# Kent Academic Repository
## Full text document (pdf)

## Citation for published version

Grainger, Catherine (2015) Self-awareness in Children and Adults with Autism Spectrum Disorder. Doctor of Philosophy (PhD) thesis, University of Kent.

## DOI

## Link to record in KAR

http://kar.kent.ac.uk/51629/

## Document Version

UNSPECIFIED

# SELF-AWARENESS IN CHILDREN AND ADULTS WITH AUTISM SPECTRUM DISORDER

*Catherine Sarah Grainger*

School of Psychology

University of Kent

Thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Faculty of Social Science at the University of Kent, Canterbury. March 2015.

This thesis is dedicated to my parents, Michele and Kevin Grainger.

# Acknowledgments

There are a number of people without whom this thesis might not have been written, and to whom I am immeasurably grateful.

Above all, I would like to thank David Williams, for being the most incredible supervisor, mentor and friend. Your constant support and encouragement has been invaluable, and you have greatly inspired me. I would also like to thank Sophie Lind, for her encouragement and support over the past years. It is difficult to express how grateful I am to both of you.

I would like to sincerely thank all of the children and adults who participated in this research. They have been a pleasure to work with and have taught me a lot over the past few years. I would also like to thank all of the families and schools who helped me with this research. Without their support, this research would not have been possible. Additionally, I would like to thank the National Autistic Society, Kent Autistic Trust, and Durham University Service for Students with Disabilities for their assistance with participant recruitment, and Anna Peel, Lucy Elias, Merve Kilic and Maddie Musgrove for their assistance with data collection and coding.

Special thanks also go to my amazing colleagues and friends at both Durham University and the University of Kent. In particular I would like to thank Pete Moseley, David Smailes, Katarina Kaduk, Anna Peel, Hannah Harvey, Julie Van De Vyver, Sally Palmer, Serena Vanzan, John Sabo, Darren McGee, Anita Ruddle, Miriam Tresh and John Allen. Thank you for the inspiring conversations, the many cups of tea and coffee, for keeping me going and for making me smile. Without these friendships the past three years would have been far less enjoyable.

Finally, I would like to thank my amazing family, my Mum and Dad, my sisters Kerry and Rebecca and my wonderful partner Paul. I cannot express in words how incredible you are. Without your constant and unwavering support I would be lost. I love you.

**"We are only as strong as we are united, as weak as we are divided."**
**J.K. Rowling, *Harry Potter and the Goblet of Fire***

# Table of Content

# List of Figures

## List of Tables

# **Abstract**

The primary aim of this thesis was to investigate aspects of self-awareness in individuals with autism spectrum disorder (ASD). Given widely accepted assumptions that "the self" is not a unitary construct, but is instead multifaceted, this thesis explored the extent to which impairments in self-awareness may be "domain-specific". The nine experimental tasks reported in this thesis explored several aspects of self-awareness in children and adults with ASD, including awareness of the physical self, conceptual self, and mental self. Overall, the results of these studies suggest that individuals with ASD demonstrate selective impairments in only some aspects of self-awareness. It is suggested that the pattern of results reported in this thesis best support the suggestion that *physical* self-awareness is intact in individuals with ASD, whilst *psychological* self-awareness is impaired. The theoretical and practical implications of these findings for our current understanding of self-awareness in autism spectrum disorder, and self-awareness in typical development, are discussed.

**CHAPTER ONE**

---

**AN INTRODUCTION TO SELF-AWARENESS IN AUTISM SPECTRUM DISORDER;**

**CONCEPTS, THEORIES, AND RESEARCH**

**Autism Spectrum Disorder**

Autism spectrum disorder (ASD) is a developmental disorder diagnosed on the basis of social-communication deficits, and fixated interests and repetitive behaviours (see American Psychiatric Association, 2013). Although it is widely acknowledged that ASD is not a modern phenomenon, it was first formally recognised in the 1940s by Leo Kanner (1943) and Hans Asperger (1944). Both Kanner and Asperger's independently described a group of children who seemed unable to engage in normal social relationships and social integration. These children also demonstrated consistent patterns of rigid and repetitive behaviour, such as a desire for sameness, and presented facets of ability that are still considered fundamental aspect of ASD today. Remarkably, both Kanner and Asperger chose to describe these children using the term "autistic", a word that originates from the Greek word *autos*, and literally translates as "self"- ism. While the term was first used by Eugen Bleuler to describe characteristics of some individuals with schizophrenia, who were very socially withdrawn, it has since been used to describe individuals like the children described by Kanner and Asperger.

It is now widely acknowledged that ASD is a neurodevelopmental disorder, and is thus present from birth (though symptoms of this disorder may not manifest themselves till later on in development). Recent epidemiology studies suggest that, regardless of age or intellectual ability, around 1% of the population have ASD (e.g., Baird et al., 2006; Brugha et al., 2011; Kadesjo, Gillberg, & Hagberg, 1999). It is also

widely established that more males have ASD than females, and that on average a male to female ratio of 4:1 appears to be the case.

ASD is defined according to behavioural characteristics, as no specific biological markers for this disorder are known.  The Diagnostic and Statistical Manual of Mental Disorders (DSM-V: American Psychiatric Association, 2013) now identifies *two* core areas of impairments in individuals with ASD (in DSM-IV ASD was previously characterised by a triad of impairments; please see Figure 1 for a comparison of DSM-IV and DSM-V).  Firstly, people with ASD demonstrate impairments in social communication and social interaction across multiple contexts.  These impairments manifest themselves in behaviours such as abnormal social interactions during conversation, poor nonverbal communication (such as atypical eye contact, facial expressions and body language) and difficulties developing and understanding relationships.  Additionally, individuals with ASD demonstrate restricted and repetitive patterns of behaviour and interests.  These often manifest themselves in the form of repetitive movements, an insistence on sameness, extremely focussed interests and abnormal sensory activity.  It is widely acknowledged that these impairments manifest on a spectrum, and the severity of social-communication deficits and fixated interests and repetitive behaviours varies across individuals with this disorder.  Alongside these impairments, some individuals with ASD (though not all) will also demonstrate accompanying structural language impairments and/or accompanying intellectual impairments.

ASD is characterised by a unique profile of cognitive strengths and weaknesses, and several cognitive theories have been suggested to explain the behavioural features of the disorder.  Although several theories have been proposed over the years, some theories have taken a particularly prominent stance within the field.  The following

section will outline three particularly prominent theories suggested to explain social communication impairments seen in individuals with ASD, namely the mindblindness theory, the executive dysfunction theory, and the weak central coherence theory.
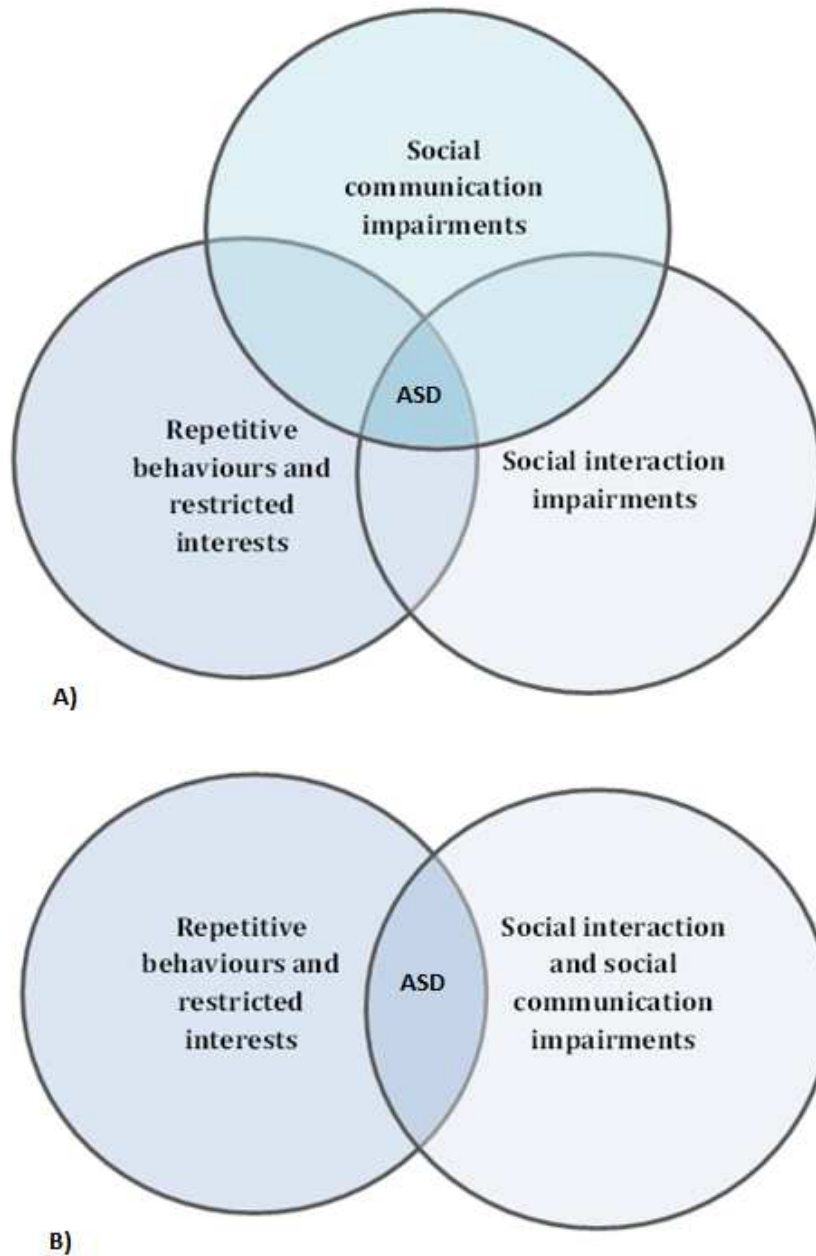


*Figure 1: Graphical representation of the core areas of impairments in autism spectrum disorder, as identified in DSM-IV (A) and DSM-V (B).*

**Cognitive accounts of ASD**

**The mindblindness theory of ASD.** According to its original definition, theory of mind (ToM) is the ability to attribute mental states, such as beliefs, desires, and intentions, to self *and* others in order to explain and predict behaviour (Premack & Woodruff, 1978). Whilst the term "metacognition" is often used to refer to the understanding of one's own mental states, "mindreading" is often used to refer to an individual's ability to infer mental states in others.

As a possible explanation for the social communication impairments found in ASD, it has been suggested that individuals with ASD are delayed (and deviant) in the development of mindreading abilities, and are thus "mindblind" to a certain extent (e.g., Baron-Cohen, Leslie, & Frith, 1985). As a consequence, individuals with ASD demonstrate difficulties understanding other people's beliefs, attitudes, and emotions. Early evidence for this theory comes from findings that many children with ASD do not pass false belief tasks, such as the Sally-Anne task. In the Sally-Anne task, one doll (Sally) places a marble in a basket and then leaves the scene. Whilst doll Sally is away a second doll (Anne) comes along and takes the marble out of the basket. Instead she places the marble into a box. When doll Sally returns, children are asked where Sally will look for the marble. To successfully pass the task children need to understand that Sally will falsely think that the marble is still in the basket, and that her belief will not represent the true state of affairs (that the marble is really in the box). Whilst four-year old neurotypical children (Wimmer & Perner, 1983), and children with other developmental disorders (who have a similar verbal mental ages to the children with ASD) typically pass false belief tasks, children with ASD are often delayed in passing the Sally-Anne task (Baron-Cohen et al., 1985). The vast majority of neurotypical children

will pass false belief tasks by the age of five (and children with other developmental disorders will pass false belief tasks by a mental age of five).  However children with ASD, regardless of their intelligence, do not appear to pass false belief tasks until a mental age of around 10 years (Happé, 1995).  Thus, although it is possible for individuals with ASD to pass tasks that rely on an understanding of others' mental states, this ability is severely developmentally delayed, and it is possible that the mindreading abilities that do develop in ASD may be acquired through atypical strategies and compensatory learning (see below).

Mindblindness accounts of ASD suggest that several characteristics of ASD can be explained by deficits in mindreading.  However, despite still demonstrating impairments in social communication that are characteristic of individuals with ASD, some individuals with ASD, particularly those with typical language abilities, do not show always impairments on typical false belief tasks.  This is not in keeping with the mindblindness theory, and needs to be explained if the mindblindness theory can viably explain social communication impairments in ASD.  However, it is critical to distinguish between undiminished *performance* on such tasks, and actual mindreading competence. Although individuals with ASD may show typical levels of performance on some tests of mindreading, this does not mean that they engage in the same underlying processes during such tasks as those employed by neurotypical individuals.  Several ASD researchers have argued that high-functioning individuals with ASD employ compensatory strategies to "hack out" solutions to mindreading tasks in the absence of true mindreading competence (Bowler, 1992; Happé, 1995).  Evidence for this hypothesis comes from a variety of sources, including, but not limited to (a) atypical patterns of performance across different tasks/conditions (e.g., Surian & Leslie, 1999; Williams & Happé, 2009b), and (b) atypically high associations between ToM task

performance and verbal ability among individuals with ASD (Fisher, Happé, & Dunn, 2005; Lind & Bowler, 2009b).  Additionally, recent evidence suggests that intellectually high-functioning individuals with ASD who pass "classic" mindreading tasks nonetheless fail to spontaneously mindread during other tasks (see Senju, 2012; Senju, Southgate, White, & Frith, 2009).  This suggests that explicit knowledge of others' mental states among individuals with ASD has been acquired in an atypical manner, and not primarily via the mindreading system.

A particular strength of the mindblindness theory is that it can make sense of the social communication difficulties present in ASD.  However it does not appear to account as easily for the "non-social" features of ASD, such as repetitive and restricted interests.  It may be that additional theories are required to explain these features of the disorder.

**Executive functioning accounts of ASD.**  The mindblindess hypothesis is considered an extremely influential account of ASD.  Of course, that is not say that there are not competing theories of ASD, and another important account of ASD is executive dysfunction accounts.  Unlike mindblindness account of ASD, executive functioning accounts can potentially make sense of restricted and repetitive interests and behaviours, stereotypical of this disorder.

Executive functioning (EF) is an umbrella term used to refer to a range of cognitive processes used in mental control and regulation, including planning, decision making, set shifting, working memory and inhibition.  Whilst, executive functions are not needed for routine actions and behaviours, they are crucial when engaging in novel actions, and shifting between separate tasks.  Executive functions are primarily underpinned by the frontal lobes (see Stuss & Knight, 2002), and patients with damage to the frontal lobes often show inflexible behaviours and cognitions, similar to those

demonstrated by individuals with ASD.  As such, this has led to suggestions that

impairments in EF may be responsible for several of the behavioural features seen in

ASD.

In keeping with executive dysfunction accounts of ASD, there is substantial

evidence to suggest that individuals with ASD demonstrate impairments in several

higher order aspects of executive functioning,  including planning ability (e.g., Ozonoff,

Pennington, & Rogers, 1991; Pellicano, 2007) and cognitive flexibility (e.g., Ambery,

Russell, Perry, Morris, & Murphy, 2006; Ozonoff et al., 1991).  However, other aspects of

executive functioning appear to be unimpaired in ASD.  For example individuals with

ASD do not appear to show impairments in inhibition, except when making prepotent

responses (Hill, 2004).  Studies have also shown that individuals with ASD tend to

demonstrate typical performance on classic measures of inhibitory control, such as

Stroop tasks (e.g., Ambery et al., 2006) and Go/No-Go tasks (e.g., Happé, Booth,

Charlton, & Hughes, 2006).  Clearly then, individuals with ASD do not demonstrate

across the board impairment in executive functioning.

One significant strength of  executive dysfunction accounts is that they offers a

credible explanation for at least some of the clinical features of ASD, in particular

restricted, repetitive behaviours and interests (RRBIs) typically demonstrated by

individuals with this disorder.  Indeed, executive dysfunction has been shown to

significantly correlate with the level of RRBIs individuals with ASD manifest (see e.g.,

Boyd, McBee, Holtzclaw, Baranek, & Bodfish, 2009; Lopez, Lincoln, Ozonoff, & Lai,

2005).  Additionally, some have also proposed that impairments in executive

functioning may also explain impaired performance on mindreading tasks in individuals

with ASD (see e.g.,Ozonoff et al., 1991).  One explanation that has been suggested is that

mindreading tasks themselves pose several executive demands, and during

mindreading individuals are required to inhibit their own belief/knowledge about reality, in order to infer another's mental state. As such, impairments in executive functioning, and the ability to inhibit one's own understanding, may be responsible for impaired performance on tasks of mindreading in ASD.

However, there are problems with executive accounts of ASD (see Hill, 2004, for a review). One difficulty is that whilst executive function difficulties are common in ASD, they are not a universal feature of the disorder. Inconsistencies in studies' findings have meant that it has been difficult to reach an overall consensus concerning which specific aspects of executive function are atypical in ASD or not. Additionally, there is little evidence for EF impairments in preschool children with ASD (see Hill, 2004), which suggests EF impairment in ASD may in fact be secondary to other cognitive deficits. As with mindblindness theories of ASD executive dysfunction accounts of ASD might best explain specific impairments found in ASD, but as yet cannot account of all aspects of this disorder.

**Weak central coherence accounts of ASD.** A third major cognitive theory of ASD suggests "weak central coherence" might explain several of the deficits associated with ASD. The term "weak central coherence" refers to detail-focused processing that has been proposed to explain many of the characteristics of ASD. Weak central coherence accounts of ASD suggest that a difference in both low-level and high-level processing means individuals with ASD fail to extract global meaning from stimuli and situations, using local processing styles instead. The weak central coherence hypothesis of autism was originally proposed by Frith (1989, 2003). Uta Frith nicely describes weak central coherence as not being able to see the wood for the trees, whereas strong central coherence, in its extreme, involves not being able to see the trees for the wood (Frith, 2003).

Another way to think about weak central coherence is in terms of a processing style that disregards context.  Context gives meaning to individual parts as a whole, and thus processing overall context can sometimes results in cognitive strengths, other times in weaknesses.  As such, weak central coherence tries to explain several aspects of ASD, in particular the fact that individuals with ASD can demonstrate remarkable ability and focus in some areas (including special interests and restricted and repetitive behaviours), yet at the same time demonstrate cognitive impairments in other areas of functioning, including impairments in social communication and interactions.

Evidence from several studies supports weak central coherence accounts of ASD. For example, a well-documented strength in individuals with ASD is good performance on block design tests.  One suggestion for why standard block design tests are thought to be difficult is that they require individuals to separate an overall design into several smaller, appropriate segments.  Thus, individuals with a strong drive towards central coherence should find these tasks particularly hard.  This idea is supported by studies that have shown that when the job of segmenting the block design into smaller components (which correspond to the appropriate blocks) is done for participants, performance in neurotypical children is drastically improved (e.g., Shah & Frith, 1993).

The idea that individuals with ASD demonstrate good performance on block design tasks (because they demonstrate weak central coherence) was also tested in this study.  Individuals with ASD performed significantly better than controls on the block design task when the designs were unsegmented, but both groups performed well when the designs were segmented (please see Figure 2).  This suggests that individuals with ASD required less effort to segment the block designs in the standard, unsegmented condition, relative to the comparison participants.  Additionally, studies have shown that individuals with ASD score above their mental age on tasks in which they are asked

to find hidden figures embedded in pictures (Jolliffe & Baron-Cohen, 1997; Witkin, Oltman, Raskin, & Karp, 1971).
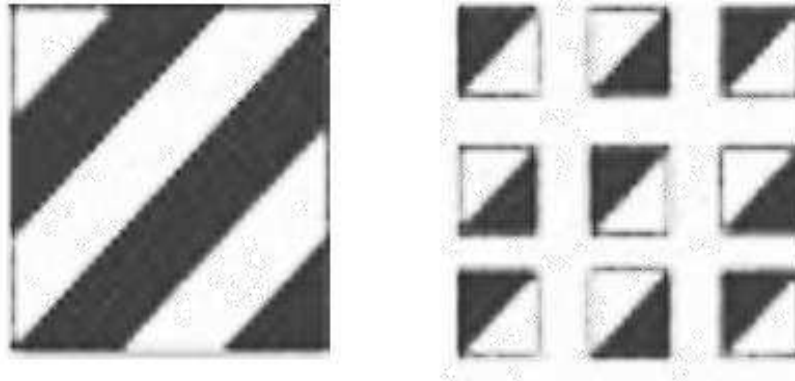


*Figure 2: Example of an unsegmented block design (left) and segmented block design (right), similar to the designs used in Shah and Frith (1993).*

Additionally, evidence for weak central coherence in ASD has been found in several different aspects of processing, not just visual processing but also auditory processing and verbal semantic processing.  For example, studies have shown that individuals with ASD benefit less from the context of meaning in sentences compare to neurotypical individuals (Happé, 1994).  This has led researchers some researchers to suggest that central coherence can be considered a cognitive style that varies within the population (between weak and strong ends of the spectrum) and that individuals with ASD tend to demonstrate cognitive styles at the weak extreme of this spectrum  (e.g., Happé, 1999).

**Fractionation of ASD characteristics**

Of course, this section has only touched on a few of the more prominent conceptual models that attempt to explain strengths and weaknesses in ASD, and is by

no means a comprehensive account of the cognitive theories proposed to explain this disorder. Indeed, it is debatable whether one theory alone can adequately explain all aspects of ASD (see e.g., Happé, Ronald, & Plomin, 2006). As things stand, none of the existing cognitive theories (including those described above) can satisfactorily explain the full range of behaviours seen in individuals with ASD. As such, research has now begun to question whether one theory can explain ASD, and there is a growing trend towards multiple explanations for ASD and suggestions that different aspects of impairments in ASD as separable. For example it has been suggests that whilst mindblindness accounts of ASD may account for social communication impairments in ASD, executive dysfunction theories may explain RRBIs (Happé & Ronald, 2008). Nevertheless, it is important to understand these theories of ASD before moving on to discuss self-awareness in ASD, and before considering what aspects of ASD might explain impairments in some aspects of self-awareness in this disorder.

**Theories of self-awareness**

Several theories of ASD predict that individuals with this disorder should demonstrate impairments in understanding the self. However, before reviewing self-awareness in ASD it is worth discussing theories concerning the typically development of self-awareness.

Within the fields of psychology, philosophy, and cognitive neuroscience the concept of "the self" has been widely discussed. The term is often used to refer to multiple different phenomena and thus a single definition of "the self" is often difficult to define. In early conceptualisations James (1890) proposed that the self could not be considered a single entity; instead the self consists of multiple different *dimensions* including the physical self, mental self, spiritual self, and the ego. Crucially, James

distinguishes between different levels of self/self-awareness; the self as the "I" and the self as the "Me".  As the "I", the self is the *subject* of experience, and is not reflexive in nature.  In contrast, the self as the "Me" is the *object* of experience, and the "Me" makes sense of the "I" acting in the present moment.  According to James, only when the self becomes the object of one's experiences (the "me") do individuals become explicitly self-*aware* (as opposed to merely implicitly *experiencing* the self).  James' early conceptualisations of the self-have influenced later theories that also distinguish between multiple aspects of the self (e.g., the aspects of self-awareness proposed by Neisser, 1988; please see below for more details).  More recently, Rochat (2003) has also suggested that throughout a child's development five levels of self-awareness unfold, developing from self-obliviousness to fully developed self-consciousness (or "meta" self-awareness).

**A framework for understanding self-knowledge in ASD**

One influential conceptualisation of the self is that outlined by Neisser (1988). Neisser suggested that people have access to five different kinds of information about themselves, each type of which *specifies* a different aspect of self; in Neisser's taxonomy, these are ecological, interpersonal, private, extended, and conceptual aspects of the self. Neisser considered these aspects of the self so distinct that he referred to them as "different *selves:* they differ in their origins and developmental history, in what we know about them, in the pathologies to which they are subject, and in the manner in which they contribute to human social experience*" (Neisser, 1988, p386).

The ecological self is the experience of the self in its physical form, and in relation to its physical environment.  The "I" in the ecological self is the person present in a physical place, or engaged in a physical activity.  The interpersonal self is the self as

experienced in social and emotional relationship with others.  The extended self can be considered the self that reflects on personal, episodic memories and also the self that is engaged when anticipating about  ones future self.  The private self encompasses the understanding that aspects of the self can only be experienced by oneself, and that one's own personal thoughts and experiences are different to other people's.  Finally, the conceptual self (which can also be thought of as an individual's self-concept) consists of the theories and assumptions that an individual believes about themself.  This can encompass awareness of social dimensions of the self (e.g., I am a mother) as well as other aspects of oneself that an individual considers significant (e.g., I am intelligent, I am attractive).

What underlies all the theories of the self that have been discussed, including Neisser's (e.g., 1988) conceptualisation of self-knowledge, is that there are different aspects of the self, which can be experienced (pre-reflectively) and known (reflectively) at any one time.  Similarly it is clear that there are multiple types of self-awareness.  This implies that aspects of self-awareness can be selectively impaired (see Zahavi, 2010).  Importantly Neisser's conceptualisation of aspects of self-knowledge clearly defines five distinct aspects of self-awareness, and draws distinctions between certain aspects of the self that might in fact be distinct in ASD.  As such, Neisser's model of the self is used in the following section, to structure a review of the literature on self-awareness in ASD.

**Self-Awareness in Autism Spectrum Disorder**

Now that a theoretical framework for considering the nature of self-awareness has been outlined, it is possible to consider how this might be applied to the case of ASD.  The following section will review the literature surrounding what is known about self-

awareness in ASD to date. It is important to note that the research reported in this thesis does not systematically explore all aspects of the self, and self-awareness, in ASD. Thus the following review does focuses on some aspects of self-awareness in ASD more than other, in particularly focussing on the current literature surrounding private and ecological self-awareness, given the strong relevance this literature has to the empirical studies reported in this thesis.

**Ecological self-awareness.** First we can begin by looking at the evidence surrounding whether ecological self-awareness is impaired in ASD. Neisser defines the ecological self as "the self as perceived with respect to the physical environment: I am the person here in this place, engaged in this particular activity" (Neisser, 1988, p.386). Awareness of the ecological self allows an individual to perceive their location in space, but also allows one to become aware of one's ongoing interaction with the environment.

*Action monitoring.* One important aspect of ecological self-awareness is the ability to monitor one's own actions. Russell and Hill (2001, p.317) define action monitoring as, "the mechanisms that ensure that agents know, without self-observation, (a) for which changes in perceptual input they are responsible and (b) what they are currently engaged in doing". As such, action monitoring allows an individual to distinguish those changes in perceptual experience that are self-caused from those that are externally-caused. Thus, action monitoring gives rise to the experience of agency.

Action monitoring ability is often assessed through tasks that examine an individual's ability to monitor and correct their own errors. Typically, individuals are able to correct errors so rapidly that they cannot simply be relying on visual feedback alone. Instead correcting errors at this speed is thought to depend on monitoring so called "efference copies" of motor plans. This enables errors to be corrected before a motor command for the particular action is even completed. Error correction problems

are normally interpreted as reflecting diminished action monitoring ability, and are found frequently in studies of schizophrenia (see Frith, Blakemore, & Wolpert, 2000). Thus, it is striking that one early study observed slowed error correction among individuals with ASD (e.g., Russell & Jarrold, 1998). This may suggest diminished action monitoring/ecological self-awareness in ASD.

