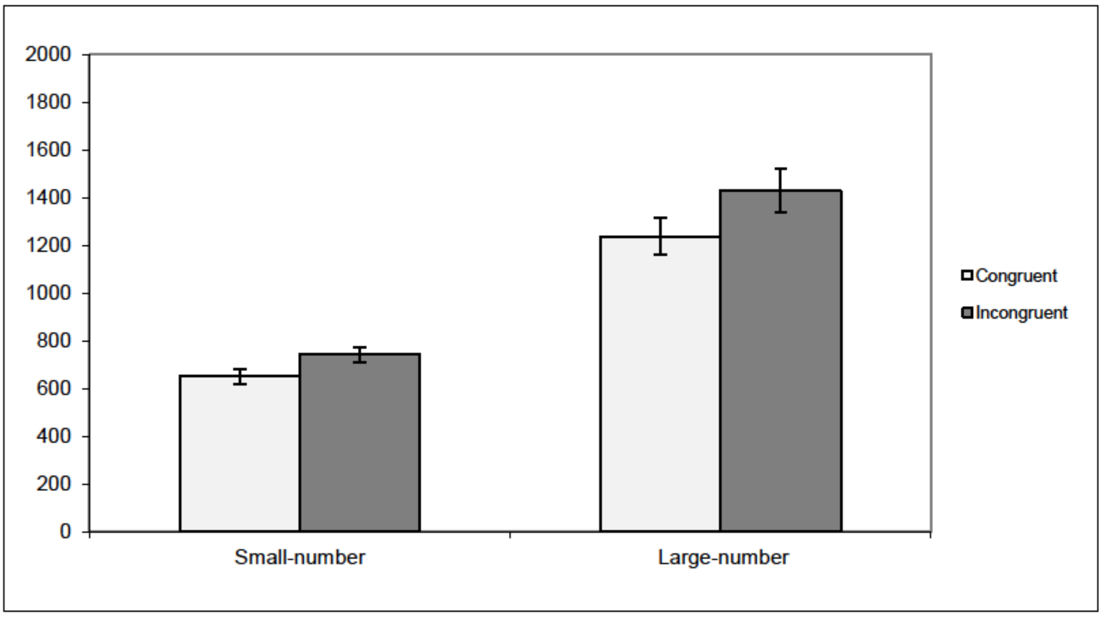


Figure 3.8 Processing cost for Experiment 3 'yes' response other trials. Error bars represent standard errors.

For the ‘no’ response trials, another 2 x 2 repeated-measures ANOVA was conducted, with congruency (congruent & incongruent) and number size (small-number & large-number) as factors. There were significant main effects of congruency, $F(1, 19) = 12.39$, $p = .002$, $\eta_p^2 = .395$, which shows that the incongruent conditions required greater processing cost compared to the congruent conditions (*M*s = 1141.04 and 1008.86, respectively), and number size, $F(1, 19) = 227.64$, $p < .001$, $\eta_p^2 = .923$, which confirms that large-numbers demand greater processing cost to compute than small-numbers (*M*s = 1388.00 and 761.90, respectively). There was no significant interaction between the two factors, $F(1, 19) = 0.77$, $p = .392$, $\eta_p^2 = .039$. Planned comparisons revealed a significant effect of small-number egocentric intrusion, $t(19) = 3.87$, $p = .001$ (other-small-number-congruent = 727.42, other-small-number-incongruent = 815.99), with a significant effect of large-number egocentric intrusion, $t(19) = 3.43$, $p = .022$ (other-large-number-congruent = 1307.92, other-large-number-incongruent = 1468.09; see Figure 3.9). Additional comparison between the small-number and the large-number conditions revealed no difference in the congruency effects, $t(19) = 0.88$, p = .392 (other-large-number = 160.17, other-small-number = 104.19). This suggests that when participants produce a ‘no’ response, they do not suffer more or less egocentrism as a result of varied enumerative demand.

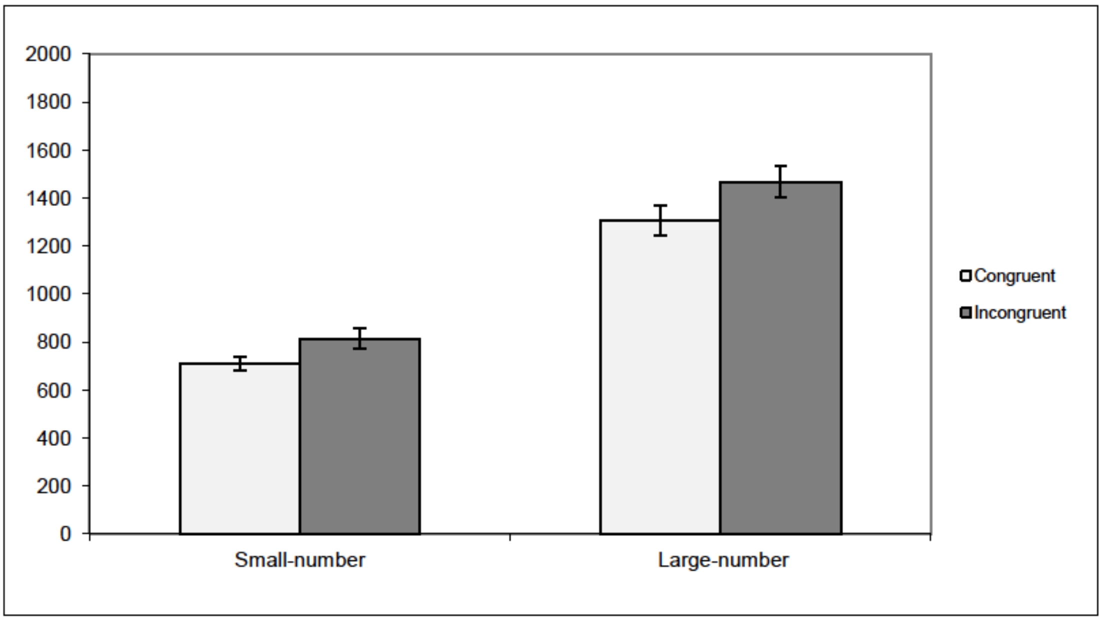


Figure 3.9 Processing cost for Experiment 3 'no' response other trials. Error bars represent standard errors.

3.4.2.1.2 *Self Condition*

For the 'yes' response trials, a 2 x 2 repeated-measures ANOVA was carried out with congruency (congruent & incongruent) and number size (small-number & large-number) as the factors. There were significant main effects of congruency, $F(1, 19) = 12.53$, $p = .002$, $\eta_p^2 = .397$, which shows that the incongruent conditions required greater processing cost compared to the congruent conditions (*M*s = 1322.75 and 1123.56, respectively), and number size, $F(1, 19) = 420.17$, $p < .001$, $\eta_p^2 = .957$, which confirms that large-numbers demand greater processing cost to compute than small-numbers (*M*s = 1674.61 and 771.71, respectively). There was a significant interaction between the two factors, $F(1, 19) = 7.68$, $p = .012$, $\eta_p^2 = .288$. Planned comparisons showed a significant effect of small-number altercentric intrusion, $t(19) = 2.40$, $p = .027$ (self-small-number-congruent = 727.42, self-small-number-incongruent = 815.99) with a significant effect of large-number altercentric intrusion, $t(19) = 3.43$, $p = .003$ (self-large-number-congruent = 1519.70, self-large-number-incongruent = 1829.52; see Figure 3.10). Additional comparison between the small-number and the large-number conditions revealed greater altercentric intrusion in the large-number condition compared to the small-number condition, $t(19) = 2.77$, p = .012 (self-large-number = 309.81, self-small-number = 88.56). This suggests that participants suffer more interference from the avatar's perspective when making 'yes' responses.

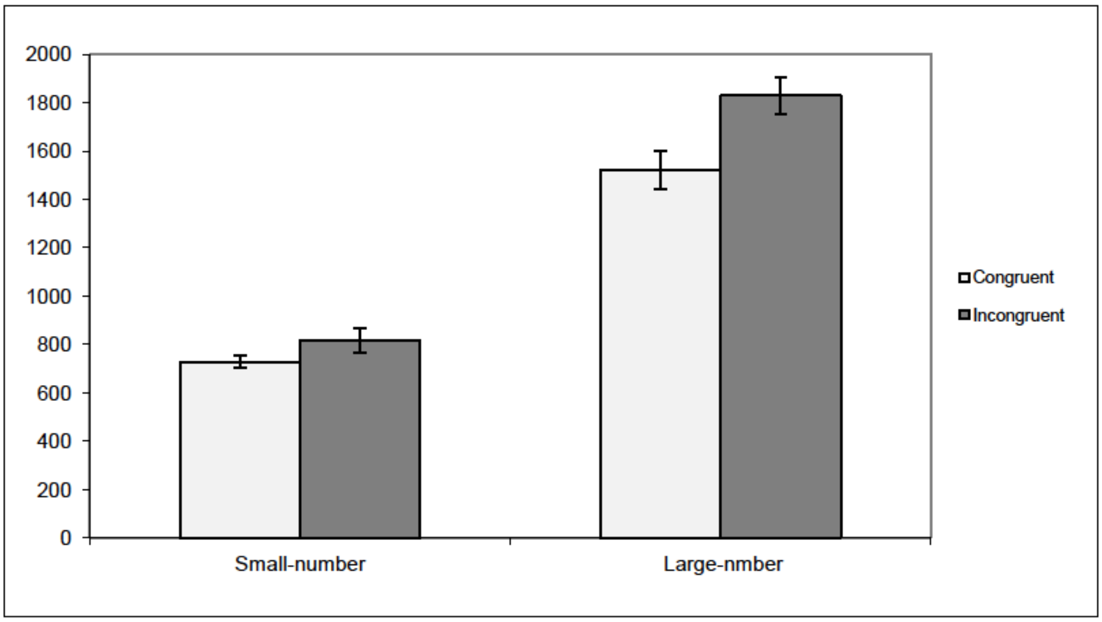


Figure 3.10 Processing cost for Experiment 3 'yes' response self trials. Error bars represent standard errors.

For the 'no' response trials, a 2 x 2 repeated-measures ANOVA was carried out with congruency (congruent & incongruent) and number size (small-number & large-number) as the factors. There was no main effect of congruency, $F(1, 19) = 0.96$, $p = .339$, $\eta_p^2 = .048$, with a significant main effect of number size, $F(1, 19) = 214.57$, $p < .001$, $\eta_p^2 = .919$, confirming that large-numbers demand greater processing cost to compute than small-numbers (*M*s = 1475.05 and 722.98, respectively), and no interaction between the two factors, $F(1, 19) = 1.57$, $p = .225$, $\eta_p^2 = .076$. Planned comparisons showed a significant effect of small-number altercentric intrusion, $t(19) = 3.43$, $p = .003$ (self-small-number-congruent = 687.66, self-small-number-incongruent = 758.21), with no significant effect of large-number altercentric intrusion, $t(19) = 0.06$, $p = .954$ (self-large-number-congruent = 1473.26, self-large-number-incongruent = 1476.83; see Figure 3.11). Additional comparison between the small-number and the large-number conditions revealed no difference between the congruency effect observed in the large-number and the small-number condition, $t(19) = 1.25$, p = .225 (self-large-number = -3.57, self-small-number = 70.55). This suggests that when participants make ‘no’ responses, they do not suffer more or less interference from the avatar’s perspective as a result of varied enumerative demand.

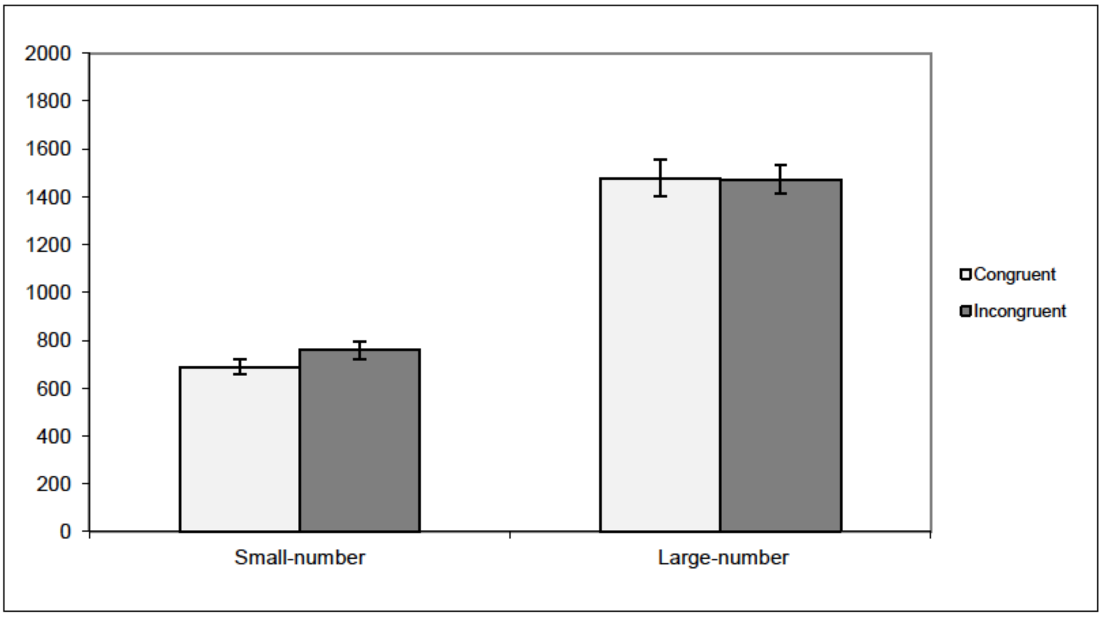


Figure 3.11 Processing cost for Experiment 3 ‘no’ response self trials. Error bars represent standard errors.

3.4.2.2 *Altercentric Probe Type*

A 2 x 2 x 2 repeated-measures ANOVA was conducted with probe-type (avatar-probe & novel-probe), congruency (congruent & incongruent), and number size (small-number & large-number) as factors. There was a significant main effect of probe-type, $F(1, 19) = 20.04$, $p < .001$, $\eta_p^2 = .513$, showing that the avatar-probe demanded more processing cost than the novel-probe (*M*s = 1222.47 and 1016.62, respectively), with no main effect of congruency, $F(1, 19) = 1.17$, $p = .293$, $\eta_p^2 = .058$, and a significant main effect of number size, $F(1, 19) = 195.97$, $p < .001$, $\eta_p^2 = .912$ that shows large-numbers demand greater processing cost to compute than small-numbers (*M*s = 1506.46 and 732.63, respectively). There was a significant interaction between number size and probe-type, $F(1, 19) = 12.45$, $p = .002$, $\eta_p^2 = .396$. No other interactions between any of the three factors were found ($F < 2.83$, $p > .108$). As a follow-up, I conducted a set of 2 x 2 ANOVA for each number size. Although this analysis was not warranted by the omnibus analysis, exploring these effects is essential to address the *a priori* hypotheses of the current experiment.

3.4.2.2.1 *Small Number Condition*

A 2 x 2 repeated-measures ANOVA with congruency (congruent & incongruent) and probe-type (avatar-probe & novel-probe) as the factors, revealed a significant main effect of congruency, $F(1, 19) = 8.70$, $p = .008$, $\eta_p^2 = .314$, which shows that the incongruent conditions required greater processing cost compared to the congruent conditions (*M*s = 687.06 and 778.19, respectively), with no main effect of probe-type, $F(1, 19) = 2.05$, $p = .169$, $\eta_p^2 = .097$, and a significant interaction between the two factors, $F(1, 19) = 8.11$, $p = .010$, $\eta_p^2 = .299$. Planned comparisons demonstrated a significant effect of small-number altercentric intrusion in the avatar-probe condition, $t(19) = 3.04$, $p = .007$ (man-small-number-congruent = 664.34, man-small-number-incongruent = 851.57), but not in the novel-probe condition, $t(19) = 0.25$, $p = .806$ (novel-small-number-congruent = 709.78, novel-small-number-incongruent = 704.82; see Figure 3.12), replicating the findings from Experiment 2.

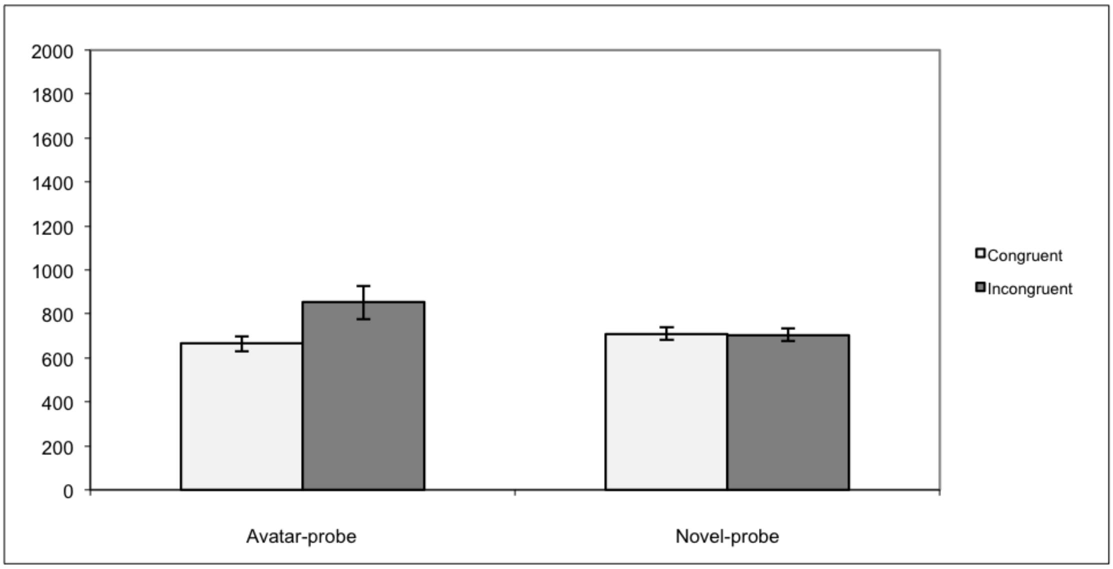


Figure 3.12 Processing cost for 'no' response small-number self trials. Error bars represent standard errors.

3.4.2.2.2 *Large Number Condition*

A 2 x 2 repeated-measures ANOVA with congruency (congruent & incongruent) and probe-type (avatar-probe & novel-probe) as the factors, revealed no main effect of congruency, $F(1, 19) = 0.03$, $p = .874$, $\eta_p^2 = .001$, a significant main effect of probe-type, $F(1, 19) = 19.04$, $p < .001$, $\eta_p^2 = .501$[6], which shows that the avatar-probe required more processing cost than the novel-probe (*M*s = 1686.98 and 1325.94, respectively), and no interaction between the two factors, $F(1, 19) = 0.01$, $p = .940$, $\eta_p^2 < .001$; see Figure 3.13. Planned comparisons demonstrated no effect of large-number altercentric intrusion in the avatar-probe condition, $t(19) = 0.14$, $p = .893$ (man-large-number-congruent = 1678.29, man-large-number-incongruent =

[6] As described earlier, the avatar-probe and novel-probe conditions contained identical displays and were matched across all but one number probe. However, since the novel-probe could not correspond to either the participants' or the avatar's perspective content, the reallocation of number probes leads to a marginally bigger numerical distance between the number probe and the participants' perspective content. Hence participants were able to respond to the novel-probe conditions with relatively little processing cost. Importantly, this numerical distance was identical across the novel-probe-congruent and novel-probe-incongruent conditions. Therefore, the overall smaller processing cost in the novel-probe conditions should not affect the current interpretation of perspective computation and representation.

1695.67), with no effect of large-number altercentric intrusion in the novel-probe condition, $t(19) = 0.13$, $p = .896$ (novel-large-number-congruent = 1321.96, novel-large-number-incongruent = 1329.91).

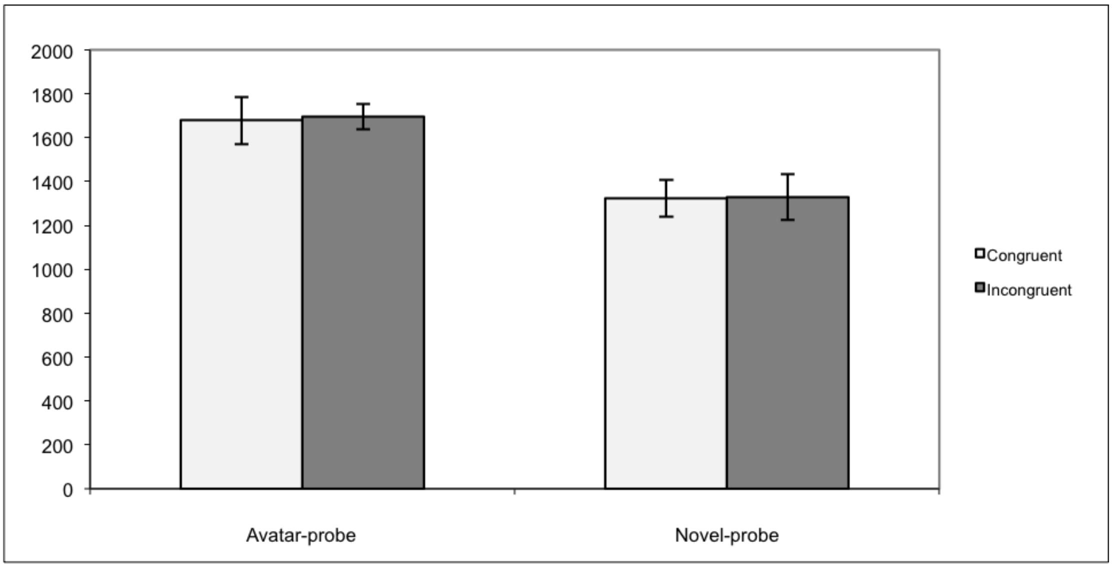


Figure 3.13 Processing cost for Experiment 3 'no' response large-number self trials. Error bars represent standard errors.

3.4.3 *Discussion*

The effects of both egocentric intrusion and altercentric intrusion were found in the 'yes' response small-number conditions, replicating the findings from Experiments 1 and 2, as well as Samson et al. (2010). The same sets of effects were also found in the 'yes' response large-number conditions, replicating the findings from Experiment 1. For the 'no' responses, effects of small-number egocentric intrusion and altercentric intrusion were found, with altercentric intrusion exclusively shown in the avatar-probe condition, replicating the exact findings from Experiment 2. An effect of large-number egocentric intrusion on the 'no' response trials was also found. This suggests that participants' own perspective interferes with their judgements about the avatar's perspective even when they were placed under high enumerative demand. However, no effect of large-number altercentric intrusion was observed from the ‘no’ response trials. The current findings suggest that participants did not hold any representation for the avatar’s large-number perspective contents. One explanation for the absence of effects of large-number altercentric intrusion in the ‘no’ response condition is that participants utilised an alternative computation route in this condition. An effective strategy is to reject number probes that are either the same or smaller than either subset of dots without completing the full calculation for participants’ own perspective contents. Since participants were always cued with the avatar’s perspective content in the avatar-probe condition, and with the content on the wall behind the avatar in the wall-probe condition, this strategy could be an efficient way of reaching an accurate rejection response. Interestingly, this strategy is only effective for ‘no’ responses; on ‘yes’ response trials, participants would have to complete summing up both subsets of dots in order to make a confirmation response about their own perspective. Moreover, the same strategy for responding to ‘no’ response trials would have been equally effective in the small-number conditions. Nevertheless, results from Experiments 2 and 3 showed that participants not only failed to utilise such a strategy, but they also made an unnecessary computation and bound the avatar’s small-number perspective contents with the avatar. This suggests that participants’ implicit computation of the avatar’s perspective is likely to be restricted by the high enumerative load and their responses to reject mismatched probes and displays.

3.5 General Discussion

3.5.1 *Summary of Current Findings*

The current investigations revealed effects of egocentric intrusion and altercentric intrusion when participants made explicit judgements about perspective contents that could be subitized. This replicates the findings of Samson et al. (2010).

The effect of egocentric intrusion indicates that participants' own perspective interferes with their explicit judgements of the avatar's perspective. The effect of altercentric intrusion suggests that when participants make explicit judgements about their own perspectives, they implicitly compute the avatar's perspective to some extent. Experiment 1 demonstrated that automatic perspective-computation not only occurs when the perspective contents are subitizable, but also when the perspective contents demanded greater enumerative effort. The findings of Experiment 1 revealed that automatic visual perspective-computation operates with some flexibility for processing perspective contents that are not subitizable. Experiment 2 further demonstrated that when participants make explicit judgements about their own subitizable visual perspectives, their implicit computation of the avatar's perspectives entails binding the avatar's perspective contents with the avatar. Interestingly, Experiment 3 showed that when participants made 'no' responses on trials that require explicit judgements about their own large-number perspectives, they no longer compute or represent the avatar's perspective contents. The findings from Experiments 2 and 3 reveal a clear distinction between participants' representation of small number perspective contents and large number perspective contents. This distinction reflects the limited amount of flexibility that is embedded in automatic visual perspective-computation.

3.5.2 *Altercentric Intrusion*

The current findings showed that when participants make explicit judgements about their own subitizable perspectives, they also implicitly compute an avatar's small-number perspective contents. When participants made 'yes' responses to their own large-number perspective contents, the avatar's large-number perspective was also processed. However, when participants made 'no' responses to number sets in their own perspective beyond the subitizable range, they showed no computation or representation of the avatar’s perspective contents. As discussed earlier, it is not surprising that the effects of altercentric intrusion were found in both small-number and large-number perspective judgements when participants made ‘yes’ responses. Participants were required to attend to the subsets both in front and behind the avatar in order to make an accurate ‘yes’ response about their own perspective contents. However, for ‘no’ responses, there was an obvious and effective strategy that participants could apply. Interestingly, the data suggest that participants did not utilise this strategy when making small-number perspective judgements. In contrast, the absence of ‘no’ response large-number altercentric intrusion indicates that participants may apply this strategy when greater enumerative demand is placed upon them. Furthermore, Experiment 3 revealed that even when the small-number and large-number trials were mixed within an experiment, participants show distinct representation and computation of perspective contents within the subitizable range and perspective contents beyond the subitizable range. This indicates that the form of computation and representation are likely to be determined by the computational demands on a trial-by-trial basis rather than a global decision or strategy. The current findings suggest that automatic visual perspective-computation operates with a limited amount of cognitive flexibility, and that participants’ computation and representation of the avatar’s perspectives are likely to be contingent on the enumerative load and the processes necessary for generating correct responses.