However, according to Russell (Russell & Jarrold, 1999), tasks which require individuals to discriminate online between their own actions and actions initiated by something/someone else provide a more direct measure of action monitoring ability. The Squares task, an "online" action monitoring task, has also been used to assess action monitoring ability in individuals with ASD (Russell & Hill, 2001; Williams & Happé, 2009a). During the Squares task participants are required to judge which of several different coloured moving squares on a computer screen is the square that is directly under their intentional control (through mouse movements; note; the participant's hand is covered meaning that success on this task requires monitoring of efference copies). In contrast the other squares on the screen are "distractor" squares, and the movements of these squares are controlled by the program, and not by the movements of the mouse. In order to judge correctly which square is under their own control, participants need to monitor their own efference copies of the movements they are generating, and compare them to their visual scene (the movements of the squares on the computer screen). Using this task Russell & Hill (2001) found that individuals with ASD were as good as neurotypical individuals at determining which of the squares was under their own control. This suggests that action monitoring may be unimpaired in ASD.

That being said, one problem with drawing decisive conclusions from this study is that all but five of the participants with ASD appeared to show either strong floor or

ceiling effects on the task (see Williams & Happé, 2009a). Following Russell and Hill (2001), Williams and Happé (2009a) employed a version of the squares task. This version was more incremental in terms of difficulty compared to that employed by Russell and Jarrold, and was designed to avoid the floor/ceiling effects that affected performance of participants in this study. Additionally, as well as employing a typical "self" version of the task (as described above) this study also included a second "other-person" condition. In this condition participants placed their hands on the computer mouse, but the movements of the mouse were controlled by the experimenter. Thus, in this condition, participants experienced no motor intentions for the movements of the mouse in the other condition, and so could not rely on feelings of agency to determine which of the stimuli is being controlled by the mouse. For an individual with an unimpaired sense of their own agency, this condition should be significantly more challenging than the self-condition. In contrast, if individuals are unable to accurately monitor their own actions then it should not matter who controls the mouse (because in both cases participants cannot rely on an experience of agency to perform the task, and instead can only rely on their ability to match felt actions with the observed consequences of these actions). Williams and Happé (2009a) did not observe any significant between-group difference in either the level or pattern of performance shown by individuals with and without ASD on the task. These results again suggest that ecological self-awareness is relatively spared in individuals with ASD.

Findings that suggest individuals with ASD appear to demonstrate typical online action monitoring ability, are in keeping with other studies in the broader action monitoring literature in ASD (e.g., Blakemore et al., 2006; David et al., 2006). These results are also in keeping with those from an early study by Frith and Hermelin (Frith & Hermelin, 1969) which also suggested children with ASD appeared to be *better* at

monitoring their own efference copy, relative to comparison participants. In this study participants were required to move a stylus along a track that had been cut into a piece of Perspex. Participants were then asked to complete the task again, this time without the aid of visual cues. Frith and Hermelin (1969) found that participants with ASD completed the task significantly fast than comparison participants, and concluded that these findings were consistent with *enhanced*, rather that impaired, action monitoring.

  ***The enactment effect.*** Another source of evidence concerning whether individuals with ASD show impairments in ecological self-awareness comes from studies that have assessed relative memory for self-performed actions versus memory for observed actions. It is well established that neurotypical individuals show reliably superior memory for actions that they themselves have performed than actions that they have observed other people perform (e.g., Baker-Ward, Hess, & Flannagan, 1990; Engelkamp, 1998). Superior memory for self-performed actions over other-performed actions is referred to as the "enactment effect" and is thought to result from additional motoric components involved in performing an action leading to those actions being more deeply encoded than observed actions (e.g., Engelkamp & Zimmer, 1989).

  Several studies have explored the enactment effect in individuals with ASD and reported finding typical enactment effects in individuals with this disorder (e.g., Hare, Mellor, & Azmi, 2007; Lind & Bowler, 2009c; Maras, Memon, Lambrechts, & Bowler, 2012; Summers & Craik, 1994; Williams & Happé, 2009a). This lends support to the view that action monitoring (and ecological self-awareness, more generally) is undiminished in ASD. However, whilst several studies report typical enactment effects in ASD, several studies have also reported reduced or absent enactment effects in ASD (Farrant, Blades, & Boucher, 1998; Hala, Rasmussen, & Henderson, 2005; Millward, Powell, Messer, & Jordan, 2000; Russell & Jarrold, 1999; Wojcik, Allen, Brown, &

Souchay, 2011; Zalla et al., 2010).  That being said, it is difficult to draw conclusions from the studies by Farrant et al., (1998), Hala et al., (2005), Millward et al., (2000) and Wojcik et al., (2011) due to methodological problems that are arguably inherent in the design of each study (for more details please see chapter six, which discusses methodological concerns with these studies in considerable detail).  Out of all these studies only Russell and Jarrold (1999), and Zalla et al.,  (2010) appear to report reduced or absent enactment effects in individuals with ASD, using sound experimental designs. Interestingly, Williams & Happé (2009a) could not replicate the results of Russell and Jarrold (1999).  As such it appears that (withholding the results of Zalla et al., 2010), the results from studies of the enactment effect are in keeping with those from studies of action monitoring, and suggest that ecological self-awareness is intact in individuals with ASD.  Overall, the majority of studies of the enactment effect and action monitoring in ASD thus suggest that ecological self-awareness is intact in this disorder.

      **Interpersonal self-awareness.**  The interpersonal self is conceptualized as the self as engaged in social interactions with other individuals.  These interactions are similar to what Trevarthen terms "primary intersubjectivity".  ASD can be considered a prototypical disorder of interpersonal self-awareness.  Neisser himself even suggested that failures in the development of the interpersonal self are associated with ASD:

> The successful achievement of intersubjectivity depends not only on the
> operation of the perceptual and motor systems but on some additional,
> specifically human mechanism that permits us to relate to members of our own
> species.  The mechanism can fail, and it has often been suggested that the
> dramatic condition called infantile autism, characterised from the outset by a
> total lack of interest in relationships with people, results from just such a failure.
> (1988, p.394)

It is clear that interpersonal self-awareness is impaired in ASD, and such impairments are a fundamental aspect of this disorder. However, recently, support for this suggestion has come from research exploring mirror neuron activity in ASD.

*Mirror neurons.* The discovery of mirror neurons may provide a neural basis for the interpersonal self. Mirror neurons are neurons in the brain that are activated by performing *and* observing an action being performed (e.g., Dipellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). A fundamental characteristic of such neurons is that they have been shown to fire both when an individual (or primate) performs an action, and when they observe someone else performing a similar action. Whilst mirror neurons are primarily thought to be involved in understanding and interpreting actions, mirror neurons have also been implicated in several other social-communication processes, including imitation (e.g., Iacoboni et al., 1999), mindreading (Gallese & Goldman, 1998), and empathy (Iacoboni, 2005). From this perspective, several features of ASD, including impairments in social communication and the capacity to understand others, match functions thought to be mediated by the mirror neurons. Thus, evidence for damaged mirror neuron systems in ASD would provide evidence of the neuro basis of impairments in the interpersonal self in ASD.

It has been proposed that the social-cognitive difficulties seen in ASD may be the results of atypical activation in the mirror neuron system, an idea that is often termed the "broken mirror" hypothesis (Ramachandran & Oberman, 2006). Evidence to support the idea that the mirror neuron system is "broken" in ASD comes from electroencephalography (EEG), transcranial magnetic stimulation (TMS), and functional magnetic resonance imaging (fMRI) studies that have shown atypical activation in the mirror system in ASD (Dapretto et al., 2006; Nishitani, Avikainen, & Hari, 2004;

Oberman et al., 2005; Théoret et al., 2005). For example, employing an fMRI study

Dapretto and colleagues (Dapretto et al., 2006) found that whilst imitating and

observing emotional expressions, children with ASD demonstrated no mirror neuron

activity in the inferior frontal gyrus (a key part of the mirror system). In contrast,

neurotypical children demonstrated typical mirror activation in this area whilst

imitating facial expressions (Dapretto et al., 2006). Interestingly, Dapretto et al., (2006)

also found that activity in the inferior frontal gyrus was inversely related to children's

functioning in the social domain. These results suggest that children with ASD appear

to process actions performed by others differently from neurotypical children, and

support suggestions that differences in mirror neuron activity in ASD may explain

several of the social-communication impairments seen in this disorder.

However, not all studies exploring mirror neuron activity in ASD have found

evidence of atypical activation in ASD (see e.g., Avikainen, Kulomäki, & Hari, 1999;

Dinstein et al., 2010). Additionally, if ASD is characterised by atypical activation in the

mirror neuron areas, studies should find evidence of behavioural impairments in

imitation ability in ASD, alongside atypical mirror neuron activity. However, there is

considerable variability in imitation ability in individuals with ASD (with studies not

always finding impairments in imitation ability. It has even been noted (see Southgate

& Hamilton, 2008) that several of the studies that report atypical mirror neuron activity

in individuals with ASD fail to find behavioural differences in imitation ability in their

ASD groups, relative to neurotypical participants (e.g., Dapretto et al., 2006; Nishitani et

al., 2004; Williams et al., 2006). This itself suggests that atypical activations in mirror

neuron brain areas do not in fact appear to relate to an individual's imitation ability.

Such findings have led some researchers to doubt the broken mirror hypothesis (e.g.,

Southgate & Hamilton, 2008), or suggest that perhaps only some aspects of mirror

neuron activity are impaired in ASD (see Boria et al., 2009; Rizzolatti & Fabbri-Destro, 2010). Whilst it appears unlikely that mirror neuron theories alone can explain cognitive impairments in ASD, studies exploring mirror neuron activity do provide at least some neurological support for the idea that interpersonal self-awareness is impaired in ASD.

**Conceptual self-awareness.** A third aspect of self-awareness defined by Neisser is that of conceptual self-awareness. The conceptual self can be considered the self as defined in terms of the theories and assumption an individual holds about themselves. More simply, the conceptual self can also be thought of as one's own self-concept, or what one brings to mind when thinking about themselves. There is some indication that individuals with ASD conceptualise themselves differently to individuals without ASD. Studies have shown that individuals with ASD tend to define themselves less in terms of social self-concepts (Lee & Hobson, 1998; Tanweer, Rathbone, & Souchay, 2010) and mental terms (Kristen, Rossmann, & Sodian, 2014) relative to neurotypical individuals, defining themselves more in terms of abstract concepts (Tanweer et al., 2010) and physical terms (Lee & Hobson, 1998). This suggests that at least some aspects of conceptual self-awareness are atypical in ASD.

As well as examining conceptual self-awareness in ASD by explicitly asking individuals to define themselves in terms of self-concepts, evidence of conceptual self-awareness in ASD can also be taken from studies of pronoun use, self-reference effects, and self-recognition. The presence of such behaviours is typically taken as behavioural indications of conceptual self-awareness. The following section reviews studies that explore pronoun use, self-referencing and self-recognition in ASD.

***Pronoun use.*** Early manifestations of conceptual self-awareness can be seen in the language neurotypical children begin to use at around two years of age, when

children start to use and understand personal pronouns.  The use of personal pronouns (such as "I", "Me" and "You" etc.) is a relatively unambiguous expressions of self-awareness.  By using words such as "I" or "You" an individual demonstrates an explicit understanding/awareness of the distinction between self and other.

In Kanner's early descriptions of childhood autism, he noted that children with ASD demonstrated abnormal personal pronoun use (Kanner, 1943).  Kanner observed that children with ASD tended to repeat pronouns "as heard, with no change to suit the altered situation" (Kanner, 1943, p.244).  Bosch (1970) also provided early clinical illustrations of unusual pronoun use in individuals with ASD, including example of pronoun reversal mistake, but also examples of children with ASD referring to themselves as "he/she" or using their proper name.  These observations suggest that individuals with ASD use pronouns incorrectly in utterances that cannot be sufficiently explained by echolalia (for example using third person pronouns like "he" to refer to oneself).  Instead, Bosch suggested that mistakes in pronoun use arise from diminished self/other distinctions in individuals with ASD.  Both Kanner's and Bosch's clinical observations have been supported by a number of empirical studies that suggest personal pronoun use is atypical in ASD (Hobson, Lee, & Hobson, 2010; Jordan, 1989; Lee, Hobson, & Chiat, 1994; Lind & Bowler, 2009a; Loveland & Landry, 1986).  For example, Jordon (1989) reported that, compared to matched comparison children, children with ASD demonstrated atypical pronoun use when an experimenter asked them questions about what had happened during a game.  In general children with ASD tended to refer to themselves or the experimenter using proper names, or pronouns in the incorrect case (e.g., saying "I" instead of "me").  This pattern of results implied children with ASD were not simply echoing pronouns, but made atypical pronoun utterances because of diminished self- and other referencing.  Alongside other studies

(Hobson et al., 2010; Lee et al., 1994; Lind & Bowler, 2009a; Loveland & Landry, 1986) such findings support the suggestion that conceptual self-awareness is diminished in ASD.

      ***The self-reference effect.***  Evidence concerning conceptual self-awareness in ASD also comes from studies exploring self-referencing.  It has been suggested that the self acts as a structure for memory, and an individual's autobiographical knowledge base is organised in relation to the self (see Conway, 2005).  Conway's self-memory system (SMS: Conway & Pleydell-Pearce, 2000) emphasises the interconnectedness between memory and the self, and this framework proposes that the integration of the self within memory is essential for typical memory functioning.  Typically, individuals show superior memory for information that is considered self-relevant, or that has been encoded in relation to the self (see Symons & Johnson, 1997).  For example, when presented with descriptive words, individuals typically remember words they have considered in relation to themselves (e.g., "Are you quiet?"), than words they have considered in relation to others (e.g., "Is your mother quiet?").  This "self-reference" effect is thought to occur because encoding information in relation to the self facilitates deeper encoding of this information within memory, in turn making this information more likely to be retrieved.  This effect can be thought of as a depth-of-processing effect (Craik & Tulving, 1975).  Craik and Tulving proposed that the deeper or more elaborately information is processed during encoding, the more likely such information is to be retrieved from memory.  For example, you are more likely to retrieve information you have processed semantically than information you have processes phonologically.  By extension, information you have processed in relation to the self is thought to be encoded deeply, and thus is more likely to be retrieved relative to information not processed in relation to the self.  As such, if the concept of the self is

diminished among individuals with ASD, then individuals with ASD should show diminished self-reference effects, relative to neurotypical individuals.

Several studies have examined the self-reference effect in individuals with ASD (Henderson et al., 2009; Lombardo, Barnes, Wheelwright, & Baron-Cohen, 2007; Toichi et al., 2002). For example, Lombardo and colleagues investigated whether adults with ASD demonstrated typical self-reference effects, compared to a group of matched neurotypical adults (Lombardo et al., 2007). In this study individuals were presented with psychological trait adjective and asked to rate on a scale of 1-6 the extent to which the word described a) themselves, b) a close friend, or c) Harry Potter. Another condition also asked participants to assess the number of syllables in the word. After a delay, participants were then given a surprise recognition test, in which they were presented with the words from the previous part of the tasks alongside new lure words, and asked to judge on a scale of 1-6 how confident they were that a word was "old" (had been presented in the previous part of the task). Judgments on a scale of 1-3 were considered "new" judgements and judgements on a scale of 4-6 were considered "old" judgements. In this study $d'$ was then used to analyse recognition performance on the task. Both groups showed depth of processing and self-reference effects (Syllable < Harry Potter < Close Friend < Self). However the magnitude of the self-reference effect demonstrated in each group was different. When difference scores were calculated between recognition for words processed self-referentially than words processed in relation to Harry Potter, analysis indicated that the self-reference effect shown in the ASD group (*Cohen's d* = 0.92) was diminished relative to the effect shown in the neurotypical group (*Cohen's d* = 1.32). The results of this study are in keeping with the results of an earlier study conducted by Toichi and colleagues (Toichi et al., 2002).

Using a similar design Henderson et al., (2009) extended these findings, and demonstrated that children with ASD also fail to show typical self-reference effects. In this study children were asked to process psychological trait adjectives in three conditions. Participants were asked to judge a) whether the word described themselves (Self condition), b) whether the word describe Harry Potter (Other condition) or c) whether the word was longer than seven letters long (Featural condition). Recognition memory for the words in all three conditions was then tested and assessed using $d'$. Children in both the neurotypical and ASD groups showed depth of processing effects (Featural< Other/Self condition). However, only the neurotypical group showed a self-reference effect (Other<Self). Individuals in the ASD group did not differ in their memory for words in the self-condition relative to the other condition. As such, both Lombardo et al., (2007) and Henderson et al., (2009) indicate that both children and adults with ASD do not show typical self-referential processing. However, whilst Lombardo et al., (2007) found that adults with ASD did demonstrate a *diminished* self-reference effect, and were thus capable of self-referential encoding to some degree, Henderson et al., (2009) implies that children with ASD show no evidence at all of processing information in relation to the self.

Furthermore, results from Lombardo et al., (2007) indicated that impairments in recognition memory on the task were not restricted purely to the self-condition. Individuals with ASD also recognised significantly fewer words from the Close friend condition relative to the neurotypical group. This suggests that individuals with ASD failed to encode and structure information within memory in relation to the self and in relation to others. One interpretation of these finding is that individuals with ASD hold diminished concepts of both self *and* other. Arguably, social-communicative impairments in ASD may hinder an individual's ability to acquire *both* a typical concept

of the self, and also typical knowledge about others.  In contrast, individuals with ASD did not show a diminished ability to encode information in relation to Harry Potter, reflecting the fact that they showed typical, psychological concept of the character.  However, acquiring knowledge of Harry Potter does not depend on any social-communicative experiences and instead depends on reading (or watching a film).  As such, this may explain why individuals with ASD did not show diminished recognition performance in the Harry Potter condition, but did show diminished recognition of words in the Self condition and Close friend condition.

*Mirror self-recognition.*  Mirror self-recognition tasks (e.g., Amsterdam, 1972; Gallup, 1970) are commonly used to measure higher-order self-awareness.  In a "classic" mirror recognition task (such as the "rouge test"), during the beginning of the task, an experimenter discretely places a coloured mark on a participant's nose/face.  After this has been done, participants are then shown their own reflection in the mirror.  If an individual proceeds to touch the mark on their face, this is taken as evidence that they possess an objective awareness of their own body (a physical self-concept).  Although it is debated exactly what mirror self-recognition tasks measure (e.g., Hobson, 1990), it is almost universally agreed that touching ones nose during the task can be taken as evidence that an individual has at least a basic physical self-concept  (i.e., a mental representation of what they physically look like).  Mirror self-recognition has been used to asses conceptual self-awareness in both infants and animals (e.g., Gallup, 1970), and neurotypical children tend to pass mirror recognition tasks around the age of 18 months (e.g., Courage, Edison, & Howe, 2004).

To date, four studies have assessed mirror self-recognition in children with ASD (Dawson & McKissick, 1984; Ferrari & Matthews, 1983; Neuman & Hill, 1978; Spiker & Ricks, 1984).  These studies have fairly consistently found that a large proportion of the

children with ASD tested successfully recognise their own image in the mirror (Williams (2010) reports that across these four studies, an average of 74% of the children with ASD successfully recognised their own reflection). This suggests that mirror self-recognition is largely unimpaired in ASD, and suggests that at least physical self-concepts appear intact in children with ASD. However, one major problem with these studies is that none include a comparison group of neurotypical participants that are closely matched for age *and* mental age to the ASD group. This makes it difficult to claim decisively that mirror self-recognition is entirely typical in ASD. That being said, it is interesting that in both Ferrari and Matthews' study (1983) and Dawson and McKissick' s study (1984) the children with ASD who failed the task were also children who demonstrated signs of developmental delay (see Williams 2010). In Ferrari's study (1983) the children who failed the task had an average mental age within the developmental time period neurotypical children also fail mirror self-recognition tasks. Additionally, the two children who failed to recognise their reflection in Dawson and McKissick' s study (1984) also failed to show developmental signs of stage V/VI object permanence. As such, this does suggest that the children with ASD who failed the mirror task in these studies did so not because of deficits specific to ASD, but because of their delays in their developmental level. This highlights the importance of assessing performance in children with ASD relative to children matched for mental ages, as well as chronological age.

As things stand, mirror self-recognition appears to be a relative strength in ASD (even among low functioning individuals). This stands in contrast to the evidence discussed concerning both pronoun use and self-referencing, which suggests that conceptual self-awareness is impaired in ASD. However, one important distinction between the conceptual self-awareness demonstrated by mirror recognition, compared

to the conceptual self-awareness demonstrated by typical pronoun use and self-referencing is that mirror recognition relies on having intact higher-order *physical* self-awareness, but does not necessarily rely on higher-order *psychological* self-awareness (but see Lewis, 2003 for a counter argument).  It is possible that individuals with ASD have intact physical self-concepts, but impaired psychological self-concepts. Interestingly, Lind  (2010) highlights the fact that all existing studies of self-referencing in individuals with ASD have explored self-referencing in relation to *psychological* trait adjectives alone and thus only examine whether individuals encode information in relation to psychological aspects of an individual's self-concept.  Indeed, Lind (2010) directly predicts that individuals with ASD would show *typical self-reference effects* when encoding information in relation to the physical self.  This distinction between physical and psychological self-awareness (not just in conceptual self-awareness but also concerning extended self-awareness) is discussed in more detail in sections to come.

**Extended self-awareness.**  Extended self-awareness involves an awareness of the self that encompasses one's present self, past self and future self.  As such extended self-awareness involves the understanding that several alternative representations of the self can reflect different representations of the same enduring self (across time).  One aspects of cognition that extended self-representations rely on is one's autobiographical memories (see Povinelli, 2001), which several studies have explored in ASD.

***Episodic memory and episodic future thinking.***  Autobiographical memory refers to an individual's memory for information concerning themselves.  Two distinct types of autobiographical memory exist; autobiographical *semantic* memory, which

refers to an individual's memory for knowledge about themselves (e.g., their age or nationality) and autobiographical *episodic* memory, which refers to an individual's memory for events they have experienced (e.g., what they did on their last birthday). Whilst both aspect of autobiographical memory are arguably related to aspects of the self, autobiographical episodic memory alone specifically involves extended self-awareness.

The majority of evidence suggests that individuals with ASD demonstrate impaired autobiographical episodic memory (Bruck, London, Landa, & Goodman, 2007; Crane & Goddard, 2008; Goddard, Howlin, Dritschel, & Patel, 2007; Klein, Chan, & Loftus, 1999; Lind & Bowler, 2010; Lind, Williams, Bowler, & Peel, 2014; Lind, Williams, Raber, Peel, & Bowler, 2013; Losh & Capps, 2003), whilst episodic semantic memory appears relatively spared, at least in adults with ASD (see e.g., Klein et al., 1999; Crane & Goddard, 2008).  Additionally, individuals with ASD appear to demonstrate impaired episodic future thinking (see Crane, Lind, & Bowler, 2013; Lind & Bowler, 2010; Lind, Williams, et al., 2014; Lind et al., 2013; Terrett et al., 2013) Episodic future thinking refers to the ability to imagine event that might plausibly happen to oneself at a future time and several links between episodic future thinking and autobiographical episodic memory have been made within the literature.  Both episodic future thinking and episodic memory emerge at the same time during typical development (Suddendorf, 2010) and it has been suggested that both rely on the same underlying cognitive mechanism (see e.g., Lind, Williams, et al., 2014). Impairments in both processes in ASD suggest that individuals with this disorder demonstrate a diminished sense of extended psychological self and an impaired understanding of themselves throughout time.

**Delayed self-recognition.**  Evidence of extended self-awareness in ASD can also be explored through studies of delayed self-recognition.  As discussed above, although

the evidence is not unequivocal, mirror self-recognition appears to be intact in children

with ASD. However, two studies have also used an extension of the mirror task to

assess *extended* self-awareness in children with ASD. During a delayed, video self-

recognition task, a video is taken while the experimenter surreptitiously places a sticker

on the back of the participants head (whilst the participant completes a distraction

task). After a short period the participant (who still has the sticker on his/her head) is

then shown this video. Whilst viewing the video, if a participant proceeds to reach up

and remove/ the sticker on the back of their head, this is typically taken as evidence that

they possess an extended self-representation (i.e., they understand that the individual

in the video is the same individual as currently watching the video (themselves), and

will thus recognise that presently they will have a sticker on their own head.

Performance on this task measures one's extended *physical* self-awareness but not

necessarily one's temporarily extended *mental* self-awareness. In keeping with this

idea studies have shown that performance on delayed video tasks does not relate to

performance on mindreading tasks (Suddendorf, 1999). Two studies have assessed

delayed video self-recognition in children with ASD (Lind & Bowler, 2009a; Nielsen,

Suddendorf, & Slaughter, 2006). Lind and Bowler (2009a) report that, amongst

participants who all demonstrated intact mirror self-recognition, the majority of

participants with ASD, as well as the majority of age- and ability-matched neurotypical

participants, passed a delayed video self-recognition task as well. This is in keeping

with the results reported in Nielsen et al., (2006), who used a similar task to explore

delayed video self-recognition in high-functioning children with ASD. As such, these

studies suggest that extended awareness of one's physical self appears to be intact in

ASD. Interestingly, Lind & Bowler (2009a) also asked participants to identify who it

was that they could see in the video and found that participants with ASD were

significantly more likely to refer to themselves using their proper name, rather than referring to themselves as "me". This is in keeping with previous findings that pronoun use is impaired in ASD (see p.21 above), and again highlights a dissociation between aspects of self-awareness in ASD. Whilst the children with ASD in this study appeared to show intact extended self-awareness (at least concerning their physical self) they still demonstrated impairments when using pronouns, evidence of impaired conceptual self-awareness.

Concerning extended self-awareness overall, it appears that individuals with ASD may demonstrate selective impairments in extended self–awareness. Whilst results from tasks of episodic memory and episodic future thinking suggests that extended psychological self-awareness is impaired in ASD, results from studies of delayed self-recognition suggest that an extended awareness of the physical self appears intact in ASD.

**Private self-awareness.** As outlined above, private self-awareness is an awareness of the self that cannot be shared with anyone else. This encompasses awareness of one's own thought processes, awareness of one's own epistemic mental states (such as beliefs, desires, and intentions), and awareness of one's own emotions. Until quite recently, there has been an absence of clear research investigating whether individuals with ASD demonstrate impaired self-awareness of their own private self. However recently there has been a growing body of studies exploring private self-awareness in ASD. Classic approaches to the study of private self-awareness in ASD have often used "self" versions of classic mindreading tasks to assess "theory of own mind" (see below). However, more recently private self-awareness in ASD has also been assessed by studies exploring metacognition.

Metacognition can be broadly defined as "thinking about thinking". More specifically, metacognition refers to an individual's beliefs and knowledge about cognitions (often referred to as *metacognitive knowledge*) and an individual's ability to assess and control their own cognitive processes (often referred to as *metacognitive skill*). Within the context of metacognitive skills metacognitive *monitoring* refers to an individual's awareness of their own current, online mental states and cognitive activity, whereas metacognitive *control* to refers to an individual's ability to regulate and control their own cognitive processes. Private self-awareness can be thought of as synonymous with metacognitive skill, therefore.