3.5.3 *Egocentrism*

The current study demonstrated effects of egocentric intrusion when participants made both ‘yes’ and ‘no’ responses and when they were under both high and low enumerative demands. The constant observation of egocentrism echoes a number of previous studies, which have suggested that adults' judgements of others' perspectives are easily affected by their own perspectives (e.g., Birch & Bloom, 2007; Keysar et al., 2003). Furthermore, Experiment 3 revealed a greater egocentric intrusion when participants were under higher enumerative demand. This is consistent with Epley et al. (2004), which indicate that overcoming interference from one's own perspective is a demanding process, such that when there are concurrent demands, adults show increased egocentrism. The relation of egocentrism with the findings from Chapter 4, and the broader mindreading literature will be further discussed in Chapter 8.

3.5.4 *Conclusion*

The current study demonstrated that automatic visual perspective-computation operates with some flexibility for enumerating perspective contents that cannot be subitized, and for binding subitizable perspective contents to the avatar. However, the current investigation also showed that there are limitations to these computations and representations. The current findings are consistent with the account of Apperly and Butterfill (2009), which proposes an exchange of flexibility for efficiency in the operation of efficient mindreading. In Chapter 8, I will discuss this issue on the context of the broader social cognition literature where a limited flexibility within various efficient processes has also been demonstrated.

CHAPTER 4 AUTOMATIC LEVEL-1 VISUAL PERSPECTIVE-TAKING: CAPACITY FOR SELECTION OF RELEVANT INFORMATION

4.1 Introduction

In the current chapter, I aim to explore whether automatic Level-1 visual perspective-computation operates with some flexibility to select relevant information. There are two main reasons to examine the extent that individuals are able to select relevant information and ignore irrelevant information. Firstly, in order to successfully engage in fluent social interactions, it is crucial for one to be able to select the relevant information that is compatible with the communicative context. Difficulty in identifying the correct referents of others' perspective, may impair even simple socio-functioning (e.g., making sense of others' finger pointing). Secondly, as described earlier, automatic processes like visual perspective-computation comprise limited cognitive flexibility. However, whether automatic visual perspective-computation includes ecologically valuable processes, such as selection of relevant information, is unclear. The current investigation aims to address this issue by introducing additional demand on the selection of relevant information whilst participants engage in visual perspective-computation.

The visual attention literature has examined processing costs associated with the selection of non-social information. Studies indicate that participants are slower to detect target items even when they anticipate that irrelevant items will be included in the display (see Simons, 2000, for review). This indicates that participants suffer additional processing cost through the spontaneous allocation of their attention to irrelevant visual stimuli.

In the current study, participants were presented with a variation of Samson et al.’s (2010) visual perspective-taking task, in which distractor items (blue squares) were added into their own and the avatar’s visual perspective contents (see Figure 4.1 for samples of experimental stimuli). Participants were given explicit instructions to ignore the distractors. If automatic visual perspective-computation operates with the flexibility to select relevant information, then participants should differentiate between the target stimuli (red dots) and the distractors, and only compute the target stimuli. However, if automatic visual perspective-computation does not have the capacity for selection of relevant information, then participants would fail to make distinctions between the target stimuli and the distractors, showing undifferentiated computation between the conditions illustrated in Figure 4.1A and 4.1B. Experiments 4 and 5 examined this issue.

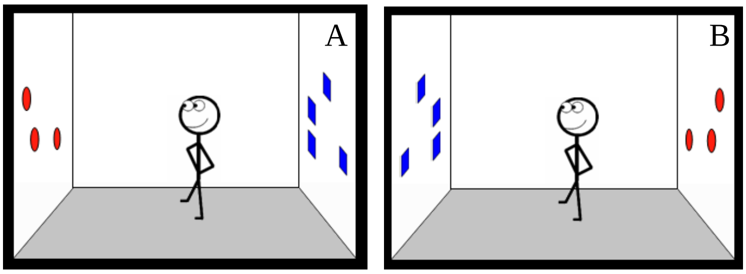


Figure 4.1 Examples of the experimental stimuli for Experiments 4 and 5. The red dots were target stimuli, the blue squares were distractors.

4.2 Experiment 4

4.2.1 *Method*

4.2.1.1 *Participants*

Sixteen students (10 female, mean age 19.29 years, age ranged from 19 to 28) from the University of Birmingham participated in this study in return for study credits. One participant’s data were replaced for failing to perform above chance in the self condition.

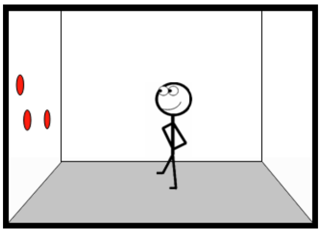
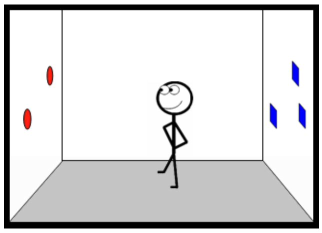
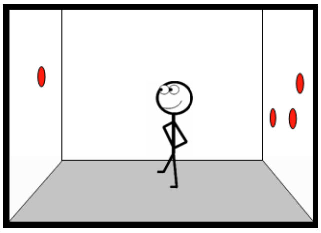
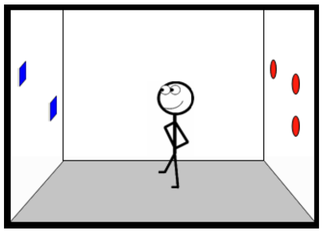
4.2.1.2 *Design and Procedure*

The current experiment employed the trial sequence from Chapter 3 (see Figure 3.1 from Chapter 3). Participants first saw a perspective cue, which indicated whether they should judge their own or the avatar's visual perspective. This was followed by a number probe, which represented the number of target stimuli in either the participants' or avatar's perspective content. At the end of the trial, participants saw a display picture depicting an avatar in a room with various numbers of target stimuli and distractors on the walls. Participants judged whether the perspective cue and the number probe correctly described the display picture. On half of the trials, participants only saw the target stimuli (red dots). On the other half of the trials, the display also contained distractors (blue squares). Critically, the distractors were distinct from the target items in both colour and shape. The maximum number of items in any subset was 4. This ensured that all subsets could be subitized either on their own (Kaufman et al., 1949) or in parallel with another subset (Wender & Rothkegel, 2000). This rule was applied for both the target stimuli sets and the distracters sets. Participants were informed that they would never be asked about the distractors and that they should only focus on the target stimuli. Participants made yes/no judgements by clicking a computer mouse, pressing the left button for 'yes' responses and the right button for 'no' responses. All participants were instructed to respond as accurately and as quickly as possible.

A 2 x 2 x 2 within-participant design was constructed with perspective (other, self), congruency (congruent, incongruent), and distractor (without distractor, with distractor) as the factors (see Table 4.1). All trial types were mixed within blocks, presented in a pseudo-random order so that participants never encountered three trials of the same condition in a row. The experiment was presented using E-prime

(Schneider et al., 2002a; 2002b). A total of 207 trials were presented across one practice block of 15 trials and four testing blocks of 48 trials each.

Table 4.1 Experimental conditions for Experiment 4. Congruent and incongruent conditions were defined with respect to the target stimuli only.

	Without-Distractor	With-Distractor
Congruent		
Incongruent		

4.2.1.3 *Predictions*

If automatic visual perspective-computation does not operate with the flexibility to select relevant information, then participants should fail to extract the target stimuli from the distractors. Since all the conditions were defined purely with respect to the target stimuli, participants should show identical processing cost for the with-distractor-congruent condition and the with-distractor-incongruent condition. That is, no effect of egocentric intrusion or altercentric intrusion should be observed in the with-distractor conditions. Alternatively, if automatic visual perspective-computation operates with the flexibility to select relevant information, then participants should be able to extract the target stimuli from the distracters. As a result, participants should only suffer interference from the target stimuli in the with-distractor conditions. Therefore, effects of egocentric intrusion and altercentric intrusion in the with-distractor conditions should be found. A further interesting comparison is between the other-with-distractor-congruent condition and the other-without-distractor-incongruent condition. In both conditions, participants judge the avatar's perspective when it contains target stimuli. However, the items behind the avatar are either 'truly incongruent' with his perspective, or 'merely distracting'. If participants distinguish the target stimuli from the distractors, then they should suffer more processing cost in the other-without-distractor-incongruent condition than in the other-with-distractor-congruent condition. However, if the distractors and the target stimuli cause the same amount of interference, then processing costs should be equivalent across these two conditions.

4.2.2 *Results*

Following the procedure of Experiment 1 and Samson et al. (2010), only the 'yes' response trials were included in the analysis. Analysis of response time data only included the trials to which participants responded correctly, 3.29% of the data were eliminated for this reason. Data points that were more than two standard deviations away from the overall mean were removed, 2.73 of the correctly responded trials were eliminated due to slow responses. Analysis was conducted with processing costs, which were calculated by dividing each participant's response time by the proportion of trials to which participants responded correctly in each condition. This score provided a concise summary of both speed and accuracy (for separate results of response time and error rate, see Appendix B). The self condition and the other condition were analysed separately, as the current investigation included no *a priori* hypothesis regarding differences in the judgements for these two perspectives.

4.2.2.1 *Other Condition*

A 2 x 2 repeated-measures ANOVA was conducted on the other condition, with congruency (congruent & incongruent) and distractor (without distractor & with distractor) as factors. There were significant main effects of congruency, $F(1, 15) = 57.91$, $p < .001$, $\eta_p^2 = .794$, which shows that the incongruent conditions required greater processing cost compared to the congruent conditions (*M*s = 883.35 and 691.77, respectively), and distractor, $F(1, 15) = 18.74$, $p = .001$, $\eta_p^2 = .555$, which confirms that the with distractor condition requires more processing cost compared to the without distractor condition (*M*s = 836.14 and 738.98, respectively). There was a significant interaction between congruency and distractor, $F(1, 15) = 8.52$, $p = .011$, $\eta_p^2 = .362$; see Figure 4.2. Planned comparisons revealed a significant effect of egocentric intrusion when no distractor was present, $t(15) = 4.40$, $p = .001$ (other-without-distractor-congruent = 684.68, other-without-distractor-incongruent = 793.27), as well as when the distractors were present, $t(15) = 5.76$, $p < .001$ (other-with-distractor-congruent = 698.86, other-with-distractor-incongruent = 973.43). Further planned comparisons showed a significantly higher processing cost in the other-without-distractor-incongruent condition compared to the other-with-distractor-congruent condition, $t(15) = 4.53$, $p < .001$ (other-with-distractor-congruent = 698.86, other-without-distractor-incongruent = 793.27). The current results revealed that participants successfully distinguished between the target stimuli and the distractors. Furthermore, effect of egocentric intrusion is likely to be driven by ‘true discrepancy’ between the avatar’s and participants’ perspectives.

Additionally, congruency effects were calculated by subtracting the processing cost in the congruent conditions from the matching incongruent conditions. Comparison between the without-distractor and the with-distractor conditions showed a significantly bigger congruency effect in the with distractor condition, $t(15) = 2.92$, p = .011 (other-with-distractor = 274.57, other-without-distractor = 108.59). This suggests that egocentrism is likely to be a function of the demand on the selection of relevant information.

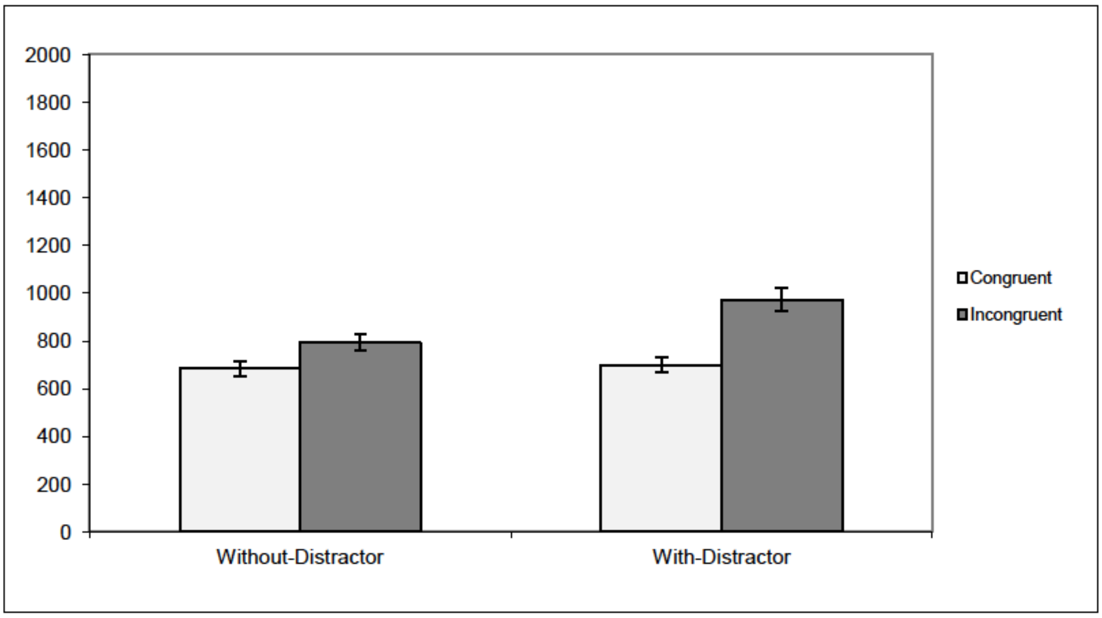


Figure 4.2 Experiment 4 other condition processing cost. Error bars represent standard errors.

4.2.2.2 *Self Condition*

A 2 x 2 repeated-measures ANOVA was carried out on the self condition, with congruency (congruent & incongruent) and distractor (without distractor & with distractor) as factors. There were significant main effects of congruency, $F(1, 15) = 28.62$, $p < .001$, $\eta_p^2 = .656$, which shows that the incongruent conditions required greater processing cost compared to the congruent conditions (*M*s = 824.12 and 710.56, respectively), and distractor, $F(1, 15) = 5.77$, $p = .030$, $\eta_p^2 = .278$, which confirms that the with distractor condition requires more processing cost compared to the without distractor condition (*M*s = 792.02 and 742.66, respectively). There was no interaction between congruency and distractor, $F(1, 15) = 1.34$, $p = .265$, $\eta_p^2 = .082$. Planned comparisons showed a significant effect of altercentric intrusion when no distractor was present, $t(15) = 4.21$, $p = .001$ (self-without-distractor-congruent =

695.30, self-without-distractor-incongruent = 790.02) as well as when the distractors were present, $t(15) = 4.35$, $p = .001$ (self-with-distractor-congruent = 725.81, self-with-distractor-incongruent = 858.23; see Figure 4.3). This indicates that participants successfully selected target stimuli and ignored distractors when making explicit judgements about their own visual perspective.

Additionally, congruency effects were calculated by subtracting the processing cost in the congruent conditions from the incongruent conditions. Comparison between the without-distractor and the with-distractor conditions did not show differed congruency effects, $t(15) = 1.16$, p = .265 (self-with-distractor = 132.42, self-without distractor = 94.72). This suggests that the amount of altercentric intrusion did not vary with the demand on the selection of relevant information.

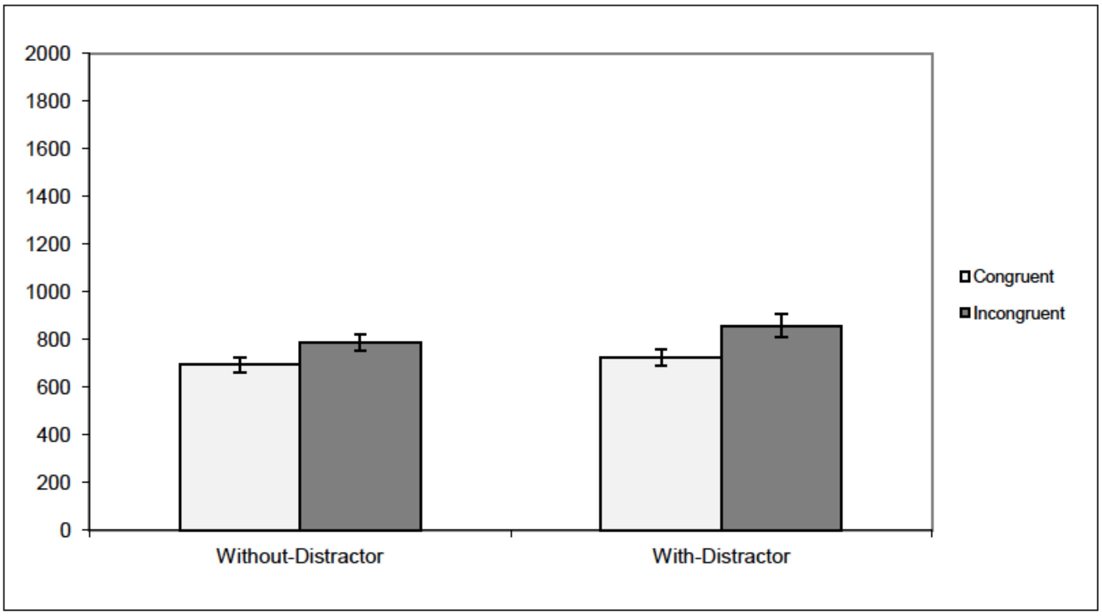


Figure 4.3 Experiment 4 self condition processing cost. Error bars represent standard errors.

4.2.3 *Discussion*

The current findings demonstrated the effects of both egocentric intrusion and altercentric intrusion when no distractors were present, replicating findings from Experiment 1 and Samson et al. (2010). These effects suggest that participants automatically compute both their own and an avatar's visual perspectives when it is unnecessary and disadvantageous for their task performance. When the distractors were present in the display, participants showed greater processing costs. Additional processing costs associated with the selection of relevant information has also been found in the visual attention literature (e.g. Simons, 2000). Despite this, effects of egocentric intrusion and altercentric intrusion were observed in the with-distractor conditions. This indicates that participants successfully selected the relevant information whilst computing both the avatar's and their own perspectives. The present findings suggest that automatic visual perspective-computation has the requisite cognitive flexibility to accommodate the selection of relevant information. Interestingly, there was a clear distinction between the processing cost participants suffered in the other-without-distractor-incongruent condition and the other-with-distractor-congruent condition. This effect indicates that egocentric intrusion arises from genuine discrepancies between participants' own perspective and the avatar's perspective, and is not merely caused by the distractors.

Before concluding that automatic visual perspective-computation can selectively attend to relevant information, it is necessary to consider the disproportionally large processing cost in the other-with-distractor-incongruent condition. Participants' success in selecting relevant information and differentiating a true perspective discrepancy from a mere distractor indicate that the large processing cost in the other-with-distractor-incongruent condition is likely to be driven by the demand to select relevant information. Nevertheless, two other potential accounts may also explain the large processing cost in the other-with-distractor-incongruent condition, and in doing so, account for the effect of egocentric intrusion observed in the with-distractor conditions. Firstly, the other-with-distractor-incongruent condition was the only condition where the avatar's perspective content merely contained distractors. Having the avatar face towards the distractors could have led participants to allocate more attention towards the distractors in this condition in comparison to the rest of the conditions. Secondly, as the avatar's perspective solely contained distractors, this condition was also the only one where participants had to translate subsets of distractors into perspective content of 'zero'. Evidence indicates that zero is conceptually more difficult to represent than other numerals (Bialystok & Codd, 2000; Wellman & Miller, 1986). Therefore, having an empty entity as the avatar's perspective content could also have demanded greater processing cost in the other-with-distractor-incongruent condition. Experiment 5 critically examines these two accounts.

4.3 Experiment 5

4.3.1 *Method*

4.3.1.1 *Participants*

Sixteen female students (mean age 20.69 years, age range 18 to 37) from the University of Birmingham participated in this study in return for study credits. All participants performed above chance on both types of perspective judgements, therefore no participants' data were excluded from the analysis.

4.3.1.2 *Design and Procedure*

A 2 x 2 x 2 x 2 within-participant design was constructed with perspective (other, self), congruency (congruent, incongruent), distractor (without distractor, with distractor), and content (number, zero, see Table 4.2) as factors. The current experiment followed identical trial sequence and blocking procedures to that of Experiment 4. The current experiment did, however, contain twice as many trials as Experiment 4 in order to maintain statistical power amongst the additional conditions.

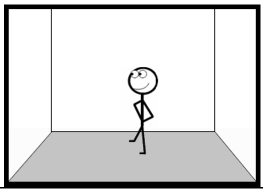
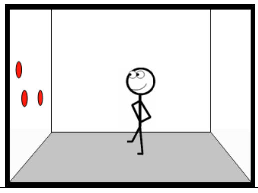
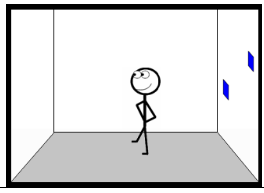
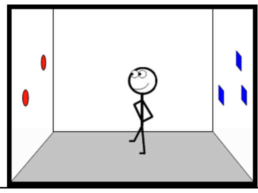
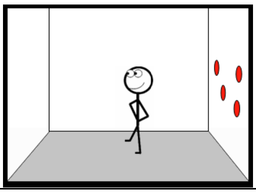
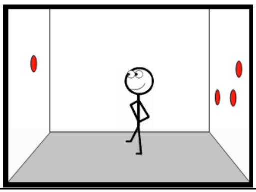
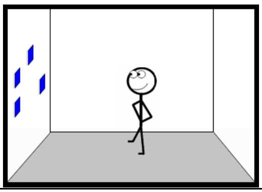
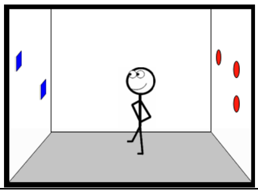
Participants were informed that they would never be asked about the distractors and that they should focus on the target stimuli only.

4.3.1.3 *Predictions*

If participants allocate extra attention to the distractors included in the avatar's perspective, then an effect of 'egocentric intrusion' should be found in the other-with-distractor-zero conditions (Pair C, see Table 4.2). In Pair C, both the participants and the avatar hold a zero perspective for the target items, eliminating any discrepancies amongst their perspective contents. Furthermore, the conditions in Pair C were labelled according to the prediction of a greater processing demand for judging the avatar's perspective when he faces the distractors. Therefore, any effect of 'egocentric intrusion' would suggest that having distractors as the avatar's perspective contents demands more processing resources. However, if the two conditions in Pair C do not differ from one another, then it would indicate that the effect of egocentric intrusion in the with-distractor conditions from Experiment 4 could not be explained by the fact that the avatar faces towards the distractors.

Secondly, if participants do find it more difficult to represent the avatar's zero perspective content compared to his number perspective content, then an overall greater processing cost should be observed in the other-without-distractor-zero conditions (Pair A). Moreover, if representing the avatar's zero perspective is demanding, then it this operation is unlikely to be included in the automatic computation of visual perspectives. In that case, there would be no effect of altercentric intrusion in Pair A, as it is unlikely that the avatar's zero perspective would interfere with participants' judgements about their own number perspective. If neither of the effects described above is observed, then this would indicate that the effect of egocentric intrusion in the with-distractor conditions from Experiment 4 could not be explained by the demand on participants to represent the avatar’s zero perspective.

Table 4.2 Experimental conditions for Experiment 5. All conditions apart from Pair C were defined with respect to the target stimuli. Pair C was defined by the predicted processing cost under the accounts of attention allocation and zero perspective representation.

	Without-Distractor		With-Distractor	
	(A) Zero	(B) Number	(C) Zero	(D) Number
Congruent				
Incongruent				

Should neither of the accounts described above prove to be correct, this would indicate that the large processing cost in the other-with-distractor-incongruent condition from Experiment 4 is not driven by the incidental elements of the design. Rather, the effect of egocentric intrusion in the with-distractor conditions in Experiment 4 indicates that automatic visual perspective-computation operates with the flexibility to select relevant information. Furthermore, the greater effect of egocentric intrusion in the with-distractor condition shows that egocentrism is likely to be a function of information selection load.

4.3.2 *Results*

For response time data, trials incorrectly responded to (4.06% of the data) and trials responded correctly but slower than two standard deviations from the mean response time (4.17% of the data) were eliminated from the data set.

4.3.2.1 *Other Condition*

A 2 x 2 x 2 repeated-measures ANOVA was conducted with congruency (congruent & incongruent), distractor (without-distractor & with-distractor), and content (number & zero) as factors. There were significant main effects of congruency, $F(1, 15) = 83.66$, $p < .001$, $\eta_p^2 = .848$, which shows that the incongruent conditions required greater processing cost compared to the congruent conditions (*M*s = 784.50 and 652.91, respectively), and distractor, $F(1, 15) = 5.42$, $p = .034$, $\eta_p^2 = .265$, which confirms that the with distractor condition requires more processing cost compared to the without distractor condition (*M*s = 743.39 and 694.02, respectively), with no significant main effect of content, $F(1, 15) = 0.00$, $p = .967$, $\eta_p^2 < .001$. There was a significant interaction between congruency and content, $F(1, 15) = 6.72$, $p = .020$, $\eta_p^2 = .309$, with no interaction between congruency and distractor, $F(1, 15) = 0.26$, $p = .619$, $\eta_p^2 = .017$, and a significant interaction between distractor and content, $F(1, 15) = 4.64$, $p = .048$, $\eta_p^2 = .236$. A significant three-way interaction was found between congruency, distractor, and content, $F(1, 15) = 15.25$, $p = .001$, $\eta_p^2 = .504$. The absence of a main effect of content indicates that the avatar’s zero perspective content does not demand more processing resources than the number perspective content. Therefore the account whereby zero perspective content is especially difficult to represent is likely to be incorrect. Two sets of 2 x 2 ANOVA with the number condition and the zero condition were conducted to explore the three-way interaction.

A 2 x 2 repeated-measures ANOVA was conducted with congruency (congruent & incongruent) and distractor (without-distractor & with-distractor) as factors in the number condition. There were significant main effects of congruency, $F(1, 15) = 78.73$, $p < .001$, $\eta_p^2 = .840$, which shows that the incongruent conditions required greater processing cost compared to the congruent conditions (*M*s = 806.91 and 631.22, respectively), and distractor, $F(1, 15) = 10.85$, $p = .005$, $\eta_p^2 = .420$, which confirms that the with-distractor condition requires more processing cost compared to the without-distractor condition (*M*s = 761.45 and 676.69, respectively). There was a significant interaction between the two factors, $F(1, 15) = 4.82$, $p = .044$, $\eta_p^2 = .243$. Planned comparisons revealed a significant effect of egocentric intrusion in the number condition when no distractor was present, $t(15) = 4.42$, $p < .001$ (see Table 4.2 Pair B: other-without-distractor-number-congruent = 613.92, other-without-distractor-number-incongruent = 739.45) as well as when the distractors were present, $t(15) = 6.84$, $p < .001$ (see Table 4.2 Pair D: other-with-distractor-number-congruent = 648.53, other-with-distractor-number-incongruent = 874.37, also see Figure 4.4). These results replicate findings from Experiment 4, suggesting that participants successfully select target stimuli and ignore distractors when making explicit judgements about an avatar's number perspectives. Additionally, comparison between the congruency effect in the with- and without-distractor conditions showed a significantly greater egocentric intrusion in the with-distractor conditions, $t(15) = 2.20$, p = .044 (other-with-distractor = 225.84, other-without-distractor = 125.53). This suggests that egocentrism is likely to be a function of the demand on the selection of relevant information.

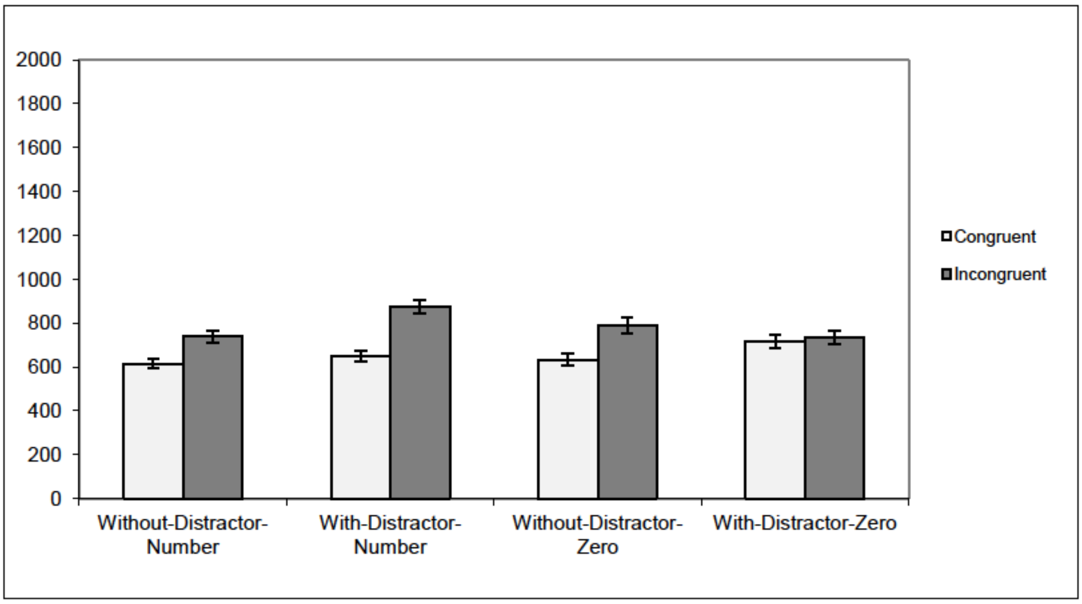


Figure 4.4 Experiment 5 other condition processing cost. Error bars represent standard errors.

A second 2 x 2 repeated-measures ANOVA was conducted with congruency (congruent & incongruent) and distractor (without-distractor & with-distractor) as factors in the zero condition. There was a significant main effect of congruency, $F(1, 15) = 12.74$, $p = .003$, $\eta_p^2 = .459$, which shows that the incongruent conditions required greater processing cost compared to the congruent conditions (*M*s = 762.09 and 674.60, respectively), no significant main effect of distractor, $F(1, 15) = 0.25$, $p = .623$, $\eta_p^2 = .016$, and a significant interaction between the two factors, $F(1, 15) = 8.07$, $p = .012$, $\eta_p^2 = .350$. Planned comparisons revealed a significant effect of egocentric intrusion in the zero condition when no distractor items were present, $t(15) = 4.90$, $p < .001$ (see Table 4.2 Pair A: other-without-distractor-zero-congruent = 633.46, other-without-distractor-zero-incongruent = 789.26). This indicates that the participants' number perspective interferes with their judgements about the avatar's zero perspective. No effect of egocentric intrusion was found when participants were solely presented with distractor items, $t(15) = 0.52$, $p = .609$ (see Table 4.2 Pair C: other-with-distractor-zero-congruent = 715.75, other-with-distractor-zero-incongruent

= 734.92). This indicates that whether the avatar gazed at distractor items or a blank wall had no effect on the amount of processing costs required when both the participants' and avatar's perspective contents for the target items were zero. Additional comparison on the scope of congruency effects observed in the zero conditions was not conducted. Since both the participants and the avatar see zero in the with-distractor-zero conditions, this does not provide a comparable effect to the without-distractor-zero conditions, where participants and the avatar held discrepant perspective contents.

4.3.2.2 *Self Condition*

A 2 x 2 x 2 repeated-measures ANOVA was carried out in the self condition, with congruency (congruent & incongruent), distractor (without-distractor & with-distractor), and content (number & zero) as factors. There were significant main effects of congruency, $F(1, 15) = 12.83$, $p = .003$, $\eta_p^2 = .461$, which shows that the incongruent conditions required greater processing cost compared to the congruent conditions (*M*s = 733.99 and 665.10, respectively), as well as distractor, $F(1, 15) = 28.32$, $p < .001$, $\eta_p^2 = .654$, which confirms that the with-distractor condition requires more processing cost compared to the without-distractor condition (*M*s = 737.49 and 661.60, respectively), and content, $F(1, 15) = 11.94$, $p = .004$, $\eta_p^2 = .443$, showing that zero conditions demand more processing cost than number conditions (*M*s = 728.76 and 670.33, respectively). There was no interaction between congruency and distractor, $F(1, 15) = 0.60$, $p = .452$, $\eta_p^2 = .038$, and no interaction between congruency and content, $F(1, 15) = 0.40$, $p = .535$, $\eta_p^2 = .026$. A significant interaction between content and distractor was found, $F(1, 15) = 9.30$, $p = .008$, $\eta_p^2 = .383$, with a trend towards a significant three-way interaction between congruency, content, and distractor, $F(1, 15) = 3.76$, $p = .072$, $\eta_p^2 = .200$. Although the three-way interaction was only marginally significant, two sets of 2 x 2 ANOVA were conducted with the number condition and the zero condition to examine the *a priori* hypothesis regarding the effects of altercentric intrusion under number perspective contents and zero perspective content.

A 2 x 2 repeated-measures ANOVA was conducted with congruency (congruent & incongruent) and distractor (without-distractor & with-distractor) in the number condition. There were significant main effects of congruency, $F(1, 15) = 19.35$, $p < .001$, $\eta_p^2 = .563$, which shows that the incongruent conditions required greater processing cost compared to the congruent conditions (*M*s = 709.78 and 630.89, respectively), as well as distractor, $F(1, 15) = 6.81$, $p = .020$, $\eta_p^2 = .312$, which confirms that the with-distractor condition requires more processing cost compared to the without-distractor condition (*M*s = 690.25 and 650.42, respectively). There was no interaction between the two factors, $F(1, 15) = 1.62$, $p = .222$, $\eta_p^2 = .097$. Planned comparisons revealed a significant effect of altercentric intrusion when no distractor was present, $t(15) = 3.28$, $p = .005$ (see Table 4.2 Pair B: self-without-distractor-number-congruent = 621.04, self-without-distractor-number-incongruent = 679.80), as well as when the distractors were present, $t(15) = 3.45$, $p = .004$ (see Table 4.2 Pair D: self-with-distractor-number-congruent = 640.73, self-with-distractor-number-incongruent = 739.77; see Figure 4.5). These results replicate the findings of Experiment 4, suggesting that participants can successfully select target stimuli and ignore distractors when making explicit judgements about their own number perspectives. Additionally, comparison between the without-distractor and the with-distractor conditions did not show differed congruency effects, $t(15) = 1.27$, p = .222 (self-with-distractor = 99.04, self-without-distractor = 58.76). This suggests that the amount of altercentric intrusion did not vary with the demand on the selection of relevant information.

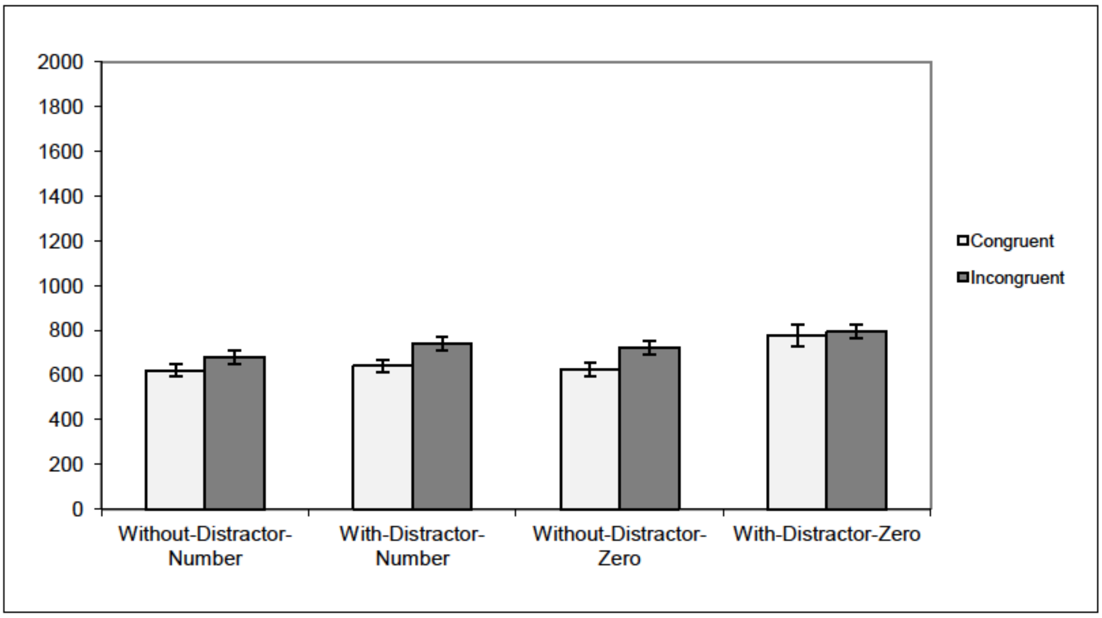


Figure 4.5 Experiment 5 self condition processing cost. Error bars represent standard errors.

A further 2 x 2 repeated-measures ANOVA was conducted with congruency (congruent & incongruent) and distractor (without-distractor & with-distractor) in the zero condition. There was a marginally significant main effect of congruency, $F(1, 15) = 3.79$, $p = .071$, $\eta_p^2 = .202$, which shows that the incongruent conditions required marginally greater processing cost than the congruent conditions (*M*s = 758.20 and 699.32, respectively), as well as a significant main effect of distractor, $F(1, 15) = 27.62$, $p < .001$, $\eta_p^2 = .648$, which confirms that the with-distractor condition requires more processing cost compared to the without-distractor condition (*M*s = 784.73 and 672.78, respectively). There was no interaction between the two factors, $F(1, 15) = 2.83$, $p = .113$, $\eta_p^2 = .159$. Planned comparisons revealed a significant effect of altercentric intrusion when no distractor was present, $t(15) = 2.66$, $p = .018$ (see Table 4.2 Pair A: self-without-distractor-zero-congruent = 623.29, self-without-distractor-zero-incongruent = 722.28). This indicates that the avatar’s zero perspective is automatically computed and interfered when participants judge their own number perspective. No effect of altercentric intrusion was found when participants were only presented with distractor items, $t(15) = 0.47$, $p = .643$ (see Table 4.2 Pair C: self-with-distractor-zero-congruent = 775.35, self-with-distractor-zero-incongruent = 794.11). This indicates that when making explicit judgements about one's own perspective, having an avatar gaze towards a number of distractor items does not generate more interference than having an avatar face a blank wall. Additional statistical test was not carried out to compare the congruency effects observed in the zero conditions due to the incomparability between the with-distractor-zero conditions and the without-distractor-zero conditions.

4.3.3 *Discussion*

The current results replicate findings from both Experiment 4 and Samson et al. (2010) by demonstrating the effects of egocentric and altercentric intrusion when participants' perspective and an avatar's perspective contain only target stimuli. Effects of egocentric intrusion and altercentric intrusion in the with-distractor-number conditions indicated that automatic visual perspective-computation accommodates the selection of relevant information. However, two alternative accounts have attempted to explain the effect of egocentric intrusion observed in these conditions. The current experiment investigated the possibility that participants might allocate additional attention to distractor items faced by the avatar. It also examined the possibility that others' zero perspective content is more difficult to represent than number perspective contents.

Firstly, if participants allocate additional attention to distractors within the avatar's perspective contents, then one would expect an effect of egocentric intrusion in the with-distractor-zero conditions. However, there is no effect of egocentric intrusion in the with-distractor-zero condition in the current results. Therefore this explanation could not account for the effect of egocentric intrusion in the with-distractor conditions.

Secondly, if the avatar's zero perspective content was difficult to represent, then these conditions should have produced an overall greater processing cost. Furthermore, one would expect no effect of altercentric intrusion when participants held a number perspective and the avatar held a zero perspective. The current results are inconsistent with this view. The other-without-distractor-zero condition did not produce a greater processing cost compared to the rest of the conditions. Moreover, there was a significant effect of altercentric intrusion in the self-with-distractor-zero conditions[7]. These findings indicate that participants' representation of the avatar's zero perspective content was no different from their representation of the avatar's number perspective content. Moreover, the avatar's zero perspective was computed so efficiently that it caused interference with participants' own number perspective.

Finally, the results of the current experiment are consistent with that of Experiment 4. Both set of results reveal an interaction between distractor and congruency in the with-distractor-number conditions. When the scope of egocentric intrusions from the with- and without-distractor conditions were directly compared, both Experiments 4 and 5 revealed significantly greater egocentrism when participants were under the information selection load. The elimination of two

[7] An alternative account for the effect of altercentric intrusion in this condition would be that in the congruent condition, participants computed no dots, whereas in the incongruent condition, participants had to compute one set of dots. Hence the larger processing cost in the incongruent condition could purely reflect the processing cost for enumerating one set of dots, rather than the present interpretation of altercentric intrusion. To address this account, an additional comparison between the self-without-distractor-number-congruent condition and the self-without-distractor-zero-incongruent condition was conducted. Both conditions required participants to enumerate a single set of dots for computing their own perspective. The comparison revealed that the self-without-distractor-zero-incongruent condition demanded significantly more processing cost than the self-without-distractor-number-congruent condition $t(15) = 3.48$, $p < .005$. This suggests that the avatar's zero perspective did in fact impair participants' judgements of their own number perspective.

alternative explanations leads to the interpretation of egocentrism to be a function of the demand to select relevant information. This is consistent with Epley et al. (2004), which indicate that overcoming interference from one's own perspective is a demanding process, such that when there are concurrent demands, adults show increased egocentrism.

4.4 General Discussion

4.4.1 *Summary of Current Findings*

Two experiments in the current chapter demonstrate that automatic visual perspective-computation operates with the flexibility to select relevant information. Experiment 4 indicated that participants were successful in extracting target stimuli from distractor items. Experiment 5 eliminated alternative accounts that have been offered in explanation of the effect of egocentric intrusion in the with-distractor conditions. The results of Experiment 5 reveal that having an avatar face towards a number of distractor items does not produce more processing cost than having an avatar face away from distractor items. Furthermore, the avatar's zero perspective does not produce a greater processing cost. The elimination of the two alternative accounts confirmed that automatic visual perspective-computation accommodates selective computation of relevant perspective contents.

In the current chapter, participants always responded to a single category of stimuli (i.e., red dots) whilst constantly ignoring the other category (i.e., blue squares). Therefore one could argue that the demand required to produce one consistent information selection strategy only shows a relatively limited amount of flexibility. This argument shows consistency with the prediction that automatic visual perspective-taking only operates with a limited amount of flexibility. Nonetheless, further investigation with increased cognitive demand is necessary to establish a clear operational limit in automatic visual perspective-taking. This could be achieved by switching the category that participants are required to judge on a trial-to-trial basis, eliminating the possibility of employing one information selection strategy for the entire task.

4.4.2 *Altercentric Intrusion*

The current investigation revealed that the effect of altercentric intrusion cannot be explained by the spatial layout of participants' perspective-computation. A standard design for the incongruent condition features two subsets of dots, presenting a disrupted gestalt structure in the array of dots in participants' perspective-computation. This is in contrast to the layout in the congruent condition, where participants only compute a single subset of dots. Experiment 5 showed that a condition in which participants see a single set of dots and the avatar sees a blank wall did not eliminate the effect of altercentric intrusion. This is despite the fact that the disturbed gestalt was eliminated. A similar argument has also been made by Samson et al. (2010). In their third experiment, a non-social stick figure replaced the avatar, producing the same disruption to the gestalt structure and identical spatial layout to the standard incongruent condition. However, the stick figure did not generate any effect of 'altercentric intrusion'. Samson et al. suggest that merely having a disturbed gestalt structure is not sufficient to produce effects of altercentric intrusion. The current findings further indicated that even with an undisturbed gestalt structure, participants still show automatic computation of the avatar's visual perspective.

4.4.3 *Conclusion*

In summary, the current chapter showed that automatic visual perspective-computation operates with the flexibility to select relevant information. The current findings not only demonstrate that automatic visual perspective-computation indeed accommodates processes with ecological values; it also provides insight into the real world operation of efficient mindreading.

CHAPTER 5 INTRODUCTION TO VISUAL WORKING MEMORY ENCODING FOR AGENT AND OBJECT INFORMATION

5.1 *Rationale for Chapter 6 and 7*

In Chapters 3 and 4, I demonstrated some limitations as well as the flexibility of automatic Level-1 visual perspective-computations. Participants revealed higher processing costs when they held a different perspective from that of the avatar. This was the case when participants judged their own perspective as well as when they judged the avatar's perspective. The automatic computation of others' visual perspective raises the possibility that individuals might process visual content observed by other agents differently from visual content unobserved by other agents. As briefly described in Chapter 1, automatic processes such as shared task representation, gaze attentional cueing, and visual perspective-computation are activated even in scenarios where their activation hinders concurrent task performance. These findings suggest that participants spontaneously allocate attention to social stimuli, such as eye gaze. However, this raises the question of whether the modulated attention allocation might also affect encoding of information that contains social stimuli. Based on previous findings, one prediction is that the additional attention allocated to objects observed by agents will lead to improved encoding of that information. Conversely, given that eye gaze commands participants' attention, it is also possible that it will not only fail to enhance further processing of gazed upon objects, but could in fact impair the processing of such objects and hinder performance on other social cognitive tasks.

5.2 *Sociocommunicative Cues*

5.2.1 *The Development of Reading Sociocommunicative Cues*

As social beings, humans pay special attention to social stimuli from a very young age. Studies have demonstrated that human faces and human face-like stimuli attract infants' attention very shortly after birth. One-month-olds turned their heads to greater angles to follow a face stimulus compared to a scrambled face stimulus that contains an identical collection of face features rearranged within the same space (Johnson, Dziurawiec, Ellis, & Morton, 1991; Slater & Butterworth, 1997). At five to seven weeks of age, when infants scan a face, they show more fixations in the eye region compared to other regions of a face, such as nose and mouth (Maurer & Salapatek, 1976). This preference for the eye region is consistent across static, dynamic, and talking face stimuli (Haith, Bergman, & Moore, 1977). When four-month-old infants observe objects that they have not seen for themselves but were merely seen by another person, they show shorter looking time that indicate familiarity to the objects (Reid & Striano, 2005). This implies infants' sensitivity to agents' object-oriented gaze. From their first birthday, infants appear to understand the relation between an agent and the object she looks at. Infants displayed surprised looking response when an agent no longer looks at the object she originally looks at. This demonstrates infants' sensitivity to the relations between agents and objects (Woodward, 2003). Furthermore, Moll and Tomasello (2006) showed that when an adult who could only see one of two objects asked 24-month-olds 'Where is it? I cannot find it. Can you give it to me?', they handed over the object that was occluded from the adults' sight instead of another object that could be seen by both of them. O'Neill (1996) demonstrated that when 24-month-old infants requested an adult for help to retrieve an object, they produced more information about the object (name the name of the object, name or gesture the new location of the object) when the adult was naïve of the object's new location. These studies suggest that at 24-month of age, infants are already able to adjust their own behaviour according to others' visual access and knowledge state. There is abundant evidence which speaks to human infants' special attention to sociocommunicative stimuli. The early-development and continued practice of such operations might provide explanations for the automatic attention allocation to sociocommunicative stimuli in adulthood.

5.2.2 *Adults' Processing of Sociocommunicative Cues*

Friesen and Kingstone (1998) demonstrated that adults reflexively shift attention to uninformative eye gaze in a Posner cueing paradigm (Posner, 1980). In this study, participants were required to detect, localise, or identify alphabet letters that came on a screen whilst fixating their gaze on a central face stimulus. The face either gazed straight ahead or gazed to one side such that eye gaze direction was either congruent or incongruent with the location of the target letter. Participants were informed that the eye gaze direction was not predictive of the target letter's location. Friesen and Kingstone found that when the gaze direction was congruent with the target letter's locations, participants responded faster in detection, localisation, and identification tasks. Driver et al. (1999) found that participants were slower to respond to incongruent trials even when a target letter was four times more likely to appear on the incongruent than congruent side. Other sociocommunicative cues have also been found to produce similar effects. For example, head orientations (Langton & Bruce, 1999) and non-human social symbols, such as arrows (Tipples, 2002), produce quicker responses when congruent with the target object's location.

In terms of the scope of gaze cueing effects, objects which are merely gazed at by a central face stimulus are rated more preferable by participants than objects not gazed at (Bayliss et al., 2006). Furthermore, the size of the preferential rating effect is modulated by the positive and negative emotional expressions displayed on the central face stimulus. Importantly, these modulatory effects are only observed when the gaze directions are congruent with object locations but not when they are incongruent with object locations (Bayliss et al., 2007). Despite abundant evidence demonstrating automatic attention shift in relation to gaze processing, little is currently known about the encoding and maintenance of the information attended. To address this, one must look to the literature on visual working memory and in particular to paradigms that might make it possible to test the effects of gaze on information encoding.

5.3 *Visual Working Memory*

Visual working memory is typically regarded as a bridge between perceptual inputs and the formation of conceptual representations, and as such is thought to play a crucial role in joining high- and low-level processes (Jiang, Makovski, & Shim, 2009). Visual working memory in adults has a capacity of four items (Scarborough, 1972; Sperling, 1960), with ceiling performances for set sizes up to four items, and a drastic drop when the set sizes exceed four items (Pashler, 1988). Change detection paradigms are commonly used to test visual working memory. Trial sequences begin with a brief sample stimulus (typically ranging from 100 ms to 1000 ms), in which participants have the opportunity to encode information from the visual scene. The sample stimulus is followed by a retention interval that lasts up to 10 seconds. Finally, participants see a test display, which is either the same or different from the sample stimulus. Participants must judge whether the test display contains any changes from the sample display. Correct responses require participants to accurately encode the sample display, retain the visual information in their visual working memory, and compare the encoded information with the test display.

Using a change detection paradigm, Luck and Vogel (1997) demonstrated that participants are as accurate to detect conjunction of features as they are to detect single features. In the single-feature condition, participants were instructed to look for changes in the object colours or object orientations amongst an array of four objects. The changes in object colours or object orientations were specified before hand, hence participants were only required to encode four features, one from each object. In the conjunction condition, participants were instructed that *either* object colours or object orientation could change; however, it was not specified beforehand whether object colour or orientation would change. Participants were required, therefore, to encode eight features of the four objects. Luck and Vogel found that participants were as accurate in the single-feature condition as in the conjunction condition for array sizes of two, four, and six objects. Furthermore, participants were as accurate in the conjunction condition as in the single-feature condition even when they were required to encode up to four different features of each object. This suggests that visual working memory allows for each of the encoded objects to comprise four different features. Luck and Vogel argued that participants automatically encode multiple features of the same object as one integrated perceptual unit rather than encode all individual features separately. However, others argue that the binding of features with objects may not be an entirely effortless process. Wheeler and Treisman (2002) found that when participants were explicitly instructed to bind multiple features with associated objects, their accuracy was worse than that of encoding the same number of features which were not bound with specific objects. Wheeler and Treisman concluded that binding separate features with the same object does produces additional processing costs in comparison to merely encoding the same features without binding them with the associated objects. Although the current thesis does not directly speak to this debate, I will examine adults' encoding for agent and object information both with and without explicit instructions to bind agent and object information in order to ensure the consistency across results.

5.4 *Visual Working Memory in Social Domain*

Wood (2008) investigated visual working memory for agent and action information using a variant of the change detection paradigm. Participants were presented with a sample sequence comprising three different agents each performing a different action. Subsequently, participants were shown a test display with an agent performing an action. Participants were required to judge whether certain aspects of the test display were identical to the sample sequence. Similar to Wheeler and Treisman's (2002) design, there were three main conditions: single condition, either condition, and binding condition. In the single condition, participants were instructed to detect changes in one aspect of the display—either the agents' identity (i.e., colours of their clothing) or the actions performed by the agents—therefore three features had to be encoded. In the either condition, participants were instructed to detect changes in either agents or actions, therefore a total of six features had to be encoded. In the binding condition, participants were instructed that agent and action information could be mismatched in the test display and that they should encode the correct combination of agents and the associated actions. This required participants to not only encode six features just as in the either condition, but also to bind agents and actions in the correct pairs. The results indicated that participants' visual working memory capacity was compromised when agent and action information had to be paired, suggesting that this information is not naturally encoded as an integrated unit. However, this effect was ameliorated when 'external visual input' was provided (i.e., agents occupying distinct spatial positions as opposed to all appearing in the centre of the screen; short temporal gaps between the presentations of each pair of agent and action; agents with physical features more distinct than colours of their clothing). Under these condition, participants' visual working memory capacity in the binding condition did not significantly differ from the either condition, suggesting that external cues, as shown by Wood: agents' spatial positions, agents' physical features, and temporal gaps, modulate participants' encoding and binding of information in the social domain.