***Self-awareness and mindreading: One-system or two?*** Before reviewing the literature surrounding private self-awareness in ASD it is worth considering theories surrounding the relationship between mindreading (which is commonly thought to be impaired in ASD), and private self-awareness. One question, currently debated within the literature is whether aspects of self-awareness rely on the same neuro- cognitive mechanism as the ability to represent others' mental states (henceforth termed mindreading). According to one perspective (e.g., Carruthers, 2009; Frith & Happé, 1999) the ability to represent one's own mental states (metacognition) relies on the same underlying metarepresentational mechanism as the ability to understand mental states in others (mindreading). One version of this ("theory-theory") approach suggests that metacognition results from turning our mindreading capacities on ourselves (Carruthers, 2009). Thus, according to Carruthers, mindreading is both ontogenetically and phylogenetically prior to metacognition. Crucially, according to this argument, no dissociation should exist between mindreading and metacognition, because a single faculty governs both processes.

However, the one mechanism theory of mentalising is far from undisputed, and this position has been contested by several alterative theories. According to one version of the "simulation theory", the ability to read others' minds stems from our ability to directly introspect the contents of our own mind, and then use this information to mentally simulate the contents of another's mind (e.g., Goldman, 2006). From this perspective, metacognition is both ontogenetically and phylogenetically prior to mindreading. According to a third theory, proposed by Nichols and Stitch (2003), mindreading and metacognition are underpinned by separate mechanisms; one "monitoring mechanism" is responsible for access to/awareness of one's own mental states, whereas a separate "mindreading mechanism" is responsible for processing information about others' mental states. As with Goldman's simulation theory, this "two mechanisms theory" shares the intuition that we possess direct, non-inferential access to our own mental states, whereas we have to infer mental states in others on the basis of behaviour. However, it departs from simulation theory by suggesting that information gained from introspection of our own mental states is *not* foundational for mindreading. The crucial implication stemming from both the simulation theory and the two mechanisms theory is that there should be some people who manifest diminished mindreading abilities, despite undiminished metacognition. Indeed, both Goldman (2006), and Nichols and Stich (2003) explicitly suggest that autism spectrum disorder (ASD) provides just such a case.

**Private self-awareness in ASD**. The study of private self-awareness in ASD can thus inform the current debate concerning the underlying mechanisms involved in mindreading/metacognition. The following section reviews studies of private self-awareness in individuals with ASD, with the former section reviewing "classic" approaches to the study of theory of own mind, and the later part focussing specifically

on studies that have used classic tests of metacognition to explore private self-awareness in individuals with ASD.

*Awareness of one's own thought processes.* Hurlburt and colleagues (Hurlburt, Happé, & Frith, 1994) used an experience sampling method to examine whether three high-functioning adults with ASD could self-report their own thought processes on a daily basis. During this study, participants were asked to carry around a small device which beeped at random intervals, and were asked to write down an account of their thoughts whenever the device beeped. Interestingly, Hurlburt found that at least one participant found it extremely difficult to report what their own thought processes had been when they heard the device beeping. Although we should bear in mind that this study only consisted of an extremely small sample of participants, this does imply that perhaps some individuals with ASD demonstrate diminished private self-awareness.

*Awareness of one's own intentions/desires.* Studies exploring the awareness of one's own intentions and desires in individuals with ASD have produced inconsistent findings. In one study Philips and colleagues (Phillips, Baron-Cohen, & Rutter, 1998) found that children with ASD confused their own *intentions* with their own *desires.* Whilst similar, intentions and desires are not the same mental states; for example it is distinctly possible to desire an outcome without having any intention of carrying it out, or to carry out an intended action that you have no desire to do. It is this distinction that children with ASD struggled with in Phillips et al., (1998). During this study children with ASD played a rigged target-shooting game, in which they attempted to shoot particular targets, only half of which contained a desirable prize. Phillips found that when children with ASD mistakenly hit a target that they had not intended to hit, but which contained a prize which they had desired, they incorrectly reported that they had intended to hit that target in the first place. Whilst neurotypical children matched

for age and intelligence were able to distinguish between the targets they intended to hit, and those they desired to hit, children with ASD struggled with this task.  This provides support for suggestions that individuals with ASD are less aware of their own private selves.

However, using a very similar procedure to that used in Phillips' study, Russell and Hill (2001 ; Experiment 2) failed to replicate these results, suggesting that children with ASD do not show impairments reporting their own (failed) intentions. Additionally, in another experiment Russell and Hill found only mixed evidence for diminished awareness of intentions in children with ASD (Russell & Hill, 2001, Experiment 3).  This experiment employed the "transparent intentions task" (Russell, Hill, & Franco, 2001) in which children were asked to complete a drawing on a piece of transparent paper (e.g., draw a handle on a cup).  However, during the task a second transparent piece of paper was placed on top of the first, which displayed another unfinished drawing (e.g., a head missing an ear) that aligned perfectly with the drawing on the first transparency.  When asked to complete the drawing on the bottom transparency children actually unintentionally completed the other drawing on the "hidden" transparency, placed on top of the picture the children thought they were completing.  As such, children unknowingly completed the top drawing rather than the one they had intended to complete.  Children were then asked what they had meant to draw (the "Mean" question), and what they had thought they were drawing throughout the task (the "Think" question).  When the Mean question was asked first, before the think question, children with ASD showed diminished performance on the task. However, performance was not diminished in the ASD group relative to the neurotypical group when the Mean question was asked second, nor on the Think question (regardless of question order).  In contrast to the results of Phillips et al.,

(1998), these findings do not support the suggestion that awareness of one's own intentions is impaired in ASD.  Nonetheless, some caution should be taken when interpreting these results.  Williams and Happé (Williams & Happé, 2010), having raised several methodological concerns with Russell & Hill's study (see William's & Happé, 2010b for details), attempted to replicate these results.  This study found that children with ASD *were* in fact significantly poorer at identifying their mistaken intentions, relative to comparison participants (Williams & Happé, 2010b, Experiment 2).  Experiment 1 of this study also explored awareness of one's own intentions in children with ASD, testing whether children were able to identify correctly automatic knee-jerk reflexes as being unintentional.  Compared to matched typically-devolving children, children with ASD were significantly less likely to recognise that their knee-jerks were not intentional actions.  Both findings can be taken as an indication that individuals with ASD find it difficult to accurately represent their own intentions.

As such, although not unequivocal, the evidence within the literature suggests that individuals with ASD show a diminished awareness of their own intentions and beliefs.  This supports the idea that individuals with ASD demonstrate impairments understanding their own mental states (Williams, 2010), and demonstrate diminished private self-awareness.

*__Awareness of one's own beliefs.__*  One of the tasks most widely used to assess understanding of mental states, in both self and others, is the unexpected contents false belief task (Hogrefe, Wimmer, & Perner, 1986).  Although several versions of the task exist, the original version assesses children's belief about the content of a smarties tube.  During the "smarties task" children are shown a smarties tube and asked what they think is inside.  Children will commonly respond that they think there are smarties (sweets/chocolate) inside the tube.  Children are then shown that, in fact, the tube does

not contain smarties but contains a pencil instead. Children are then asked two questions; 1) what was it they previously thought was inside the smarties tube (Self condition), and 2) what would another person think was inside the smarties tube, if they had not been shown the contents (Other condition). Typically, children of four or five years of ages will consistently pass this test, and can accurately report what false belief they had previously held, as well as what false belief another person would have. Notably, neurotypical individuals also tend to show similar performance on the task in both the Self and Other condition of the task, finding each condition of equivalent difficulty (Wellman, Cross, & Watson, 2001). Several studies have used the smarties task to assess awareness of one's own and others' beliefs in children with ASD (e.g., Baron-Cohen, 1992; Fisher et al., 2005; Leslie & Thaiss, 1992; Perner, Frith, Leslie, & Leekam, 1989; Williams & Happé, 2009b). In two early studies (Leslie & Thaiss, 1992; Perner et al., 1989) children with ASD showed stereotypical impairments in understanding other people's false belief, but were able to report accurately what their own previously held false beliefs were during the task. At face value this would imply that individuals with ASD demonstrate a typical understanding of their *own* mental states, despite impairments in understanding others' beliefs. However, methodological issues associated with the Self condition in these studies leave these conclusions somewhat questionable (Williams & Happé, 2009b). Namely, it is possible that the children with ASD in these studies passed the smarties task simply by remembering what they had previously *said* was in the tube, allowing them to successfully pass the task despite a despite a diminished understanding of their previous belief.

However, some researchers have suggested that even when children are asked to report what they *thought* was in the tube, the smarties task may still over-estimate children ability to understand their own mental states. In each of the studies previously

mentioned children with ASD were always asked to explicitly state their false belief before they were asked the Self question. It is possible that the processes of asking children with ASD to verbally report their false belief during the task might itself lead to inflated levels of performance on the task.

To test this idea, Williams & Happé (2009b) developed a version of a false beliefs task in which children were not asked to verbalise their beliefs about the content of a container. Instead, a plasters box and two other containers were places within reach of participants, but not the experimenter. During the task, the experimenter pretended that they had cut their finger, and asked the child if they would pass him a plaster. Participants reached for the plasters box and unexpectedly found that it contained candles instead of plasters. Children were then asked 1) what they had thought was inside the box (Self condition) and 2) what someone else would think was inside the box if they had not seen its contents (Other condition) . By selecting the plasters box, participants undeniably demonstrated their false belief that the box contained plasters, without ever explicitly verbalising this belief. As such, during the self-condition of the task, children were not able to pass the task simply by reporting what they had stated was in the box, but could only pass the self question if they recognised and recalled their false belief. Williams & Happé (2009b) found that children with ASD found it significantly *more* difficult to report what their own false belief had been on the task, compared to what another person's false belief would be. In contrast, participants without ASD and participants with developmental disabilities performed consistently on both the self and other question. These results suggest that even when children with ASD cannot rely on recalling their previous statement, they show a diminished ability to understand their own mental states.

***Awareness of one's own emotions.*** Recently research has also explored whether individuals with ASD are able to understand their own emotional feelings. Whilst it is well documented that individuals with ASD demonstrate impairments in empathy and understanding others' emotions, studies have also suggested that this may also be the case with understanding their own emotions. The term alexithymia refers to the inability to accurately identify and describe one's own emotions, and is a subclinical phenomenon that affects roughly 10% of the population (Linden, Wen, & Paulhus, 1995). However, studies of alexithymia in ASD have suggested that perhaps almost 50% of individuals with this disorder demonstrated difficulties identifying and describing emotions, impairments that are typical of alexithymia (Hill, Berthoz, & Frith, 2004). Silani and colleagues investigated this issue by asking individuals with ASD, and matched neurotypical participants to complete the Toronto Alexithymia Scale (TAS-20), a commonly used measure of alexithymia characteristics (Silani et al., 2008). Alongside completing this questionnaire participants were asked to rate a series of photographs on how emotionally arousing they considered them, during an fMRI experiment. In keeping with the results of Hill et al., (2004) participants with ASD reported significantly higher scores on the TAS-20, indicating higher self-reported levels of alexithymia in the ASD group relative to the control group. Silani and colleagues also found a strong relationship between TAS-20 scores in both groups and activity in the anterior insula, when participants were asked to assess their feelings towards unpleasant pictures. Studies have shown that the anterior insula appears to be involved during mentalising (e.g., Modinos, Ormel, & Aleman, 2009). Additionally, when asked to introspect on their feelings the ASD group showed atypical activity in several brain regions typically associated with mentalising, relative to the neurotypical group. These areas included the mPFC, ACC, precuneus, and cerebellum, areas that previous studies

have indicated are activated during self-reflection and mentalising (Frith & Frith, 2006).

Such results support the suggestion that emotion awareness is atypical in individuals

with ASD.

Another study that has explored awareness of emotions in ASD was carried out

by Ben Shalom and colleagues (Ben Shalom et al., 2006), who asked both ASD and

neurotypical children to report how emotional they found a series of pictures, alongside

taking physiological indicators of participant's emotional arousal when viewing these

pictures. What this study found was that participants' physiological responses to

emotional pictures did not appear to differ between participants in the ASD and

neurotypical group. However the study did find some differences in the emotion ratings

participants in the ASD group gave relative to the neurotypical group. Namely, children

were asked to rate pictures (pleasant, unpleasant and neutral pictures) on ratings of

"pleasantness" and rating of "interestingness". Whilst the neurotypical group

significantly differed in the pleasantness and interestingness ratings they gave for

unpleasant and neutral pictures, children with ASD didn't rate pictures in any of the

conditions differently depending on whether they were rating pleasantness or

interestingness. The authors tentatively pose that due to impairments expressing or

understanding conscious feelings of emotions the children with ASD may have

employed a compensatory strategy on the task (Ben Shalom et al., 2006). They suggest

it is possible that the children with ASD rated the pictures simply on a generic

"goodness" rating rather than distinguishing separately between feelings of

pleasantness and interest. Whilst there is no explicit evidence for this suggestion, this

idea does suggest that children with ASD do not show a typical understanding of their

own emotions, or at least do not rate emotional pictures in a similar way to neurotypical

participants. However, again caution should be taken when interpreting these results.

Group difference in distinguishing between ratings of pleasantness and interestingness (which are both relatively abstract concepts) could potentially be explained by large group difference in VIQ (with the ASD group having an average VIQ of 27 IQ points less than the average VIQ in the neurotypical group).

Overall, the literature exploring awareness of one's own emotions in individuals with ASD is sparse, relative to the literature exploring an understanding of other's emotions in ASD. The few studies exploring this do appear to suggest impairments in self-reported awareness of emotions in ASD, which supports the suggestion private self-awareness, is impaired in ASD. However, the results supporting this are far from conclusive, given the limited research in this area.

***Awareness of one's own knowledge states.*** Similarly, only a few studies have directly tested awareness of one's own knowledge in individuals with ASD. In a study by Perner and colleagues (Perner et al., 1989) children with ASD and neurotypical children were shown a series of boxes. During the experiment children were informed that each box contained a different object. On some trials the participant was allowed to look in the box to see the contents and on other trials a confederate "participant" looked in the other box (instead of the participant). After either the participant or confederate looked in the box children were asked whether they knew what was in the box (Self condition) or whether the confederate knew what was inside the box (Other condition). Perner found that whilst the neurotypical group was able to accurately judge both their own knowledge states and those of the confederate, the children with ASD showed significantly poorer performance in both the self and other conditions. The children with ASD tended to overestimate their own knowledge and that of the confederate, suggesting more often that the person who had not looked inside the box would know what the contents were.

Kazak and colleagues also assessed awareness for knowledge states in children with ASD.  Kazak and colleagues (Kazak, Collis, & Lewis, 1997) found no significant difference between children with and without ASD on their ability to understand the difference between knowing something and guessing it.  Again this suggests that private self-awareness appears to be unexpectedly intact in children with ASD.  However, it has been argued (Williams, 2010) that this could be mainly due to poor performance in Kazak's control group.  The neurotypical 4-year-old participants performed at below chance level on the task, demonstrating atypically poor ability.  As such, it is unclear whether the children with ASD in the study showed intact understanding of knowledge, or whether the control children just showed impairments.

**Problems with self-versions of mindreading tasks.**  One potential difficulty interpreting self-versions of classic mindreading tasks is that test questions in such studies require participants to recall their *prior* mental states, rather than report their *current* mental states.  Simulation and two mechanisms theories claim that only current mental states are directly accessible without the need for mindreading.  Thus, arguably, the results from the above studies do not necessarily show that metacognition/private self-awareness is impaired in ASD, because these tasks require inferences to be drawn about *past* mental states (but see Williams, 2010, for a counter-argument).  By contrast, it is widely agreed that metacognitive judgements are based on awareness of *current* mental states.  As such, studying classic tests of metacognition overcomes problems associated with the majority of studies of private self-awareness in ASD discussed.

**Metamemory.**  One important component of metacognition is metamemory, an individual's ability to monitor and control their own memory.  Nelson and Narens' (1990) influential model of metamemory divides metamemory processes into two levels; the "object-level" and the "meta-level" (please see Figure 3 for a graphical

representation). The object-level consists of first-order memory processes, whilst the

meta-level consists of dynamic second-order representations of the object-level.

Through monitoring, the meta-level can acquire information about the state of the

object-level and can change object-level representations accordingly. Whilst research

has extensively investigated impairments at the "object-level" (Nelson & Narens, 1990)

of cognitive processes in ASD, little research has examined the "meta-level" of such

processes.
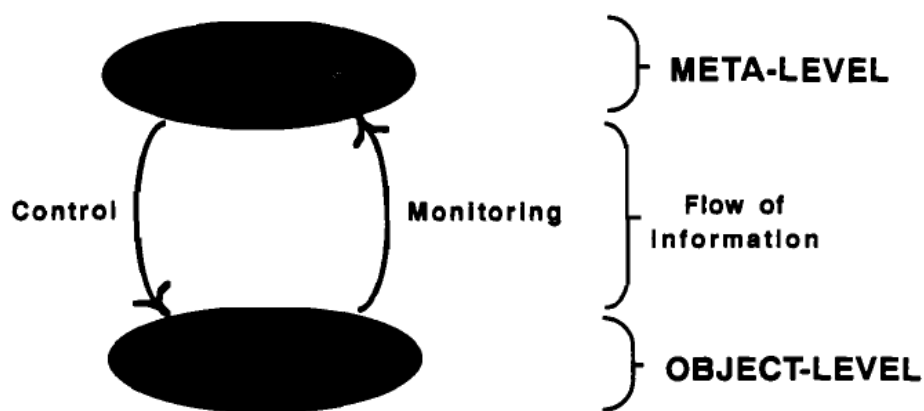


*Figure 3: A graphical representation of Nelson and Narens metamemory model.*

*Figure taken from Nelson & Narens, (1990).*

***Metamemory judgments.*** Research exploring metamemory abilities employs a

variety of different paradigms to test monitoring accuracy. In general these either ask

participants to make metamemory judgements concerning a future memory event, or

concerning a past memory event. Whilst *prospective* metamemory judgements refers to

judgements in which participants assess their confidence about a future memory event

(e.g., feeling of knowing judgments, tip of the tongue judgments, judgements of learning)

*retrospective* metamemory judgements refers to judgements in which participants

assess a previous memory event (e.g., confidence judgements). It has been suggested

that prospective and retrospective judgments rely on different sources of information (Fleming & Dolan, 2012).

Two paradigms widely used to assess metamemory monitoring accuracy involve making judgements-of-learning (JOL; Arbuckle & Cuddy, 1969) and/or feelings-of-knowing (FOK; Hart, 1965). During a standard JOL task participants are asked (during a learning phase) to memorise a series of stimuli pairs (e.g., pairs of words, such as "pen-key", "computer-elephant" etc.). After the learning phase, participants are presented with one stimulus from the pair (the cue; e.g., "pen") and asked to make a judgement on the likelihood that, at a later point, they will be able to recall the stimulus pair (the target; i.e., "key"). Participants are then presented with one stimulus from a word pair, the cue, and asked to recall the missing target word. The accuracy of participants' judgements is then measured by comparing participant's judgments about their future recall performance with their actual recall performance. Generally, neurotypical adults are able to make accurate JOLs, although accuracy is influenced by the length of the delay between the learning phase and when participants make a JOL (e.g., Nelson & Dunlosky, 1991). Nelson and Dunlosky (1991) demonstrated that individuals make much more accurate JOLs when there is a short delay between learning stimulus pairs and making a JOL about them (a delayed-JOL) than when JOLs are made immediately after leaning (an immediate-JOL). Nelson and Dunlosky suggested that this effect may be explained by the suggestion that when individuals make delayed JOLs their judgements are based more on information recollected from long-term memory than immediately accessible information from short-term memory (e.g., Nelson & Dunlosky, 1991). This information is thought to be a better indicator of future memory performance than information available immediately after stimuli are learnt. The delayed-JOL effect may also be explained by the fact that the context delayed JOLs are

made in are less similar to the context stimuli are learnt, and more similar to the context in which recall is tested (than immediate JOLs).

During a typical FOK task, participants are asked to make judgments about whether they will be able to recognise previously learned stimuli pairs, which they have currently failed to recall. Typically, participants are firstly asked to memorise a series of stimulus pairs (a learning phase). Participants are then presented (during a recall test-phase) with one stimulus from each pair, the cue, and asked to recall its pair (i.e., the target). Importantly, on trials in which participants *fail to correctly recall* the target they are then asked to judge the likelihood that, at a later point, they would be able to *recognise* the missing word. Finally, participants are then presented with one word from each pair (the cue), and are asked to select the stimulus's pair (the target) from several options (a recognition test-phase). Again, the accuracy of participants' judgments is measured by comparing participants' predictions about their future ability to recognise the correct target with their actual recognition performance.

Another paradigm that has been widely used to measure metamemory ability in neurotypical individuals is a judgment of confidence (JOC) task. Unlike both FOK and JOL paradigms, which involve making prospective metamemory judgements, JOC tasks ask participants to make retrospective judgments concerning their memory ability. Studies assessing judgments of confidence typically involve participants answering questions about recently-studied material or stored semantic knowledge, and then reporting their confidence in the answers they provided. If an individual's meta-monitoring ability is high, then their confidence judgements should discriminate accurately between correct and incorrect answers. This aspect of self-monitoring is perhaps particularly important, because confidence judgements are often used by individuals to control their behaviour (see Koriat & Goldsmith, 1996).

Different heuristic-based theories have been proposed to explain what type of information typically informs individuals' metamemory judgements. The cue-familiarity hypothesis (e.g., Metcalfe, Schwartz, & Joaquim, 1993; Reder, 1987) suggests that individuals rely on "familiarity" with the cue to judge the future memorability of a missing target. Familiarity in this sense refers to a general feeling of memory that varies in strength, but does not contain any information about the context in which that knowledge was acquired. This contrasts with "recollection", which involves the additional incorporation of contextual information in memory. As such, the cue-familiarity hypothesis suggests that when making a metamemory judgement (for example on the stimuli pair, "pen-key"), individuals rely on how familiar they find the cue ("pen"), and not on how much information they can recollect about the target ("key"). In contrast, the accessibility hypothesis (e.g., Koriat, 1993) suggests that individuals base their metamemory judgements on the extent to which they can retrieve partial or related information about the *target*, at the time they make a metamemory judgment. For example, a person may make a more positive JOL if they can recall some information about the missing target (e.g., its first letter, how many letters it consisted of, its semantic category etc.).

*Metamemory in ASD.* As discussed above, whilst the relationship between understanding one's own and other's mental state is currently debated, to the extent that both processes rely on the same mechanisms then there is reason to predict that metacognition (private self-awareness) should be severely impaired in ASD. However, despite an abundance of studies examining whether individuals with ASD show impaired mindreading little research has been carried out directly investigating metacognition in individuals with ASD.

In an early study Farrant and colleagues attempted to directly investigate metamemory in children with ASD across a series of five experiments (Farrant, Boucher, & Blades, 1999). The experiments carried out tested children's knowledge of how different variables (task difficult, individuals' age and strategy use during learning) influence memory. Farrant found that on all five metamemory tasks the children with ASD demonstrated undiminished performance (although they were less likely than controls to make spontaneous use of memory strategies involving other people). As such the study proposed that metacognition was unimpaired in ASD. However, Farrant and colleagues' study only tested children's metacognitive knowledge. It is quite possible that different processes are involved in acquiring metacognitive knowledge and the ability to accurately assess one's own mental states. Thus, this conclusion only applies to metamemory knowledge in ASD.

***Judgments of confidence (JOC).*** More recently Wilkinson and colleagues asked children with ASD to make judgments of confidence (JOCs) concerning how certain they were that they had correctly recognised faces (Wilkinson, Best, Minshew, & Strauss, 2010). JOCs involve making retrospective judgements, either after a recall or recognition task, on how accurate one's performance was. This study suggested that confidence judgments made by children with ASD were less accurate (i.e., less in keeping with actual recognition performance) than those made by neurotypical children, indicating that to some extent children with ASD were not able to accurately judge what information they knew. However, this difference was not replicated in adults. Although the study found subtle differences between memory awareness in adults with and without ASD, no significant between-group difference was found between adults overall memory awareness. Additionally, in both adult and children, memory awareness in this study was assessed during a facial recognition task. Given

that research has shown unequivocally that individuals with ASD show impairments in face processing (e.g., Hauck, Fein, Maltby, Waterhouse, & Feinstein, 1998; Williams, Goldstein, & Minshew, 2005) there could be concern that impairments at the basic level of this task could have confounded individuals' confidence judgements.

Wojcik and colleague (Wojcik et al., 2011) also used a JOC paradigm to assess whether children with ASD were able to make accurate confidence judgements surrounding whether they had accurately performed a series of actions correctly. In contrast to Wilkinson's findings (2010), in this study children with ASD appeared to be as accurate as neurotypical children in judging whether they had correctly recalled a series of actions (Wojcik et al., 2011). In keeping with this study, Elmose and Happé report typical JOC accuracy in children with ASD (Elmose & Happé, 2014). Taken together these studies demonstrate inconsistent findings in studies examining whether individuals with ASD show impairments in their ability to monitor their own memory.

Finally, Sawyer employed a JOC task that assessed both monitoring *and* control in adults with ASD, the only study thus far to explore the accuracy of metamemory control processes in individuals with ASD (Sawyer, Williamson, & Young, 2014). Whereas metacognitive monitoring refers to one's awareness of one's own mental states, metacognitive control refers to the ability to regulate one's own current, online mental states and cognitive activity (Flavell, 1979). In this study, participants were asked to complete an emotion recognition task involving facial stimuli. Participants were instructed that the aim of the study was to submit as many correct responses as possible. For each emotion recognition judgement, participants rated how confident they were that they had selected the correct response. Participants were then given the opportunity to submit each answer towards their total score (and gain a point for each correct answer), or discard the answer (and avoid losing a point for getting an answer

wrong). This provided a measure of metacognitive control. In a second experiment, the same procedure was used but participants' judgements concerned their answers to general knowledge questions, rather than emotion recognition.

Across both experiments, Sawyer et al., (2014) reported no significant between-group differences in JOC accuracy, implying undiminished meta-monitoring ability in ASD. However, the between-group difference in JOC accuracy *on the general knowledge task* was of borderline significance (associated with a one-tailed *p* value of .06), potentially implying a subtle monitoring impairment in ASD.

In terms of metacognitive control, Sawyer et al., (2014) found no between-group differences on their key index (*d'*), implying undiminished metacognitive control in ASD. However, Sawyer et al., performed additional *post-hoc* tests, which suggested that a significantly higher proportion of ASD participants (*n* = 12) than neurotypical participants decided not to withhold *any* answers. This could imply that these 12 ASD participants were not showing any metacognitive control at all. Alternatively, it could reflect a mere failure to understand the task demands among these participants. As such, the extent to which metacognitive control is diminished in ASD is still not entirely clear.

**Feeling of knowing (FOK) judgments.** Only one study so far has directly examined whether individuals with ASD are able to accurately monitor their own memory to predict their future memory performance whilst making FOK judgements (Wojcik, Moulin, & Souchay, 2013). Wojcik and colleagues assessed children's metamemory monitoring ability using two FOK tasks, one asking individuals to assess their memory for information stored episodically and one assessing memory for information stored semantically. Wojcik reported that children with ASD were

significantly poorer than neurotypical children at making accurate FOK judgements, but only when assessing their episodic memory.

***Judgments of learning (JOL).*** One final study has also recently published the results of two experiments exploring metamemory in individuals with ASD using JOL paradigms. Recently Wojcik and colleagues (Wojcik, Waterman, Lestie, Moulin, & Souchay, 2014) reported the results of two JOL tasks in children with ASD, and reported finding no metacognitive deficits in the accuracy of children's JOL assessments.