5.5 *Current Work on Agent and Object Information Encoding*

The main aim of Chapters 6 and 7 is to verify whether agents' eye gaze affects visual working memory encoding and maintenance of agent and object information. Previous evidence suggests that participants automatically shift attention in response to agent eye gaze directionality. It is not known, however, whether information encoding is facilitated or hindered by the altered attention allocation. Both Chapters 6 and 7 will address this question. In Chapter 6, a variant of the change detection paradigm (Wood, 2008) is employed to test whether participants naturally encode sequentially presented agents and objects in pairs. Moreover, I examine whether agents' object-oriented gaze affects binding in visual working memory. In Chapter 7, Luck and Vogel's change detection paradigm (1997) is employed in examining a different temporal sequence with agents' participant-oriented gaze integrated into the design. Evidence suggests that humans are sensitive to both object-oriented and participant-oriented gaze early in development[8]. It is likely, therefore, that both object-oriented gaze and participant-oriented gaze may modulate information encoding in adulthood.

[8] The literature on participant-oriented gaze will be reviewed in Chapter 7, as this literature is not directly relevant to the investigation in Chapter 6.

CHAPTER 6 VISUAL WORKING MEMORY FOR RETAINING PAIR-WISE PRESENTATION OF AGENT AND OBJECT INFORMATION

In Experiments 6, 7, and 8, I adopt Wood's (2008) design to examine participants' encoding and maintenance of agent and object information. The first aim of these experiments is to examine whether the distinction between agents and objects can serve as the basis for separate representation in visual working memory. A second aim is to examine whether the eye gaze of agents can affect agent-object binding. These hypotheses will be described in further detail in the predictions section of Experiment 6 (Section 6.1.1.3).

6.1 Experiment 6

This experiment examines participants' visual working memory encoding accuracy under conditions in which they are required to encode pair-wise presentations of agent and object information where the agents are gazing towards and gazing away from the objects. The motivation for conducting such an experiment is that if agents' eye gaze influences the binding of agent and object information, then participants may show stronger binding of agent and object information when they observe agents gazing towards objects.

6.1.1 *Method*

6.1.1.1 *Participants*

Twenty-four students (20 female, mean age 19.38, age range 18 to 24) from the University of Birmingham participated in this study in return for study credits. All participants had normal colour vision and normal or corrected-to-normal visual acuity. All of the participants' visual working memory proportion correct scores lay within

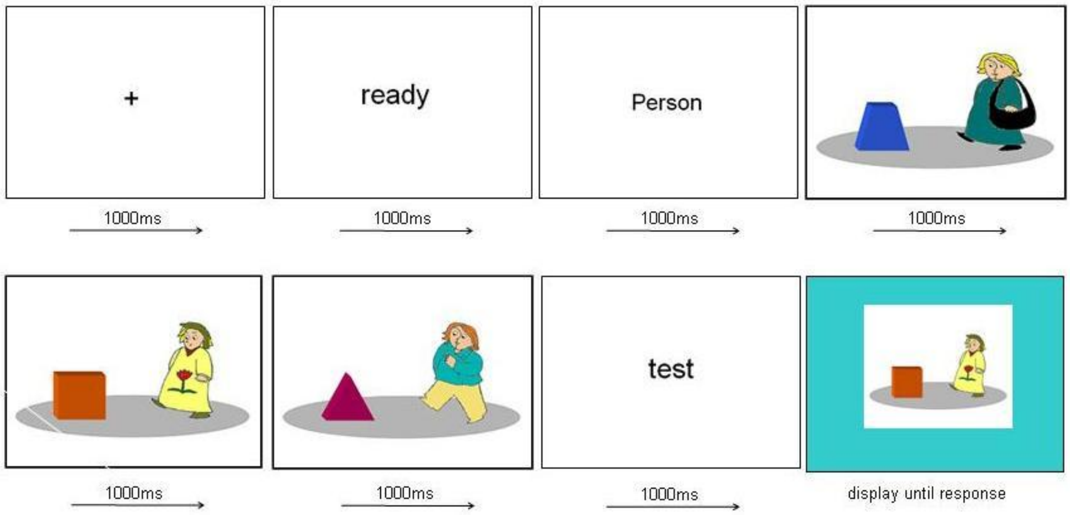
2.5 standard deviations of the mean score in the two gaze conditions, therefore no participant's data were excluded from the analysis.

6.1.1.2 *Design and Procedure*

A 5 x 2 within-participant design was constructed with condition (agent-only, object-only, either-agent, either-object, pairing) and gaze (gaze, no-gaze) as the factors. The condition factor corresponded to the aspects of the test picture that could differ from the sample pictures. The gaze factor indicated whether the agents looked towards or away from the objects.

In each trial, participants were presented with three pairs of agents and objects (see Figure 6.1), matching the set size of Wood's (2008) design. The displays subtended 5.5º (height) x 7.2º (width) in the centre of a computer screen surrounded by black backgrounds. Participants observed three different agents presented with three different objects. The agent-object pairings were presented serially. Participants' task was to judge whether a given aspect of the test picture was different from one of the three sample pictures shown. Matching the procedure Wood employed, participants engaged in a concurrent articulation task to rule out any potential use of verbal encoding strategies (Besner, Davies, & Daniels, 1981). At the start of each trial, a 'ready' cue appeared. This was followed by a condition cue, which explicitly indicated the aspects of the test picture that might differ from the sample pictures. The cues included 'Person', 'Object', 'Either', and 'Pairing'. Three sample pictures then followed this condition cue, each picture depicting pair-wise presentation of an agent and an object. Participants then saw a 'test' cue, which was followed by a test picture. Each of the displays described above was onset for 1000 ms, except for the test picture, which was displayed until participants made a response.

(a)



(b)

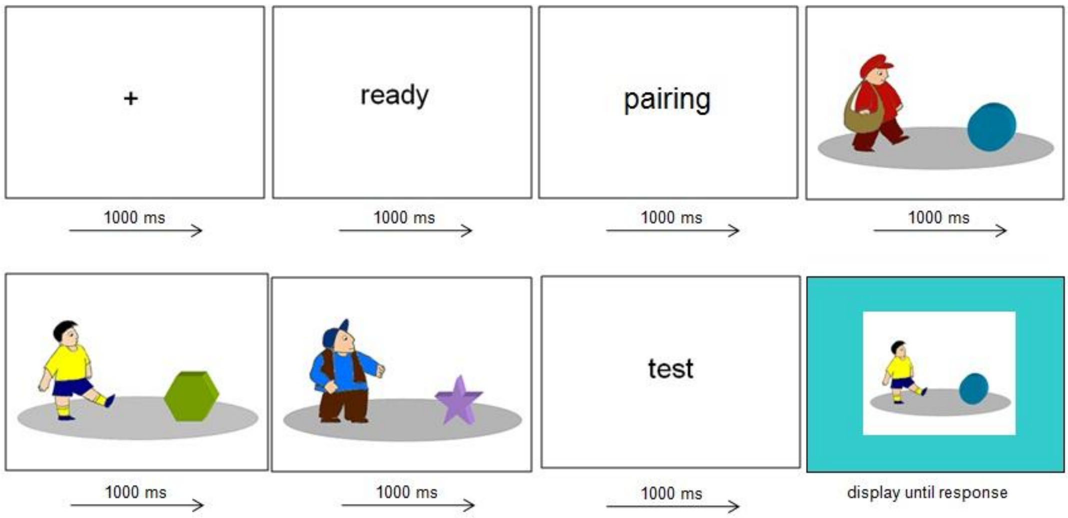


Figure 6.1 Trial sequence in Experiment 6. The agents were selected from a pool of eight different agents all with distinct physical features. Agents of both genders were included. All agents featured the same set of eyes. The objects were also selected from a pool of eight objects featuring distinct shapes and colours. (a) An example of a no-change trial from gaze-agent condition. (b) An example of a change trial from the gaze-pairing condition.

On half of the trials, the test picture contained a change, on the other half of the trials, the test picture was identical to one of the three sample pictures from the same trial. Amongst trails that contained a change, the person, object, either, and pairing trials occurred with equal frequency. Following Wood’s design, half of the change trials with 'either' changes contained changes to an agent (either-agent trials), the other half contained changes to an object (either-object trials). On half of the trials, the agents looked towards the objects (gaze condition); on the other half of the trials, the agents looked away from the objects (no-gaze condition). The three sample pictures within a trial were always from the same gaze condition. Each participant completed 250 trials in total across a practice block of 10 trials and eight testing blocks of 30 trials. Across all trials the target sample picture was presented either first, second or third with equal probability, controlling for a recency effect (e.g., Baddeley & Hitch, 1993). All conditions were mixed within blocks, presented in a pseudo-random order with the caveat that participants never encountered three trials from the same condition in a row. Participants indicated whether the test picture was one of the sample pictures by clicking on a computer mouse, using their dominant hand. To ensure that participants do not employ verbal encoding strategies, a concurrent articulation task was implemented (Besner et al., 1981). Participants responded to the concurrent articulation task with their non-dominant hands on a computer keyboard. They were required to judge whether two letters shown at the end of a trial were the letters they had been rehearsing from the start of the trial.

6.1.1.3 *Predictions*

If agents and objects are represented separately in visual working memory, then these stimuli will not compete for the same memory store even when participants are required to attend to both simultaneously (Wood, 2008). Under this scenario, participants should show similar encoding accuracies in the agent-only condition (in which participants only needed to attend to agents) and the either-agent condition (in which they had to attend to agents and objects). For the same reasons there should be no difference between the object-only condition and the either-object condition.

Furthermore, the encoding accuracy in the pairing condition should be reduced in comparison to the either condition. This follows from Woods' demonstration that binding information together that is not by default encoded in a bound form requires cognitive effort. Such demand is often reflected in a compromised encoding accuracy in the pairing condition (Wheeler & Treisman, 2002; Wood, also see Sections 5.3 & 5.4 in the present thesis).

Alternatively, if participants bind agent and object information together in visual working memory, then they should show similar encoding accuracies in the pairing condition and the either condition, as binding would no longer be effortful (Wood, 2008). Moreover, bound representation of information leads to competition for the same memory store between the two types of information, therefore a compromised encoding accuracy should be found in the either condition compared with the only condition. Specifically, if agents' object-oriented gaze leads to the binding of agent and object information, then the pattern of results for binding should be more pronounced in the gaze condition compared to the no-gaze condition. I have labelled the hypothesised binding of agents and objects through gaze direction as 'social binding effect'.

6.1.2 *Results*

Proportion correct for each condition was computed. Participants responded to 0.76 of the change trials correctly, and 0.71 of the no-change trials correctly. As described above, certain differences across the five conditions would indicate whether participants represented agent and object information separately or jointly when this information is retained in visual working memory. Differences between the two gaze conditions would indicate whether agents' gaze lead to stronger binding of agent and object information.

A 5 x 2 repeated-measures ANOVA was conducted with condition (agent-only, object-only, either-agent, either-object, & pairing) and gaze (gaze & no-gaze) as factors. There was a significant main effect of condition, $F(4, 92) = 7.98$, $p < .001$, $\eta_p^2 = .258$, with no main effect of gaze, $F(1, 23) = 0.75$, $p = .396$, $\eta_p^2 = .032$, and no interaction between condition and gaze, $F(4, 92) = 0.47$, $p = .761$, $\eta_p^2 = .020$; see Figure 6.2. Planned comparisons were carried out on the main effect of condition in order to establish whether agent and object information was represented separately or jointly in visual working memory store. A significant difference between the agent-only and the either-agent conditions was found, $t(23) = 4.69$, $p < .001$ (agent-only = .826, either-agent = .723), with a marginally significant difference between the object-only and the either-object conditions, $t(23) = 2.25$, $p = .035$[9] (object-only = .763, either-object = .717). There was no significant difference between the either-agent and the pairing conditions, $t(23) = 1.58$, $p = .128$ (either-agent = .723, pairing = .755), and no significant difference between the either-object and the pairing conditions, $t(23) = 1.90$, $p = .086$ (either-object = .717, pairing = .755). This pattern of results was largely consistent with that predicted by the binding hypothesis. However, the effect observed was not driven by agents' object-oriented gaze.

[9] Bonferroni correction was applied to adjust the significance level for four pairs of comparisons to .0125. This same correction was applied for Experiments 6, 7, and 8.

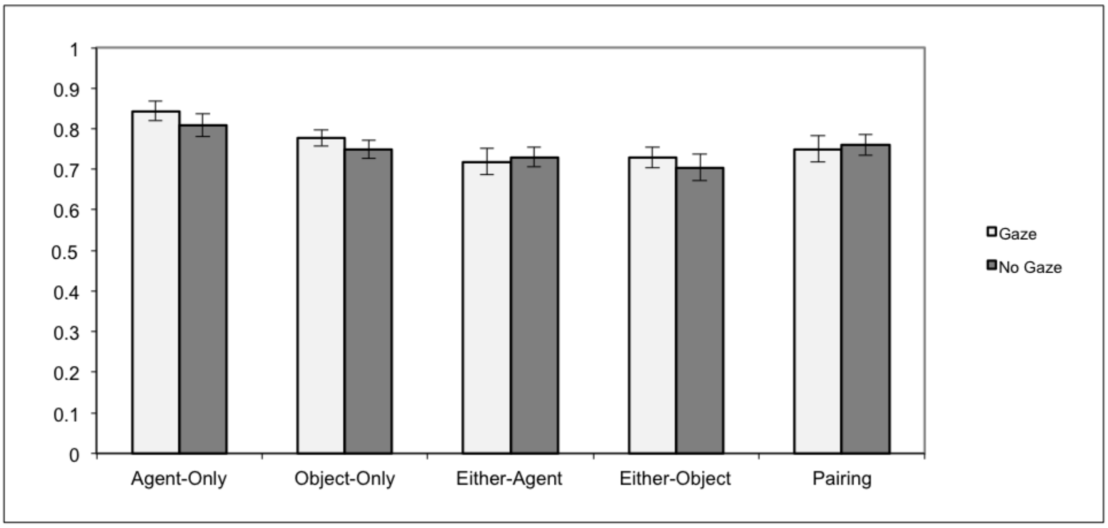


Figure 6.2 Experiment 6 results divided by condition and gaze. The figure was graphed with proportion correct. The error bars represent standard errors.

6.1.3 *Discussion*

The current results indicated that participants retained pair-wise presented agent and object information in a bound form. Participants' encoding accuracies largely matched the descriptions of binding above, showing differences between the agent-only and the either-agent conditions, a near significant difference between the object-only and the either-object conditions. This indicates that agent and object information is likely to compete for the same memory store. Furthermore, the either-agent, either-object, and the pairing conditions showed similar levels of encoding accuracy, suggesting that little effort is required to bind agents with objects.

The reduced difference between the object-only and either-object conditions, compared to the agent-only and either-agent conditions, could be explained by a general trend of lower encoding accuracy for object information compared to agent information. A similar trend was also observed in Wood (2008), where participants consistently showed greater encoding capacity for agent information than action information. Although the current study proposes no specific predictions regarding participants' encoding performance for social versus non-social information, the trend of greater encoding accuracy for agent information is consistently observed across the present chapter.

The absence of a social binding effect indicates that although participants represent agent and object information in a bound form, this effect is not driven by agents' eye gaze. One possibility is that participants simply encoded the pair-wise presentation of agent and object information together. The other possibility is that the agents' distinct physical features prompted participants to integrate agents and objects as bound units. Wood (2008) demonstrated that when agents were merely differentiated by the colour of their clothing, participants represent agent and action information separately. However, when the agents had distinctive individual identities, participants bound agent and action information together. In order to see whether the bound agent and object information might be obscuring a social binding effect, agents' distinct physical features were removed from the design of Experiment 7.

6.2 Experiment 7

In the present experiment, agents were differentiated solely by the colour of their clothing. If the binding effect observed in Experiment 6 was not caused by external cues such as agents' distinct identity, then agent and object information should be represented in a bound form. However, if the binding effect previously observed was not a result of the pair-wise presentation of agents with objects, then the removal of external cues should reveal separate representation of agent and object information. Furthermore, if social binding effect was previously obscured by bound representation of agent and object information, then the present manipulation may reveal a more pronounced binding effect in the gaze condition. In addition to the rationale above, a simple priming procedure was employed to highlight the agents' gaze directions. This was to ensure that the absence of a social binding effect in Experiment 6 was not merely due to the manipulation of eye gaze going unnoticed.

6.2.1 *Method*

6.2.1.1 *Participants*

Sixteen students (11 female, mean age 22.81, age range 18 to 34) from the University of Birmingham took part in the study in return for either study credits or a small honorarium. All participants had normal colour vision and normal or corrected-to-normal visual acuity. All participants had scores within 2.5 stand deviations of the overall means in the two gaze conditions, therefore no participant's data were excluded from the analysis.

6.2.1.2 *Design and Procedure*

The conditions and procedure in the current experiment were identical to those of Experiment 6 with the following exceptions. Firstly, in order minimise the contribution of external cues on participants' information binding, agents' identities were now coded solely by the colour of their clothing, and objects were differentiated solely by their shapes. Secondly, to ensure the current procedures are closely matched to Wood's (2008) design, the trials were now presented in blocks of agent condition, object condition, either condition, and pairing condition. Finally, since the conditions were blocked, the trial sequence was also altered so that participants were presented with agent eye gaze direction cues (either 'Looking Towards' or 'Looking Away', see Figure 6.3) where they had previously been presented with a condition cue.

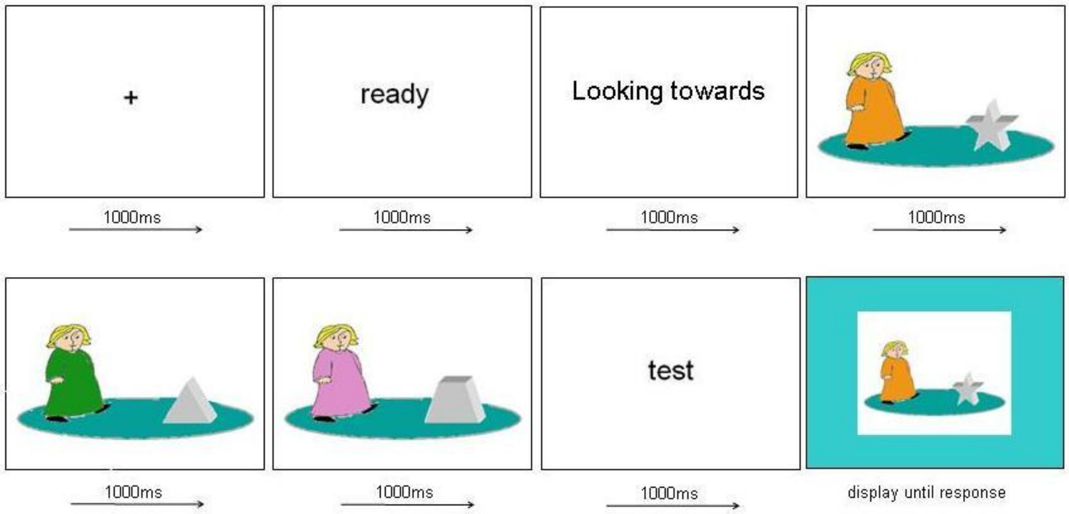


Figure 6.3 Experiment 7 trial sequence. This is an example of a no-change trial from gaze condition.

6.2.1.3 *Predictions*

Firstly, if the binding effect observed in Experiment 6 is not a function of external cues, then participants' encoding accuracies will correspond to that predicted by the binding hypothesis. However, if the binding effect observed in Experiment 6 is a function of external cues, then spontaneous binding of agent and object information should not occur in the current experiment. Secondly, if the absence of a social binding effect in Experiment 6 was due to participants failing to notice agents' gaze directions, then the current priming manipulation should help overcome this by highlighting agents' gaze directions. In which case, the gaze condition should produce a stronger binding effect than the no-gaze condition. However, if agents' eye gaze plays no role in the binding of agents with objects, then there should be no difference between the gaze and the no-gaze conditions even when emphasis was placed on agents' eye gaze direction with a prime.

6.2.2 *Results*

Participants responded to 0.82 of the change trials correctly, and 0.80 of the no-change trials correctly. A 5 x 2 repeated-measures ANOVA was conducted with condition (agent-only, object-only, either-agent, either-object, & pairing) and gaze (gaze & no-gaze) as factors. There was a significant main effect of condition, $F(4, 60) = 18.30$, $p < .001$, $\eta_p^2 = .550$, with no main effect of gaze, $F(1, 15) = 0.53$, $p = .478$, $\eta_p^2 = .034$, and no interaction between condition and gaze, $F(4, 60) = 0.88$, $p = .483$, $\eta_p^2 = .055$; see Figure 6.4. Planned comparisons were conducted to explore the main effect of condition. A significant difference between the agent-only and the either-agent conditions was found, $t(15) = 5.35$, $p < .001$ (agent-only = .943, either-agent = .734), with a significant difference between the object-only and the either-object conditions, $t(15) = 4.77$, $p < .001$ (object-only = .888, either-object = .667). There was no significant difference between the either-agent and the pairing conditions, $t(15) = 0.20$, $p = .847$ (either-agent = .734, pairing = .742), and no significant between the either-object and the pairing conditions, $t(15) = 1.87$, $p = .082$ (either-object = .667, pairing = .742). This pattern of results was consistent with that expected under the binding hypothesis.

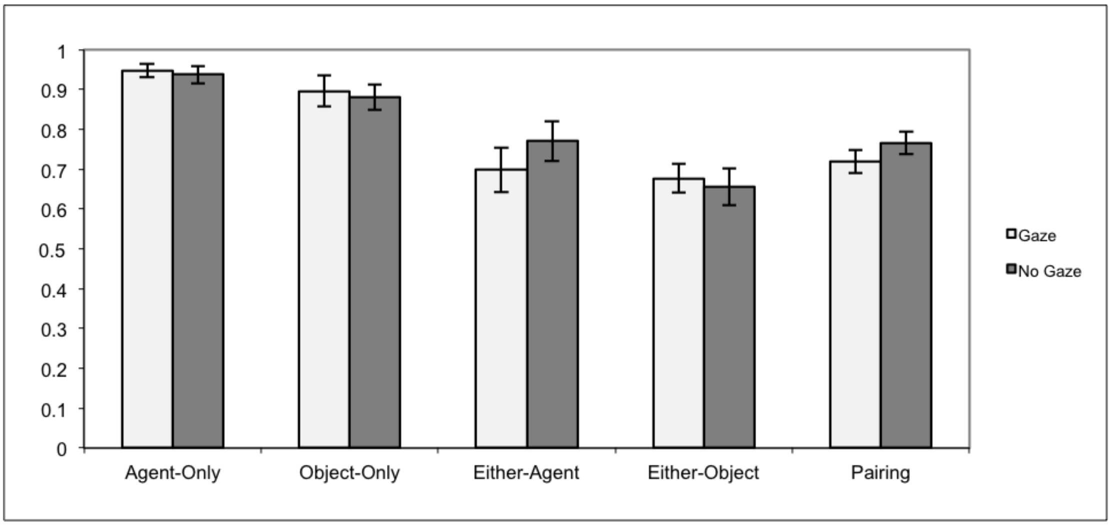


Figure 6.4 Experiment 7 results divided by condition and gaze. The figure graphed memory encoding accuracy. The error bars represent standard errors.

6.2.3 *Discussion*

The current experiment once again showed bound representation of agent and object information. The removal of external cues did not prevent participants from binding agents with objects. This indicates that participants are likely to represent the pair-wise presentation of agent and object information as a bound unit. Furthermore, the highlighting of the agents' gaze directions did not lead to stronger binding in the gaze condition. This suggests that the lack of a social binding effect in Experiment 6 was unlikely to result from participants failing to notice the eye gaze direction of the avatar. Before concluding that agents' gaze does not lead to binding of agent and object information, a further experiment is required to address two low-level explanations for the absence of a social binding effect. Firstly, the agent stimuli employed in the current experiment had realistic body features, including proportionally small eye-features. A related study on the visual attentional processes of participant-oriented gaze found a different pattern of results using realistic human stimuli with realistic eye-features than when using cartoon-like avatars which contained exaggerated eye-features (see Note 4 in Conty, Tijus, Hugueville, Coelho, & George, 2006). Relatedly, a potential problem with the current agent stimuli is that they might provide more salient information about agents' body orientation rather than gaze direction. In order to check for stimulus-specific effects, in Experiment 8, a matchstick figure with salient eye-features and less prominent body parts replaced the current realistic agents. Secondly, although the priming manipulation did not produce a social binding effect, participants may have failed to process both the primes and the agents' gaze. In order to ensure that agents' gaze direction is not ignored, participants were required to explicitly judge agents' gaze directions in Experiment 8.

6.3 Experiment 8

In the present experiment, a matchstick figure with salient eye-features and less prominent body parts replaced the agent-stimuli from Experiments 6 and 7. The purpose of this was to investigate the possibility of stimulus-specific effects *a la* Conty et al. (2006). To ensure that participants process agents' gaze directions, they explicitly judged the gaze direction of each agent in each display. These manipulations rule out low-level explanation for the absence of a social binding effect under the current paradigm.

6.3.1 *Method*

6.3.1.1 *Participants*

Twenty-four students (20 female, mean age 20.42, age range 18 to 31) from the University of Birmingham participated in this study in return for study credits. All participants had normal colour vision and normal or corrected-to-normal acuity. The scores from two participants were greater than 2.5 stand deviations away from the mean in the two gaze conditions. Their data were replaced before the analysis was carried out.

6.3.1.2 *Design and Procedure*

The conditions and procedure of Experiment 8 were identical to those of Experiment 7 with the following exceptions. Firstly, the agents' gaze target was manipulated within trial. This was achieved by having the objects appear on a wall either in front or behind the agent (see Figure 6.5). The agent's body orientation remained the same throughout each trial, with the agent facing towards the left and towards the right equally frequently. Secondly, participants were required to judge whether the agent sees an object as soon as each display was presented. Participants made a computer keyboard response using their non-dominant hands. If a response was not detected within 1500 ms of the onset of a display, then the display would timeout and a beep noise would inform participants of the judgement they should have made. The concurrent articulation task was excluded from the procedure to match the task demand of Experiments 6 and 7.

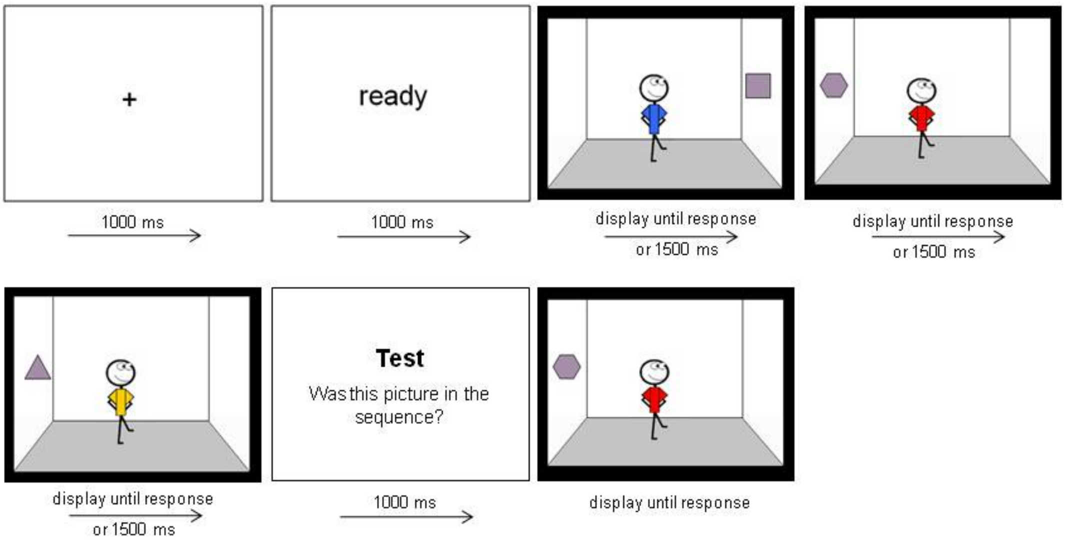


Figure 6.5 Experiment 8 trial sequence. This is an example of a no-change trial.

6.3.1.3 *Predictions*

As before, pair-wise presented agents and objects were expected to be represented as a bound unit in visual working memory. In addition, should a social binding effect have been obscured by the nature of the specific agent-stimuli previously used, then the new agent-stimuli with exaggerated eye-features may be expected to produce a social binding effect. Finally, if the absence of a social binding effect in Experiments 6 and 7 was due to participants failing to process the agents' gaze direction, then the stipulation that participants should judge whether or not the agents see the object should result in a social binding effect.

6.3.2 *Results*

Participants responded to 0.79 of the change trials correctly, and 0.81 of the no-change trials correctly. A 5 x 2 repeated-measures ANOVA was conducted with condition (agent-only, object-only, either-agent, either-object, & pairing) and gaze (gaze & no-gaze) as factors. A significant main effect of condition was found, $F(4, 92) = 6.27$, $p < .001$, $\eta_p^2 = .214$, with a significant main effect of gaze, $F(1, 23) = 9.42$, $p = .005$, $\eta_p^2 = .291$, revealing more accurate encoding in the gaze condition compared to the no-gaze condition (*M*s = .804 and .754, respectively). There was no interaction between condition and gaze, $F(4, 92) = 1.24$, $p = .300$, $\eta_p^2 = .051$; see Figure 6.6. The significant main effect of gaze indicates that agents' object-oriented gaze did promote overall encoding accuracy. However, the absence of a significant interaction between gaze and condition suggests that this effect does not generate a stronger binding effect in the gaze condition. Therefore, planned comparisons were carried out as previously, on the main effect of condition in order to establish whether agent and object information was represented separately or jointly in visual working memory store. A significant difference between the agent-only and the either-agent conditions was found, $t(23) = 3.23$, $p = .004$ (agent-only = .848, either-agent = .756), with a significant difference between the object-only and the either-object conditions, $t(23) = 3.23$, $p = .004$ (object-only = .832, either-object = .705). There was no significant difference between the either-agent and the pairing conditions, $t(23) = 0.07$, $p = .944$ (either-agent = .756, pairing = .755), and no significant difference between the either-object and the pairing conditions, $t(23) = 1.80$, $p = .085$ (either-object = .705, pairing = .755). This pattern of results is consistent with that predicted by the binding hypothesis.

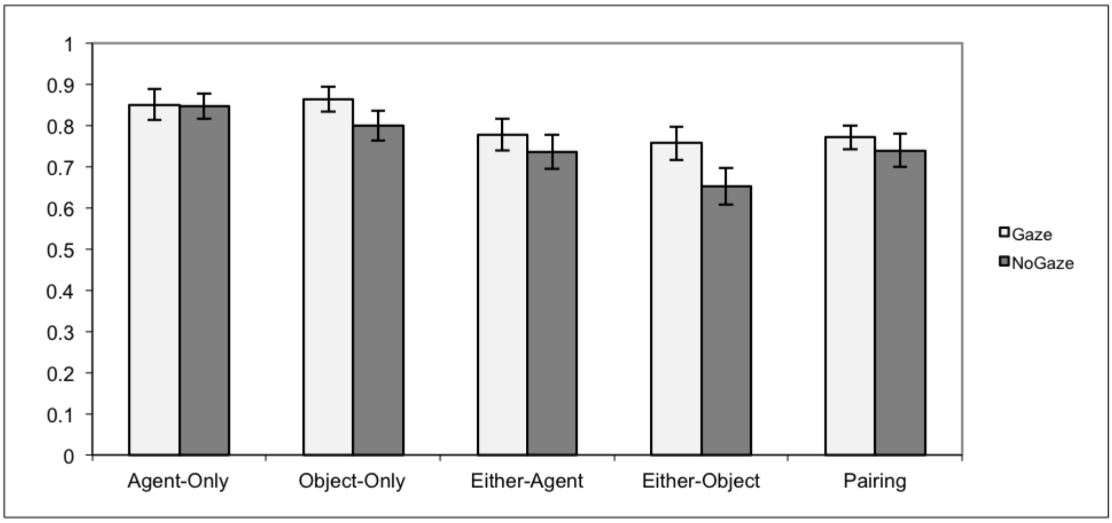


Figure 6.6 Experiment 8 results divided by condition and gaze. The figure graphed memory encoding accuracy. The error bars represent stand errors.

6.3.3 *Discussion*

The current experiment revealed an overall more accurate visual working memory encoding in the gaze condition compared to the no-gaze condition. However, this did not lead to a stronger binding in the gaze condition than the no-gaze condition. This finding does not lend support to the social binding hypothesis. Nevertheless, the current findings indicate that firstly, agents' object-oriented gaze promotes participants' visual working memory encoding accuracy. Although there was no effect of social binding, the current results provide incentive to explore the role of agents' object-oriented gaze with a different paradigm that is not restricted to the measurement of binding. This informs the rationale of Chapter 7. Secondly, the present findings indicate that the latest manipulations to emphasise agents' eye gaze generated sufficient contrast between agents' gazing towards and away from objects. The discrepant results between the present experiment and Experiments 6 and 7 led to two potential explanations. One possibility is that in Experiments 6 and 7, participants fail to process the relationship between agent and object, hence the present requirement to make explicit judgements about agents' gaze helped overcome this issue. This could imply that agents' object-oriented gaze has no observable effect on encoding accuracy when passively viewed; however, when participants were required to make explicit judgements based on this information, agents' object-oriented gaze facilitates encoding accuracy. The other possibility is that effects of agents' eye gaze are likely to be sensitive to agent stimuli employed (Conty et al., 2006). In Experiments 6 and 7, the agent stimuli might have provided more salient information about agents' body orientation rather than gaze direction. The present experiment overcame this issue by featuring new agents with salient eye-features and less prominent body parts. The current findings provided valuable information about the role of agents' object-oriented gaze in visual working memory and about methodologies for examining effects of agents' gaze. However, the current paradigm demonstrates no social binding effect, hence it is unlikely that agents' object-oriented gaze leads to stronger binding of agent and object information in visual working memory.

6.4 General Discussion

6.4.1 *Binding of Agent and Object Information*

Three experiments demonstrated that participants represent pair-wise presented agent and object information as a bound unit in visual working memory. Participants' memory encoding accuracies across the five conditions revealed patterns largely consistent with studies that have previously demonstrated binding effects (Wheeler & Treisman, 2002; Wood, 2008). However, participants' spontaneous binding of agents and objects might have obscured a binding effect that is sensitive to agents' eye gaze. One possibility is that the pair-wise presentation of agents and objects led participants to spontaneously bind this information. Therefore, to minimise participants' spontaneous binding and its potential concealment of a social binding effect, a further examination which eliminates any pairing cue is required. This informs the rationale of Chapter 7.

6.4.2 *Social Binding Effect*

Three experiments in the current study failed to demonstrate a stronger binding effect in the gaze condition than the no-gaze condition, predicted by the social binding hypothesis. Furthermore, only one of the three experiments, which increased emphasis on agents' eye gaze, showed encoding performance sensitive to agents' gaze directions. Nonetheless, this finding is consistent with studies that demonstrate adults automatically allocate their attention to objects gazed upon by agents (e.g., Frischen et al., 2007; Langton et al., 2006). It is possible that in the gaze condition, participants' attention is directed to objects looked at by agents, whereas in the no-gaze condition, participants' attention is directed away from objects. Having participants' attention focused on the agents and objects in a display likely leads to improved encoding of the display. One could further predict that objects that receive agents' gaze also receive more attention from the participants; therefore these objects should be encoded with greater accuracy compared to objects that do not receive agents' gaze[10]. That said, it is clear that the current paradigm is not optimum for measuring agent and object binding. Beyond issues discussed in Sections 6.3.3 and 6.4.1, the current paradigm also suffers from a relatively long trial sequence. This is in contrast to the short sequence typically employed in gaze attentional cueing studies (e.g., Frischen et al., 2007). To ensure that any transient effect can be captured as well as to match the sequence to gaze cueing studies more closely, in Chapter 7, I employed a one-shot

[10] Exploratory comparisons reveal that the Experiment 8 show trends to support this prediction; the differences between the encoding accuracies in the gaze and no-gaze conditions were marginally significant in the object-only conditions, $t(23) = 2.36$, $p = .027$, and the either-object conditions, $t(23) = 2.79$, $p = .010$.

change detection paradigm that presents a shorter temporal sequence (e.g., Luck & Vogel, 1997).

6.4.3 *Conclusion*

The current study demonstrates that participants spontaneously integrate pair-wise presentation of agent and object information in their visual working memory. Furthermore, the present results reveal that although agents' object-oriented gaze promotes overall encoding accuracy, it does not lead to enhanced binding of agent and object information. The next chapter examines whether agents' gaze affects the encoding of agent and object information in a one-shot change detection paradigm (e.g., Luck & Vogel, 1997).

CHAPTER 7 VISUAL WORKING MEMORY ENCODING FOR UNPAIRED AGENT AND OBJECT INFORMATION

7.1 Introduction

7.1.1 *Overview*

In Chapter 5, I reviewed evidence suggesting that individuals automatically shift their attention to the directions of agents' object-oriented gaze even when it does not facilitate task performance and demands greater processing cost (e.g., Frischen et al., 2007). In the current chapter, I will continue the examination of the influence that agents' object-oriented gaze has on visual working memory, begun in Chapter 6. An efficient one-shot change detection paradigm (e.g., Luck & Vogel, 1997) will be employed to test any transient effect that was not captured in Chapter 6 with the variant of the change detection paradigm (Wood, 2008). The current chapter shall also investigate the role of agents' participant-oriented gaze (typically known as *direct-gaze*) in visual working memory. I will begin by reviewing studies of participant-oriented gaze, which demonstrate both its facilitative role in children's and adults' sociocommunicative operations (e.g., Farroni, Csibra, Simion, & Johnson, 2002) and also the disruptive effect it can have on performance in various tasks (e.g., Conty, Gimmig, Belletire, George, & Huguet, 2010).

Experiment 9a has two aims. The first aim is to examine whether agents' object-oriented gaze leads to more efficient encoding of agent and object information. The second aim is to examine whether agents' participant-oriented gaze hinders visual working memory encoding. Recent studies indicate that participant-oriented gaze hinders performance on concurrent visual attention tasks (Senju, Hasegawa, & Tojo, 2005) and on a Stroop task (Conty et al., 2010). However, there has been little or no evidence indicating whether participant-oriented gaze produces similar effects on visual working memory tasks. The current study will address this issue. Experiment 9b will further investigate whether individuals' encoding accuracy is more sensitive to agents' participant-oriented gaze or object-oriented gaze. Experiments 10a and 10b will examine whether different types of visual inputs following participant-oriented gaze stimuli affects encoding accuracy.

7.1.2 *Participant-Oriented Gaze Facilitates Sociocommunication and Socio-function*

Participant-oriented gaze has been referred to as a sociocommunicative signal. Direct eye contact with others often indicates the intention to communicate with others or attract others' attention. A number of studies have revealed that humans are sensitive to gaze of others from an early age. From two to five days after birth, young infants discriminate between eye gaze that are direct at them and eye gaze that are directed away from them. Furthermore, at this age, infants prefer to look at faces that display eye contact compared to faces that looks away (Farroni et al., 2002). Four-month-olds exhibit enhanced neural processing for upright faces accompanied by participant-oriented gaze (Farroni et al.), particularly for faces displaying an angry expression (Striano, Kopp, Grassmann, & Reid, 2006). From six months of age, infants show significantly more gaze following when an adult makes initial eye contact with them compared to when an adult makes no eye contact before shifting her gaze directions (Senju & Csibra, 2008). These studies suggest that the sensitivity to others' eye contact signal develops early in life, and is likely shaped by particular evolutionary pressures.

Participant-oriented gaze also affects individuals' face-related processing, such as identification of emotional expressions (e.g., Adams & Franklin, 2009) and face recognition (Adams, Pauker, & Weisbuch, 2010). Adams and Franklin suggest that adults are more efficient in identifying facial expressions when they are accompanied by congruent gaze directions compared to incongruent gaze directions (approach-oriented emotions are congruent with participant-oriented gaze; avoidance-oriented emotions are congruent with gazing away). Furthermore, Adams et al. revealed that the cross-race memory effect (Meissner & Brigham, 2001) is modulated by gaze direction. The cross-race memory effect is the finding that individuals have poor face-recognition memory for faces of other-race individuals compared to faces of individuals from their own race. Adams et al. found that cross-race memory effect only occurs when the faces display participant-oriented gaze and not when the faces gaze away. These studies indicate that participant-oriented gaze plays an essential role in face-related processing.

Visual attention studies indicate that adults (Conty et al., 2006), typically developing children, and autistic children (Senju et al., 2005) are all quicker to detect participant-oriented gaze stimuli compared to gaze-away-stimuli in a visual search task. Interestingly, Senju et al. demonstrated that when the face stimuli were inverted, typically developing children showed a reduced search advantage for participant-oriented gaze stimuli. In contrast, the face inversion manipulation did not affect autistic children’s performance. This suggests that autistic children likely use the local features of others’ eyes in processing participant-oriented gaze stimuli. Typically developing children, however, are likely to process participant-oriented gaze along with the configuration of a whole face (Senju et al., 2005; Senju, Kikuchi, Hasegawa, Tojo, & Osanai, 2008). Furthermore, Akechi et al. (2010) revealed that typically developing children performed better on an emotion discrimination task when the emotional expressions displayed on the faces were paired with congruent gaze directions. However, autistic children's performance was unaffected by the congruency of gaze direction with emotion expression. These studies demonstrate the socio-functional and sociocommunicative values of eye contact signal. The amount of visual attention participant-oriented gaze attracts may be beneficial for social interaction and communication. Nonetheless, in contrast with the studies described above, other evidence suggests that participant-oriented gaze plays a disruptive role in cognitive tasks. These studies will be reviewed in the following section.

7.1.3 *Participant-Oriented Gaze Disrupts Cognitive Processes*

Evidence indicates that face-to-face conversation reduces speech fluency (e.g., Beattie, 1981). Some authors have argued that this is because maintaining eye contact with others occupies capacity in one's visuospatial sketchpad (Doherty-Sneddon, Bruce, Bonner, Longbotham, & Doyle, 2002). Thus gestures (e.g., Goldin-Meadow, Wein, & Chang, 1992) and gaze aversion (Doherty-Sneddon et al.) may aid communicative fluency by diverting cognitive load. This is known as the *cognitive load hypothesis* (Glensberg, 1997). Although these authors demonstrate that participant-oriented gaze produces a higher perceptual load than other gaze directions, they provide little explanation of how the excessive demand for maintaining eye contact emerges. One account is that a vast amount of visual attention is attracted to participant-oriented gaze (Conty et al., 2006; Senju et al., 2005), this distracts individuals from the concurrent communication or cognitive tasks. Conty et al. (2010) showed that when individuals perform a Stroop task, the presence of a pair of irrelevant eye stimuli displaying participant-oriented gaze leads to an exaggerated Stroop interference effect. Individuals' performance is unaffected, however, when the irrelevant eye stimuli gaze away or when the eyes are closed. Conty et al. further demonstrated that a black-and-white grating stimulus with matching contrast to the participant-oriented gaze stimulus has no effect on the amount of Stroop interference individuals suffer. This suggests that saliency alone cannot account for the effect of participant-oriented gaze upon Stroop performance. These studies indicate that in certain circumstances, participant-oriented gaze disrupts online communication and concurrent cognitive processes. This contrasts with earlier reviewed evidence that highlights the facilitative role eye contact plays in certain social interactions. It is not known whether participant-oriented gaze has a primarily facilitative or disruptive effect on individuals' visual working memory encoding of social information. The current study will investigate this issue.

7.1.4 *Rationale for Current Study*

The current study examines whether agents' object-oriented gaze leads to more efficient encoding of agent and object information. Building on the work from Chapter 6, the current study employs a one-shot change detection paradigm (Luck & Vogel, 1997). This enables the use of a relatively short temporal delay between encoding and recognition, which increases the chance of detecting a transient effect that was not observed in Chapter 6. If agents' object-oriented gaze leads to more efficient encoding of agent and object information, then this should lead to a higher encoding accuracy compared to when agents are looking away from objects. The present study also examines the role of participant-oriented gaze in visual working memory encoding, as there is currently little evidence to indicate whether participant-oriented gaze has a facilitative or disruptive effect on visual working memory encoding. Finally, the current study will also examine whether different types of visual input modulate the effect of agents' eye gaze.

7.2 Experiments 9a & 9b

7.2.1 *Experiment 9a*

The current experiment compares participants' encoding accuracy for displays in which agents directly look towards participants and displays in which agents look towards an object. Participants' task is to encode the colours of the agents' clothing and the shapes of the objects. The agents are not differentiated by their eye-features, therefore this is not relevant to participants' encoding task. Nonetheless, if agents' eye gaze affects participants' visual working memory encoding, then one of the following hypothesis may be true. Firstly, if agents' object-oriented gaze leads to more efficient encoding of agent and object, then participants should show a higher encoding accuracy for displays that contain agents looking towards objects. Secondly, if participant-oriented gaze hinders performance in visual working memory encoding, then participants should show lower encoding accuracy for displays that contain agents looking out at them.

7.2.1.1 *Method*

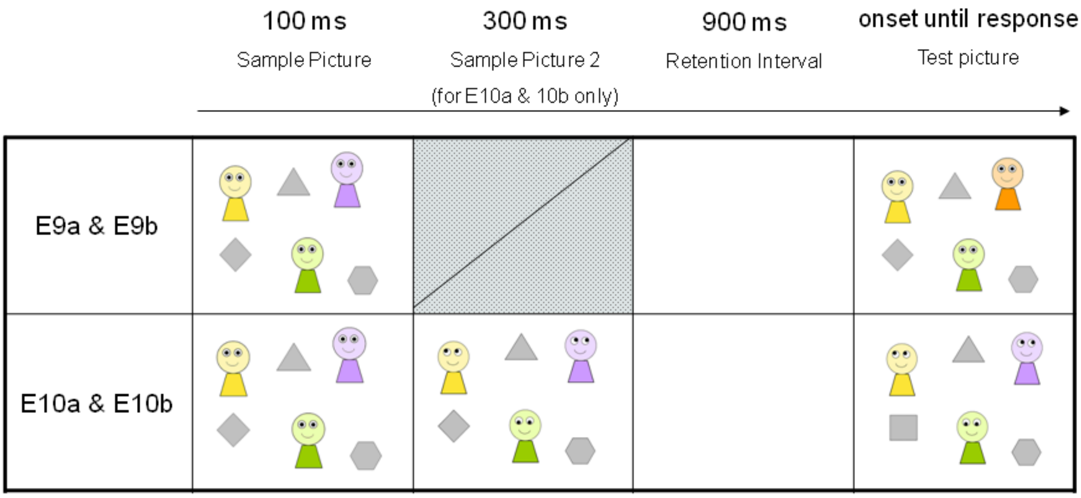
7.2.1.1.1 *Participants*

Sixteen students (14 female, mean age 20.06 years, age range 18 to 23) from the University of Birmingham took part in this experiment in return for a small honorarium or course credits. All participants had normal colour vision and normal or corrected-to-normal visual acuity. All participants' encoding accuracies were within two-standard deviations from the means in all four conditions, therefore no participants' data were excluded from the analysis.

7.2.1.1.2 *Design and Procedure*

A 2 x 2 within-participant design was constructed with gaze direction (look-at-you, look-at-object) and change element (agent-change, object-change) as the factors. Each display contained either three or four agents along with a matching number of objects. For each set size, 25 displays containing different combinations of agents and objects were generated from a pool of six different agents and six different objects. In the look-at-object condition, the agents always looked towards an object. In the look-at-you condition, the agents always looked straight ahead as if they were looking out at the participants. A one-shot change detection paradigm (e.g., Luck & Vogel, 1997; also see Rensink, 2002, for a review) was employed. Each trial began with a sample picture, presented for 100 ms, during which participants were to encode the information. This was followed by a 900 ms retention interval. At the end of the trial, a test picture was displayed until participants made a response (See Table 7.1).

Table 7.1 Examples of trials sequences from Experiments 9a, 9b, 10a, and 10b. Sequences begin with the displays on the left and progress towards to the right. All experiments contained two gaze direction conditions and two change element conditions. This table does not include all displays included in each experiment; the displays present here are selected as examples.



Luck and Vogel (1997) have shown that individuals performed identically in a one-shot change detection paradigm both with and without a concurrent articulation task; therefore it was considered unnecessary to include an additional concurrent task to suppress verbal encoding. It is likely that the display time in the one-shot change detection paradigm is too brief for participants to employ verbal encoding strategy.

All displays in the current study subtended 7.3º (height) x 9.8º (width) in the centre of a computer screen. For half of the trials, the test pictures were identical to the sample pictures, for the other half, they contained changes from the sample pictures. The change element was either the shape of one of the objects or the colour of one of the agents' clothes. The two types of changes occurred equally frequently. Participants responded by producing a left-click on the mouse when they saw a change in the test picture from the sample picture, and a right-click when they saw no change. Agent-change trials and object-change trials were blocked; therefore participants anticipated changes to occur either amongst the agents or amongst the objects. A total number of 384 test trials were presented in four test blocks of 96 trials. Each test block was preceded by an additional four practice trials from the same condition. The experiment was presented with E-prime (Schneider et al., 2002a; 2002b).

7.2.1.1.3 *Predictions*

Firstly, if the agents' object-oriented gaze leads to more efficient encoding of agent and object information, then participants should show a higher encoding accuracy in the look-at-object condition compared to the look-at-you condition. Secondly, if participant-oriented gaze disrupts visual working memory encoding, then participants should show lower encoding accuracy in the look-at-you condition compared to the look-at-object condition. Alternatively, if neither object-oriented gaze or participant-oriented gaze plays a significant role in visual working memory encoding, then no differences between the look-at-you condition and the look-at-object condition should be observed (see Panels A & B of Figure 7.1 for picture samples).

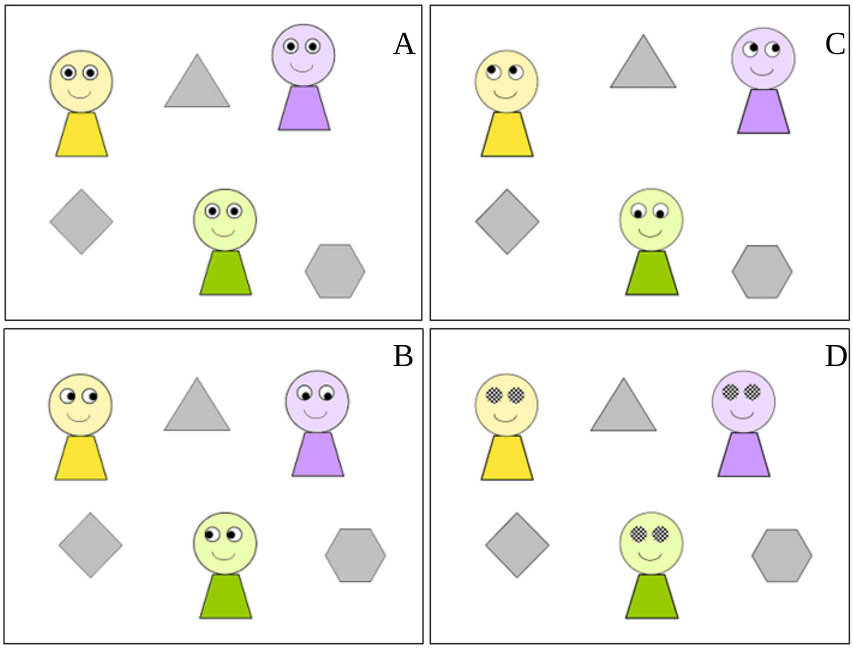


Figure 7.1 Examples of experimental displays. (A) Look-at-You (appeared in E9a, E10a, & E10b). (B) Look-at-Object (appeared in E9a, E9b, & E10a). (C) Look-Away (appeared in E9b & E10a). (D) Mask (appeared in E10b)

7.2.1.2 *Results*

Proportion correct for each condition was computed. Participants responded to 0.52[11] of the change trials correctly, and 0.79 of the no-change trials correctly. That the no-change trials were responded to more accurately is known as the *fast-same effect* (for an overview, see Farrell, 1985). It is a common phenomenon for performance on the no-change trials to be better than performance on the change trials. Only the change trials were analysed, as no manipulation on change element could be applied to the no-change trials. A 2 x 2 repeated-measures ANOVA was

[11] Although participants' overall performance on the change trials was near chance level, this was a combination of an encoding accuracy significantly above chance in the look-at-object-agent-change condition, and an encoding accuracy significantly below chance in the look-at-you-object-change condition (see Figure 7.2). This pattern indicates that participants were not merely responding at chance, or only producing 'no change' responses. Participants' low encoding accuracy could be due to the complexity of the current stimuli and the relatively short encoding time. In later Experiments 10a & 10b, when participants are allowed more encoding time, their encoding accuracy becomes significantly above chance in all conditions. Nonetheless, across all experiments, participants were consistently worse in encoding the object-change conditions than the agent-change conditions.

conducted with gaze direction (look-at-you & look-at-object) and change element (agent-change & object-change) as factors. There were significant main effects of gaze direction, $F(1, 15) = 19.90$, $p < .001$, $\eta_p^2 = .570$, which shows that the look-at-object condition was encoded more accurately than the look-at-you condition (*M*s = .566 and .472, respectively), and change element, $F(1, 15) = 7.46$, $p = .015$, $\eta_p^2 = .332$ demonstrating more accurate encoding for the agent-change condition compared to the object-change condition (*M*s = .573 and .465, respectively). There was no interaction between the two factors, $F(1, 15) = 1.65$, $p = .219$, $\eta_p^2 = .099$ (see Figure 7.2).