**Problems with study of metamemory in ASD thus far**

Table 1 provides a summary of the existing studies of metamemory monitoring accuracy in ASD thus far. From Table 1 it is clear that there are several inconsistencies within the literature concerning whether monitoring accuracy is impaired in ASD or not. Whilst some studies report large group difference in monitoring accuracy others report no group difference at all. Such inconsistencies do not appear to be explained by differences in the specific type of metamemory judgements being made, or whether the metamemory judgements being made are prospective or retrospective in nature. However, methodologically there is some concern surrounding the existing studies of metamemory in ASD thus far, which may explain why the results from studies of metamemory are not in keeping with other studies in the literature that fairly consistently suggest private self-awareness is impaired in ASD.

*Table 1: Summary of previous studies of metamemory monitoring in individuals with ASD.*

| Study | Metacognitive judgements assessed | Prospective or retrospective judgement | Key meta-monitoring performance | Potential confounds |
|---|---|---|---|---|
| Wilkinson et al., (2010) | Judgments of confidence (JOC) | Retrospective | ASD children< Neurotypical children<br>ASD adults= Neurotypical adults (with some group difference moderate in size) | Object level task (face processing) may have confounded meta-level task performance in the ASD group. |
| Wojcik et al., (2011) | Judgments of confidence (JOC) | Retrospective | ASD= Neurotypical controls | PIQ and VIQ for groups not reported (only FSIQ reported) |
| Wojcik et al., (2013) | Feeling of knowing judgements (FOK) | Prospective | ASD< Neurotypical controls (but only for episodic material, not semantic material) | Groups not matched for VIQ |
| Wojcik et al., (2014) | Judgements of learning (JOL) | Prospective | ASD= Neurotypical controls (in both experiments) | Groups not matched for VIQ |
| Sawyer et al., (2014) | Judgments of confidence (JOC) | Retrospective | ASD = Neurotypical Controls | Groups not matched for Age and VIQ. Additionally the object level task (emotion recognition) may have confounded task performance in the ASD group. |
| Elmose & Happé, (2014) | Judgements of learning (JOL) | Prospective | ASD = Neurotypical Controls | Groups are not matched for VIQ Judgments of confidence were made on a limited rating scale. |
| | Judgments of confidence (JOC) | Retrospective | ASD = Neurotypical Controls | |

One potential explanation for discrepancies between the results of metamemory studies in ASD could be that individuals with ASD only demonstrate impairments in metacognitive monitoring alongside impairments in mindreading performance. This suggestion is in keeping with one mechanism theories that predict both processes rely on the same underlying mechanisms. If one mechanism theories are correct, you would only predict metamemory impairments in individuals impaired in mindreading task performance. To date only one study has directly explored metamemory abilities in ASD alongside mindreading abilities (Farrant et al., 1999). As discussed the results of Farrant et al., suggested that children with ASD no not demonstrate metamemory impairments. However, an unexpectedly small number of the children in their ASD group showed deficits on the false belief task carried out in this study, indicating that mindreading was relatively unimpaired in their sample. Of course, it is possible that the children with ASD in this study passed the false belief task by hacking out a solution, despite impaired mindreading ability. Nevertheless if it is the case that the majority of the children in the ASD group passed the false belief task because of genuinely intact mindreading ability then it is unsurprising that they also demonstrated unimpaired performance on the metacognitive (if the prediction that mindreading and metacognition rely on the same underlying processes is to be believed). Several empirical chapters in this thesis explore this issue in more depths, assessing metacognition alongside mindreading ability in individuals with ASD.

Another particular methodological difficulty affecting several of the studies discussed above is that often the ASD and neurotypical groups were not matched for participant characteristics, specifically verbal IQ (e.g., Sawyer et al., 2014; Wojcik et al., 2013; Wojcik et al., 2014) and age (Sawyer et al., 2014). Matching for intellectual abilities is essential in such studies, because differences between groups in this respect

can potentially entirely explain between-group differences in experimental task performance (see Mervis & Klein-Tasman, 2004). Although some studies of metacognition have recognised this limitation and tried to overcome it using an ANCOVA to "control" for group differences in VIQ (see e.g., Wojcik et al., 2013), this does not, in fact, solve this problem (see Miller & Chapman, 2001). As such, as things stand, we cannot determine whether group differences in these studies were driven by diagnostic status or by differences in age/intellectual ability.

Finally, another potential confound of both Sawyer et al., (2014) and Wilkinson et al., (2010) is the fact that in both studies metamemory ability was assessed on tasks in which participants with ASD stereotypically demonstrate difficulties with at the object level. Sawyer et al., (2014) required participants to make judgments of confidence surrounding emotion recognition judgements, which are stereotypically impaired in ASD (see e.g., Harms, Martin, & Wallace, 2010; Hobson, Ouston, & Lee, 1988) and Wilkinson et al., (2010) asked participants to make confidence judgements during a face processing task, a known difficulty in ASD (see e.g., Hauck et al., 1998; Rouse, Donnelly, Hadwin, & Brown, 2004). Crucially, in studies of metacognition it is essential that groups are equated for object level ability on the task. If this is not the case group differences in metamemory monitoring accuracy may be simply the result of difference in overall memory processes, rather than specifically due to differences in monitoring ability alone (see Dunlosky & Metcalfe, 2008).

## Conclusions

To summarise, it appears that whilst aspects of self-awareness appear to be impaired in ASD, individuals with this disorder clearly do not demonstrate completely across the board impairments in self-awareness. For example ecological self-awareness

and some aspects of both conceptual and extended self-awareness appear relatively unimpaired in this disorder. One question that remains to be answered within the literature is why some aspects of self-awareness are spared in ASD, whilst others appear to be considerably impaired. The empirical studies carried out in this thesis explore several aspects of self-awareness in ASD, and attempt to better understand where impairments in self-awareness lie. This thesis contributes to the existing literature by using novel methods to explore self-awareness in ASD, in particular private self-awareness and its relationship to mindreading abilities. It will be argued that findings from the literature, and the results of the studies reported in the thesis, support the idea that individuals with ASD demonstrate impairments in understanding their own *mental* selves, but not impairments understanding their *physical* selves.

**CHAPTER TWO**

**GENERAL METHODOLOGY**

The studies conducted in this thesis explore self-awareness in both children and adults with ASD, and neurotypical children and adults. This chapter outlines aspects of methodology and procedure that were used during all experiments. It provides an overview of the inclusion and exclusion criteria that were used to select participants for all studies, and discusses how ASD and neurotypical groups were matched in the experiments. Additionally, several tasks were used in multiple experiments throughout the thesis. An outline of the procedure used in these tasks is provided here, instead of being individually reported in each chapter. Finally, general details concerning how data were analysed and reported in the empirical chapters are outlined.

**Participant recruitment**

**Adults.** Adult participants with ASD were recruited through the National Autistic Society and Durham University Service for Students with Disabilities. Neurotypical adults were recruited from the University of Durham and from advertisements in local newspapers. Additionally, both adults with ASD and neurotypical adults were recruited from an existing database of individuals with autism spectrum disorder and neurotypical individuals (at Durham University), who had expressed an interest in taking part in research. Participants all gave written, informed consent before participating.

**Children.** Children with ASD were recruited through mainstream schools and specialist schools for children with ASD. Children participants were also recruited through parent support groups organised by the Kent Autistic Trust (KAT).

Neurotypical children were all recruited from mainstream schools around Kent. Parents of children with ASD and neurotypical children provided informed consent, agreeing to their child's participation. Additionally, children all gave written, informed consent before participating. At the time of testing, the children were also asked if they are happy to take part and the nature of the testing session was explained to them. All participants were aware that they could stop participation in any experiments, at any time.

**Participant diagnoses**

Children and adults who were included in the ASD groups had all received formal diagnoses of autistic disorder or Asperger's disorder according to DSM-IV-TR/ICD-10 criteria (American Psychiatric Association, 2000; World Heath Organisation, 1993). No participants, in either the ASD or neurotypical group, reported using any psychotropic medication or any history of neurological or psychiatric disorders (apart from ASD).

**Adults.** Firstly, to confirm diagnoses, adults with ASD provided a copy of a medical statement, outlining the details of their diagnosis. Additionally, in order to assess current ASD features, 15 of the 18 adults with ASD who participated in the studies reported in this thesis completed the Autism Diagnostic Observation Schedule-Generic (ADOS; Lord et al., 2000) assessments. The ADOS is a semi-structured, standardized assessment of communication, social interaction, and imaginative use of materials that can be used to help diagnose autism spectrum disorder (ASD). All participants who completed the ADOS received a total score ≥7, the defined cut-off for ASD (Lord et al., 2000). The remaining three participants declined to complete the ADOS, as they did not feel comfortable being filmed. The three participants who did not

complete the ADOS had rigorous diagnoses and scored above the cut-off on the Autism-spectrum Quotient (see immediately below).

All adults with ASD also completed the Autism-spectrum Quotient (AQ; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001), a self-report questionnaire that assesses ASD/ASD-like features. Fifteen out of 18 participants with ASD scored above the defined cut-off for ASD on the AQ (total score ≥26; Woodbury-Smith, Robinson, & Baron-Cohen, 2005). Only three participants missed this cut-off. However, all three of these participants scored well above the defined ASD cut-off on the ADOS (all ADOS scores among these three participants were ≥ 12). Comparison participants also completed the AQ. All neurotypical adults scored below the defined cut-off for ASD on the AQ.

**Children.** For children with ASD, details of each child's diagnosis were acquired through their special educational needs (SEN) statements. To assess severity of ASD features, parents of participants with ASD also completed the Social Responsiveness Scale (SRS; Constantino et al., 2003). Parents of the neurotypical children also completed the SRS. The SRS alone cannot be used as a diagnostic tool for ASD, and thus participants were not automatically excluded from participating in any experiments because they did not meet the criteria on this questionnaire. However, to ensure that including these children's results did not affect the results of any of the experiments they participated in, all analyses were re-run, excluding any children who did not meet the recommended cut-off on the SRS (in both the ASD and neurotypical group). After removing these participants from analyses, none of the results (nor study conclusions) reported in any chapter changed substantively (i.e., no $p$ value changed from significant to non-significant or vice versa, and no effect size changed category – small, moderate, large). As such, these children's results were included in analyses. It is clearly noted in

each chapter how many participants met the diagnostic cut off for ASD, in both the ASD and neurotypical groups.

**Participant matching**

Appropriate ASD/comparison group matching is fundamental to any study of cognition in ASD (see the 2004 special issue on matching of the Journal of Autism and Developmental Disorders, Vol. 34). As such, the participant groups in all experiments were closely equated for verbal and non-verbal ability (please see each chapter for specific detail of group characteristics). For both children and adults participants, Verbal IQ (VIQ), performance IQ (PIQ), and full-scale IQ (FSIQ) were assessed using the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999). Groups were also closely equated for chronological age.

Additionally, for experiments investigating metacognitive ability on a task (see chapters three, four, and five), it was essential that the ASD group did not demonstrate impairments at the object-level of the task (i.e. basic memory performance). Theoretically, impairments at the object-level of a metamemory task may confound performance at the meta-level of the task, and potential group differences in metacognition could be explained by group differences overall on task performance. Thus it was essential that participants groups were also matched so that ASD groups showed similar recall/ recognition ability on memory task, relative to comparison participants.

**Participation in multiple experiments**

Data collection for several of the studies reported in this thesis was often carried out during the same batch of testing, and thus participants with ASD and neurotypical

participants sometimes participated in more than one of the experimental studies. However, it was not the case that all participants completed all of the experiments they were suitable for. This was mainly due to restrictions that occurred during data collection (e.g., participants missing testing sessions, participants having to leave early, schools only agreeing for children to participate in particular studies etc.). As such, participant numbers vary between experimental tasks.

**Measures of mindreading (ToM) ability**

In several studies, participants' mindreading ability was also assessed, in order to determine whether the individuals in ASD groups demonstrated stereotypical mindreading impairments, relative to comparison, neurotypical individuals. It was important to assess participants' mindreading ability, as many of the predictions made in the thesis were based on the assumption that mindreading would be impaired in individuals with ASD (e.g., according to the one-mechanism theory, metacognitive impairments should only be apparent if mindreading impairments are also present). To assess mindreading ability in adults with ASD, a version of the animations task (Abell, Happé, & Frith, 2000) was used. To assess mindreading ability in children two well established mindreading tasks were employed; the strange stories task (Happé, 1994) and also the animations task (Abell et al., 2000).

**The animations task.** The animations task is a widely-used measure of mindreading ability, and several studies have found that individuals with ASD demonstrate impaired performance on the task (Abell et al., 2000; Campbell et al., 2006; Castelli, Frith, Happé, & Frith, 2002; Jones et al., 2011; Lind, Williams, et al., 2014; Salter, Seigal, Claxton, Lawrence, & Skuse, 2008; White, Coniston, Rogers, & Frith, 2011; Zwickel, White, Coniston, Senju, & Frith, 2011). Additionally, neuroimaging studies

have also shown that the network of brain regions typically activated during

mindreading (see, for example, Perner, Aichhorn, Kronbichler, Staffen, & Ladurner,

2006; Samson, Apperly, Chiavarino, & Humphreys, 2004; Saxe & Powell, 2006) are

activated when neurotypical adults watch the animations (see Castelli et al., 2002;

Castelli, Happé, Frith, & Frith, 2000).
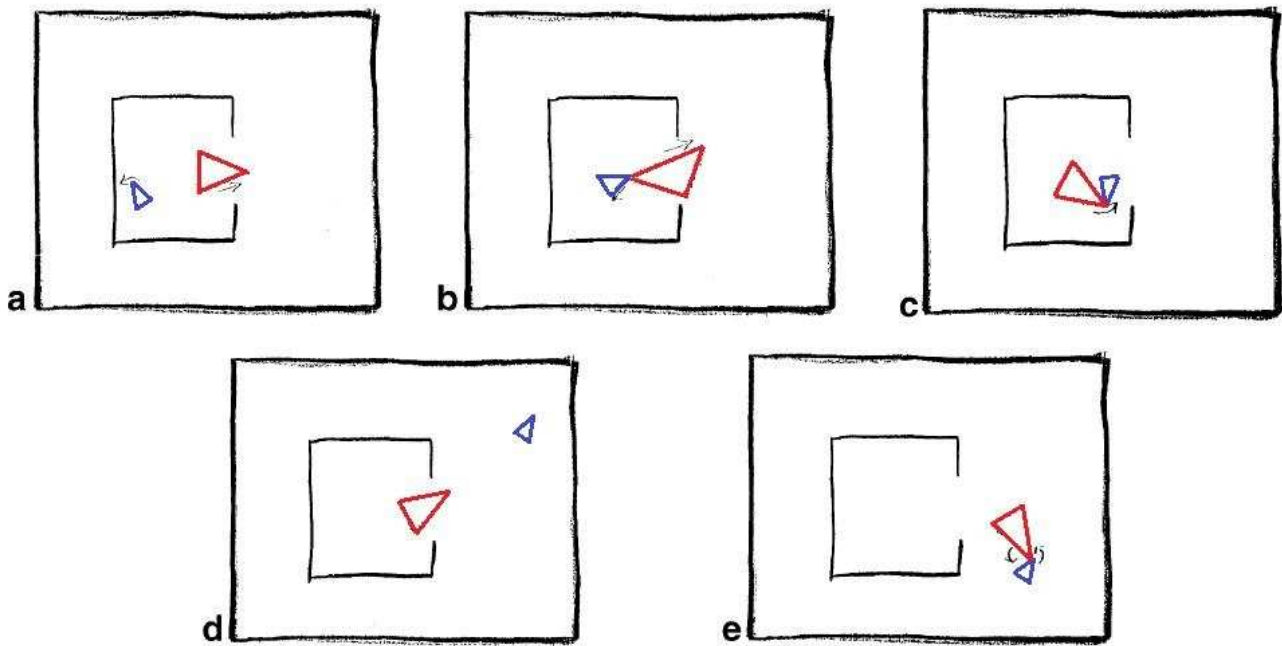


Figure 4: Graphical representation of the animations task (taken from Abell, et al.,
2000). These five stills are *taken from the "Coaxing" animations (mother and child): a)*
Mother tries to interest child in going outside. b) Child is reluctant to go out. c) Mother
gently nudges child towards the door. d) Child explores outside. e) Mother and child play
happily together.

During the animations task, participants were required to provide a verbal

description of silent video clips, each of which displayed an interaction between a large

red triangle and a small blue triangle (please see Figure 4). These clips were taken

directly from Abell et al., (2000). For adults, in four of the clips, an adequate

explanation of the triangles' interaction required the attribution of propositional attitudes, such as beliefs, intentions, and/or desires. As in Abell et al.'s study, these clips comprised a "mentalising" condition (assessing higher-level mindreading). In the remaining four clips, an adequate explanation of the triangles' interaction required the attribution of goal states, but did not necessarily require the attribution of propositional attitudes/epistemic mental states. As in Abell et al., (2000), these clips comprised a "goal-directed" condition. Child participants completed a shorter version of the animations task, in which only two clips (the "Coaxing" and "Tricking" animations) comprised the mentalising condition, and two clips (the "Fighting" and "Following" animations) comprised the goal-directed condition.

Each clip was presented to participants on a laptop computer. Before undertaking the experimental trials, participants (both children and adults) also completed two practice trials, to familiarise themselves with the task (one goal-directed and one mentalising). During practice trials, participants were asked to describe the behaviour displayed by the triangles in each of the video clips, and the experimenter gave feedback after each description. For each of the experimental animations, participants watched each clip twice. First, participants watched the clip through once in silence and were told to "watch the clip and see how the triangles are interacting". Participants then watched the clip again and were asked "as you watch the clip again I would like you to tell me how the triangles are interacting". Participants provided a running commentary on the triangles' interactions during this second presentation of the clip. For the experimental trials, a digital audio recording of participants' responses was made for later transcription. No feedback was given on the experimental trials.

Voice recordings of participants' commentaries were transcribed verbatim. These transcriptions were then scored by a rater who was blind to the diagnostic status

of the participants, according to the scoring criteria outlined in Abell et al., (2000; please see appendix one for a detailed copy of the scoring criteria). Participants' descriptions of each animation were given a score of zero, one, or two according to their level of accuracy, defined as the extent to which the participant's description captured the intended meaning of the animation. As such the total score achievable in each condition (Mentalising/Goal-directed) was between zero and eight for adults, and zero and four for children. Inter-rater reliability for scores across the four animations was assessed by *Cronbach's α,* and was extremely reliable (inter-rater reliability for animations scores in is reported separately in each chapter).

**The strange stories task.** The Strange Stories task (Happé, 1994) was also used as a second measure of mindreading ability, for child participants. The strange stories task is widely used in ASD research and has been shown to be a sensitive measures of mindreading (e.g., e.g., Happé, 1994; Joliffe & Baron-Cohen, 1999). During the Strange Stories task participants were presented with four short vignettes (two mentalising stories and two physical stories). These stories were taken from Happé (1994). First, each story was individually presented on cards to participants. Children either read the story aloud or, if children did not feel comfortable reading aloud, the experimenter read each story to them. After each story had been read the experimenter produced a second card and presented participants with a question about the content of the story (e.g., "Why did the prisoner say that?"). The experimenter read this question aloud, and participants provided a verbal response to each question. Before undertaking the experimental trials, participants also completed one practice trial, to familiarise themselves with the task (one mentalising story). During practice trials, participants read a story and provided an answer to the practise question. The experimenter then gave feedback after participants' response to the practise question. For the

experimental trials, a digital audio recording of participants' responses was made for later transcription, and no feedback was given.  Voice recordings of participants' answers to each question were transcribed verbatim.  These transcriptions were then scored by a rater who was blind to the diagnostic status of each participant, according to the scoring criteria outlined in White, Hill, Happé and Frith (2009; please see appendix two for a detailed copy of the scoring criteria).

Participants' answers to each question were given a score of zero, one, or two according to their level of accuracy, defined as the extent to which the participant's answers correctly answered the questions.  The score achievable in each condition (Mentalising/Physical) was between zero and four.  Again, inter-rater reliability for scores across the four stories was assessed by Cronbach's $\alpha$, and was reliable.  Inter-rater reliability for strange stories scores is reported separately in each chapter that uses this task.

**Measures of metacognitive monitoring accuracy.**

There are several approaches used to assess metacognitive monitoring accuracy in metamemory paradigms, including the FOK, JOC and JOL paradigms used in chapters three, four, and five.  The most commonly used measure of metacognitive monitoring accuracy is a Gamma correlation (Goodman & Kruskal, 1954).  Gamma correlations are a non-parametric measure of association, and assess the extent to which participants' object-level task performance is associated with their meta-level judgements on the task.  As such they provide an index of overall judgement accuracy during a metamemory task.  Gamma correlations are recommended by Nelson (1984) as the most appropriate way to analyse metacognitive monitoring accuracy, and are commonly used to analyse FOK tasks (Kelemen, Frost, & Weaver, 2000; Nelson &

Narens, 1990; Nelson, Narens, & Dunlosky, 2004; Wojcik et al., 2013), JOL tasks (Wojcik et al., 2014) and JOC tasks (e.g., Sawyer et al., 2014; Roebers, Schmid & Roderer, 2009). Gamma correlations were used to assess monitoring accuracy in the experiments reported in chapters three, four, and five. Gamma scores were calculated in one of two ways, depending on whether participants made dichotomous metamemory judgements (e.g., Yes/No) judgements, or continuous metamemory judgements.

**Dichotomous metamemory judgements.** For the adult FOK and JOL tasks (reported in chapters three and four), participants made dichotomous Yes/No memory judgements, and gamma scores were calculated using the formula $G = (ad - bc)/(ad + bc)$. In this equation (a) represented the number of correct "Yes" predictions an individual made, (b) the represented number of incorrect "Yes" predictions, (c) represented the number of incorrect "No" predictions, and ($d$) represented the number of correct "No" predictions. Gamma scores range between + 1 to -1, where a score of 0 indicates chance-level accuracy, a large positive value indicates a good degree of accuracy, and a large negative value indicates less than chance-level performance on the task. However, when calculating gamma scores this way, the score cannot be calculated when two or more of the prediction rates (a, b, c, or d) are equal to zero. As such, the raw data were adjusted by adding 0.5 onto each prediction frequency and dividing by the overall number of FOK/JOL judgements made (N) plus 1 (N+1). This correction is recommended by Snodgrass and Corwin (1988) and is routinely used when calculating gamma scores on metamemory tasks (Bastin et al., 2012; Wojcik et al., 2013).

**Continuous metamemory judgements.** In the remaining metamemory experiments, participants did not make simple Yes/No judgements concerning their

memory ability, but made metamemory judgements on a likert scale (e.g., in the JOC task reported in chapter five participants made confidence judgements on a scale ranging between one to seven, where one indicates very little confidence in their answer, and seven indicated that participants were extremely confident in their answer). For these experiments gamma correlations were calculated using the formula below, where *Na* represents the total number of agreements between an individual's metamemory judgements and *Ni* represents the total number of disagreements between metamemory judgements. Gamma calculations were calculated individually for each participant, using SPSS.

$$G = \frac{Na - Ni}{Na + Ni}$$

**Power analyses**

G*Power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007) was used to conduct power analyses, to determine the sample size required to detect *between-group* differences in experimental task performance for each experiment (using Cohen's 1992 criterion). However, *a priori* power analyses were not conducted to determine the necessary power to detect reliable associations between variables. As such, in the General Discussion, issues surrounding power are discussed, specifically issues surrounding whether some of the exploratory correlational analyses run were adequately powered to detect significant correlations.

**Statistical considerations**

A standard alpha level of .05 was used to determine statistical significance. All reported significance values are for two-tailed tests, unless otherwise indicated. Where

ANOVAs were used, $\eta_p^2$ values are reported as measures of effect size ($\geq$ .01 = small effect, $\geq$ .06 = moderate effect, $\geq$. 14 = large effect; Cohen, 1969). Where *t*-tests were used, Cohen's *d* values are reported as measures of effect size ($\geq$.0.20 = small effect, $\geq$ 0.50 = moderate effect; $\geq$ 0.80 = large effect; Cohen, 1969). When correlations were used $\geq$ .30 was considered a small effect, $\geq$ .50 was considered a moderate effect and $\geq$.70 was considered a large effect (Cohen, 1969).

**CHAPTER THREE**

---

**METACOGNITION, METAMEMORY, AND MINDREADING IN HIGH-FUNCTIONING**

**ADULTS WITH AUTISM SPECTRUM DISORDER**[1]

Metacognition can be broadly defined as "thinking about thinking".  More specifically, it refers to an individual's awareness of cognitions and encompasses *"metacognitive knowledge"*, "*metacognitive monitoring", and "metacognitive control".* Metacognitive knowledge refers to one's beliefs and factual knowledge about cognitive processes in general (in self and others), whereas metacognitive monitoring and control refer respectively to one's awareness of and ability to regulate one's own current, online mental states and cognitive activity (Flavell, 1979).

One extensively studied component of metacognition is metamemory, which refers to an individual's knowledge of memory processes, and ability to monitor and control their own memory.  Nelson and Narens' (1990) influential model of metamemory divides metamemory (monitoring and control) processes into two levels: the "object-level" and the "meta-level".  The object-level consists of first-order memory processes (i.e., memory itself), whilst the meta-level consists of dynamic, second-order representations of the object-level.  This model is supported by neuropsychological (e.g., Janowsky, Shimamura, & Squire, 1989; Shimamura & Squire, 1986) and psycho-pharmacological (e.g., Dunlosky et al., 1998) data, which highlight a dissociation

---

[1] This chapter has been adapted from Grainger, C., Williams, D. M. & Lind, S. E. (2014). Metacognition, Metamemory, and Mindreading in High-Functioning Adults with Autism Spectrum Disorder. *Journal of Abnormal Psychology, 123* (3), pp. 650-609.  It has been modified to fit the format of this thesis.

between memory and metamemory.  According to Nelson and Narens' model, through

metamemory *monitoring* individuals create a meta-representation of the object-level

(Nelson & Narens, 1990).  Additionally, metamemory *control* processes use information

held at this meta-level to feedback to the object-level, allowing individuals to alter

object-level processes and implement different strategies during learning (e.g., by

allocating more study time to information that one believes one has not learnt).  It is

partly for this reason that metamemory is considered essential for adaptive functioning,

allowing one to tailor one's behaviour according to one's strengths and weaknesses in

object-level memory.  As such, if an individual's metamemory monitoring is inaccurate

the strategies they implement during learning are likely to be ineffective.

### Metamemory judgments

One of the most commonly-used and classic paradigms to assess metamemory

monitoring involves asking people to make feeling-of-knowing (FOK; Hart, 1965)

judgements.  During a typical FOK task, participants are asked (during a study phase) to

memorise a series of stimulus pairs (e.g., pairs of words, such as "pen-key", "computer-

elephant" etc.).  Participants are then presented (during a cued-recall test phase) with

one stimulus from each pair (the cue; e.g., "pen"), and asked to recall its missing pair

(the target; e.g., "key").  Importantly, on trials in which participants fail to correctly

*recall* the target they are asked to judge the likelihood that, at a later point, they would

be able to *recognise* it.  Finally, participants are then presented with the cue and are

asked to select the unrecalled target from several options (a recognition test phase).

The accuracy of participants' judgments on metamemory tasks is typically assessed

using Gamma correlations (Goodman & Kruskal, 1954), which measure the association

between individuals' predictions about their future ability to recognise the correct

target with their actual subsequent recognition performance (see p.63 for a detailed description of how Gamma correlations are calculated).

**Metacognition as "applied theory of mind"**

Theory of mind (ToM) is the ability to attribute mental states, such as beliefs, desires, and intentions, to self and others in order to explain and predict behaviour (Premack & Woodruff, 1978). While most research into ToM focuses on awareness of *other* minds (henceforth called "mindreading"), research into metacognition focuses on awareness of one's own mind. Indeed, given the potential role of metacognition in self-regulation, Flavell (2000) considered metacognition an example of "applied ToM".

Several different perspectives have been proposed to explain the potential relation between mindreading and metacognition. According to one perspective (e.g., Carruthers, 2009; Frith & Happé, 1999), the ability to represent one's own mental states (metacognition) relies on the same underlying metarepresentational mechanism as the ability to understand mental states in others (mindreading). Crucially, according to this one-mechanism theory, no dissociation should exist between mindreading and metacognition ability; individuals who demonstrate mindreading impairments should also demonstrate impaired metacognition. However, this proposal has been disputed. According to a version of the "simulation theory", our ability to read other minds stems from our ability to directly introspect the contents of our own mind, and then use this information to mentally simulate the contents of another's mind in imagination (e.g., Goldman, 2006). From this perspective, metacognition is both ontogenetically and phylogenetically prior to, and foundational for, mindreading. According to a third theory, proposed by Nichols and Stich (2003), mindreading and metacognition are underpinned by separate mechanisms; the "monitoring mechanism" is responsible for

access to/awareness of one's own mental states, whereas a separate "mindreading mechanism" is responsible for processing information about others' mental states. Crucially, both of these latter two theories imply that there should be some people who manifest diminished mindreading abilities, despite undiminished metacognition. Indeed, both Goldman, and Nichols and Stich explicitly suggest that people with autism spectrum disorder (ASD) present precisely this pattern of impaired mindreading, but intact metacognition.

**Metacognition in Autism Spectrum Disorder**

Autism spectrum disorder (ASD) is a developmental disorder diagnosed on the basis of social-communication deficits, and fixated interests and repetitive behaviours (American Psychiatric Association, 2013). It is widely acknowledged that ASD is characterised by diminished mindreading ability (see Yirmiya, Erel, Shaked, & Solomonica-Levi, 1998). However, until recently the question of whether metacognition is diminished among people with ASD has remained largely unexplored.

The study of metacognition in ASD could have important implications for educational practice among individuals with ASD. Metacognition in general and, more specifically, metamemory play key roles in aspects of learning and decision-making that we know people with ASD have difficulties with. According to Nelson and Narens' (1990) metamemory model, information gained by monitoring one's own memory feeds back to memory functioning, allowing individuals to *control* their learning efficiently. As such, having a good awareness of what one has learnt can improve an individual's subsequent learning ability. For example, when revising for an exam, if an individual can accurately assess what information they already know, they are able to spend their time effectively, revising the topics they do not know. This issue may be

particularly relevant for intellectually high-functioning people with ASD, given that many of these individuals show significantly lower academic achievement than would be expected on the basis of their intelligence, which in turn impacts negatively on their life chances (see Estes, Rivera, Bryan, Cali, & Dawson, 2011). Indeed, the educational domains in which people with ASD frequently under-achieve are just those in which learning is known to be fostered by metacognitive training. Such training has been shown to remediate difficulties in reading comprehension (see Brown & Campione, 1996), writing (e.g., Sitko, 1998) and mathematical reasoning (e.g., Fuchs et al., 2003). In each of these domains, individuals with ASD show statistically significant under-achievement, relative to IQ (see Estes et al., 2011; Jones et al., 2009). It is possible that diminished metacognitive monitoring contributes to the lower-than-expected levels of academic achievement in ASD in these areas.

Thus, for several reasons it is important to establish the extent to which individuals with ASD show diminished metacognitive ability. In a seminal paper, Frith and Happé (1999) argued explicitly that individuals with ASD are *as* impaired at metacognition as they are at mindreading. More recently, Williams (2010) has taken up this idea, citing evidence that individuals with this disorder are as impaired at recognising their own and others' thought processes (Hurlburt et al., 1994), emotions and specific mental states, such as beliefs and intentions (Williams & Happé, 2010), as they are at recognising these states in others. Evidence from "self" versions of classic mindreading tasks (e.g., Williams & Happé, 2009b), in which participants are asked to report their own previously held (now false) belief, also suggests that individuals with ASD demonstrate diminished awareness of their own beliefs. Each of these findings suggests that metacognition is impaired in individuals with ASD, which appears in keeping with the view that mindreading and metacognition rely on the same underlying

mechanism. As such, it has been argued that the evidence from studies of mental state attribution in ASD provides support for the one-mechanism account.

However, some have argued that there is a critical limitation with these types of studies that prevents definitive conclusions being drawn about metacognitive ability in ASD (see Carruthers, 2009; Nichols & Stich, 2003). The potential difficulty is that test questions in self versions of classic mindreading tasks require participants to recall their *prior* mental states, rather than report their *current* mental states. Simulation and two mechanisms theories claim that only current mental states are directly accessible without the need for mindreading. Thus, arguably, the results from the above studies do not necessarily show that metacognition is impaired in ASD, because these tasks require inferences to be drawn about past mental states (but see Williams, 2010, for a counter-argument).

By contrast, it is widely agreed that metamemory monitoring judgements are based on awareness of *current* mental states. As such, if the accuracy of metamemory monitoring is diminished among people with ASD, this would provide strong support for the suggestion that metacognition is diminished in ASD, contrary to the predictions that follow from the simulation/two-mechanisms theory. In this regard, a seminal study by Farrant, Boucher and Blades (1999) reported no metamemory impairment in ASD. This study was used by Nichols and Stich (2003) to support the suggestion that metamemory is unimpaired in individuals with ASD, and thus to support their two-mechanisms theory. However, an issue with this study is that Farrant et al., assessed metamemory *knowledge.* The one-mechanism account proposes that metacognitive *monitoring/control,* rather than metacognitive knowledge, necessarily relies on the same metarepresentational mechanism as mindreading. As such, Farrant et al.'s study cannot be taken as conclusive evidence that all aspects of metamemory are typical in

individuals with ASD.  At most, it suggests that the metamemory *knowledge* may be intact – the study did not assess metamemory monitoring or control.

In order to unambiguously test whether metacognition is impaired in ASD, evidence is instead required from studies of metacognitive *monitoring* (or control). Performance on FOK tasks relies on individuals monitoring *current* internal memory states.  Only one study to date has examined metamemory in ASD using a FOK task (Wojcik et al., 2013).  Wojcik and colleagues assessed children's metamemory monitoring ability using two FOK tasks, one asking individuals to assess their memory for information stored episodically and one assessing memory for information stored semantically.  Wojcik reported that children with ASD were significantly poorer than neurotypical children at making accurate FOK judgements, but only when assessing their episodic memory.  However, there is a particular methodological difficulty affecting Wojcik et al.'s (2013) study that arguably prevents valid conclusions from being drawn.  The difficulty is that the ASD and neurotypical groups were not matched for verbal IQ (VIQ).  Matching for VIQ is essential in such studies, because differences between groups in this respect can potentially entirely explain between-group differences in experimental task performance (see Mervis & Klein-Tasman, 2004). Wojcik et al., (2013) recognised this limitation and tried to overcome it using an ANCOVA to "control" for group differences in VIQ.  However, ANCOVA does not, in fact, solve this problem (see Miller & Chapman, 2001)  and, thus, we cannot determine whether group differences were driven by diagnostic status or by VIQ differences.  In the current study, FOK accuracy was explored among ASD and comparison groups that were closely matched for VIQ, as well as for age, PIQ, and FSIQ.  If, as predicted, between-group differences in FOK accuracy were apparent, this would provide the first definitive evidence of a diminution of this ability among individuals with ASD.

**The Current Study**

The aim of this study was to explore the extent to which individuals with ASD are able to accurately monitor their own memory. To examine this, a classic FOK task was employed. It was predicted that participants with ASD would make significantly less accurate FOK judgments than comparison participants. During the FOK task different types of errors can lead to inaccurate FOK judgements; individuals can make over-confident errors (in which individuals incorrectly predict they will recognise a word that they subsequently fail to recognise) and also under-confident errors (in which individuals fail to predict their subsequently successful recognition of a target word). The type of error made by people with ASD during metacognitive monitoring tasks has not been explored previously, but it was predicted that individuals with ASD would make more FOK judgement errors overall, but would not be specifically biased towards over-confident or under-confident errors.

Additionally, the Meta-cognitions Questionnaire (MCQ; Cartwright-Hatton & Wells, 1997) was also used, as a self-report measure of participants' beliefs about their own metacognitive ability. To date, no study has previously assessed metacognitive ability in individuals with ASD using a self-report questionnaire. It was predicted that individuals in the ASD group would report diminished confidence in and awareness of and their own thoughts, as reflected by lower scores on the cognitive self-consciousness sub-scale and higher scores on the cognitive confidence sub-scale of the MCQ.

A measure of mindreading ability was also included in the current study. It was important to assess participants' mindreading ability, because according to the one-mechanism theory, metacognitive impairments should only be apparent if mindreading impairments are also present. To assess mindreading ability, a version of the

animations task (Abell et al., 2000) was employed. During this task, individuals are asked to view a series of clips in which animated triangles interact with one another. Participants are asked to provide descriptions of/explanations for the patterns of interaction between the triangles in each clip. An adequate explanation of the triangles' interactions requires the attribution of mental states (e.g., intentions, desires). Two conditions of the animations task were employed; namely a mentalising condition and a goal-directed condition. Both of these conditions appear to rely on the mindreading system, although performance on the mentalising condition is thought to rely on mindreading to a greater extent than the goal-directed condition. Based on the findings from previous studies (e.g., Abell et al., 2000; Lind, Williams, et al., 2014), it was predicted that participants with ASD would show diminished overall performance on the animations task, but not a group (neurotypical/ASD) by condition (mentalising/goal-directed) interaction on the task.

**Method**

*A priori* **power analysis**

Prior to commencing the study, G*Power 3.1 (Faul et al., 2007) was used to conduct a power analysis to determine the sample size required to detect the predicted group differences in gamma correlation on the FOK task. It can be argued that no valid studies of FOK accuracy have been conducted among individuals with ASD. Thus, for the purpose of this power analysis, an effect size for the between-group difference in FOK accuracy could not be predicted based on effect sizes found in previous studies. Therefore, based on theoretical inclinations toward the one-mechanism view, it was predicted that metacognitive impairments in ASD should be of a similar magnitude to the magnitude of mindreading impairments in this disorder. As such, the prediction for

the effect size associated with between-group difference in FOK accuracy in the current

study was based on the effect size found for between-group differences in mindreading

ability in studies of ASD.  In a meta-analysis exploring mindreading ability in individuals

with ASD relative to neurotypical individuals, Yirmiya and colleagues reported an

average Cohen's $d$ of 0.88 (Yirmiya et al., 1998).  Thus, assuming $d$ = 0.88 for between-

group differences in metamemory accuracy and α = .05, it was established that a total

sample size of $n$ = 17 participants per group would achieve Cohen's (1992)

recommended power of .80.

**Participants**

Ethical approval for this study was obtained from Durham University ethics

committee.  Eighteen adults with ASD (13 males, 5 females) and 18 neurotypical

comparison adults (11 males, 7 females) took part, all of whom gave written, informed

consent before participating.  One participant with ASD completed the MCQ incorrectly,

and so that participant's data for this questionnaire could not be used.  Participants in

the ASD group had all received formal diagnoses of autistic disorder ($n$ = 4) or

Asperger's disorder ($n$ = 14), according to DSM or ICD criteria (American Psychiatric

Association, 2000; World Heath Organisation, 1993).

In order to assess current ASD features, 15 of the 18 participants in the ASD

group completed Autism Diagnostic Observation Schedule-Generic (ADOS; Lord et al.,

2000) assessments.  The remaining three participants declined to complete the ADOS,

as they did not feel comfortable being filmed.  The three participants who did not

complete the ADOS had rigorous diagnoses and scored above the cut-off on the Autism-

spectrum Quotient (see immediately below).  All participants who completed the ADOS

received a total score ≥7, the defined cut-off for ASD (Lord et al., 2000).  All participants

completed the Autism-spectrum Quotient (AQ; Baron-Cohen et al., 2001), a self-report questionnaire that assesses ASD/ASD-like features. Fifteen out of 18 participants with ASD scored above the defined cut-off for ASD on the AQ (total score ≥26; Woodbury-Smith et al., 2005). Only three participants missed this cut-off. However, all three of these participants scored well above the defined ASD cut-off on the ADOS (all ADOS scores among these three participants were ≥ 12). All comparison participants scored below the defined cut-off for ASD.

*Table 2: Participant Characteristics (Means, Standard Deviations and Inferential Statistics).*

|  | Group | | t | p | Cohen's d |
|---|---|---|---|---|---|
|  | ASD (n = 18) | Neurotypical (n = 18) |  |  |  |
| Age (years) | 28.96 (10.28) | 30.43 (14.59) | 0.35 | .730 | 0.12 |
| VIQ | 111.67 (14.66) | 112.28 (10.87) | 0.14 | .888 | 0.05 |
| PIQ | 109.67 (15.75) | 114.50 (10.96) | 1.07 | .293 | 0.36 |
| FSIQ | 112.33 (15.00) | 114.94 (10.50) | 0.61 | .549 | 0.20 |
| AQ Total Score | 33.39 (9.24) | 13.00 (6.22) | 7.77 | <.001 | 2.59 |
| ADOS Social + Communication Score* | 11.93 (2.19) |  |  |  |  |

AQ: Autism-spectrum Quotient; ADOS: Autism Diagnostic Observation Schedule; PIQ = performance IQ; FSIQ = full scale IQ; VIQ = verbal IQ

*Based on 15/18 participants

No participants, in either group, reported using any psychotropic medication or any history of neurological or psychiatric disorders (apart from ASD). The participant groups were closely equated for verbal and non-verbal ability (see Table 2 for

participant characteristics).  Verbal IQ (VIQ), performance IQ (PIQ), and full-scale IQ

(FSIQ) were assessed using the full (four subtest) version of the Wechsler Abbreviated

Scale of Intelligence (WASI; Wechsler, 1999).  Groups were also closely equated for

chronological age.

## Materials and Procedures

**Feeling-of-knowing task.**  The stimuli used in the FOK task were 80 word pairs,

comprising of 160 concrete nouns (80 cue words and 80 target words).  Cue words

were matched with the target words for syllable length and word frequency (Kucera &

Francis, 1967), as reported in the MRC psycholinguistic database (Coltheart, 1981).  The

adequacy of this matching was confirmed by a non-significant effect of word type

(cue/target) in a multivariate ANOVA (using Wilks' Lambda criterion) that included

syllable length and word frequency as the dependent variables, $F$ (2, 157) = 0.68, $p$ = .93.

The procedure for the FOK task consisted of a study phase, a cued-recall test

phase (during which FOK judgements were also made; see below), and a recognition

test phase (see Figure 5 below for a graphical representation of one trial of the task).

The task was run on an LG desktop computer and lasted approximately 25 minutes.

Before completing the task participants completed a practice version of the entire

procedure, consisting of five word pairs.  As such, individuals knew before the study

phase that their memory for the word pairs would be tested, both by a cued-recall test

and a recognition test.

***Study phase.***  During the study phase, participants were presented with

individual word pairs (e.g., "bear-bridge"), each consisting of a cue word ("bear") and a

target word ("bridge").  Each word pair was presented individually for four seconds.

After the study phase, there was a five minute break, during which participants filled in

the MCQ (see subsection below).  After this break participants immediately completed the cued-recall test phase.

  ***Cued-recall and FOK phase.***  During the cued-recall phase, participants were shown individually presented cue words, in a random order, and were asked to recall the missing target word associated with each cue.  Immediately after each recall attempt (i.e., on a trial-by-trial basis), participants were asked to make a FOK judgement as to whether they thought they would be able to *recognise* the missing target word at a later point (either "Yes" or "No").  As such, participants made FOK judgements for all cue words, regardless of whether their recall of the target word had been accurate or not.  However, in the statistical analyses of FOK accuracy, only judgements made on trials in which participants *failed* to recall the target were included.  This procedure is common to studies of FOK ability among typically and atypically developing populations.  The procedure is designed to test participants' ability to judge the likelihood that they will be able to recognise information they have failed to recall.

  ***Recognition phase.***  Immediately after the cued-recall phase, participants completed the recognition test phase.  During the recognition test, participants were individually presented with all 80 cue words, in a random order, and were asked to identify the correct target word in a four-alternative, forced-choice recognition test.  On each trial, participants were asked to click (using the computer's mouse) the word they thought had been previously paired with the cue, from a selection of four options; the correct target word, an incorrect target word (that had previously been paired with a different cue word), and two novel distractor words not previously used in the task.  Importantly, for a given cue word, all participants were shown the same four options to choose from.  Once participants had clicked on a response the next trial began.  During the recognition test phase a target word only appeared as an option twice; once on a

trial in which it was the correct target word and once on a trial as an incorrect target word. The same target word (appearing either as the correct or incorrect option) never appeared on two consecutive trials.
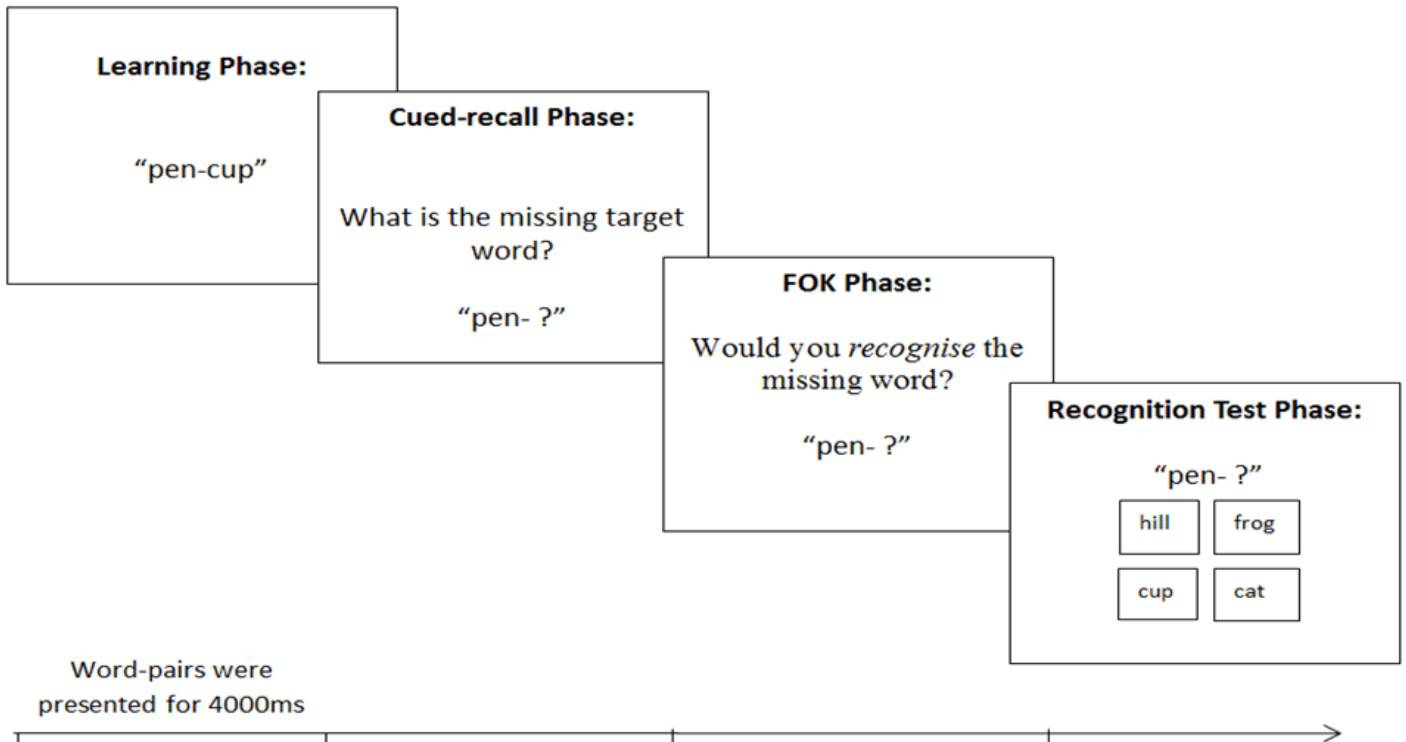


Figure 5: Graphical representation of the procedure used during the FOK task.

**Meta-Cognitions Questionnaire.** The Meta-Cognitions Questionnaire (MCQ; Cartwright-Hatton & Wells, 1997) was used to assess participants' beliefs about their own thoughts, and the efficacy of different thought processes. The MCQ presents participants with individual statements (e.g., "I have little confidence in my memory for words and names") and participants were asked to decide the extent to which they agreed with each statement, responding on a 4-point likert scale, ranging from do not agree, agree slightly, agree moderately, to agree very much. The questionnaire consists of 65 items comprising five subscales. The study was interested in two of these subscales specifically. The Cognitive confidence and Cognitive self-consciousness

subscales each address participants' awareness of their own thought processes and their confidence in their own cognitions, which are of particular relevance to this study. In contrast, the remaining subscales addressed issues about worrying and the effects intrusive negative thoughts may have on one's functioning, which seemed less related to the aims of the study.

**Animations task.** The animations task (Abell et al., 2000) was administrated, as a measure of mindreading ability. A full description of this task is provided in chapter two, p.59.

## Scoring

**Feeling-of-knowing task.** Two measures of participants' basic object-level memory performance were calculated on the FOK task. Recall ability was calculated as the proportion of target words participants correctly recalled during the cued-recall-stage. Similarly, recognition ability was calculated as the proportion of target words participants correctly recognised during the recognition test phase of the task. Gamma scores (Goodman & Kruskal, 1954) were calculated to provide an index of overall *FOK judgement accuracy*. This analysis is recommended by Nelson (1984) and is commonly used to analyse FOK tasks (Kelemen et al., 2000; Nelson & Narens, 1990; Nelson et al., 2004; Wojcik et al., 2013). Gamma scores are a non-parametric measure of association (between predictions and actual performance) and were calculated by comparing the number of correct predictions that each individual made with the number of incorrect predictions they made. To calculate gamma scores the formula $G = (ad - bc)/(ad + bc)$ was used, with (a) representing the number of correct "Yes" predictions an individual made, (b) the number of incorrect "Yes" predictions, (c) the number of incorrect "No" predictions, and (*d*) the number of correct "No" predictions. Gamma

scores range between + 1 to -1, where a score of 0 indicates chance-level accuracy, a large positive value indicates a good degree of accuracy, and a large negative value indicates less than chance-level performance on the task. However, when calculating gamma scores, the score cannot be calculated when two or more of the prediction rates (a, b, c, or *d*) are equal to 0. As such, the raw data were adjusted by adding 0.5 onto each prediction frequency and dividing by the overall number of FOK judgements made (N) plus 1 (N+1). This correction is recommended by Snodgrass and Corwin (1988) and is routinely used when calculating gamma scores on metamemory tasks (Bastin et al., 2012; Wojcik et al., 2013).

The number of errors made by participants in each group was calculated for two different *types of errors in FOK predictions*. The number of under-confident errors participants made was calculated as the number of incorrect "No" predictions, in which individuals failed to predict their subsequently successful recognition of a target word. The number of over-confident errors participants made was calculated as the number of incorrect "Yes" predictions made, in which individuals inaccurately predicted that they would recognise a word that they subsequently failed to recognise.

**Meta-Cognitions Questionnaire.** MCQ Sub-scale scores were calculated for the Cognitive confidence subscale and the Cognitive self-consciousness subscale. Lower total scores on the Cognitive confidence sub-scale indicated a greater confidence in one's own cognitions, whilst higher total scores on the Cognitive self-consciousness sub-scale indicated a higher reported awareness of one's own thought processes.

**Animations task.** Voice recordings of participants' commentaries were transcribed verbatim. These transcriptions were then scored by a second, independent rater (who was blind to the hypotheses of the study and the diagnostic status of the

participants) on the basis of scoring criteria outlined in Abell et al., (2000). Inter-rater reliability for scores across the eight animations was almost perfect, *Cronbach's α* = .98.

## Results

**Feeling of knowing task**

**Memory (object-level) performance.** Group differences in object-level memory performance were examined using independent-samples *t*-tests (see Table 3 for descriptive and inferential statistics). These indicated that individuals in the ASD group *recalled* significantly fewer target words than comparison participants in the FOK task. However, no significant group difference was found in the proportion of target words correctly *recognised* in the FOK task.

**Metamemory performance.** Group differences in metamemory monitoring accuracy were examined (see Table 3 for descriptive and inferential statistics). An independent-samples *t*-test indicated that there was a significant difference in gamma scores between the ASD and neurotypical group. Thus, in accordance with predictions, participants with ASD were significantly poorer at predicting their own memory performance than were neurotypical participants. Nonetheless, one-sampled *t*-tests indicated that gamma scores were significantly above chance (i.e. significantly greater than 0) in ASD and neurotypical groups, all $t$s > 2.97, all $p$s < .009.

An additional analysis was also carried out to investigate whether the significant group difference in object-level recall of target words confounded performance at the meta-level of the task (i.e., FOK judgements). For the purpose of this analysis, two participants from each group were excluded to create ASD and neurotypical groups that were matched closely for recall ability, $t$ (30) = 1.14, $p$ = .26, $d$ = 0.41. Groups also remained matched for chronological age, VIQ, PIQ, and FSIQ (all $p$s > .33, all $d$s < 0.35).

An independent-samples $t$-test indicated that even when groups were equated closely for recall ability, FOK gamma scores were still significantly lower in the ASD group ($M$ = .09, $SD$ = .16) than in the neurotypical group ($M$ = .25, $SD$ = .18), $t$ (30) = 2.60, $p$ = .014, $d$ = 0.94.

Group differences in the specific type of errors participants made on the FOK task were also examined.  Independent samples $t$-tests indicated that participants in the ASD group made significantly more under-confident FOK errors than participants in the neurotypical group (see Table 3 for statistics).  There was no significant group difference in the number of over-confident FOK errors made (see Table 3 for statistics).

**Self-report measure of metacognitive ability**

Table 3 shows the means and standard deviations for the two key MCQ subscale scores in the ASD and neurotypical group.  A significant between-group difference was found in scores on the Cognitive self-consciousness subscale, indicating that participants in the ASD group believed they were *superior* at monitoring their own thoughts, and more aware of their own thought processes relative to comparison adults. There was no significant between-group difference in scores on the Cognitive confidence subscale.