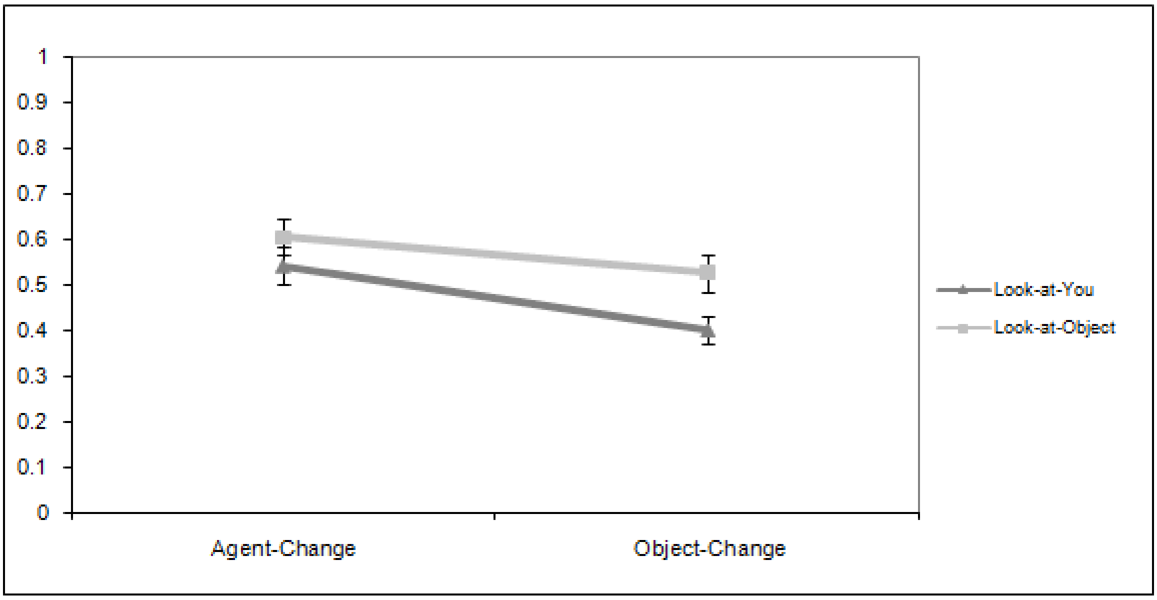


Figure 7.2 Proportion correct for Experiment 9a. Error bars represent standard errors from each condition.

7.2.1.3 *Discussion*

The current results revealed that participants encoded displays containing agents looking at objects more accurately compared to displays containing agents looking at out them. There are two potential interpretations of this finding. One is that the agents' object-oriented gaze led to more efficient encoding of agent and object information. Hence participants' encoding accuracy benefited from the relationship between the agents and the objects. This interpretation is consistent with the findings that agents' object-oriented gaze promotes overall encoding accuracy from Experiment 8. The other possibility is that the agents' participant-oriented gaze impaired participants' encoding performance, similar to its effect on other concurrent cognitive tasks (e.g., Conty et al., 2010). This led to lower encoding accuracies for displays containing participant-oriented gaze stimuli.

To distinguish between these two accounts, the participant-oriented gaze element was removed from Experiment 9b. Instead I contrasted agents' object-oriented gaze and their gaze directed away from objects. If agents' object-oriented gaze led to the higher encoding accuracy in the current experiment, then the look-at-object condition should be encoded more accurately than a look away from object condition. However, if the agents' participant-oriented gaze disrupted participants' visual working memory encoding in Experiment 9a, then no differences should be found between the two conditions in Experiment 9b.

7.2.2 *Experiment 9b*

The current experiment aims to distinguish between the potential effects of participant-oriented gaze and object-oriented gaze observed in Experiment 9a. The current design excludes any participant-oriented gaze component, replacing the look-at-you condition with the look-away condition. In the look-away condition, the agents display eye gaze similar to object-oriented gaze, but directed away from the objects (see Panel C of Figure 7.1). If the agents' object-oriented gaze leads to more efficient encoding of agent and object information, then a higher encoding accuracy should be observed in the look-at-object condition than the look-away condition. However, if difference between the look-at-object and look-at-you conditions was a consequence of a hindering effect of participant-oriented gaze, then no differences should be found between the two conditions in the current experiment.

7.2.2.1 *Method*

7.2.2.1.1 *Participants*

Fifteen students (13 female, mean age 18.93 years, age range 18 to 23) from the University of Birmingham took part in this experiment in return for course credits. All participants had normal colour vision and normal or corrected-to-normal acuity. All participants' encoding accuracies were within two-standard deviations of the means in all four conditions, therefore no participants' data were excluded from the analysis.

7.2.2.1.2 *Design and Procedure*

A 2 x 2 within-participant design was constructed with gaze direction (look-away, look-at-object) and change element (agent-change, object-change) as factors. The rest of the design and procedure were identical to those of Experiment 9a.

7.2.2.2 *Results*

Proportion correct for each condition was computed. Participants responded to 0.54[12] of the change trials correctly, and 0.80 of the no-change trials correctly. Only the change trials were analysed. A 2 x 2 repeated-measures ANOVA was conducted with gaze direction (look-away & look-at-object) and change element (agent-change & object-change) as factors. No significant main effect of gaze direction was found, $F(1, 14) = 0.23$, $p = .643$, $\eta_p^2 = .016$, with no significant main effect of change element, $F(1, 14) = 1.23$, $p = .286$, $\eta_p^2 = .081$, and no interaction between the two factors, $F(1, 14) = 0.03$, $p = .859$, $\eta_p^2 = .002$ (see Figure 7.3).

[12] Participants performed significantly above chance in both the look-away-agent-change condition and the look-at-object-agent-change condition.

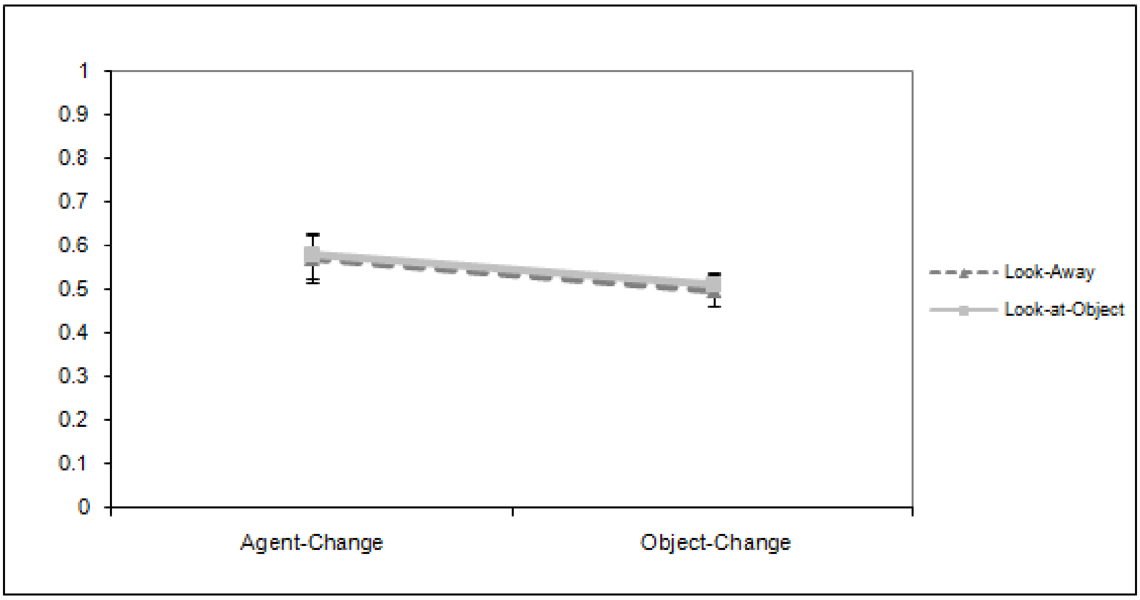


Figure 7.3 Proportion correct for Experiment 9b. Error bars represent standard errors.

7.2.2.3 *Discussion*

Participants performed virtually identically in the two conditions in the current experiment. This suggests that the agents' object-oriented gaze did not lead to more efficient encoding of agent and object information. The present finding might appear to be inconsistent with that of Experiment 8. However, it is noteworthy that in Experiment 8, participants were required to make explicit judgements about agents' gaze, whereas the current experiment did not implement this procedure. Given that both the current experiment and Experiment 8 employed agent stimuli with salient eye-feature, it is possible that the explicit requirement to judge agents' eye gaze in Experiment 8 enforced attention to be allocated to agents' eye region that does not necessarily occur when displays are merely viewed passively.

The difference between the look-at-object condition and the look-at-you condition in the current experiment was likely to be caused by a hindering effect of participant-oriented gaze. However, the current experiment did not show a significant main effect of change element, which was observed in Experiment 9a. Hence this interpretation remains tentative. As described earlier, participant-oriented gaze is a sociocommunicative cue that attracts visual attention (e.g., Conty et al., 2006; Senju et al., 2005). Conty et al. (2010) found that participants automatically shift attention towards participant-oriented gaze stimuli even when this hinders concurrent task performance. The present results are consistent with Conty et al.'s (2010) findings, and further demonstrate the disruptive effect participant-oriented gaze has on visual working memory tasks. It is likely that participants were not successful in applying top-down strategies to avoid processing the irrelevant participant-oriented gaze stimuli. However, it is possible that bottom-up visual inputs may enable participants to shift their attention away from participant-oriented gaze stimuli, and hence recover from the disrupted encoding. This was the rationale for Experiments 10a and 10b.

7.3 Experiments 10a & 10b

The present experiments aim to investigate whether bottom-up visual input may allow participants to overcome the hindering effect of participant-oriented gaze. In an extension of the sequence used in Experiment 9a, participants were presented with a new display after the agents' participant-oriented gaze. The new display was either be object-oriented gaze or gaze away from objects (which may enable participants to reinterpret agents' direct eye gaze as merely the start point for a shift of gaze), a low-level visual mask (which covers over agents' participant-oriented gaze), or a blank screen (which gives participants the same amount of time as the mask presentation). This will allow us to investigate whether participants' encoding accuracy for displays containing participant-oriented gaze may recover with the aid of various bottom-up visual input. It will also allow us to test whether a reinterpretation of the agents' eye gaze is essential for participants to recover their encoding, or whether a low-level mask or additional processing time will also aid participants' recovery.

7.3.1 *Experiment 10a*

The current experiment was informed by gaze attentional cueing studies (see Frischen et al., 2007 for a review), which indicate that individuals involuntarily shift their attention to the location gazed at by others even when it is not advantageous for task performance. As briefly described in Chapter 5, these studies mostly employ the following trial sequence. A trial begins with a human avatar directly gazing at participants, followed by their gaze averting to one of two locations. A target object then appears in the location either congruent or incongruent with the direction of the avatar's gaze. Participants' task is to report the object's location. These studies consistently reveal that when the avatar's gaze direction is incongruent with the object's location, participants are slower to respond to the object's location. As described earlier, related studies on participant-oriented gaze indicate that participant-oriented gaze attracts a vast amount of visual attention (e.g., Conty et al., 2006; Senju et al., 2005). Presenting a participant-oriented gaze stimulus at the start of a gaze cueing sequence should ensure that participants' attention dwells on the avatar's eye region before the avatar shifts her eye gaze. The current experiment investigates whether employing a similar trial sequence in visual working memory tasks will facilitate participants to shift their attention away from agents' participant-oriented gaze. Participants will be presented with an object-oriented gaze stimulus or a gaze-away stimulus after a participant-oriented gaze stimulus, giving the impression that the agents have shifted gaze direction. If the gaze cueing sequence allows participants to reinterpret agents' participant-oriented gaze, then the current experiment should show restored encoding of agent and object information that was previously found to be disrupted by the agents' participant-oriented gaze.

7.3.1.1 *Method*

7.3.1.1.1 *Participants*

Sixteen students (15 female, mean age 19.06 years, age range 18 to 20) from the University of Birmingham took part in this experiment in return for course credits. All participants had normal colour vision and normal or corrected-to-normal acuity. One participant's data were replaced prior to analysis due to an accuracy score two-standard deviations below the mean in two of the four conditions.

7.3.1.1.2 *Design and Procedure*

The experimental factors were identical to that of Experiment 9b, with gaze direction (look-at-object, look-away) and change element (agent-change, object-change) as factors. In the current experiment, all trials began with 100 ms of participant-oriented gaze display, which was identical to the sample pictures from Experiment 9a. There were either three or four agents gazing towards participants, and a matching number of objects around them. This display was followed by a 300 ms display, in which the agents either looked towards the objects or looked away from the objects. These displays were identical to the displays used in Experiment 9b. When the new gaze displays appeared, participants only saw the agents' gaze directions shift from participant-oriented gaze to either object-oriented gaze or gazing away from objects, the rest of the displays remained identical across the two displays. This was then followed by a 900 ms retention interval. At the end of each trial, a test picture onset until a response was detected (see Table 7.1). For half the trials the test pictures were exactly the same as the new gaze displays, while for the other half the test pictures contained a change. The changes were either in the colour of one of the agents' clothing or the shape of one of the objects.

7.3.1.1.3 *Predictions*

Firstly, if the gaze cueing sequence allows reinterpretation of agents' participant-oriented gaze, then participants should recover from the hindering effect of participant-oriented gaze. That is, they should show a higher encoding accuracy in the current experiment compared to Experiment 9a, where agents' participant-oriented gaze caused disruption to participants' encoding. Secondly, in line with the original hypothesis in Chapters 5 and 6, the current experiment provides another opportunity to examine whether agents' object-oriented gaze leads to more efficient encoding of agent and object information. As the gaze cueing sequence may reinforce participants' attention on the agents' eye region, the potential effects of agents' object-oriented gaze will be re-examined in the current experiment.

7.3.1.2 *Results*

7.3.1.2.1 *Experiment 10a*

The proportion of correct responses for each condition was computed. Participants responded to 0.66 of the change trials correctly, and 0.82 of the no-change trials correctly. Only the change trials were analysed. A 2 x 2 repeated-measures ANOVA was conducted with gaze direction (look-at-object & look-away) and change element (agent-change & object-change) as factors. No main effect of gaze direction was found, $F(1, 15) = 0.03$, $p = .877$, $\eta_p^2 = .002$, with a significant main effect of change element, $F(1, 15) = 23.80$, $p < .001$, $\eta_p^2 = .613$, demonstrating more accurate encoding for the agent-change condition compared to the object-change condition (*M*s = .736 and .585, respectively). There was no interaction between the two factors, $F(1, 15) = 0.02$, $p = .904$, $\eta_p^2 = .001$ (see Figure 7.4).

7.3.1.2.2 *Comparisons between Experiments 9a & 10a*

In order to make a fair comparison between Experiments 9a and 10a, participants' performance in the common condition – the look-at-object condition –

was used as a baseline. Differences between the look-at-object condition and 'the other' condition (the look-at-you condition in Experiment 9a, the look-away condition in Experiment10a) were computed, and these differences were subject to analysis. A 2 x 2 mixed ANOVA was conducted with change element (agent-change & object-change) and experiment (Experiment 9a & Experiment 10a) as factors. No main effect of change element was found, $F(1, 30) = 1.29$, $p = .264$, $\eta_p^2 = 041$, with a significant main effect of experiment, $F(1, 30) = 14.13$, $p = .002$, $\eta_p^2 = .320$, revealing more accurate encoding in Experiment 10a compared to Experiment 9a (*M*s = .094 and .002, respectively). There was no interaction between the two factors, $F(1, 30) = 1.01$, $p = .324$, $\eta_p^2 = .032$. Results revealed that participants' encoding accuracy in the current experiment was significantly higher than that of Experiment 9a.

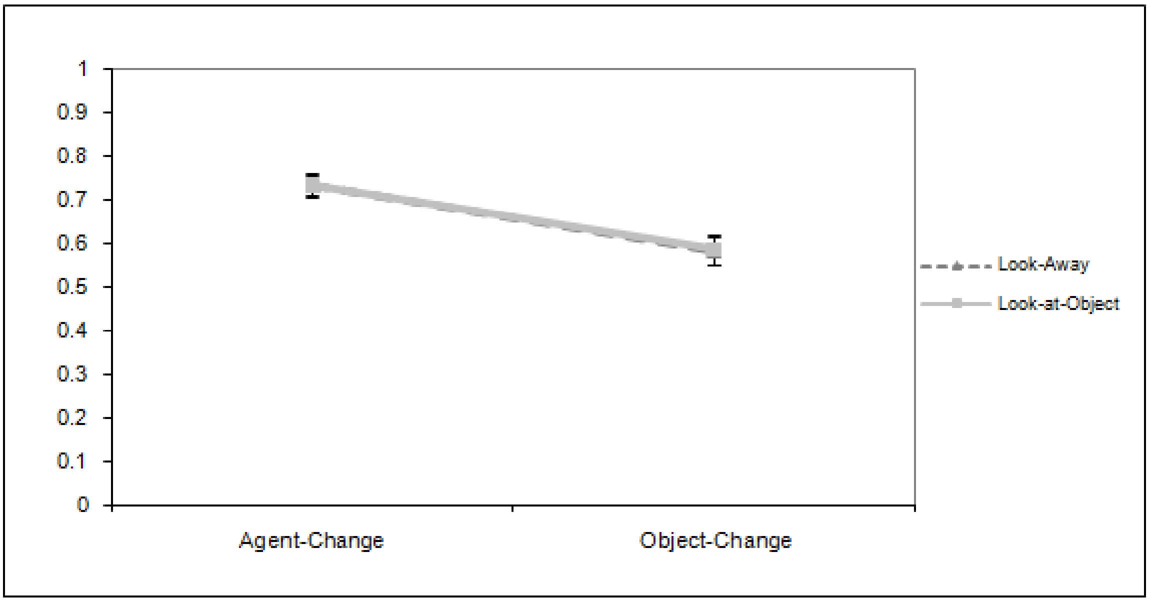


Figure 7.4 Experiment 10a results for change trials in proportion correct. Error bars represent standard errors.

7.3.1.3 *Discussion*

The current findings revealed that when participants saw agents' participant-oriented gaze followed by a new gaze direction, they were able to recover from the hindering effect of participant-oriented gaze. This indicates that the gaze cueing sequence allows participants to shift their attention away from agents' participant-oriented gaze. That said, a recovery of encoding may not require reinterpretation of agents' eye gaze. One possibility is that simply masking over agents' participant-oriented gaze with other types of visual input may produce the same effect. Another possibility is that simply having 300 ms of additional time is sufficient for participants to recover from the disruption caused by participant-oriented gaze. Experiment 10b aims to test these two possibilities. In a separate vein, the current findings indicate that participants encode displays containing agents looking towards objects and displays containing agents looking away from objects with equal accuracy. The gaze cueing sequence did not produce a different pattern of results from that of Experiment 9b[13]. Both the current chapter and Chapter 6, provide clear evidence that mere observation of agents' object-oriented gaze does not affect participants' encoding performance. Examination on agents' object-oriented gaze will not be pursued further here.

7.3.2 *Experiment 10b*

The current experiment investigates whether recovery of encoding requires reinterpretation of agents' eye gaze. All trials will start with agents' participant-oriented gaze, but will be followed by either a low-level visual mask or a blank screen for the same length of time. The comparisons between the current experiment and Experiment 10a should reveal whether seeing a low-level visual mask or having an equivalent amount of time with no additional visual input allows participants to recover their encoding to the same extend as seeing agents displaying new gaze direction. If reinterpretation of agents' eye gaze is necessary for a recovery of

[13] An unreported experiment was conducted with 500 ms of participant-oriented gaze display, followed by 500 ms of new gaze direction. The unreported experiment showed identical results to that of Experiment 10a. This indicates that encoding time does not directly affect participants' encoding performance. A similar argument has also been made by Luck and Vogel (1997); they showed that participants performed equally well when given 100 ms as when given 500 ms of encoding time.

encoding, then participants should not show the same amount of recovery as in Experiment 10a.

7.3.2.1 *Method*

7.3.2.1.1 *Participants*

Sixteen students (12 female, mean age 22.19 years, age range 18 to 37) from the University of Birmingham took part in this experiment in return for course credits or a small honorarium. All participants had normal colour vision and normal or corrected-to-normal visual acuity. All participants' encoding accuracies were within two-standard deviations of the mean in all four conditions, therefore no participants' data were excluded from the analysis.

7.3.2.1.2 *Design and Procedure*

A 2 x 2 within-participant design was constructed with gaze direction (look-at-you + mask, look-at-you) and change element (agent-change, object-change) as factors. Each trial began with 100 ms of participant-oriented gaze -stimulus, where the agents gazed towards the participants. A new input then followed and was presented for 300 ms. The new input for the look-at-you + mask condition was a checkerboard pattern covering over the agents' eye regions (see Panel D of Figure 7.1). The new input for the look-at-you condition was a blank display. This was followed by a 900 ms retention interval. At the end of a trial, a test picture was displayed until a response was detected (see Table 7.1).

7.3.2.1.3 *Predictions*

Firstly, if reinterpretation of agents' eye gaze is necessary for overcoming the hindering effect of participant-oriented gaze, then participants should show lower encoding accuracies in the look-at-you + mask condition and the look-at-you condition compared to the two conditions from Experiment 10a. However, if low-level visual input following agents' participant-oriented gaze facilitates recovery, then the look-at-you + mask condition should be encoded as accurately as the two conditions from Experiment 10a. Alternatively, if an additional 300 ms is sufficient for participants to recover from the disruptive effects of participant-oriented gaze, then the look-at-you condition should be encoded as accurately as the two conditions from Experiment 10a. However, if an additional 300 ms does not produce a recovery in encoding, then participants' encoding accuracy for the look-at-you condition should be lower than the two conditions from Experiment 10a. Finally, reinterpretation of agents' eye gaze may not be essential for producing recovery from encoding. Nonetheless, there may be a difference between the amount of recovery produced by a low-level visual mask and that produced by additional time, if, for example, a minimal visual input may be necessary for covering over participant-oriented gaze stimuli. In this case, one would expect the look-at-you + mask condition to be encoded more accurately than the look-at-you condition.

7.3.2.2 *Results*

7.3.2.2.1 *Experiment 10b*

The proportion of correct responses in each condition was computed. Participants responded to 0.59 of the change trials correctly, and 0.82 of the no-change trials correctly. Only the change trials were analysed. A 2 x 2 repeated-measures ANOVA was conducted with gaze direction (look-at-you & look-at-you + mask) and change element (agent-change & object-change) as factors. There were significant main effects of gaze direction, $F(1, 15) = 13.71$, $p = .002$, $\eta_p^2 = .478$, revealing more accurate encoding in the look-at-you + mask condition compared to the look-at-you condition (*M*s = .622 and .564, respectively) and change element, $F(1, 15) = 21.92$, $p < .001$, $\eta_p^2 = .594$, demonstrating more accurate encoding for the agent-change condition compared to the object-change condition (*M*s = .641 and .546, respectively), with a significant interaction between the two factors, $F(1, 15) = 13.63$, $p = .002$, $\eta_p^2 = .476$. Planned comparisons revealed a significant difference between the look-at-you and look-at-you + mask conditions in the agent-change condition, $t(15) = 4.34$, $p = .001$ (look-at-you-agent-change = .581, look-at-you + mask-agent-change = .701), but not in the object-change condition, $t(15) = 0.15$, $p = .880$ (look-at-you-object-change = 547, look-at-you + mask-object-change = .544; see Figure 7.5).

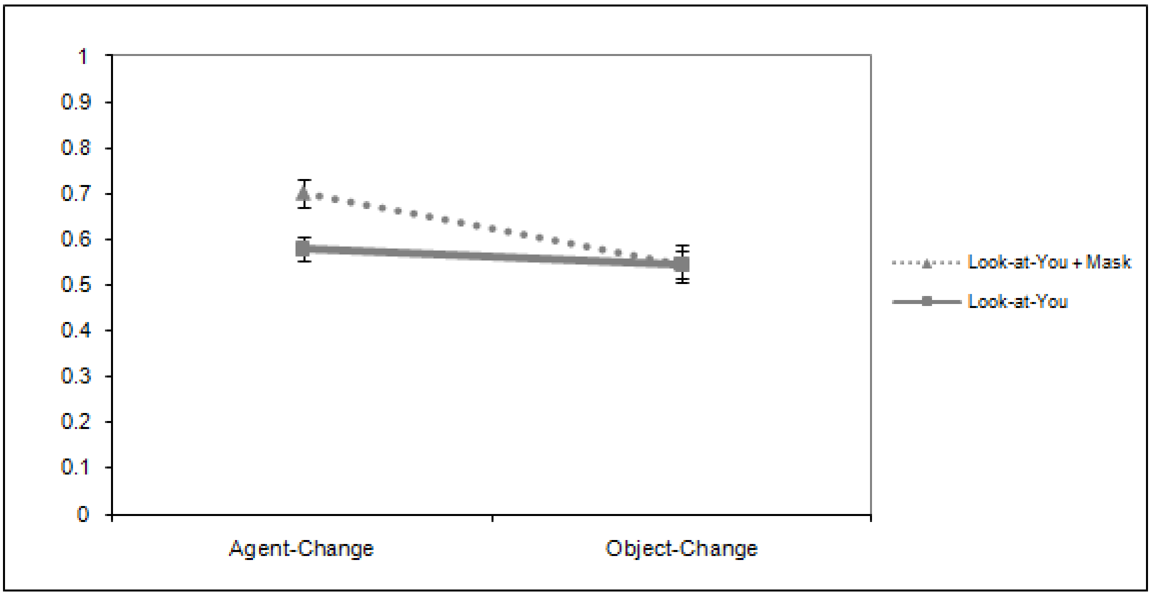


Figure 7.5 Experiment 10b results for change trials in proportion correct. Error bars represent standard errors.