*Table 3: Means (SDs) and inferential statistics for group differences in performance on the FOK task, MCQ, and*

*animations task.*

| Experimental Measure | | Group | | t | p | Cohen's d |
|---|---|---|---|---|---|---|
| | | ASD | Neurotypical | | | |
| FOK Task: Object-level memory performance | Proportion of targets recalled | .18 (.15) | .31 (.22) | 2.04 | .049 | 0.69 |
| | Proportion of targets recognised | .65 (.23) | .73 (.19) | 1.20 | .240 | 0.38 |
| FOK Task: Metamemory performance | Gamma score* | .11 (.15) | .27 (.18) | 2.90 | .007 | 0.97 |
| | Number of over-confident judgments | 6.50 (6.56) | 6.89 (7.14) | 0.17 | .866 | 0.06 |
| | Number of under-confident judgments | 22.50 (7.41) | 14.22 (6.59) | 3.54 | .001 | 1.18 |
| MCQ | Cognitive self-consciousness subscale | 21.06 (3.73) | 16.89 (4.31) | 3.05 | .004 | 1.03 |
| | Cognitive confidence subscale | 19.00 (4.30) | 19.83 (5.17) | 0.52 | .609 | 0.17 |
| Animations task | Mentalising condition | 3.78 (1.70) | 4.89 (1.71) | 1.96 | .059 | 0.65 |
| | Goal-directed condition | 5.83 (1.50) | 7.22 (0.73) | 3.52 | .001 | 1.18 |

Note: ASD = autism spectrum disorder; FOK = Feeling of knowing; MCQ = Meta-cognitions Questionnaire

*Gamma scores index metamemory monitoring accuracy

**Animations task**

Table 3 shows the means and standard deviations for performance on the animations task. A mixed-model ANOVA was carried out on these data with Group (neurotypical/ASD) entered as the between-subjects variable, and Animation Type (mentalising/goal-directed) entered as the within-subject variable. There was a significant main effect of Group on animations scores, reflecting the fact that participants with ASD performed significantly less well than comparison participants on the task overall, $F$ (1, 34) = 9.02, $p$ = .005, $\eta_p^2$ = .21. There was also a significant main effect of Animation Type, indicating that, across both groups, scores were higher in the goal-directed condition than the mentalising condition, $F$ (1, 34) = 72.82, $p$ < .001. $\eta_p^2$ = .68. There was no significant Group by Animation Type interaction, $F$ (1, 34) = 0.29, $p$ = .59, $\eta_p^2$ = .01, suggesting that individuals in the ASD group were impaired at both higher- and lower-level mindreading, relative to individuals in the neurotypical group.

**Exploratory correlation analyses: Associations between metamemory ability, and mindreading ability and self-reported metacognitive skill**

A series of correlational analyses was carried out to explore the relation between performance in each condition of the animations (mindreading) task and performance on the FOK (metacognition) task. It should be noted that, although the current study was sufficiently powered to detect predicted group differences in FOK accuracy, it was not sufficiently powered to detect moderately-sized correlations (r = .30) between FOK accuracy and mindreading ability (please see the discussion for further information regarding study power). The following correlation analyses should, thus, be considered exploratory. In summary, neither FOK accuracy (gamma score), nor the number of under-confident FOK errors made, nor the number of over-confident FOK errors made

was associated significantly with performance in the mentalising condition of the animations task, or performance in the goal-directed condition of the animations task, among ASD or comparison participants, all $r$s ≤ -.32, all $p$s ≥ .201. Additionally, neither FOK accuracy (gamma score), nor the number of under-confident FOK errors made, nor the number of over-confident FOK errors made was associated significantly with scores on either of the MCQ sub-scales, among ASD or comparison participants, all $r$s ≤ -.43, all $p$s ≥ .077.

## Discussion

Until now, no study has established the extent to which individuals with ASD are able to accurately monitor their own memory by judging feelings-of-knowing. As such, the primary aim of this study was to establish this. In terms of the central experimental finding, the study found that participants with ASD showed significantly diminished FOK accuracy. This diminution was associated with a large effect size ($d$ = 0.97), indicating a substantial difficulty with metamemory monitoring.

This result is in keeping with predictions that individuals with ASD would show impairments in metamemory monitoring. However, there are several potential explanations for the observation of diminished gamma scores in the ASD group. One possibility is that individuals with ASD demonstrated a "positive illusory bias" during the task. The concept of a positive illusory bias refers to a tendency for an individual to self-assess their perceived competence as greater than their actual ability. This bias has been observed among individuals with attention deficit hyperactivity disorder (see Owens, Goldfine, Evangelista, Hoza, & Kaiser, 2007). More importantly, some studies have indicated that individuals with ASD tend to self-report their own social functioning more positively than parents will report (e.g., Lerner, Calhoun, Mikami, & De Los Reyes,

2012), and will self-report the level of their own autistic traits as less severe than parents will report (e.g., Johnson, Filliter, & Murphy, 2009). These studies have been interpreted as suggesting that individuals with ASD may also show a tendency to manifest a positive illusory bias. Demonstrating a positive illusory bias may indeed partly explain findings that participants with ASD self-reported (on the MCQ) greater awareness of their own mental states than neurotypical comparison participants reported. This self-reported superior awareness among participants with ASD stood in direct contrast to their diminished performance on an objective, well-established measure of metamemory monitoring ability. As such, the idea that some individuals with ASD manifest a positive illusory bias provides a plausible explanation for the MCQ findings.

However, it is not apparent that a positive illusory bias can explain the central finding of diminished FOK accuracy among participants with ASD. Individuals who manifest a positive illusory bias would, by definition, overestimate their memory ability and would, thus, be expected to make more over-confident errors when making FOK judgements. In other words, diminished FOK accuracy among people whose judgements were driven by a positive illusory bias would be driven by over-confidence. Yet, participants with ASD did not specifically make significantly more *over-confident* errors than comparison participants. Rather, individuals with ASD made significantly more under-confident errors than comparison participants. As such, it appears that demonstrating a positive illusory bias cannot explain the specific pattern of results shown in this study

The finding that participants with ASD made significantly more errors of the under-confident type (i.e., they tended to recognise targets that they judged they would not recognise), but not the over-confident type, was contrary to the prediction that

between-group differences in monitoring accuracy would be driven by an increase of both types of error among participants with ASD. This suggests that diminished performance on the FOK task among participants with ASD was driven by a relative lack of awareness of existing knowledge, rather than a belief in the possession of knowledge that does not, in fact, exist.

These results have several potential practical and clinical implications. Ultimately, if an individual has a reduced ability to accurately assess what information they know, and what they do not know, this may have several consequences. From an educational perspective, studies have shown that several outcomes (such as exam performance) can be predicted by metacognitive monitoring accuracy (e.g., Hartwig, Was, Isaacson, & Dunlosky, 2012; Thiede, Anderson, & Therriault, 2003). Findings that individuals with ASD show impaired metamemory monitoring need to be taken into account in educational environments, and should inform intervention efforts designed to remediate cognitive impairments in ASD. Studies in typical development have also shown that cognitive impairments can be remediated by fostering metacognition (e.g., Dunlosky, Kubat-Silman, & Hertzog, 2003). Indeed, training metacognitive skills has been shown to remediate difficulties in reading, writing and mathematical reasoning (see Brown & Campione, 1996; Fuchs et al., 2003; Sitko, 1998) in typical development. The results of the current study make it plausible to suggest that diminished metacognitive monitoring ability contributes to educational underachievement in these areas among people with ASD. If this turns out to be correct, it could have revolutionary effects on educational practices for people with ASD. It is important for future research to build upon the current results by exploring the extent to which metacognitive impairments contribute to educational success among individuals with ASD.

As well as having important educational implications, the central finding of reduced FOK accuracy in ASD also has theoretical implications. The central findings of diminished FOK accuracy alongside diminished mindreading ability are in keeping with the predictions of the one mechanism theory of the relation between metacognition and mindreading. Of course, the results do not definitively prove the theory, but certainly they are not in keeping with a key prediction made by either the simulation theory or the two-mechanisms theory that metacognition is unimpaired in ASD. As such, the main results of this study provide some support for the one-mechanism account. Having said this, the study did not find a significant positive association between FOK accuracy and performance in either the mentalising or goal-directed conditions of the animations task. The one-mechanism account would have predicted such associations between metamemory and mindreading, so the current results did not support the theory in this respect. However, caution should be taken when interpreting the results of the correlation analyses. The exploration of associations between FOK task performance and animations (mindreading) task performance was carried out as exploratory analysis, and no *a priori* power analysis was conducted to establish that the study had adequate power for this secondary aim. A subsequent power analysis (after completion of the study) was conducted with a view to determining what sample size would have been necessary to detect meaningful, statistically significant associations between metacognitive monitoring ability and mindreading ability. Assuming a moderate association ($r = 0.30$) and $\alpha = .05$, a total sample size of $n = 67$ participants would be needed to achieve Cohen's (1992) recommended power of .80 for the correlational analyses. Thus, this study was under-powered to detect a meaningful association between these two abilities. This represents a limitation of this study and, as such, caution should be taken when interpreting the findings from these correlation analyses.

Future studies using larger sample sizes are warranted to further investigate relations between metacognitive monitoring and mindreading ability.

What is clear is that the current study was sufficiently powered to detect predicted group differences in FOK accuracy and that results indicated participants with ASD showed a substantial diminution of metamemory monitoring. Of course, there are other forms of judgement that can be used to assess metamemory, namely judgements of learning and judgements of confidence. It remains possible that people with ASD will show undiminished accuracy in these judgements. Judgments of learning involve assessing how well one thinks one has learnt a piece of information, and judgements of confidence involve making retrospective judgments about how certain one is in one's knowledge about a piece of information. The literature on typical development suggests that metamemory accuracy is only modestly correlated across different types of metamemory judgement (Kelemen et al., 2000; Leonesio & Nelson, 1990). This has led to suggestions that different metamemory judgments may be based on different sources of information. Metamemory judgements are thought to be based on mnemonic cues and it is possible that different judgements are based on different cues (see Koriat, 1993; Metcalfe et al., 1993). Although it was predicted that individuals with ASD will demonstrate impairments across different metamemory judgements, this may not turn out to be the case. So far there have been only two published studies of judgment of confidence accuracy (Wilkinson et al., 2010; Wojcik et al., 2011). Results from these studies have been inconsistent; whereas Wilkinson et al., (2010) report that confidence judgments made by children with ASD were less accurate than those made by neurotypical children, Wojcik and colleague report no impairments in JOC accuracy in children with ASD (Wojcik et al., 2011). Thus, the study of metacognitive monitoring in

ASD is in its infancy and a sustained study of metamemory and its neurocognitive basis in ASD would be fruitful.

Future research should address these issues, and should also aim to address whether it is possible to foster metacognitive skills in individuals who do show impairments. In our view, a comprehensive investigation of metacognition in ASD is essential, given the consequences that impaired metacognitive monitoring and regulation may have on an individual's cognitive performance. It is hoped that alongside future research the findings from this study will help to establish a more definitive account of metacognitive ability in ASD, and that a greater understanding of this area will eventually contribute to successful remediation of cognitive and behavioural impairments in this disorder.

**CHAPTER FOUR**

## JUDGMENT OF LEARNING ACCURACY IN HIGH-FUNCTIONING CHILDREN AND ADULTS WITH AUTISM SPECTRUM DISORDER

Metacognition is typically assessed in the context of metamemory, an individual's awareness of their own memory and their ability to regulate their own memory processes during learning. Employing a traditional metamemory task, chapter three investigated whether individuals with ASD demonstrate metamemory impairments, when asked to judge the extent to which they feel they will know a piece of information in the future (a FOK judgements). The results of this chapter indicated that individuals with ASD demonstrate impairments in metacognitive monitoring processes, and are less accurate at assessing their own states of knowledge than comparison participants.

There are of course several other forms of judgements that can be used to assess metamemory, most notably judgements of learning and judgements of confidence. Although chapter three indicated that individuals with ASD demonstrate impaired metamemory on a FOK task, it remains possible that people with ASD will show undiminished accuracy in these other judgements. The literature on typical development suggests that metamemory accuracy is not necessarily consistent across different types of metamemory judgements (Kelemen et al., 2000; Leonesio & Nelson, 1990). This has led to suggestions that different metamemory judgments may be based on different sources of information. As such, although theoretically impairments on other metamemory tasks are predicted, it is possible that metamemory monitoring is not universally impaired in ASD.

**Metacognitive judgments-of-learning**

The ability to monitor one's own cognitions has repeatedly been assessed by asking people to make judgement of learning (JOL; Arbuckle & Cuddy, 1969) assessments.  During a standard JOL task participants are asked (during a learning phase) to memorise a series of stimuli pairs (e.g., pairs of words, such as "pen-key", "computer-elephant" etc.).  After the learning phase, participants complete the JOL phase.  During this phase, participants are presented with one stimulus from the pair (the cue; e.g., "pen") and asked to make a judgement on the likelihood that, at a later point, they will be able to recall the stimulus' pair (the target; i.e., "key").  Finally, during a recall phase, participants are presented with each cue word in turn and asked to recall the corresponding missing target word.  The accuracy of participants' metacognitive judgements is measured by comparing participant's judgments about their future recall performance with their actual recall performance.

Thus, in a standard procedure, the JOL phase involves making so-called "cue-alone" judgements (Dunlosky & Nelson, 1992), and participants are presented only with the cue word and are asked the judge the likelihood that they will later recall the corresponding target.  However, an important variant to this standard procedure involves participants making so-called "cue-target" judgements (Dunlosky & Nelson 1992).  In this version, which involves a manipulation of the JOL phase, individuals are asked to determine the future retrievability of the target when presented with *both* the cue and the target.  Whilst both JOL tasks ask individuals to assess their future memory ability, it is distinctly possible that individuals rely on different types of information and cues during such tasks.  Typically, individuals demonstrate better accuracy on cue-alone JOL tasks than cue-target JOL task (e.g., Dunlosky & Nelson, 1992, 1997).  One

explanation for this finding is that when making each type of JOL individuals retrieve

information about the to-be-remembered stimuli from both their short term memory

(STM) and long term memory (LTM). Information from LTM is thought to be a better

indicator of future memory performance than information from STM, as it is more

similar to the information participants will rely on during the recall phase of a JOL task.

Delayed cue-alone JOLs are thought to be based mostly on information recollected from

long-term memory, than from information immediately accessible n from short-term

memory. However, when making delayed cue-target JOLs information directly

accessibly from STM (reactivated into STM by the presentation of both the cue and

target word) is thought to add "noise" to the information retrieved from LTM, thus

reducing the accuracy of individuals cue-target predictions (Dunlosky & Nelson, 1992,

1997). Alternatively it has been suggested that individuals are less accurate at making

cue-target JOL assessments, as the presence of the target word during the JOL phase

means participants are not given the opportunity to experience retrieving the target

word during the JOL phase ,which presumably occurs when individuals make cue-alone

JOL assessments (see Dunlosky & Nelson, 1997).

Additionally, judgements of learning are not always made on an individual, item-

by-item basis. When making an *aggregate JOL* individuals are asked to make an overall

estimate of the number or percentage of items they think they will remember, and

predictions about future performance are made after individuals have completed all

learning trials. It has been proposed that aggregate JOLs may provide individuals with a

different opportunity for self-monitoring, compared to item-by-item JOLs (e.g.,

Schneider, Visé, Lockl, & Nelson, 2000). Individuals tend to make different types of

errors on aggregate JOL tasks, compared to item-by-item tasks; whereas neurotypical

individuals tend to demonstrate overconfidence in their judgements on item-by-tem JOL

tasks, individuals tend to demonstrate less overconfidence on aggregate JOL tasks (e.g., Mazzoni & Nelson, 1995). Individuals also typically demonstrate better accuracy on aggregate JOL tasks than item-by-item JOL tasks (e.g., Schneider et al., 2000). This is thought to be because individuals base aggregate judgements on prior experience retrieving similar information and on how deeply they feel they have encoded all memory items, rather on specific mnemonic cues associated with individual memory items.

**JOL accuracy in individuals with ASD**

There is a growing body of research that suggests individuals with ASD manifest diminished awareness of their own mental states (e.g., Williams & Happé, 2009b; Williams & Happé, 2010; Wojcik et al., 2013). In keeping with the results from these studies, the results reported in chapter three indicated that adults with ASD showed significantly diminished feeling of knowing accuracy. These results suggest that, at least when making on one type of metamemory judgment, individuals with ASD are less accurate at monitoring their own mental states.

At the time of designing and running Experiment 1, no studies had been published using a JOL task to assess metamemory in either children or adults with ASD. Thus, based on the results of previous studies, as well as theoretical predictions that metacognition should be impaired in individuals who demonstrate mindreading impairments, predictions of impaired monitoring accuracy on a JOL task were unequivocal. However, recently Wojcik and colleagues (Wojcik et al., 2014) reported the results of two JOL tasks in children with ASD, and reported finding no metacognitive deficits in the accuracy of children's JOL assessments. In their study children with ASD, and neurotypical children, were presented with word pairs, and were either

immediately asked to judge whether they thought they would be able to remember the target words (during an immediate cue-alone JOL task) or made JOL decisions after a delay (during a delayed cue-alone JOL task). Wojcik et al., (2014) found that individuals with ASD were as accurate as neurotypical participants at judging their future memory performance, across both the immediate and delayed JO tasks. That being said, there was some indication that metamemory was impaired in individuals with ASD in this study. Wojcik et al., (2014) found that on the immediate JOL task, on average, the ASD group's JOL accuracy was not significantly better than chance level performance on the task (i.e., not significantly greater than 0). As such, JOL predictions made by individuals with ASD were in no way associated with their subsequent memory performance on the task. Although the study found that across both tasks the ASD group did not make significantly less accurate JOL predictions than the neurotypical group, on the immediate JOL tasks the difference between gamma scores in the ASD group (*Mean* = .05, *SD* = .11) and neurotypical group (*Mean* = .27, *SD* = .11) was very large (*Cohen d* = 2.00). As such, at least on the immediate JOL task used in this study, it did appear that individuals with ASD demonstrated poor metamemory monitoring accuracy.

Additionally, there are several potential methodological issues with Wojcik et al.'s (2014) study that suggest caution should be taken when interpreting the study results. Firstly, no measure of mindreading ability was used in this study. Theoretically, predictions that individuals with ASD will demonstrated an impaired understanding of their own mental states are based on findings that individuals with ASD demonstrate impairments understanding mental states in others. It is possible that ASD participants in this study did not demonstrate impairments in metamemory accuracy because they showed similar mindreading ability to the comparison individuals. Although this is purely speculative, without directly measuring mindreading ability we cannot be sure

that mindreading was impaired in the sample of ASD participants, relative to the

comparison group.

Another alternative explanation for Wojcik et al.'s finding that individuals with

ASD are not impaired when making JOL assessments is suggestion that individuals with

ASD used an atypical strategy during the JOL phase of the task. In other words, it might

be possible that participants with ASD in Wojcik et al's study *performed* well on the

delayed JOL task, despite diminished underlying metacognitive monitoring *competence.*

It has been speculated that relatively accurate JOLs could be made on cue-*alone* JOL

tasks simply by judging whether one can bring to mind the target word at the time a JOL

is made. For example, during cue-alone JOL tasks, if presented with the cue word (e.g.,

"bear-?") and asked to make a JOL about whether you will remember what the missing

target word is at a later point, individuals might adopt the strategy of simply answering

"yes" if, at the point they make the JOL, they can remember the target word, and "no" if

they cannot. As such, when making JOLs, individuals with impaired metacognition

might still be able make relatively accurate JOLs, simply by employing this strategy.

This may be a potential explanation for why Wojcik et al., (2014) found seemingly intact

monitoring ability on two cue-alone JOL tasks, which seems inconsistent with other

studies of metacognition in ASD (Grainger, Williams, & Lind, 2014; Wojcik et al., 2013).

Participants cannot use this strategy to compensate for impairments in metacognition

during cue-target JOL tasks (in which both cue and target words are presented during

the JOL phase of the task), aggregate JOL tasks, or other metamemory tasks (such as

FOK tasks). For example, in cue-target JOL tasks individuals are not given the

opportunity to retrieve the target word during the JOL phase, and thus they cannot rely

on this experience when making their judgements. Similarly, in a standard FOK task

(such as the task employed in chapter three) judgements are made on items

participants failed to recall, and thus again individuals cannot achieve typical performance on a FOK task by relying on their current memory for missing target words.

**The Current Study**

The current study aimed to address these issues, and explored the extent to which both children and adults with ASD are able to make accurate JOL assessments on a series of JOL tasks. Three experiments were carried out, two exploring JOL accuracy in adults with ASD and one exploring JOL accuracy in children with ASD. In these experiments multiple JOL tasks were employed, and the experiments were designed to follow on from each another, as outlined below. Additionally, measures of mindreading ability were also included, firstly to confirm typical mindreading impairments in the sample of individuals with ASD being tested, and secondly to analyse the relation between performance on metamemory and mindreading tasks.

The primary aim of both experiment 1 and experiment 2 was to thoroughly assess JOL ability in adults with ASD, including the relationship between mindreading impairments and impairments on JOL tasks. Importantly, to date no study has previously assessed JOL accuracy in relation to mindreading ability, and thus it was important to establish to what extent performance on such tasks are related. The final experiment was carried out to address alternative explanations for the earlier findings reported in experiments 1 and 2, and also previous findings from the literature (Wojcik et al., 2014).

The study's main prediction was that participants with ASD would demonstrate impairments in accuracy on all the JOL tasks employed. Given the results from chapter three, it was also predicted that individuals with ASD would demonstrate more under-

confident errors whilst monitoring their memory, relative to neurotypical participants.

Additionally, mindreading ability was assessed in all experiments.  It was also important

to assess participants' mindreading ability, as theoretically it was predicted that

metacognitive impairments in ASD should only be apparent if mindreading

impairments are also present.  Throughout, mindreading ability was assessed in adults

using the Animations task (Abell et al., 2000), and in children using both the Animations

task and the Strange Stories task (Happé, 1994).  It was predicted that individuals with

ASD would demonstrate stereotypical mindreading impairments on all mindreading

tasks, relative to matched neurotypical participants.  Including measures of

mindreading also meant that the relationship between performance on the JOL tasks

and performance on mindreading tasks could be explored.  It was predicted that

individuals with ASD would demonstrate corresponding deficits on both the JOL and

mindreading tasks, and that impairments in JOL accuracy would be positively related to

impairments in mindreading accuracy.

## Experiment 1: Method

### Participants

Eighteen adults with ASD (13 males, 5 females) and 18 neurotypical comparison

adults (11 males, 7 females) took part, all of whom gave written, informed consent

before participating.  Participants in the ASD group had all received formal diagnoses of

autistic disorder ($n$ = 4) or Asperger's disorder ($n$ = 14), according to DSM-IV or ICD

criteria (American Psychiatric Association, 2000; World Heath Organisation, 1993).  In

order to assess current ASD features, 15 of the 18 participants in the ASD group

completed ADOS (Lord et al., 2000) assessments (please see chapter two for a

description of this tool).  The remaining three participants declined to complete the

ADOS, as they did not feel comfortable being filmed. The three participants who did not complete the ADOS had rigorous diagnoses and scored above the cut-off on the Autism-spectrum Quotient (AQ; Baron-Cohen et al., 2001). All participants who completed the ADOS received a total score ≥7, the defined cut-off for ASD (Lord et al., 2000). All participants completed the AQ questionnaire. Fifteen out of 18 participants with ASD scored above the defined cut-off for ASD on the AQ (total score ≥26; Woodbury-Smith et al., 2005). Only three participants missed this cut-off. However, all three of these participants scored well above the defined ASD cut-off on the ADOS (all ADOS scores among these three participants were ≥ 12). All comparison participants scored below the defined cut-off for ASD.

*Table 4: Participant Characteristics (Means, Standard Deviations and Inferential Statistics).*

|  | Group | | $t$ | $p$ | Cohen's $d$ |
|---|---|---|---|---|---|
|  | ASD (n = 18) | Neurotypical (n = 18) |  |  |  |
| Age (years) | 28.96 (10.28) | 30.43 (14.59) | 0.35 | .730 | 0.12 |
| VIQ | 111.67 (14.66) | 112.28 (10.87) | 0.14 | .888 | 0.05 |
| PIQ | 109.67 (15.75) | 114.50 (10.96) | 1.07 | .293 | 0.36 |
| FSIQ | 112.33 (15.00) | 114.94 (10.50) | 0.61 | .549 | 0.20 |
| AQ Total Score | 33.39 (9.24) | 13.00 (6.22) | 7.77 | <.001 | 2.59 |
| ADOS Social + Communication Score* | 11.93 (2.19) |  |  |  |  |

AQ: Autism-spectrum Quotient; ADOS: Autism Diagnostic Observation Schedule; PIQ = performance IQ; FSIQ = full scale IQ; VIQ = verbal IQ

*Based on 15/18 participants

No participants, in either group, reported using any psychotropic medication or any history of neurological or psychiatric disorders (apart from ASD). The participant groups were closely equated for verbal and non-verbal ability (see Table 4 for participant characteristics). Groups were also closely equated for chronological age.

**Materials and procedures**

**Judgement-of-learning task.** A delayed JOL design was employed, consisting of a learning phase, a JOL phase, and a cued-recall test phase (please see Figure 6 for a graphical representation of the task). The stimuli used during the JOL task were 80 word pairs (160 words) all of which all were concrete nouns. Each word pair was made up of a "cue word", which was used as a cue in both the JOL and recall test phase, and a "target word", which participants were not presented with during the JOL and recall phase. Cue words and target words were matched for word frequency (Kucera & Francis, 1967), as reported in the MRC psycholinguistic database (Coltheart, 1981). The adequacy of this matching was confirmed by a non-significant main effect of word type (cue/target) in an ANOVA, that included word frequency as the dependent variable, $F$ (1, 158) = 1.63, $p$ = .204, $\eta_p^2$ = .01.

Before participants completed the task the entire procedure was explained to them, and participants completed a practice of the task (consisting of five word pairs) before beginning the experimental trials. As such, before learning the word pairs, participants were aware that their memory for each word pair would be tested. The task was run on an LG desktop computer and lasted approximately 25 minutes.