7.3.2.2.2 *Comparison between Experiments 10a & 10b*

An omnibus analysis was not conducted for two reasons. Firstly, the gaze direction conditions in Experiment 10a and Experiment 10b are not exact replicas. Secondly, there was no *a priori* hypothesis regarding the comparison between participants' encoding performance for agent and object *per se*. For these reasons, *t*-tests were conducted to directly examine the four gaze direction conditions within the agent-change conditions and the object-change conditions respectively. Amongst the agent-change conditions from Experiments 10a and 10b, *t*-tests revealed a significant difference between the look-away condition and the look-at-you condition, $t(30) = 1.14$, $p < .001$[14] (look-away-agent-change = .736, look-at-you-agent-change = .581). A significant difference was also found between the look-at-object condition and the look-at-you condition, $t(30) = 3.89$, $p = .005$ (look-at-object-agent-change = .736, look-at-you-agent-change = .581). As reported in Section 7.3.2.2.1, a significant difference was found between the look-at-you condition and the look-at-you + mask condition, $t(15) = 4.34$, $p = .001$. The rest of the conditions did not differ significantly from each other (all $p > .303$). The current results indicate that participants' encoding for agent information recovered equally well when agents' participant-oriented gaze was followed by agents displaying a new gaze direction and when it was followed by a visual mask. However, having the same amount of time without any visual input did not produce the same degree of recovery for agent information encoding.

Amongst the object-change conditions from Experiments 10a and 10b, none of the gaze direction conditions was found to be significantly different from each other (all $p > .427$). The current results revealed that participants recovered their encoding for object information both when presented with a visual mask or with an equivalent amount of time after seeing agents' participant-oriented gaze. There was no distinction between the recovery for encoding object information produced by agents' new gaze directions, a visual mask, and a matching length of time.

7.3.2.3 *Discussion*

The current findings reveal that when a visual mask follows agents' participant-oriented gaze, participants recover their encoding for both agent information and object information. This indicates that reinterpretation of the agents'

[14] Six *t*-tests were conducted, therefore significance level become .008 after applying Bonferroni corrections.

eye gaze is not necessary for participants to recover from the impairment produced by participant-oriented gaze. Furthermore, when participants were merely provided with an equivalent amount of time, their encoding for object information also recovered. In contrast, participants' encoding for agent information did not spontaneously recover when additional time was provided. It is unlikely that this asymmetry is due to a more severely disrupted encoding of agent information, given that Experiment 9a demonstrated that participants' encoding of both agent information and object information were equally affected by agents' participant-oriented gaze. Rather, this asymmetry is more likely to arise in the recovery process; participants' recovery of encoding for agent information may require some visual input to override the effects of agents' participant-oriented gaze.

7.4 General Discussion

7.4.1 *Summary of Current Findings*

Chapters 6 and 7 investigated the possibility that observation of agents' object-oriented gaze might lead to more efficient encoding of agent and object information. However, even with a shortened delay between encoding and retrieval, there was no support for this hypothesis. Taken together, the experiments across Chapters 6 and 7 indicate that mere observation of agents' object-oriented gaze does not lead to more efficient encoding of agent and object information. In addition, experiments in the present chapter also investigated the potential hindering effect of participant-oriented gaze on visual working memory encoding. Experiments 9a and 9b revealed that participants' encoding accuracy was lower when agents directly gazed towards participants. This is consistent with the findings of previous studies (Conty et al., 2006; Senju et al., 2005), which suggest that participant-oriented gaze monopolizes visual attention. It is likely that the drawing of attention to participant-oriented gaze stimuli disrupts visual working memory encoding. Experiments 10a and 10b demonstrate that an apparent shift of an agents' gaze produces recovery in participants' visual working memory encoding. Furthermore, even a visual mask that does not allow reinterpretation of agents' eye gaze also produced recovery of encoding. Interestingly, when participants were merely given additional time that is equivalent to the length of the mask display after having an agent directly gaze towards them, their encoding of object but not agent information recovered. Although object information encoding is not immune from the disruptive effects of participant-oriented gaze, it recovers more efficiently than agent information encoding.

7.4.2 *The Disruptive and Facilitative Role of Participant-oriented Gaze*

The finding that participant-oriented gaze hinders visual working memory is in contrast with the findings of Adams et al. (2010). As described earlier, these authors demonstrated that the cross-race memory effect only occurs when the presented faces directly gaze towards participants. Interestingly, in Adams et al.'s study, face recognition performance was better when faces displayed participant-oriented gaze as opposed to faces that gaze away, regardless of race. This appears to contradict the findings of Experiment 9a, where agent information was encoded poorly when the agents displayed participant-oriented gaze. However, an important difference between the current study and Adams et al.'s study is that in the latter, participants' task was to encode faces, including presumably the agents' eyes, whereas in the current study, the agents' eye gaze was not directly relevant to participants' encoding task. That agents' gaze is an irrelevant component of the task is more similar to Conty et al.'s (2010) study, where irrelevant gaze stimuli were present when participants engaged in a Stroop task. The current findings are consistent with Conty et al. in demonstrating the disruptive influence of participant-oriented gaze on concurrent task performance. One possibility is that participant-oriented gaze facilitates tasks that involved processing of facial and emotional contents but hinders performance on tasks that do not directly involve ‘reading’ others’ eyes.

7.4.3 *Conclusion*

The current chapter along with Chapter 6 indicates that mere observation of agents’ object-oriented gaze does not lead to more efficient visual working memory encoding of agent and object information. Nonetheless, the current study found that participant-oriented gaze hinders visual working memory encoding; and that such effect is likely to be caused by the attentional draw of participant-oriented gaze stimuli. Importantly, reinterpretation of agents’ gaze is not essential for recovery of information encoding. The current findings are consistent with previous findings in the sociocommunication literature. Nevertheless, further investigation is necessary in order to draw with confidence conclusions from direct comparisons. In Chapter 8, I will discuss future research, which may help to resolve some outstanding issues and extend this work in new directions.

CHAPTER 8 GENERAL DISCUSSION

8.1 Overview

The aim of this thesis was to examine how the automatic processing of the visual perspectives and eye gaze of others affects the way we perceive and encode the social world. I have addressed this question in two ways. Firstly, I explored the computational capacity of automatic visual perspective-taking. Secondly, I examined the roles that participant-oriented gaze and object-oriented gaze play in visual working memory encoding. In the current chapter, I will firstly summarize and discuss the findings from the visual perspective studies in Chapters 3 and 4, and from the visual working memory studies in Chapters 6 and 7, respectively. I will then discuss the implications of the current findings for efficient social cognitive processes and the efficient mindreading system. Finally, I will highlight future directions for this line of work as well as further investigations on related topics.

8.2 Chapters 2, 3, & 4: Visual Perspective-Computation

8.2.1 *Summary*

In Chapters 3 and 4, I employed Samson et al.'s (2010) visual perspective-taking paradigm to examine the potential limitations and flexibility in automatic visual perspective-computation. Chapter 3 showed that participants' judgements about their own visual perspectives were influenced by an avatar's discrepant perspective. This influence was observed irrespective of the size of the perspective contents; even when both the participants' and the avatar's perspective contents exceeded the range of subitization, participants were still found to have automatically computed the avatar's perspective. Nonetheless, the current findings demonstrate that participants' automatic computation of the avatar's perspective only allows binding the avatar with this perspective contents when the contents could be subitized but not when the perspective contents exceed the range of subitization. This suggests that automatic computation of the avatar’s large number perspective is vaguely specified. Chapter 4 demonstrated that automatic visual perspective-computation allows selection of relevant information. When participants were instructed to merely attend to the relevant information (red dots) and ignore the distractors (blue squares), automatic visual perspective-computation was only found for the relevant information but not for the distractors. This suggests that participants were successful in selecting and computing the relevant information. Furthermore, cases in which the avatar’s perspective was ‘zero’ (the avatar either sees a blank wall or a number of blue squares, see Table 4.2) did not demand greater processing cost to compute than cases where his perspective was greater than zero, and was found to interfere with participants’ number perspective contents. The current investigation demonstrated that automatic visual perspective-computation is characterized by both flexibility and limitations. The current findings also inform our understanding of both egocentric and altercentric intrusion, and consequently advance our understanding of adults’ mindreading abilities.

8.2.2 *Egocentric Intrusion: Interference from One’s Own Perspective*

Throughout the investigations in Chapters 3 and 4, the effect of egocentric intrusion was always present regardless of the enumeration load or the demand to select relevant information. This indicates that it is difficult for participants to make judgements about others’ perspectives without suffering interference from their own perspectives. As described in Chapter 1, a number of other studies provide evidence indicating that adults’ judgements about an ignorant protagonist’s choices are often biased by their own privileged knowledge (Birch & Bloom, 2007; Mitchell et al., 1996). This bias has been shown in story-based task scenarios (Birch & Bloom, 2007; Mitchell et al., 1996), as well as in an online referential communication game scenario (Apperly et al., 2010; Keysar et al., 2003). The results of both Chapters 3 and 4 further suggest that egocentrism varies as a function of the demand to select relevant information: when participants were required to select relevant information, the effect of egocentric intrusion was greater. Previous studies also demonstrated increased egocentrism when adults were under time pressure to respond (Epley et al., 2004). Healthy adults' ability to account for the director's discrepant viewpoint was found to be correlated with their performance on executive function tests (Qureshi, 2009). Furthermore, studies with brain-injured patient W.B.A. revealed that the saliency of his own beliefs determined the level of performance on non-verbal mindreading tasks (Samson et al., 2005). The current findings agree with these previous studies in suggesting that overcoming interference from one's own perspective is an effortful process even for healthy adults.

8.2.3 *Altercentric Intrusion: Implicit Computation of Others' Perspectives*

The current investigation further adds to our understanding of adults' implicit computation of others' perspectives when it is unnecessary and disadvantageous to their task performance. Firstly, individuals are likely to automatically compute the perspective of others when the demand for computing their own perspectives is relatively small or when another's perspective contents need to be attended in order to compute one's own perspective contents. In Chapter 3, I found that altercentric intrusion was consistently present when participants made 'yes' responses, that is, when they attended to both subsets of dots on the walls for computing their own perspective contents. The effects of altercentric intrusion were also found when participants made 'no' responses on trials where they held subitizable perspective contents, but not when their own and the avatar's perspective contents exceeded the range of subitization. This suggests that there are boundaries to the operation of automatic computation of others' visual perspectives. Relatedly, Cohen and German (2009) showed that adults encode others' beliefs without overt instructions. Nonetheless, the retention of such information was restricted to a very short amount of time, suggesting that there are operational boundaries to adults' capacity to automatically compute others' perspectives. Secondly, as briefly described before, the effect of altercentric intrusion did not derive from a disturbed gestalt structure in the array of dots, or divided attention (Samson et al., 2010). As shown in Chapter 4, even when the discrepancy between participants' own perspectives and the avatar's perspectives was merely caused by one set instead of two sets of dots, participants still suffered interference from the avatar's perspective. The effect of altercentric intrusion is likely, therefore, to reflect a genuine computation of the avatar's perspective, and not merely result from lower-level computations unrelated to perspective-taking.

The effect of altercentric intrusion is comparable to effects indicative of a range of efficient social cognitive processes. A common feature of these effects is the automatic processing of social perceptual cues in the environment (e.g., eye gaze and stereotypic attributes) even when such computations often produce undesirable consequences that lead to additional processing cost. These effects will be discussed in Section 8.4 along with the relevant findings from Chapters 6 and 7.

8.3 Chapters 5, 6, & 7: Eye Gaze and Visual Working Memory Encoding

8.3.1 *Summary*

In Chapters 6 and 7, I examined the possibility that observation of others' eye gaze affects participants' visual working memory encoding. In Chapter 6, I tested whether agents' object-oriented gaze might lead to binding of agents and objects in visual working memory encoding. The current findings, however, revealed pair-wise encoding of agents and objects that was not sensitive to agents' gaze directions. In Chapter 7, I further examined the role of agents' object-oriented gaze with a further change detection paradigm that comprised a shorter temporal sequence and no spatial pairing cue. However, once again findings revealed that mere observation of agents' object-oriented gaze does not affect participants' encoding of agents and objects. The other hypothesis in Chapter 7 was that participants' encoding might be disrupted when agents directly look out at them, even when agents' gaze directions were not relevant to participants' encoding task. It was found that, as hypothesised, when the agents gazed directly towards participants, participants' encoding performance was considerably worse than when the agents gazed in other directions. Moreover, although participants appeared unable to employ a top-down strategy to avoid the hindering effect of participant-oriented gaze, bottom-up visual input did facilitate participants' recovery to normal encoding. The current investigations implicated participant-oriented gaze in visual working memory encoding but indicated a minimum role for others' object-oriented gaze in such processes. In the following section I shall explore some factors that may have shaped the present findings of a minimal effect of agents' object-oriented gaze.

8.3.2 *Visual Working Memory as a Measure of Social Perception: Factors to Consider for Further Experiments*

As described in Chapter 5, much evidence demonstrates automatic attention shifting in response to the observation of others' eye gaze (e.g., Driver et al., 1999). The current investigation extended the examination of automatic attention shifting to examine whether the automatic processing of others' gaze affects individuals' performance on visual working memory tasks. Results across two different change detection paradigms suggest that the lack of effect of agents' object-oriented gaze on participants' encoding performance is not due to the specific temporal sequence or spatial pairing cue used. There are, however, a number of other factors that may account for the minimal effect of agents' object-oriented gaze, and which could be considered when carrying out further examinations. Firstly, throughout Chapters 6 and 7, the distances between the agents and the objects gazed upon by the agents were designed to have minimal contribution towards any binding that might be observed (i.e., all agents and objects are equally distant from one another). One possibility is that in order to achieve integration of agent and object information, it is necessary that the agent object pairings are spatially close to one another. Roberts and Humphreys (in prep) found that the binding of congruently paired actions and objects required the stimuli to be in close proximity to one another. Xu (2006) also found that object-parts presented in close proximity to one another are bound together more easily than object-parts that are further apart from one another. Hence, further manipulations of the proximity between agents and objects may provide more optimised parameters for the observation of binding.

Secondly, in the current design, the differences in agents' gaze directions (either gaze towards objects or gaze away from objects) have never been presented in a single display. Participants only observe agents' varied gaze directions either between trials (Experiments 6, 7, and 9b) or across different displays within the same trial (Experiment 8). This is dissimilar to studies that employed flicker change detection paradigm (Freeth, Ropar, Chapman, & Mitchell, 2010; Langton et al., 2006). In these studies, participants observe an agent looking towards one particular object amongst an array of objects. This allows presentation of the agent's gaze towards and gaze away from objects within a single display, providing a stronger contrast of different types of agent and object relationships than the current design. It is possible that the contrast of agents' varied gaze directions within a single display may facilitate the observation of a binding effect sensitive to agents' gaze directions.

Thirdly, the experiments reported in Chapters 6 and 7 presented agents positioned in the lateral regions of the visual field. We know that when healthy adults make speeded judgements about the gaze directions displayed by a central face stimulus, their responses are unaffected by the gaze directions of a secondary distracting face stimulus presented in peripheral position (Burton, Bindemann, Langton, Schweinberger, & Jenkins, 2009). This suggests that automatic processing of eye gaze is likely to require the stimuli to be in the focus of attention, and that positioning the face stimuli centrally helps achieve this. Studies that have demonstrated automatic gaze processing (e.g., Frischen et al., 2007; Langton et al., 2006) also position the avatars in the centre of the displays. Further examination of agents' object-oriented gaze could benefit from positioning the avatars centrally in the field of vision.

Further investigations notwithstanding, Chapters 6 and 7 provide a systematic examination of the possible effect of agents' object-oriented gaze on visual working memory encoding. The finding that participant-oriented gaze impairs visual working memory, suggests that automatic processing of agents' gaze may affect performances beyond those that concern one's allocation of attention. Furthermore, as previously mentioned, there are common features amongst individuals' automatic processing of various social perceptual cues in the environment (e.g., eye gaze and stereotypic attributes). I will discuss these findings along with the relevant findings from Chapters 3 and 4 in the following sections.

8.4 Discussion of General Findings

8.4.1 *Efficient Social Cognitive Processes*

The social cognition literature provides plenty of examples demonstrating that individuals automatically process certain social stimuli regardless of the relevance of the stimuli for the task at hand. These operations have been found in the following domains: stereotyping triggered by features that are associated with the stereotyped group (e.g., Devine, 1989); representing others' task contents when sharing a task with others (e.g., Sebanz et al., 2003); allocating attention to the locations attended by others' object-oriented gaze (e.g., Frischen et al., 2007); and implicitly computing others' visual perspective when explicitly judging one's own visual perspective (Samson et al., 2010). These efficient social cognitive processes exhibit a common set of features that have been identified as critical aspects of automaticity as defined by Payne and Bishara (2009). Namely, they are stimulus-driven, cannot be interrupted once started, operate effortlessly, and run in parallel with the non-automatic processes. The aim of the following section is not to debate whether these efficient social cognitive processes are truly automatic; rather, I will attempt to use two key features of automaticity to demonstrate the limitations and constraints in these efficient social cognitive operations. On the basis of the current findings, I will also argue that although individuals have little control over the onset and offset of these efficient social cognitive and social perceptual processes, they do show some flexibility in certain aspects. In the following sections I will discuss the implications that these efficient yet predominantly inflexible processes might have on the account of dual operational systems in social cognition, including mindreading.

8.4.2 *Features of Automaticity in Efficient Social Cognitive Processes*

8.4.2.1 *'Cannot Be Interrupted Once Started'*

Evidence from a number of studies suggests that individuals cannot inhibit the activation of efficient social cognitive processes even when consciously applying overt strategies. For example, a gaze attentional cueing study demonstrated that even when the target objects were four times more likely to onset on the side opposite to an avatar gaze direction (incongruent trials), participants' responses were nevertheless slower on these trials than on congruent trials (Driver et al., 1999). This is indicative of individuals' inability to apply strategy capable of overcoming the automatic processing of eye gaze. In a similar vein, in the Samson et al. (2010) study described earlier (Section 1.4.2.4), the additional processing cost associated with the implicit computation of the avatar's visual perspective occurs even when participants are merely required to judge their own perspective throughout the entire experiment and never the avatar's.

The present thesis further illustrates individuals' inability to apply top-down strategies to prevent the automatic processing of certain social stimuli. Rather, the onset and cessation of these processes are likely to be determined by 'bottom-up' inputs as well as the capacity of automatic visual perspective-taking. Chapters 3 and 7 lend support to this interpretation. Experiment 3 in Chapter 3 showed that altercentric intrusion only occurred when participants were required to attend to both subsets of dots to judge the content of their own perspective or when their perspective contents could be subitized. The findings of Experiments 9a and 9b in Chapter 7 demonstrated that participants' proclivity for attending to participant-oriented gaze stimuli, even when it lead to reduced encoding accuracy in visual working memory. Experiments 10a and 10b revealed that despite participants' failure to employ top-down strategies to prevent processing of participant-oriented gaze, some bottom-up inputs could aid recovery from a disrupted encoding.

8.4.2.2 *'Stimulus-Driven'*

Evidence from different domains of social cognition suggests that efficient social cognitive processes are often either stimulus-driven or context-driven. For example, both object-oriented gaze (e.g., Frischen et al., 2007) and participant-oriented gaze (e.g., Conty et al., 2006) affect attention allocation. Relatedly, adults automatically compute contents viewed by agents (Samson et al., 2010). Task-sharing contexts have been found to trigger the automatic representation of others' task contents (e.g., Sebanz et al., 2003) and the mere presence of features associated with a stereotyped group activates stereotypical thoughts (Devine, 1989). Chapter 4 of the current thesis critically examines the utility of such stimulus-driven processes by studying the selection of relevant information during automatic visual perspective-taking. Without the embedded ability to select relevant information, the automatic computation of an agent's visual content would be impractical. An analogous phenomenon observed in day-to-day social interaction is the following of others' finger pointing. As with automatic visual perspective-computation, quick head turns to follow the direction of another's finger pointing is likely to be driven by the pointing to stimuli in one's sight. Despite having little conscious control over our responses to pointing stimuli, most of us find it trivially easy to successfully identify the correct referent to which others' finger pointing picks out. It is possible that whilst the following of others' finger pointing is automatised, the identification of correct referents is not, as it is likely to depend upon on the specific context. It would unlikely be so successful if it depended upon particular stimuli. However, given the two processes (finger following and referent identification) are frequently coupled, it is possible that the identification of correct referents becomes more efficient and less effortful over time. This could also account for the finding from Experiments 4 and 5 that although visual perspective-computation is likely to be stimulus-driven, it has sufficient capacity to allow for the selection of relevant information, adding utility to automatic perspective-computation. The current findings show consistency with Samson et al. Experiment 3, in which a non-social figure (a stick) failed to trigger participants' computation of the contents on the either side of the stick. Although participants appear unable to consciously perform these efficient social cognitive processes, these processes are likely to be tuned by experience to compute information that is relevant to particular social contexts.

8.4.3 *Two Systems of Social Cognition*

Efficient social cognitive processes across various domains demonstrate a number of common features that are not observed in the more flexible social cognitive operations. The two systems of social cognition have distinct characteristics for coping with different demands (Bargh, 1994; Gilbert, 1998). Variations of the two systems account has been applied in the domains of stereotyping (e.g., Devine, 1989), attitude change (Gawronski & Bodenhausen, 2006), and most recently mindreading (Apperly & Butterfill, 2009). As described in Chapter 1, the mindreading literature had focused primarily on effortful and late-developing social reasoning abilities in typically developing children (e.g., Wimmer & Perner, 1983) and in autistic children (e.g., Baron-Cohen, Leslie, & Frith, 1985). Evidence indicates that individuals' flexible mindreading abilities are closely associated with their executive function and language abilities (e.g., Carlson & Moses, 2001; Milligan et al., 2007). However, recent findings suggest that young infants have the ability to make inference about others' mental states (e.g., Onishi & Baillargeon, 2005). The infancy findings indicate that there is an efficient mindreading capacity that does not heavily depend on executive functions or language abilities. The two-system approach to mindreading

(Apperly & Butterfill, 2009) provides a framework for reconciling the findings from preschoolers (e.g., Wimmer & Perner, 1983) with the more recent findings from infants (e.g., Onishi & Baillargeon, 2005). The current thesis focused on efficient mindreading in adults, examining the amount of flexibility that the efficient mindreading system comprises.

8.4.4 *Implications for the Efficient Mindreading System*

The current thesis addressed the following issues related to the efficient mindreading system. Firstly, the efficient mindreading system is limited in its computational capacity. Chapter 3 demonstrated that participants only automatically compute an avatar’s perspectives when the concurrent enumerative load is small or when they are required to attend to a subset of dots that represent the avatar’s perspective in order to complete the computation for their own perspective contents. Secondly, in spite of the limitations of automatic visual perspective-computation, Chapter 4 found evidence for flexibility within the efficient mindreading system that allows selection of relevant information. Thirdly, Chapters 6 and 7 showed that it is unlikely that the mere observation of agents’ object-oriented gaze has enduring effects on visual working memory performance. Although the automatic processing of agents’ gaze and visual perspectives affects participants’ speeded responses, it is likely that these effects are transient, and hence do not affect participants’ visual working memory performance. This interpretation is consistent with Cohen and German’s (2009) findings that although adults encode others’ beliefs without overt instructions, the maintenance of such information is brief. This suggests that the efficient mindreading system allows the online computation of perspectives but does not extend the storage of this perspective information. Lastly, as described in the Section 8.4.2.1, Chapter 7 showed that the impaired visual working memory encoding under participant-oriented gaze conditions could not be overcome by top-down strategies. Rather, bottom-up visual inputs were found to facilitate participants to recover from a disrupted encoding. This indicates that participants have minimal flexible control over the processing of participant-oriented gaze. Moreover, the onset and offset of such processing is likely to rely on the presence of particular stimuli. On balance, the current investigation demonstrated that there is some flexibility in the efficient mindreading system combined with several aspects over which individuals have little control. The current findings support the proposal that there are two mindreading systems operating under different contexts and demands (Apperly & Butterfill, 2009). Nevertheless, the current findings indicate that the efficient mindreading system also comprises some flexibility. Therefore, the division between the two systems should not by itself depend on the associated the efficiency and flexibility characteristic of each system.

8.5 Future Directions

8.5.1 *Outstanding Questions*

In the current discussion, a number of issues have been raised which require further investigation. Firstly, the current interpretation of the hindering effect of participant-oriented gaze is that participants spontaneously allocate attention to agents' eye gaze, leaving little processing resources available for encoding the rest of the scene. The justification for this interpretation comes from studies which indicate that participant-oriented gaze attracts visual attention (e.g., Conty et al., 2006). However, this interpretation of the effect of participant-oriented gaze needs to be further verified with direct tests. One way to address this is through the use of eye-tracking techniques, which would allow us to examine navigation of visual attention. In particular, the participant-oriented gaze studies described in Chapter 7 reveal that participant-oriented gaze could either facilitate or hinder task performance. On the one hand, participants show enhanced performance on tasks that require reading others' eyes, such as face recognition and emotion processing (e.g., Adams et al., 2010). However, participants show decreased performance on tasks that do not require reading others' eyes, such as Stroop task, visual scene encoding, and making judgements about object locations (e.g., Conty et al., 2010; present thesis Chapter 7). Based on these findings, it appears that the degree to which participant-oriented gaze hinders or facilitates participants' task performance depends on the relevance of the eye stimuli for the task. As other studies reveal, participant-oriented gaze stimuli attract a vast amount of visual attention (e.g., Conty et al., 2006; Senju et al., 2005). It is likely, therefore, that tasks which involve reading others' eyes will recruit additional attention to the eye stimuli, enhancing task performance. However, when additional attention is allocated to eye stimuli that are irrelevant to the current task, such as in Chapter 7, participant-oriented gaze will hinder task performance. That said, there is currently little evidence to support the assertion that differences in attention allocation produces the discrepant findings across studies of participant-oriented gaze. Supplementing these studies with eye-tracking techniques would further our understanding of attention allocation to participant-oriented gaze across varied task contents.