*Learning phase.* Firstly participants completed the learning phase of the task. During the learning phase, participants were presented with the 80 cue-target word pairs. Word pairs were presented to participants individually and participants were

asked to memorise the word pair on the screen, and then to click the mouse whenever they were ready to see the next word-pair. Whenever the mouse was clicked the next word-pair appeared on the screen. As such, participants could take as long as they wanted to learn each word-pair. Word pairs were presented to participants during the learning phase in a fixed, randomised order.

| Learning Phase: | JOL Phase: | Cued- recall Phase: |
|---|---|---|
| "pen-cup" | "pen- ?" <br><br> *"Will you remember the target word at a later point?"* | "pen-?" <br><br> "What was the missing target word?" |
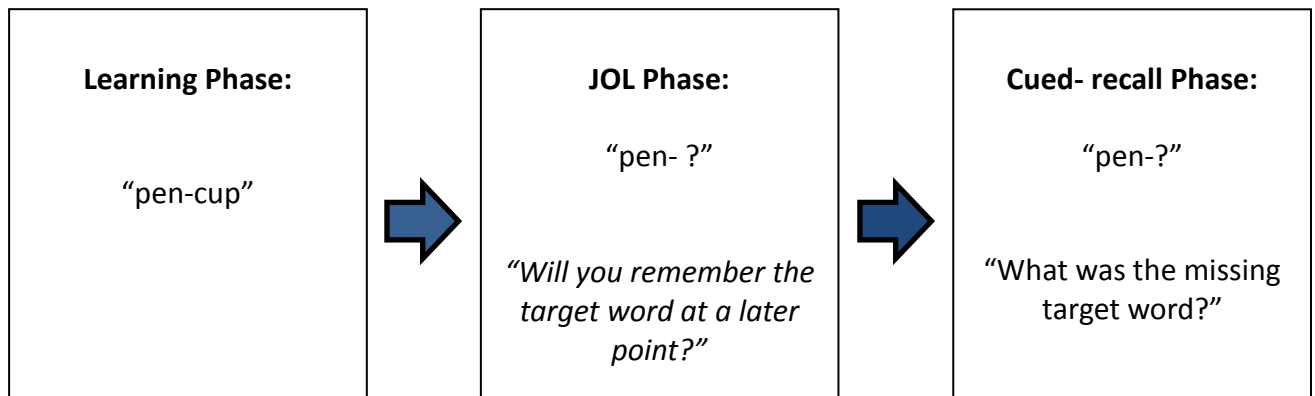
*Figure 6: Graphical representation of the JOL tasks used in Experiment 1.*

***JOL Phase.*** After the learning phase, there was a five minute break. Participants then completed the JOL phase. During the JOL phase participants were individually presented, in a random order, with the cue words alone. For example, if participants learnt the cue-target word pair *"bear-bridge"* during the learning phase then participants JOLs were cued by the presentation of the cue word *"bear - ?",* and asked to judge whether they thought they would be able to recall the correct target word (*"bridge"*). For each cue word, participants were asked to make a JOL (either "Yes" or "No") as to whether they would be able to recall the associated target word, when prompted with the cue word at a later point. Participants made their JOL response by pressing the "Y" key on the keyboard if they thought they would correctly remember the missing target word, and the "N" key if you did not think they would know the missing word.

***Cued-recall phase.*** Immediately after the JOL phase participants completed a cued-recall test. Participants were presented with the cue words again, in a random order, and were asked to recall the missing target word. Participants typed out their recall response, and submitted their response by pressing the "enter" key. Once a recall response was submitted, the next cue word appeared on the screen. There was no time limit on this part of the task.

**Mindreading task.** The animations task (Abell et al., 2000) was administrated, as a measure of mindreading ability. A full description of this task is provided in chapter two, p.59.

**Scoring**

**Judgment-of-learning task.**

***Memory (object-level) performance.*** Participants' basic object-level memory performance was calculated on the JOL task. Recall ability was calculated as the proportion of target words participants correctly recalled during the cued-recall stage. The vast majority of recall responses were unambiguously correct or not correct. However on very few occasions there was some debate as to whether a recall response should be considered correct. On such occasions recall responses were only considered correct if participants had a) recalled a plural of the target word (e.g., if the target word was "*tree",* a recall response of "*trees"* was considered correct), or b) had clearly made an typing error when entering their response (e.g., if the target word was "*tree",* a recall response of "*treew"* was also considered correct. Recall responses that were semantically similar to the target word, but were not the correct target word, were considered incorrect (e.g., if the target word was "*flask",* a recall response of "*thermos"* was considered wrong).

***Metamemory performance.*** Gamma correlation (Goodman & Kruskal, 1954) were calculated to provide an index of overall JOL accuracy. Gamma correlations were calculated based on all JOLs made. Please see chapter two, p.63, for a detailed explanation of Gamma correlations and how they were calculated.

The number of errors made by participants in each group was calculated for two different *types of error in JOL predictions*. The number of *under-confident errors* participants made was calculated as the number of incorrect "No" predictions, in which individuals failed to predict their subsequently successful recall of a target word. The number of *over-confident errors* participants made was calculated as the number of incorrect "Yes" predictions made, in which individuals inaccurately predicted that they would recall a word that they subsequently failed to remember.

**Animations task.** Voice recordings of participants' commentaries were transcribed verbatim. These transcriptions were then scored by a second rater who was blind to the diagnostic status of the participants, according to the scoring criteria outlined in Abell et al., (2000; see also appendix one). Inter-rater reliability for scores across the four animations was excellent, *Cronbach's α* = .98.

## Experiment 1: Results

### Mindreading task

Table 5 shows the means and standard deviations for performance on the animations task. A mixed-model ANOVA was carried out on these data with Group (neurotypical/ASD) entered as the between-subjects variable, and Animation Type (mentalising/goal-directed) entered as the within-subject variable. There was a significant main effect of Group on animations scores, reflecting the fact that participants with ASD performed significantly less well than comparison participants on

the task overall, $F$ (1, 34) = 9.02, $p$ = .005, $\eta_p^2$ = .21.  There was a significant main effect of

Animation Type, indicating that, across both groups, scores were higher in the goal-

directed condition than the mentalising condition, $F$ (1, 34) = 72.82, $p$ < .001. $\eta_p^2$ = .68.

There was no significant Group × Animation Type interaction, $F$ (1, 34) = 0.29, $p$ = .593,

$\eta_p^2$ = .01, suggesting that individuals in the ASD group were impaired at both higher- and

lower-level mindreading, relative to individuals in the neurotypical group.  Therefore, in

line with one-mechanism accounts of mentalising, you would expect to see

corresponding deficits in metamemory monitoring accuracy on the JOL, in individuals

with ASD.


**Judgement-of-learning task**

     **Memory (object-level) performance**.  Group differences in object-level

memory performance were examined using independent-samples $t$-tests (see Table 5

for descriptive and inferential statistics).  These indicated that individuals in the ASD

group recalled significantly fewer target words than comparison participants during the

JOL task.

     **Metamemory performance.**  Group differences in metamemory monitoring

accuracy were also examined (see Table 5 for descriptive and inferential statistics).  An

independent-samples $t$-test indicated that there was no significant difference in gamma

scores between the ASD and neurotypical group.  Thus, not in keeping with predictions,

participants with ASD were not significantly poorer at predicting their own memory

performance than were neurotypical participants, on the JOL task.  One-sampled $t$-tests

indicated that gamma scores were significantly above chance (i.e. significantly greater

than 0) in the ASD and neurotypical groups, all $ts$ > 21.16, all $ps$ < .001.  An additional

analysis was also carried out to investigate whether the significant group difference in object-level recall of target words confounded performance at the meta-level of the task (i.e., JOL accuracy). For the purpose of this analysis, two participants from each group were excluded to create ASD and neurotypical groups that were matched closely for recall ability, $t$ (30) = 1.31, $p$ = .200, $d$ = 0.47. These sub-groups also remained matched for age, VIQ, PIQ, and FSIQ (all $p$s > .52, all $d$s < 0.23). An independent-samples $t$-test indicated that even when groups were equated closely for recall ability, JOL gamma scores were still not significantly different in the ASD group ($M$ = .75, $SD$ = .13) than in the neurotypical group ($M$ = .70, $SD$ = .15), $t$ (30) = 1.15, $p$ = .261, $d$ = 0.36.

Group differences in the specific type of errors participants made on the JOL task were also examined. Independent samples $t$-tests indicated that participants in the ASD group did not make significantly more under-confident JOL errors than participants in the neurotypical group (see Table 5 for statistics). There was no significant group difference in the number of over-confident JOL errors made (see Table 5 for statistics).

**Associations between monitoring ability, and mindreading ability.**

There was no association between metamemory accuracy (Gamma scores) and performance on either the mentalising or goal-directed condition of the animations task, in the ASD group, all $r$s < -.35, all $p$s > .157. In neurotypical participants there was no association between metamemory accuracy and performance on the mentalising condition of the animations task, $r$ = .27, $p$ = .276. However there was a significant negative correlation between metamemory accuracy and performance on the goal-directed condition of the animations task, $r$ = -.47, $p$ = .047. This indicated that, in neurotypical participants, the more accurate participants JOL predictions, the poorer their performance was on the goal-directed condition of the animations task

*Table 5: Means (SDs) and inferential statistics for group differences in performance on the JOL task and animations task.*

| | | Group | | *t* | *p* | Cohen's |
| --- | --- | --- | --- | --- | --- | --- |
| | | ASD | Neurotypical | | | *d* |
| | | (n = 18) | (n = 18) | | | |
| Animations task | Mentalising condition | 3.78 (1.70) | 4.89 (1.71) | 1.96 | .059 | 0.65 |
| | Goal-directed condition | 5.83 (1.50) | 7.22 (0.73) | 3.52 | .001 | 1.18 |
| JOL Task; Object-level memory performance | Proportion of targets recalled | .30 (.26) | .49 (.25) | 2.28 | .029 | 0.74 |
| JOL Task: Meta-memory performance | Gamma scores* | .76 (.13) | .71 (.14) | 1.06 | .295 | 0.37 |
| | Number of over-confident judgments | 7.06 (4.35) | 9.06 (5.75) | 1.18 | .247 | 0.39 |
| | Number of under-confident judgments | 2.28 (3.91) | 2.22 (2.04) | 0.05 | .958 | 0.02 |

*Gamma scores index metamemory monitoring accuracy

**Experiment 1: Discussion**

The results of experiment one suggest that, despite mindreading impairments, adults with ASD appear as accurate as neurotypical individuals on the JOL task. This is not in keeping with predictions that individuals with ASD would demonstrate poorer awareness of their own mental states, nor in keeping with the results reported in chapter three (on the FOK task). Instead, it suggests that adults with ASD do not demonstrate monitoring impairments on a JOL task.

However, it is unclear from the results reported in experiment one what strategy individuals with ASD, and neurotypical individuals, used whilst making their JOL decisions. Despite item-by-item (cue-alone) JOL tasks being well established measures of metacognition, it was speculated that relatively accurate JOL assessments could have been made on the task used in experiment one using an alternative strategy; simply by judging whether one can bring to mind the target word at the time a JOL is made. As such, it is possible that individuals with ASD might have been able to achieve typical levels of performance on the task, despite underlying impairments in metacognition.

For this reason, aggregate JOL tasks may provide a better measure of monitoring ability. Aggregate JOL assessments are thought to be based on how deeply an individual feels they have encoded all the items, and an individual's knowledge of their prior experiences retrieving similar information. Importantly, individuals cannot base aggregate JOLs simply on whether they can bring to mind individual missing target words. If performance in the ASD group in experiment one was driven by such a strategy it is possible that individuals with ASD may demonstrate impairments on aggregate JOL task but not on item-by-item JOL tasks.

As such, to gain a better understanding of whether typical JOL accuracy in individuals with ASD was the result of using this alternative strategy, 12 participants with ASD and 10 neurotypical participants (all of whom participated in the first experiment) were brought back to complete an aggregate JOL task.

## Experiment 2

Experiment two reports the results of a preliminary study, exploring aggregate JOL accuracy in individuals with ASD. During this task participants were presented with two lists of words to remember, one list of high-frequency words and a list of low-frequency words. Participants were given 8 minutes to memorise as many words from both lists as possible, and could choose how much of this time they spent learning each word list. Metamemory monitoring accuracy was assessed by asking participants to make aggregate JOL predictions for each list, and participants' aggregate JOLs were then compared to their actual recognition performance on the task. The design of this study also allowed us to investigate whether JOL assessments made by individuals with ASD were influenced by specific aspects of the learning material judgements were made on. The effect of word frequency on memory has been extensively studied, and studies have typically found that individuals remember high-frequency words better than low-frequency word, but are more likely to make false-alarms for low-frequency words than high-frequency words (e.g., Gardiner & Java, 1990; Reder et al., 2000). Studies have also shown that individuals with ASD also demonstrate typical word frequency effects in recognition memory (Bowler, Gardiner, & Grice, 2000). This experiment is novel in that it explored whether individuals with ASD, and neurotypical individuals, are able to make aggregate JOL assessments that reflect this pattern, predicting better memory for the low-frequency word list than the high-frequency word list.

Additionally, the design of experiment two meant it was possible to investigate whether individuals with ASD demonstrate specific strategy use during a JOL task. Some studies suggest that individuals with ASD may demonstrate difficulties spontaneously implementing memory strategies (e.g., Bowler, Gaigg, & Gardiner, 2008; Gaigg, Gardiner, & Bowler, 2008). Given findings that individuals typically find low-frequency words harder to remember than high-frequency words (e.g., Gardiner & Java, 1990; Reder et al., 2000), the study explored whether individuals, in both the ASD and neurotypical groups, regulated their learning on the task and spent significantly more time learning the list of low-frequency words than the list of high-frequency words. As predicted in the introduction of this chapter, it was expected that individuals with ASD would demonstrate impairments in monitoring accuracy on the aggregate JOL task, and would make less accurate JOL predictions, relative to neurotypical participants. It was also predicted that individuals with ASD would not regulate their study behaviour during the learning phase of the task, to the extent that neurotypical participants would. Whilst it was expected that neurotypical participants would spend more time learning the list of low-frequency words, relative to high frequency words, it was predicted that individuals with ASD would spend an equal amount of time memorising both word lists.

## Experiment 2: Method

### Participants

Twelve adults with ASD (8 males, 4 females) and 10 neurotypical comparison adults (7 males, 3 females) took part, all of whom had taken part in the previous experiment (please see Table 6 for details of participant characteristics). Participant groups were closely equated for verbal ability, non-verbal ability and chronological age.

*Table 6: Participant Characteristics (Means, Standard Deviations and Inferential Statistics).*

| | Group | | | | |
| --- | --- | --- | --- | --- | --- |
| | ASD (n = 12) | Neurotypical (n = 10) | *t* | *p* | Cohen's *d* |
| Age (years) | 27.54 (8.83) | 32.27 (15.15) | 0.91 | .372 | 0.38 |
| VIQ | 112.25 (13.08) | 113.80 (13.68) | 0.27 | .789 | 0.12 |
| PIQ | 108.58 (16.37) | 113.70 (10.14) | 0.86 | .401 | 0.38 |
| FSIQ | 112.08 (13.96) | 115.60 (12.24) | 0.62 | .541 | 0.27 |
| AQ Total Score | 32.08 (9.26) | 13.80 (6.39) | 5.28 | <.001 | 2.30 |
| ADOS Social + Communication Score* | 11.60 | | | | |

*Based on 10/12 participants

**Material and procedure.**

**Aggregate JOL task.** An aggregate JOL design was employed, consisting of a learning phase, a JOL phase, and a recognition test phase. The stimuli used during the learning phase of the task were two lists of 80 words (160 words) all of which all were concrete nouns. List one consisted of 80 low-frequency word pairs, and list two consisted of 80 high-frequency word pairs. A *t*-test confirmed that list one contained words with significantly lower word frequencies than list two, $t$ (158) = 12.44, $p < .001$, $d = 1.97$. Word lists were matched for syllable length, as reported in the MRC psycholinguistic database (Coltheart, 1981), $t$ (158) = 0.54, $p = .581$, $d = 0.10$. Eighty lure words were also used during the recognition phase of the task. Lure words

consisted of 40 high frequency lure words, and 40 low frequency lure words. Again, *t*-tests confirmed that high frequency lure words significantly differed from low frequency lure words in word frequency *t* (78) = 7.03, *p* < .001 *d* = 1.57, but not in syllable length, *t* (78) = 0.30 *p* = .764 *d* = 0.07. Before beginning the actual experiment, the entire procedure was explained to participants and participants completed a practise of the task (consisting of two words lists, each three words long). The task was run on an LG desktop computer and lasted approximately 25 minutes.
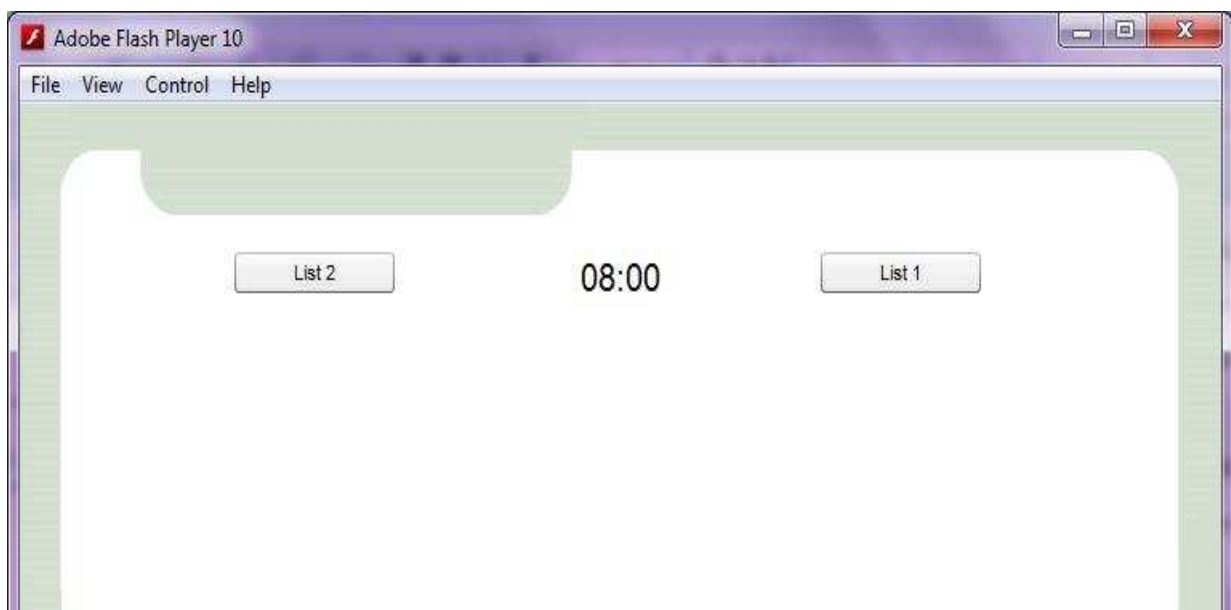


*Figure 7: Example of the start screen used in Experiment 2.*

***Learning phase.*** Firstly participants completed the learning phase of the task. During the learning phase, participants were presented with the two lists of words; one list of low-frequency words (List 1) and one list of high-frequency words (List 2). Participants were told that they had eight minutes to remember as many words from both lists as possible. The learning phase started once participants chose which list to view first, and clicked on the corresponding button (see Figure 7). Whether List 1 or List 2 appeared in the right hand side of the screen was counterbalanced. A countdown

timer appeared on the screen throughout the learning phase. This allowed participants to keep track of how much time they had spent learning each of the lists, and how much time remained.

    *JOL Phase.* After the learning phase participants then completed the JOL phase of the task. During this phase participants were asked to make three aggregate JOL assessments. Firstly participants were told that in total they had been presented with 160 words (80 on each list), and were asked to decide how many they thought they would be able to correctly recognise in total. Secondly participants were asked how many words they thought they would be able to recognise from list one (the LF word list). Finally, participants were asked to judge how many words they thought they would be able to recognise from list two (the HF word list).

    *Recognition phase.* In the final phase of the task participants' memory for words was tested using a recognition task. Participants were individually shown words, and were asked to decide whether each word had appeared in the previous lists ("Yes" or "No"). Participants were presented with all 160 words as well as the 80 novel lure words in a random order.

    **Mindreading task.** The animations task (Abell et al., 2000) was administrated, as a measure of mindreading ability. A full description of this task is provided in chapter two, p.59.

**Scoring**

    **Aggregate JOL task.**

    *Object-level performance.* Corrected hit rates were calculated to assess recognition ability on the task. Corrected hit rates were calculated using the formula H-FA, where H represents hit rate (the proportion of word participants *correctly*

identifying as having appeared previously on the lists) and FA represents false alarm rate (the proportion of new word participants *incorrectly* identifying having appeared previously on the lists).

Corrected hit rates were calculated for (a) recognition ability overall on the task, (b) recognition ability for words in list one (low-frequency words), and (c) recognition ability for words in list two (high-frequency words).  For all corrected hit rates, a false alarm rate based on all (80) lure words was used.  However, for the purpose of exploring group differences in the type of words participants made false alarms on, the proportion of low-frequency lure words participants made false alarms for was also calculated, as well as the proportion of high-frequency lure words participants made false alarms for.

***Metamemory monitoring accuracy.***  To assess the accuracy of participants' aggregate JOL assessments three difference scores were calculated; the difference between participants' aggregate JOL assessment for all 160 words (words from both lists) and their actual recognition performance for all 160 words was calculated.  The difference between participants aggregate JOL assessment for list one and their recognition ability (i.e.  corrected hit rate) for words in list one was calculated, as well as the difference between participants  JOL assessment for words in list two and their actual recognition ability (i.e.  corrected hit rate) for words from list two.

***Metamemory control processes.***  Additionally, the average time participants in both the ASD and neurotypical group spent learning list one (LF words) and list two (HF words) was calculated.  This allowed us to assess the prediction that, if participants in either the ASD or neurotypical group were employing control processes, they would spend more time learning list two (high-frequency words) than list one (low-frequency words).

**Mindreading task.** Voice recordings of participants' commentaries were transcribed verbatim and scored. These transcriptions were then scored by a rater who was blind to the diagnostic status of the participants, according to the scoring criteria outlined in chapter two. Inter-rater reliability for scores across the four animations was excellent, *Cronbach's α* = .98

## Experiment 2: Results

### Aggregate JOL task

**Memory (object-level) performance.** Table 7 shows the means and standard deviations for corrected hit rates on the JOL task, for both high-frequency words (List two) and low-frequency words (list one), and performance overall (across both lists). To explore group differences in recognition ability a mixed-model ANOVA was carried out on these data with Group (neurotypical/ASD) entered as the between-subjects variable, and List Type (low-frequency/high-frequency) entered as the within-subject variable. There was a significant main effect of List Type on recognition ability (corrected hit rate), reflecting the fact that across participant groups, participants recognised significantly more words from list one (low-frequency words) than words from list two (high-frequency words) $F(1, 20) = 11.48$, $p = .003$, $\eta_p^2 = .37$. There was no significant main effect of Group, indicating that corrected hit rates in the ASD group did not significantly differ from those in the neurotypical group, $F(1, 20) = 0.08$, $p = .778$, $\eta_p^2 = .004$. There was no significant Group × List Type interaction, $F(1, 20) = 0.10$, $p = .751$ $\eta_p^2 = .005$.

Analysis was also carried out to investigate whether individuals with ASD and neurotypical individuals demonstrated typical "mirror effects" on the task. Table 7 also shows the means and standard deviations for hit rates on the JOL task, for both high-

frequency words and low-frequency words. A mixed-model ANOVA was carried out on these data with Group (neurotypical/ASD) entered as the between-subjects variable, and List Type (low-frequency/high-frequency) entered as the within-subject variable. As expected, there was a significant main effect of List Type on hit rate scores, reflecting the fact that across participant groups, participants recognised significantly more low-frequency words (list one) than high-frequency words (list two), $F$ (1, 20) = 11.48, $p$ = .003 , $\eta_p^2$ = .37. There was no significant main effect of Group, indicating that individuals with ASD made as many hits as neurotypical individuals, $F$ (1, 20) = 1.24, $p$ = .280, $\eta_p^2$ = .06. There was no significant Group × List Type interaction, $F$ (1, 20) = 0.10, $p$ = .751 $\eta_p^2$ = .005. Overall, these results indicate that individuals with ASD showed a similar pattern of recognition ability relative to neurotypical individuals, recognising significantly more low-frequency words than high-frequency words.

Table 7 also shows the means and standard deviations for false alarm rates on the recognition phase of the JOL task, for both high-frequency words and low-frequency words. A mixed-model ANOVA was carried out on these data with Group (neurotypical/ASD) entered as the between-subjects variable, and Word Type (low-frequency/high-frequency) entered as the within-subject variable. There was a significant main effect of Word Type on false alarm rate, reflecting the fact that, as expected, across participant groups, participants made significantly more false alarms for high-frequency lure words than low frequency lure words, $F$ (1, 20) = 10.15, $p$ = .005, $\eta_p^2$ = .34. There was no significant main effect of Group, indicating that individuals with ASD made as many false alarms on the task as neurotypical individuals, $F$ (1, 20) = 2.51, $p$ = .128, $\eta_p^2$ = .112. There was no significant Group × Word Type interaction, $F$ (1, 20) = 0.02, $p$ = .897 $\eta_p^2$ = .001.

*Table 7: Means (SDs) and inferential statistics for group differences in "object" level performance on the aggregate JOL task.*

| | | Group | | *t* | *p* | Cohen's *d* |
|---|---|---|---|---|---|---|
| | | ASD (n = 12) | Neurotypical (n = 10) | | | |
| Hit rates | Hit rate for low-frequency words (list one) | .69 (.14) | .64 (.07) | 0.97 | .345 | 0.45 |
| | Hit rate for high-frequency words (list two) | .57 (.21) | .49 (.17) | 0.89 | .383 | 0.42 |
| | Hit rate overall (performance on both lists) | .63 (.15) | .57 (.09) | 1.11 | .280 | 0.49 |
| False alarm rates | False alarms for low-frequency words | .20 (.14) | .13 (.07) | 1.49 | .153 | 0.63 |
| | False alarms for high-frequency words | .29 (.15) | .21 (.13) | 1.31 | .205 | 0.57 |
| | Overall false alarm rate (for all words) | .25 (.13) | .17 (.09) | 1.59 | .129 | 0.72 |
| Corrected hit rates* | Corrected hit rate for low-frequency words (list one) | .44 (.19) | .47 (.11) | 0.43 | .673 | 0.19 |
| | Corrected hit rate for high-frequency words (list two) | .32 (.18) | .32 (.14) | 0.04 | .969 | 0.00 |
| | Corrected hit rate overall (performance on both lists) | .38 (.15) | .40 (.10) | 0.29 | .778 | 0.16 |

*\* Note: All corrected hit rates were calculated using the overall false alarm rate, based on all (80) lure words.*

Overall, these results indicate that individuals with ASD showed a similar level and pattern of recognition ability relative to neurotypical individuals on the task. Corrected hits rates (recognition ability) were similar in both groups (see Table 7). Additionally, both groups demonstrated typical "mirror effects" on the task, showing better recognition for low-frequency words, and making more false alarms on high-frequency words (see Table 7).

**Metamemory monitoring performance.**

*Table 8: Means (SDs) and inferential statistics for group differences in "meta" level performance on the aggregate JOL task.*

|  |  | Group | | | | |
|---|---|---|---|---|---|---|
|  |  | ASD (n = 12) | Neurotypical (n = 10) | *t* | *p* | Cohen's *d* |
| JOL Predictions | Aggregate JOL for low-frequency words (list one) | 29.58 (10.97) | 33.60 (14.02) | 0.76 | .459 | 0.32 |
|  | Aggregate JOL for high-frequency words (list two) | 27.50 (13.06) | 24.80 (13.47) | 0.48 | .639 | 0.20 |
|  | Aggregate JOL overall | 62.92 (25.45) | 55.00 (28.78) | 0.69 | .501 | 0.34 |
| Prediction Accuracy | Difference score for low-frequency words (list one) | .20 (.16) | .14 (.13) | 0.97 | .346 | 0.41 |
|  | Difference score for high-frequency words (list two) | .19 (.13) | .14 (.08) | 0.99 | .332 | 0.46 |
|  | Difference score overall | .20 (.15) | .16 (09) | 0.83 | .418 | 0.32 |

***JOL Predictions.*** Table 8 shows the means and standard deviations for participants JOL predictions concerning how many words they thought they would recognise from list one (low-frequency words) and list two (high-frequency words). A mixed-model ANOVA was carried out on these data with Group (neurotypical/ASD) entered as the between-subjects variable, and List Type (low-frequency/high-frequency) entered as the within-subject variable. There was a significant main effect of List Type on JOL prediction, reflecting the fact that across participant groups, participants predicted they would recognise more words from list one (low-frequency words) than from list two (high-frequency words), $F(1, 20) = 6.85$, $p = .017$, $\eta_p^2 = .26$. There was no significant main effect of Group, indicating that across both judgements individuals with ASD made as high JOL assessments as neurotypical individuals, $F(1, 20) = 0.02$, $p = .898$, $\eta_p^2 = <.01$. There was no significant Group × List Type interaction, $F(1, 20) = 2.61$, $p = .122$ $\eta_p^2 = .115$. These results indicate that both groups correctly predicted they would recognise more words from the low-frequency list than from the high-frequency list.