The application of eye-tracking techniques could also provide insight into participants' processing style when they automatically compute an avatar's perspective. In Chapter 3, it was noted that participants automatically compute an avatar's perspective only when it was necessary to attend both subsets of dots in order to make a 'yes' response about their own perspective. Furthermore, in the same chapter, I speculated that participants may use a computational strategy on 'no' response trials, which allows them to produce correct responses without completing the computation of their own perspective contents. For both yes- and 'no' responses, attending either subset of dots prior to the other subset of dots would not affect participants' computation of their own perspective. However, were automatic visual perspective-computation triggered by social stimuli, one would expect that the subset of dots seen by the avatar would be prioritised in attention. At present what can be said is that participants automatically compute others' visual perspectives, and that further examination is required to advance our understanding of the underlying attentional processes.

Chapters 6 and 7 indicate that it is unlikely that visual working memory encoding is affected by observation of others' object-oriented gaze. Two change detection paradigms with different temporal sequences were employed. The two paradigms produced similar results suggesting that the lack of effect of agents' object-oriented gaze was unlikely to be a function of the specific temporal sequences used. Nonetheless, Friesen and Kingstone (1998) found gaze cueing effects with stimulus onset asynchronies (SOA) of 105 ms. This effect was also observed in SOA of 300 ms and 600 ms, but not when SOA was 1005 ms. It is possible that agents' object-oriented gaze affects individuals' encoding performance only up to sometime between 600 ms and 1005 ms after the onset of gaze stimulus. The time frame for the decay of gaze cueing effect may account for the findings from Chapters 6 and 7; as Chapter 6 contained a 1000-ms delay between the final encoding display and retrieval, similarly Chapter 7 contained a retention interval of 900 ms. This possibility could be explored with a shortened retention interval.

8.5.2 *Further Investigations in Other Domains of Social Cognition*

The current thesis demonstrated effects that bear close resemblance to some social cognitive processes. Here I will discuss the relations between the current findings and those in domains of social cognition. In doing so, I will explore avenues for future research. Firstly, the current results indicate that individuals lack top-down control over various efficient social perceptual processes such as visual perspective-computation and processing of participant-oriented gaze. Teufel et al. (2009) revealed that when information regarding the agent's eye gaze was not available from bottom-up input, the presence of the 'gaze cueing effect' depended on instructions regarding the agent's visual access. One possibility is that the hindering effect of participant-oriented gaze demonstrated in the current experiments is also subject to top-down manipulations. For example, as per Teufel et al., the actual observation of an agents' eye gaze may not be essential for the production of the effects of participant-oriented gaze reported here.

Secondly, as previously described, adults have been shown to represent their task partner's task contents in a manner similar to their own (Sebanz et al., 2003). In Chapter 4, I demonstrated that participants were successful in selecting relevant information whilst computing both their own and an avatar's perspective. However, if participants were performing in a task-sharing context with one individual responsible for the 'relevant information' and the other responsible for the 'distractors', then a different pattern of results might be observed. Although the 'distractors' would still be irrelevant to participants' own task performance, they would be relevant to the other participant's performance. Therefore, having an agent attending to contents irrelevant to one's own task could lead to a failure to select the relevant information[15].

[15] I thank Steven Butterfill for the stimulating discussion at an EPS conference.

Thirdly, the current investigation demonstrates that participants automatically compute others’ visual perspectives based on their accessibility to visual information. Evidence suggests that adults automatically compute others’ Level-1 visual perspective contents (Samson et al., 2010) but not Level-2 visual perspective contents (Surtees, Apperly, & Samson, submitted). If the perspective information regarding accessibility (i.e., whether someone can see the dots) is processed automatically, then one might expect to observe a similar effect in modalities other than vision. For example, an individual’s decision to whisper something into another individual’s ear is directed by the intention for other individuals nearby not to hear the content. This type of processing or computation clearly involves consideration of others’ accessibility to the vocal contents and potentially others’ mental states. An investigation of this sort could help develop our understanding of cross-modal perspective-taking, which may have profound implications for day-to-day sociocommunication and socio-functioning.

8.6 Conclusion

The current thesis examined adults’ computation and encoding of social perceptual information with a visual perspective-taking paradigm (Samson et al., 2010) and visual working memory measures (Luck & Vogel, 1997; Wood, 2008). I have demonstrated that adults automatically process others’ visual perspectives and participant-oriented gaze even when it is disadvantageous to do so. On the basis of this evidence I argue that basic computation and encoding of information are affected by the social components it contains. Moreover, individuals have little control over the activation of such processes. The current investigation advances our understanding of both the limitations and flexibility of the efficient mindreading system. It also broadens our knowledge of the effects that processing of participant-oriented gaze has on social cognition beyond one's allocation of visual attention. Finally, the current findings open several avenues for future investigations in domains such as attention navigation when processing others' eye gaze and the top-down and bottom-up processes involved in social perception.

APPENDIX A: RESPONSE TIME AND ERROR RATES FROM EXPERIMENTS 1, 2, AND 3 IN CHAPTER 3

Table A.1 Response time (RT) and error rate summary for 'yes' response trials.

		Other				Self			
		Small-number		Large-number		Small-number		Large-number	
		Congruent	Incongruent	Congruent	Incongruent	Congruent	Incongruent	Congruent	Incongruent
E1	RT (ms)	716.83	876.44	1499.15	1604.21	765.66	876.45	1616.31	1660.43
	Error Rate (%)	0.42	6.15	1.67	2.22	0.56	4.72	0.00	6.90
E2	RT (ms)	531.92	593.96			597.07	621.69		
	Error Rate (%)	0.52	5.03			2.34	7.81		
E3	RT (ms)	639.25	710.38	1216.25	1338.15	721.70	767.60	1479.85	1676.20
	Error Rate (%)	1.67	3.75	1.25	5.42	0.83	5.00	2.08	7.50

Table A.2 Response time and error rate summary for 'no' response trials

		Other				Self			
		Small-number		Large-number		Small-number		Large-number	
		Congruent	Incongruent	Congruent	Incongruent	Congruent	Incongruent	Congruent	Incongruent
E2	RT (ms)	550.40	589.00			573.87	580.69		
	Error Rate (%)	1.56	5.38			5.21	10.24		
E3	RT (ms)	706.91	776.67	1210.15	1336.04	675.15	719.72	1346.96	1336.02
	Error Rate (%)	0.42	4.58	7.50	8.33	1.67	4.17	7.92	9.17

Table A.3 Response time and error rate summary for 'no' response self conditions probe-types. C:congruent condition, IC: incongruent condition.

		Small Number						Large Number					
		Avatar		Wall		Novel		Avatar		Wall		Novel	
		C	IC	C	IC	C	IC	C	IC	C	IC	C	IC
E2	RT (ms)	583.28	619.90	521.63	531.65	595.56	602.06						
	Error Rate (%)	1.56	12.50	3.12	2.08	8.33	11.25						
E3	RT (ms)	650.12	738.49			698.17	704.82	1462.91	1385.23			1239.22	1288.77
	Error Rate (%)	1.67	8.33			1.67	0.00	10.00	15.00			5.83	3.33

APPENDIX B: RESPONSE TIME AND ERROR RATES FROM EXPERIMENTS 4 AND 5 IN CHAPTER 4

Table B.1 Response time and error rate summary for number condition

		Other				Self			
		Without Distractor		With Distractor		Without Distractor		With Distractor	
		Congruent	Incongruent	Congruent	Incongruent	Congruent	Incongruent	Congruent	Incongruent
E4	RT (ms)	680.70	751.47	695.43	874.87	695.30	761.42	713.90	800.46
	Error Rate (%)	0.52	5.73	0.52	8.85	0.00	3.65	1.56	5.73
E5	RT (ms)	613.92	710.38	644.95	788.04	621.04	665.14	638.17	690.49
	Error Rate (%)	0.00	3.65	0.52	8.85	0.00	2.08	0.52	6.25

Table B.2 Response time and error rate summary for zero condition

		Other				Self			
		Without Distractor		With Distractor		Without Distractor		With Distractor	
		Congruent	Incongruent	Congruent	Incongruent	Congruent	Incongruent	Congruent	Incongruent
E5	RT (ms)	617.82	736.35	656.53	673.13	581.87	682.34	716.67	704.03
	Error Rate (%)	2.08	5.73	7.81	7.81	5.73	5.21	6.69	11.46

REFERENCES

Adams, R. B. Jr., & Franklin, R. G. J. (2009). Influence of emotional expression on the processing of gaze direction. *Motivation & Emotion*, *33,* 106-112.

Adams, R. B. Jr., Pauker, K., & Weisbuch, M. (2010). Looking the other way: The role of gaze direction in the cross-race memory effect. *Journal of Experimental Social Psychology*, *46*, 478-481.

Akechi, H., Senju, A., Kikuchi, Y., Tojo, Y., Osanai, H., & Hasegawa, T. (2010). The effect of gaze direction on the processing of facial expressions in children with autism spectrum disorder: An ERP study. *Neuropsychologia*, *48)*, 2841-2851.

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

Apperly, I. A., Back, E., Samson, D., & France, L. (2008). The cost of thinking about false beliefs: Evidence from adults' performance on a non-inferential theory of mind task. *Cognition*, *106*, 1093-1108.

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, *116,* 953-970.

Apperly, I. A., Carroll, D. J., Samson, D., Humphreys, G. W., Qureshi, A., & Moffitt, G. (2010). Why are there limits on theory of mind use? Evidence from adults' ability to follow instructions from an ignorant speaker. *Quarterly Journal of Experimental Psychology*, *63*, 1201-1217.

Apperly, I. A., Riggs, K. J., Simpson, A., Chiavarino, C., & Samson, D. (2006). Is belief reasoning automatic? *Psychological Science*, *17,* 841-844.

Apperly, I. A., Samson, D., Carroll, N., Hussain, S., & Humphreys, G. (2006). Intact first- and second-order false belief reasoning in a patient with severely impaired grammar. *Social Neuroscience, 1*, 334-348.

Apperly, I. A., Samson, D., Chiavarino, C., & Humphreys, G. W. (2004). Frontal and temporo-parietal lobe contributions to theory of mind: Neuropsychological evidence from a false-belief task with reduced language and executive demands. *Journal of Cognitive Neuroscience, 16*, 1773-1784.

Apperly, I. A., Warren, F., Andrews, B.J., Grant, J. & Todd, S. (2011). Developmental continuity in theory of mind: Speed and accuracy of belief-desire reasoning in children and adults. *Child Development*, 82, 1691-703.

Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology*, *35*, 1311-1320.

Back, E., & Apperly, I. A.. (2010). Two sources of evidence on the non-automaticity of true and false belief ascription. *Cognition*, *115*, 54-70.

Baddeley, A. D., & Hitch, G. (1993). The recency effect: Implicit learning with explicit retrieval? *Memory & Cognition*, *21*, 146-155.

Bargh, J. A. (1994). The Four Horsemen of automaticity: Awareness, efficiency, intention, and control in social cognition. In R. S. Wyer, Jr., & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., pp. 1-40). Hillsdale, NJ: Erlbaum.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition*, *21*, 37-46.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The ‘reading the mind in the eyes’ test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, *42*, 241-251.

Bayliss, A. P., Frischen, A., Fenske, M. J., & Tipper, S. P. (2007). Affective evaluations of objects are influenced by observed gaze direction and emotional expression. *Cognition*, *104,* 644-653.

Bayliss, A. P., Paul, M. A., Cannon, P. R., & Tipper, S. P. (2006). Gaze cuing and affective judgments of objects: I like what you look at. *Psychonomic Bulletin & Review*, *13,* 1061-1066.

Beattie, G. W. (1981). A further investigation of the cognitive interference hypothesis of gaze patterns during conversation. *British Journal of Social Psychology, 20,* 243–248.

Bennett, J. (1978). Some remarks about concepts. *Behavioral and Brain Sciences*, *1*, 557-560.

Besner, D., Davies, J., & Daniels, S. (1981). Reading for meaning. The effects of concurrent articulation. *Quarterly Journal of Experimental Psychology*, *33,* 415-437.

Bialystok, E., & Codd, J. (2000). Representing quantity beyond whole numbers: Some, none, and part. *Canadian Journal of Experimental Psychology*, *54,* 117-128.

Birch, S. A. J., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science*, *18,* 382-387.

Bird, C. M., Castelli, F., Malik, O., Frith, U., & Husain, M. (2004). The impact of extensive medial frontal lobe damage on ‘Theory of Mind’ and cognition. *Brain*, *127,* 914-928.

Bull, R., Phillips, L. H., & Conway, C. A. (2008). The role of control functions in mentalizing: Dual-task studies of theory of mind and executive function. *Cognition*, *107,* 663-672.

Burton, A. M., Bindemann, M., Langton, S. R. H., Schweinberger, S. R., & Jenkins, R. (2009). Gaze perception requires focused attention: Evidence from an interference task. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 108-118.

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*, 1032-1053.

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*, *92*, 73-92.

Carlson, S. M., Moses, L. J., & Claxton, L. J. (2004). Individual differences in executive functioning and theory of mind: An investigation of inhibitory control and planning ability. *Journal of Experimental Child Psychology*, *87*, 299-319.

Chaiken, S., & Trope, Y. (Eds.). (1999). *Dual process theories in social psychology*. New York: Guilford Press.

Cohen, A. S., & German, T. C. (2009). Encoding of others' beliefs without overt instruction. *Cognition*, *111*, 356-363.

Conty, L., Gimmig, D., Belletier, C., George, N., & Huguet, P. (2010). The cost of being watched: Stroop interference increases under concomitant eye contact. *Cognition, 115*, 133-9.

Conty, L., Tijus, C., Hugueville, L., Coelho, E., & George, N. (2006). Searching for asymmetries in the detection of gaze contact versus averted gaze under different head views: A behavioural study. *Spatial Vision*, *19*, 529-545.

Cozby, P. C. (2009). *Methods in behavioral research* (5th ed.). Mountain View, CA, US: Mayfield Publishing.

de Villiers, J. G., & de Villiers, P. A. (2000). Linguistic determinism and false belief. In P. Mitchell, & K. Riggs (Eds.), *Children's reasoning and the mind* (pp. 191-228). Hove, UK: Psychology Press.

Dennett, D. (1978). Beliefs about beliefs. *Behavioral and Brain Sciences, 1*, 568-570.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56,* 5-18.

Doherty-Sneddon, G., Bruce, V., Bonner, L., Longbotham, S., & Doyle, C. (2002). Development of gaze aversion as disengagement from visual information. *Developmental Psychology*, *38*, 438-445.

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

Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective Taking as Egocentric Anchoring and Adjustment. *Journal of Personality and Social Psychology*, *87,* 327-339.

Farell, B. (1985). “Same”-“different” judgments: A review of current controversies in perceptual comparisons. *Psychological Bulletin, 98*, 419-456.

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences, U.S.A, 99*, 9602-9605.

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin*, *133*, 694-724.

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

Fletcher, P. C., Happé, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S. J., & Frith, C. D. (1995). Other minds in the brain: A functional imaging study of “theory of mind” in story comprehension. *Cognition*, *57*, 109-128.

Freeth, M., Ropar, D., Chapman, P., & Mitchell, P. (2010). The eye gaze direction of an observed person can bias perception, memory, and attention in adolescents with and without autism spectrum disorder. *Journal of Experimental Child Psychology*, *105*, 20-37.

Friedman, O., & Leslie, A. M. (2004a). A developmental shift in processes underlying successful belief-desire reasoning. *Cognitive Science*, 28, 963-977.

Friedman, O., & Leslie, A. M. (2004b). Mechanisms of belief-desire reasoning: Inhibition and bias. *Psychological Science*, *15*, 547-552.

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

Frith, C. D., & Frith, U. (1999). Interacting minds- A biological basis. *Science*, *286*, 1692-1695.

Frith, C. D., & Frith, U. (2006). The neural basis of mentalizing. *Neuron*, *50*, 531-534.

Frye, D., Zelazo, P.D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Child Development*, *10*, 483-527.

Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of ‘theory of mind’. *Trends in Cognitive Sciences*, *7*, 77-83.

Gallagher, H. L., Happé, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: An fMRI study of ‘theory of mind’ in verbal and nonverbal tasks. *Neuropsychologia*, *38*, 11-21.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, *132*, 692-731.

German, T. P., & Hehman, J. A. (2006). Representational and executive selection resources in ‘theory of mind’: Evidence from compromised belief-desire reasoning in old age. *Cognition*, *101*, 129-152.

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3½-7 years old on a Stroop-like day-night test. *Cognition*, *53*, 129-153.

Gilbert, D. T. (1998). Ordinary personology. In D. T. Gilbert, S. T., Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., pp. 89-150). New York: McGraw Hill.

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences*, *20*, 1–19.

Goldin-Meadow, S., Wein, D., & Chang, C. (1992). Assessing knowledge through gesture: Using children’s hands to read their minds. *Cognition and Instruction*, *9*, 201–219.

Gopnik, A. & Astington, J. W. (1988). Children’s understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, *59*, 26-37.

Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories*. Cambridge, MA: MIT Press.

Gordon, A. C. L., & Olson, D. R. (1998). The relation between acquisition of a theory of mind and the capacity to hold in mind. *Journal of Experimental Child Psychology*, *68*, 70-83.

Grice, P. (1957). Meaning. *The Philosophical Review*, *66*, 377-388.

Haith, M. M., Bergman, T., & Moore, M. J. (1977). Eye contact and face scanning in early infancy. *Science*, *198*, 853-855.

Happé, F. G. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders*, *24*, 129-154.

Happé, F. G. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development*, *66*, 843-855.

Happé, F. G., Malhi, G. S., & Checkley, S. (2001). Acquired mind-blindness following frontal lobe surgery? A single case study of impaired 'theory of mind' in a patient treated with stereotactic anterior capsulatomy. *Neuropsychologia*, *39*, 83-90.

Hogrefe, G. J., Wimmer, H., & Perner, J. (1986). Ignorance versus false belief: A developmental lag in attribution of epistemic states. *Child Development*, *57*, 567-582.

Hughes, C. (1998). Finding your marbles: Does preschoolers' strategic behavior predict later understanding of mind? *Developmental Psychology*, *34*, 1326-1339.

Hughes, C., Russell, J., & Robbins, T. W. (1994). Evidence for executive dysfunction in autism. *Neuropsychologia*, *32*, 47–92.

Jiang, Y. V., Makovski, T., & Shim, W. M. (2009). Visual working memory for features, conjunctions, objects, and locations. In J. Brockmole (Eds.), *The visual world in memory* (pp. 33-65). Hove and New York: Psychology Press.

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*, 1-19.

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The discrimination of visual number. *The American Journal of Psychology*, *62*, 498-525.

Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): Effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology, 78*, 871-888.

Kawakami, K., Young, H., & Dovidio, J. F. (2002). Automatic stereotyping: Category, trait, and behavioral activations. *Personality and Social Psychology Bulletin*, *28*, 3-15.

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

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

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

Langton, S. R. H., O'Donnell, C., Riby, D. M., & Ballantyne, C. J. (2006). Gaze cues influence the allocation of attention in natural scene viewing. *Quarterly Journal of Experimental Psychology*, *59*, 2056-2064.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279-281.

Macrae, C. N., Bodenhausen, G. V., Schloerscheidt, A. M., & Milne, A. B. (1999). Tales of the unexpected: Executive function and person perception. *Journal of Personality and Social Psychology*, *76*, 200-213.

Maurer, D., & Salapatek, P. (1976). Developmental changes in the scanning of faces by young infants. *Child Development*, *47*, 523-527.

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law*, *7*, 3-35.

Mendoza, S. A., Gollwitzer, P. M., & Amodio, D. M. (2010). Reducing the expression of implicit stereotypes: Reflexive control through implementation intentions. *Personality and Social Psychology Bulletin*, *36*, 512-523.

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

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development*, *78*, 622-646.

Mitchell, P., Robinson, E. J., Issacs, J. E., & Nye, R. M. (1996). Contamination in reasoning about false belief: An instance of realist bias in adults but not children. *Cognition*, *59*, 1-21.

Moll, H., & Tomasello, M. (2006). Level 1 perspective-taking at 24 months of age. *British Journal of Developmental Psychology*, *24*, 603-613.

O'Neill, D. K. (1996). Two-year-old children's sensitivity to a parent's knowledge state when making requests. *Child Development*, *67*, 659-677.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*, 255-258.

Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive function deficits in high-functioning autistic individuals: Relationship to theory of mind. *Journal of Child Psychology and Psychiatry*, *32*, 1081-1105.

Pashler, H. (1988). Familiarity and visual change detection. *Perception and Psychophysics*, *44*, 369-378.

Payne, B. K., & Bishara, A. J. (2009). An integrative review of process dissociation and related models in social cognition. *European Review of Social Psychology*, *20*, 272-314.

Perdue, C. W., & Gurtman, M. B. (1990). Evidence for the automaticity of ageism. *Journal of Experimental Social Psychology, 26*, 199-216.

Peterson, C. C., & Slaughter, V. P. (2006). Telling the story of theory of mind: Deaf and hearing children's narratives and mental state understanding. *British Journal of Developmental Psychology*, *24*, 151-179.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*, 3-25.

Pratto, F., & Bargh, J. A. (1991). Stereotyping based on apparently individuating information: Trait and global components of sex stereotypes under attention overload. *Journal of Experimental Social Psychology*, *27*, 26-47.

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

Pylyshyn, Z. W. (1978). When is the attribution of beliefs justified? *Behavioral and Brain Sciences*, *1,* 592-593.

Quinn, K. A., Mason, M. F., & Macrae, C. N. (2010). When Arnold is "the terminator", we no longer see him as a man: The temporal determinants of person perception. *Experimental Psychology, 57*, 27-35.

Qureshi, A. W. (2009). The cognitive basis of mental state reasoning in adults. Unpublished doctoral thesis, University of Birmingham, UK.

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*, 230-236.

Reid, V. M., & Striano, T. (2005). Adult gaze influences infant attention and object processing: Implications for cognitive neuroscience. *European Journal of Neuroscience*, *21*, 1763-1766.

Rensink, R. A. (2002). Change detection. *Annual Review of Psychology*, *53*, 245-277.

Roberts, K. L., & Humphreys, G. W. Action relationship and binding in visual working memory. Manuscript in preparation.

Ruby, P., & Decety, J. (2003). What you believe versus what you think they believe: A neuroimaging study of conceptual perspective-taking. *European Journal of Neuroscience*, *17*, 2475-2480.

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

Samson, D., Apperly, I. A., Kathirgamanathan, U., & Humphreys, G. W. (2005). Seeing it my way: A case of a selective deficit in inhibiting self-perspective. *Brain*, *128*, 1102-1111.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in “theory of mind”. *Neuroimage*, *19*, 1835-1842.

Scarborough, D. L. (1972). Memory for brief displays of symbols. *Cognitive Psychology*, *3*, 408-429.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002a). *E-Prime user's guide.* Pittsburgh: Psychology Software Tools Inc.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002b). *E-Prime reference guide.* Pittsburgh: Psychology Software Tools Inc.

Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: Just like one's own? *Cognition*, *88*, B11-B21.

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, *18*, 668-671.

Senju, A., Hasegawa, T., & Tojo, Y. (2005). Does perceived direct gaze boost detection in adults and children with and without autism? The stare-in-the-crowd effect revisited. *Visual Cognition*. *12*, 1474-1496.

Senju, A., Kikuchi, Y., Hasegawa, T., Tojo, Y., & Osanai, H. (2008). Is anyone looking at me? Direct gaze detection in children with and without autism. *Brain and Cognition*, *67*, 127-139.

Simons, D. J. (2000). Attentional capture and inattentional blindness. *Trends in Cognitive Sciences*, *4*, 147-155.

Slater, A., & Butterworth, G. (1997). Perception of social stimuli: Face perception and imitation. In G. Bremner, A. Slater, & G. Butterworth (Eds.), *Infant development: Recent advances* (pp. 223-245). Hove, UK: Psychology Press.

Sodian, B., Thoermer, C., & Metz, U. (2007). Now I see it but you don't: 14-month-olds can represent another person's visual perspective. *Developmental Science*, *10*, 199-204.

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, *18*, 587-592.

Sperber, D., & Wilson, D. (1995). *Relevance: communication and cognition* (2nd ed.). Oxford: Basil Blackwell.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Momographs: General and Applied, 74(11)*, 1-29.

Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience*, *10*, 640-656.

Striano, T., Kopp, F., Grassmann, T., & Reid, V. M. (2006). Eye contact influences neural processing of emotional expressions in 4-month-old infants. *SCAN*, *1*, 87-94.

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science*, *18*, 580-586.

Surtees, A., Apperly, I. A., & Samson, D. (2011). Automatic processing of level-1 but not level-2 visual perspectives. Manuscript submitted for publication.

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

Tipples, J. (2002). Eye gaze is not unique: Automatic orienting in response to uninformative arrows. *Psychonomic Bulletin & Review*, *9*, 314-318.

Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review*, *101*, 80-102.

Varley, R., & Siegal, M. (2000). Evidence for cognition without grammar from causal reasoning and ‘theory of mind’ in an agrammatic aphasic patient. *Current Biology*, *10*, 723-726.

Varley, R., Siegal, M., & Want, S. C. (2001). Severe impairment in grammar does not preclude theory of mind. *Neurocase*, *7*, 489-493.

Vogeley, K., Bussfeld, P., Newen, A., Herrmann, S., Happé, F., Falkai, P., Maler. W., Shah, N. J., Fink, G. R., & Zilles, K. (2001). Mind reading: Neural mechanisms of theory of mind and self-perspective. *Neuroimage*, *14*, 170-181.

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

Wellman, H. M., & Miller, K. F. (1986). Thinking about nothing: Development of concepts of zero. *British Journal of Developmental Psychology*, *4*, 31-42.

Wender, K. F., & Rothkegel, R. (2000). Subitizing and its subprocesses. *Psychological Research*, *64*, 91-92.

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, *131*, 48-64.

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

Winter, L., & Uleman, J. S. (1984). When are social judgments made? Evidence for the spontaneousness of trait inferences. *Journal of Personality and Social Psychology*, *47*, 237-252.

Wolters, G., van Kempen, H., & Wijlhuizen, G. (1987). Quantification of small numbers of dots: Subitizing or pattern recognition? *The American Journal of Psychology*, *100*, 225-237.

Wood, J. N. (2008). Visual memory for agents and their actions. *Cognition*, *108*, 522-532.

Woodward, A. L. (2003). Infants' developing understanding of the link between looker and object. *Developmental Science*, *6*, 297-311.

Xu, Y. (2006). Understanding the object benefit in visual short-term memory: The roles of feature proximity and connectedness. *Perception and Psychophysics*, *68*, 815-828.

Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Child Development*, *11*, 37-63.