***Prediction Accuracy.*** Group differences in the *accuracy* of participants overall aggregate JOL assessment were examined using independent-samples *t*-tests (see Table 8 for descriptive and inferential statistics). This indicated that the difference between participants' overall aggregate JOL prediction and their overall recognition performance was not significantly different in the ASD group relative to the neurotypical group.

Finally, group differences in the *accuracy* of participants' aggregate JOL assessments for how many low-frequency words they would remember from List one and how many high-frequency words they would remember from List two were examined using independent-samples *t*-tests (see Table 8 for descriptive and inferential

statistics).  These indicated that the difference between participants' JOL predictions (for either the high frequency or low frequency list) and their actual recognition performance for each list was not significantly different in the ASD group relative to the neurotypical group.

**Metamemory control processes.**  Group difference in the average time individuals spent learning list one and list two were explored.  A mixed-model ANOVA was carried out on these data with Group (neurotypical/ASD) entered as the between-subjects variable, and List Type (low-frequency/high-frequency) entered as the within-subject variable.  However, there was no significant main effect of list type, no significant main effect of group, and no Group × List Type interaction, all $ps \geq .263$, $\eta_p^2$ $\geq .06$.  These results indicate that neither group spent significantly more time learning one list more than another list, and on average groups spent half their time learning list one (*mean* = 4.04 minutes)  and half their time learning list two (*mean* = 3.96 minutes).

**Mindreading task.**

Table 9 shows the means and standard deviations for performance on the animations task.  A mixed-model ANOVA was carried out on these data with Group (neurotypical/ASD) entered as the between-subjects variable, and Animation Type (mentalising/goal-directed) entered as the within-subject variable.  There was no significant main effect of Group on animations scores, reflecting the fact that participants with ASD did not performed significantly less well than comparison participants on the task overall, $F (1, 20) = 1.61$, $p = .219$, $\eta_p^2 = .07$.  There was a significant main effect of Animation Type, indicating that, across both groups, scores were higher in the goal-directed condition than the mentalising condition, $F (1, 20) = 76.14$, $p < .001$. $\eta_p^2 = .79$.  There was no significant Group × Animation Type interaction,

$F$ (1, 20) = 2.82, $p$ = .108, $\eta_p^2$ = .12. These results suggest that, on the animations task, individuals with ASD did not show mentalising impairments relative to the neurotypical individuals.

*Table 9: Means (SDs) and inferential statistics for group differences in performance on the animations task.*

|  | Group | | $t$ | $p$ | Cohen's $d$ |
|---|---|---|---|---|---|
|  | ASD (n = 12) | Neurotypical (n = 10) |  |  |  |
| Mentalising condition | 3.67 (1.72) | 3.80 (1.40) | 0.20 | .846 | 0.04 |
| Goal-directed condition | 5.83 (1.47) | 7.00 (0.67) | 2.32 | .031 | 1.02 |

**Associations between monitoring ability and mindreading ability**

Exploratory correlations were run to investigate the association between monitoring accuracy on the JOL task (the difference between participants overall JOL and their overall recognition ability) and performance on the animations task. Overall there was no significant correlation between performance in either the mentalising or goal-directed condition of the animations task and accuracy on the JOL task, in either the ASD or neurotypical group, all rs < -.22, all $p$s > .547. It should be noted that, given the small sample size of participants in each group, some caution should be taken when making conclusions regarding these correlations.

<div align="center">

**Experiment 2: Discussion**

</div>

The results of experiment two indicate that, as expected, both groups demonstrated clear "mirror effects" on that task (recognising more low-frequency

words than high-frequency words during the recognition phase of the task, and making more false alarms on high-frequency lure words than low-frequency lures). Regarding meta-level performance on the task, individuals with ASD, and comparison individuals, both successfully predicted that they would remember more words from the low-frequency word list than the high-frequency word list. Additionally, the aggregate JOL assessments made by individuals with ASD were as accurate at predicting their future memory performance on the task as those made by neurotypical individuals. The results from experiment two were in keeping with the results reported in experiment one, and suggest that adults with ASD do not show impaired metamemory monitoring accuracy on JOL tasks.

A number of methodological issues with experiment two should be considered. Firstly, a limited number of participants took part in this study (12 ASD participants, 10 neurotypical participants). As such, there is some concern that the study was not adequately powered to detect group differences on the aggregate JOL task. When looking at group differences in the aggregate JOL predictions participants made for list one (low frequency words) and list two (high frequency word) no significant group by list type interaction was found. However, participants with ASD did predict they would remember fewer words from the low frequency list (list one) than neurotypical participants predicted, and predicted they would remember more words from the high frequency list (list two), than neurotypical participants predicted. This pattern of results does suggest that participants with ASD were less accurate at predicting their future memory performance, and it is possible that this interaction effect was not significant ($p = .122$) due to a lack of power.

Additionally, in experiment two individuals with ASD did not appear to show impairments on the animation task. This suggests that, in this sample of ASD

participants, mindreading ability was not impaired relative to comparison participants. This presents a theoretical problem, as (according to the one-mechanism theory) one should only expect to find metacognitive impairments in individuals with ASD who only demonstrate mindreading impairments.

Additionally, whilst experiments one and two suggest JOL accuracy is intact in adults with ASD, impairments in JOL accuracy maybe developmental in nature and only apparent in children with ASD. Experiment three addressed this issue, and explored JOL accuracy in children with ASD, using two JOL paradigms. Children with and without ASD were asked to complete both a cue-alone JOL task and a cue-target JOL task. Additionally, to provide a more sensitive measure of mindreading ability, two mindreading tasks were employed; the Animations Task (Abell et al., 2000) and the Strange Stories Task (Happé, 1994). This allowed us to assess mindreading more comprehensively than in the two previous studies. The results of experiment one and two suggest that JOL accuracy is typical in adults with ASD. However, methodological concerns with both studies might explain intact JOL accuracy in the individuals with ASD. As such predictions regarding JOL accuracy in children with ASD, on both the cue-alone and cue-target JOL tasks, were equivocal.

### Experiment 3: Method

**Participants**

Twenty-two children with ASD and 21 neurotypical comparison children took part in this experiment, after their parents had given written, informed consent. All participants completed the JOL tasks and the animations task. However, due to restrictions that occurred during data collection three participants did not complete the strange stories task (two participants with ASD, one neurotypical participant).

Participants in the ASD group had all received formal diagnoses of autistic disorder (*n* = 17) or Asperger's disorder (*n* = 5). In all but one case, participants with ASD scored above the defined cut-off for ASD on the SRS (total score ≥ 60; Constantino et al., 2003). The remaining participant scored 56 on the SRS, which is just below the conventional ASD cut-off of 60. This participant had a formal diagnosis of Autism Spectrum Disorder.

*Table 10: Participant Characteristics (Means, Standard Deviations and Inferential Statistics).*

| | Group | | | | |
| --- | --- | --- | --- | --- | --- |
| | ASD (n = 22) | Neurotypical (n = 21) | *t* | *p* | Cohen's *d* |
| Age (years) | 13.70 (1.45) | 13.21 (1.18) | 1.21 | .234 | 0.37 |
| VIQ | 100.68 (15.48) | 98.76 (12.54) | 0.45 | .658 | 0.14 |
| PIQ | 101.41 (14.80) | 102.86 (14.11) | 0.33 | .744 | 0.10 |
| FSIQ | 100.95 (14.06) | 101.14 (13.68) | 0.04 | .965 | 0.01 |
| SRS Total Score | 83.14 (9.93) | 47.29 (11.66) | 10.87 | <.001 | 3.31 |

SRS: Social Responsiveness Scale (Constantino et al., 2003); VIQ = verbal IQ; PIQ = performance IQ; FSIQ = full scale IQ.

Parents of the neurotypical children also completed the SRS. All but four participants in the neurotypical group scored below the defined cut-off for ASD. The remaining participants' SRS scores ranged from 60 to 73. To ensure that including these participants in the overall sample did not affect the results of the study all analyses in the paper were re-run, excluding these four participants and the participants with ASD who scored below the recommended cut-off on the SRS. After

removing these participants, none of the experimental results reported in the paper changed, from being previous significant to non-significant. Both groups were equated closely for VIQ, PIQ, FSIQ, and chronological age. Participant characteristics are presented in Table 10.

**Materials and procedures**

**Judgment-of-learning tasks.** Two sets of 22 word-pairs (44 words) were used as stimuli for the JOL tasks. Both sets were matched for mean syllable length and word frequency (Kucera & Francis, 1967), as reported in the MCR psycholinguistic database (Coltheart, 1981). To check that the words used in each set were adequately matched, a multivariate analysis of syllable length and word frequency across both sets was carried out. There was no main effect of set, as established by Wilks' Lambda criterion, $F$ (2, 85) = .152, $p$ = .859, $\eta_p^2$ = .004. Participants were tested individually on both tasks during two separate testing sessions (please see Figure 8 for a graphical representation of both JOL tasks). To control for any tasks effects, the order participants completed each JOL task was counterbalanced. Before completing either task participants completed a practice block, consisting of five word pairs.

*Cue-alone JOL Task.* The procedure employed during the judgement-of-learning task used a delayed JOL design, consisting of a learning phase, a JOL phase and a cued-recall test phase. The task was run on a Sony VAIO laptop, and lasted approximately 15-20 minutes. During the learning phase participants were individually presented with 22 cue-target word pairs for 8 seconds each. Participants were told that their memory for each word pair would be tested at a later point, with the presentation of the *cue word alone.* After the learning phase participants then completed the JOL phase of the task, in which they were presented, in a random order, with cue words alone (i.e., if

participants learnt the cue-target pair "bear-bridge" then the JOL for this word pair was cued by the presentation of "bear - ?"). The only previous study of JOL accuracy (Wojcik et al., 2014), alongside Experiment 1, asked participants to make dichotomous (Yes/No) JOL assessments. In both studies individuals with ASD did not demonstrate impairments in monitoring accuracy, relative to neurotypical individuals. However, it is possible that categorical judgements might not provide the variation necessary to observe group difference in JOL accuracy. As such, in this experiment participants were presented with each cue word individually for 5 seconds, and were asked to make a JOL on a scale of 1-5. It was explained to participants that a JOL of 1 indicated that they thought they would definitely not be able to remember the missing target word, and a JOL of 5 indicated they thought they would definitely be able to remember the missing target word. Immediately after the JOL phase participants completed a cued-recall test. Participants were presented again with cue words alone, in a random order, and were asked to recall the missing target word. Participants were not limited in the amount of time they had to recall the target word for a given cue word.

  ***Cue-target JOL Task.*** The cue-target JOL tasks followed the same procedure as the cue-alone JOL task. However, during the JOL phase instead of being presented with cue-words alone, participants were presented with the **entire word pairs** again (i.e., if participants learnt the cue-target pair "bear-bridge" then the JOL for this word pair was cued by the presentation of "bear - bridge"). Again, participants were asked to make a JOL on a scale of 1 to 5. Apart from this difference, the procedure for each JOL task was exactly the same.

  **Mindreading tasks.** Two measures of mindreading ability were administered: the animations task (Abell et al., 2000) and the strange stories task (Happé, 1996). A full description of both these tasks is provided in chapter two, p.59-62.
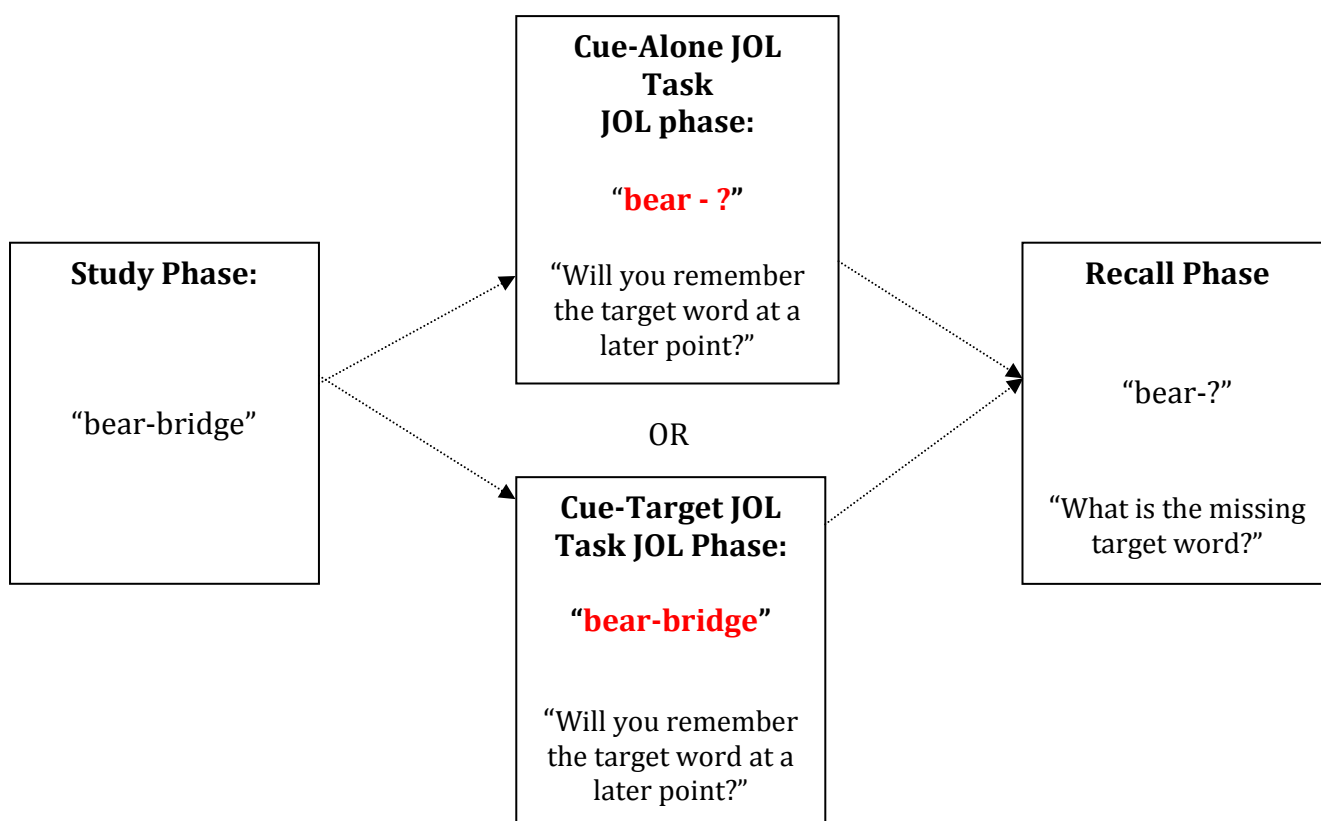
*Figure 8: Graphical representations of the cue-alone and cue-target JOL tasks used in Experiment 3.*

**Scoring**

**Judgment-of-learning task.**

*Memory (object-level) performance.* Participants' basic object-level memory performance was calculated on the JOL task. Recall ability was calculated as the proportion of target words participants correctly recalled during the cued-recall-stage. Target words were considered correct if participants had a) correctly recalled the target word, b) participants recalled a plural of the target word (e.g, if the target word was "*cup*", a recall response of "*cups*" was considered correct), or c) participants had clearly made an typing error when entering their response (e.g., if the target word was "*cup*", a

recall response of "*cupz*" was also considered correct. Recall responses that were semantically similar to the target word, but were not the correct target word, were considered incorrect (e.g., if the target word was "*flask*", a recall response of "*thermos*" was considered incorrect).

      ***Metamemory performance.*** Gamma correlations (Goodman & Kruskal, 1954) were calculated to provide an index of overall JOL accuracy. Gamma correlations were calculated based on all JOLs made. Please see chapter two, p.63, for a detailed explanation of Gamma correlations, and how they were calculated.

      **Animations task.** Voice recordings of participants' commentaries were transcribed verbatim. These transcriptions were then scored by a second rater who was blind to the diagnostic status of the participants, according to the scoring criteria outlined in chapter two. Inter-rater reliability for scores across the four animations was excellent, Cronbach's α = .93.

      **Strange Stories Task.** Voice recordings of participants' answers to each question were transcribed verbatim. These transcriptions were also then scored by a second rater who was blind to the diagnostic status of each participant, according to the scoring criteria outlined in chapter two. Again, inter-rater reliability for scores across the four stories was excellent, *Cronbach's α* = .92.

### Experiment 3: Results

**Judgment-of-learning tasks**

      **Memory (object-level) performance.** Figure 9 shows the means and standard deviations for recall performance on the cue-alone JOL task, and recall performance on the cue-target JOL task (see also Table 11). A mixed-model ANOVA was carried out on these data with Group (neurotypical/ASD) entered as the between-subjects variable,

and JOL Type (stimuli-repose/cue-alone) entered as the within-subject variable.  There

was a significant main effect of JOL Type on recall ability, reflecting the fact that, across

participant groups, children recalled significantly more target words in the cue-target

JOL task than in the cue-alone JOL task, $F$ (1, 41) = 31.14, $p$ < .001, $\eta_p^2$ = .43.  Better recall

performance in the cue-target JOL task (*Mean* = .49, *SD* = .21) than in the cue-alone JOL

task (*Mean* = .34, *SD* = .16) was expected, given that on the cue-target JOL task children

were presented with the target words twice, compared to only once during the cue-

alone JOL task.  There was no significant main effect of Group, $F$ (1, 41) = 0.01, $p$ = .917,

$\eta_p^2$ < .001, nor Group × JOL type interaction, $F$ (1, 41) = 1.08, $p$ = .305 $\eta_p^2$ = .03.  Thus,

children with ASD demonstrated similar levels and patterns of recall to the neurotypical

children.
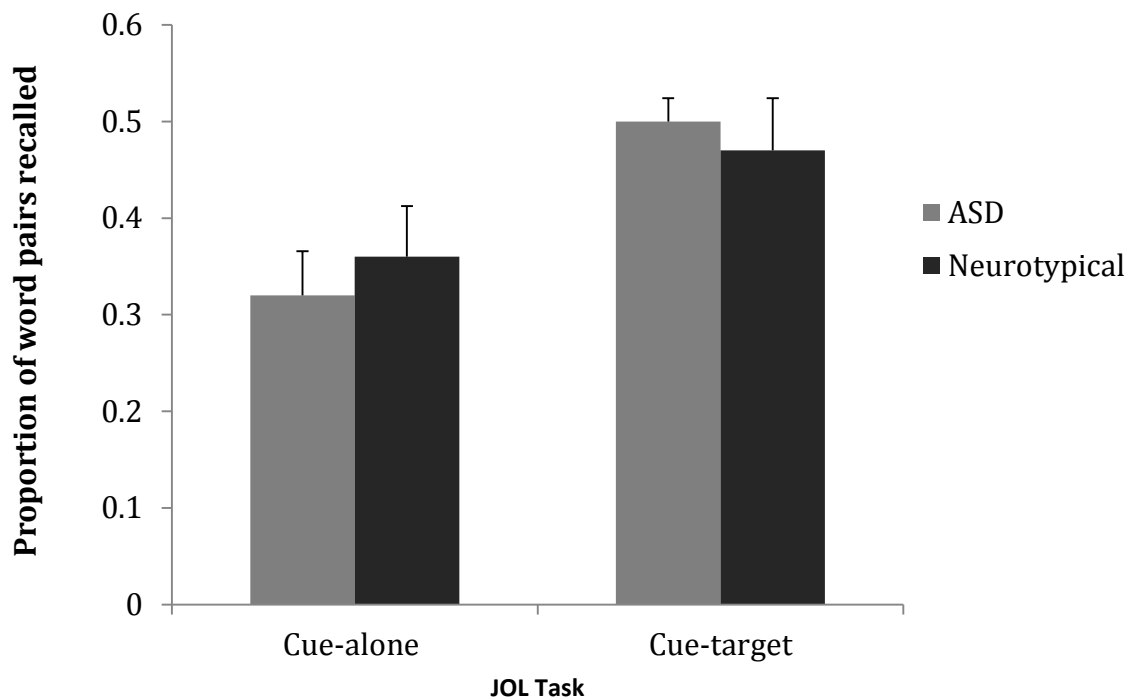
Figure 9*: "Object level" performance on the cue*-alone and cue-target JOL tasks, in

the ASD and neurotypical group. *Error bars represent standard error of the mean.*

**Metamemory performance.** Figure 10 also shows the means and standard

deviations for Gamma correlations on the cue-alone JOL task, and recall performance on

the cue-target JOL task, for both ASD and neurotypical participants (see also Table 11).

A mixed-model ANOVA was carried out on these data with Group (neurotypical/ASD)

entered as the between-subjects variable, and JOL Type (stimuli-repose/cue-alone)

entered as the within-subject variable. There was a significant main effect of JOL type

on gamma scores, reflecting the fact that children in both groups had significantly lower

gamma scores (i.e., made significantly less accurate JOL assessments) on the cue-target

JOL task than on the cue-alone JOL task, $F(1, 41) = 42.62$, $p < .001$, $\eta_p^2 = .51$. There was

no significant main effect of group, indicating that children with ASD did not have lower

gamma scores across both JOL tasks, relative to neurotypical children $F(1, 41) = 0.46$, $p$

$= .504$, $\eta_p^2 = .01$. There was also no significant Group JOL× Type interaction, $F(1, 41) =$

$0.14$, $p = .706$, $\eta_p^2 < .01$, suggesting that children in the ASD group did not demonstrate

impairments in JOL accuracy specifically on one of the JOL tasks.
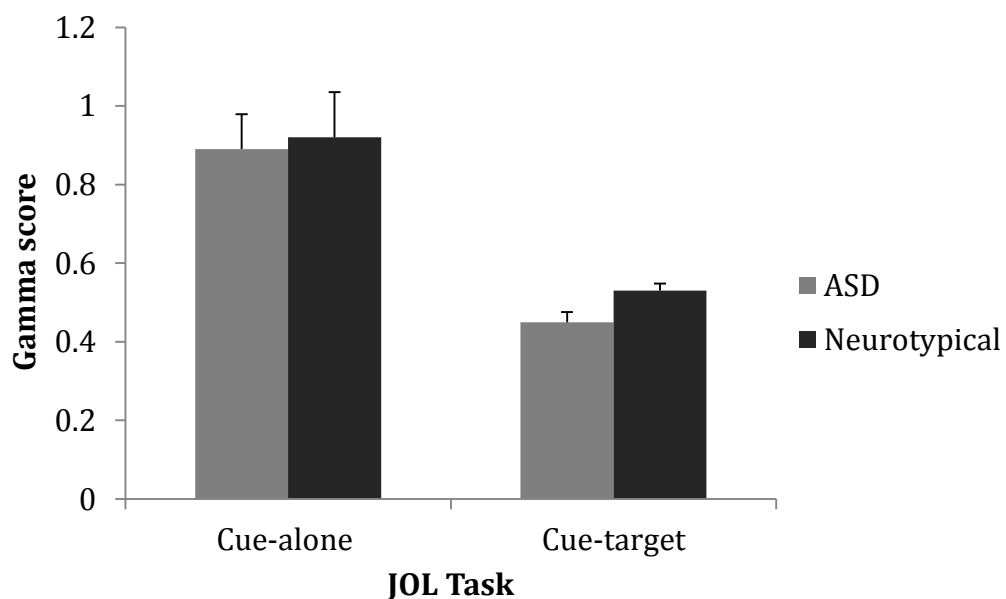
*Figure 10: "Meta level" performance (gamma scores) on the cue-alone and cue-target JOL tasks, in the ASD and neurotypical group.  Error bars represent standard error of the mean.*

**Mindreading tasks**

**Animations task.**  Table 11 shows the means and standard deviations for performance on the animations task.  A mixed-model ANOVA was carried out on these data with Group (neurotypical/ASD) entered as the between-subjects variable, and Animation Type (mentalising/goal-directed) entered as the within-subject variable.  There was no significant main effect of Group on animations scores, reflecting the fact that children with ASD did not perform less well than neurotypical participants on the task overall, $F(1, 41) = 0.04$ , $p = .851$ , $\eta_p^2 = .001$.  There was a significant main effect of Animation Type, indicating that, across both groups, scores were higher in the goal-directed condition than the mentalising condition, $F(1, 41) = 100.57$, $p < .001$ $\eta_p^2 = .76$.  There was no significant Group × Animation Type interaction, $F(1, 40) = .02$, $p = .888$, $\eta_p^2 <.001$.  These results do not support predictions and suggest that the ASD group did not show diminished mindreading performance in either condition of the task, relative to the neurotypical group.

**Strange Stories task.**  Table 11 shows the means and standard deviations for performance on the strange stories task.  A mixed-model ANOVA was carried out on these data with Group (neurotypical/ASD) entered as the between-subjects variable, and Story Type (mentalising/physical) entered as the within-subject variable.  There was a significant main effect of Story Type, indicating that, across both groups, scores were higher in the physical condition than the mentalising condition, $F(1, 37) = 23.22$, $p < .001$ $\eta_p^2 = .39$.  There was no significant main effect of Group on strange stories scores, $F(1, 37) = 2.33$, $p = .136$, $\eta_p^2 = .06$.  However, there was a marginally significant Group ×

Story Type interaction, $F$ (1, 37) = 3.83, $p$ = .058, $\eta^2_p$= .09.  Independent samples $t$-tests

indicated that the ASD group performed significantly worse that neurotypical

participants in the mentalising condition of the task, but did not significantly differ from

neurotypical participants in the physical condition (see Table 11 for descriptive and

inferential statistics).  Paired samples $t$-tests indicated that participants in the ASD

group performed significantly worse in the mentalising condition of the task (*Mean* =

1.63, *SD* = 1.26) than in the goal-directed condition of the task (*Mean* = 3.05, *SD* = 0.71),

$t$ (18) =5.20, $p$ < .001, $d$ = 1.39.  However, in the neurotypical group, there was no

difference in performance on the mentalising (*Mean* = 2.45, *SD* = 1.28) and goal-directed

(*Mean* = 3.05, *SD* = 0.89) conditions of the task, $t$ (19) =, 1.88, $p$ = .076, $d$ = 0.54.  These

results suggest that, as predicted, the ASD group demonstrated impairments specifically

in the mindreading condition of the strange stories task.

**Associations between mindreading ability and metacognitive monitoring ability.**

A series of correlational analyses was carried out to explore the relation between

performance in each condition of the animations (mindreading) task and performance

on the JOL (metacognition) tasks.  In summary, JOL accuracy (gamma score) on either

the cue-alone JOL task or the cue-target JOL task, were not significantly associated with

performance in the mentalising condition of the animations task, or performance in the

goal-directed condition of the animations task, among ASD or comparison participants,

all $r$s ≤ -.32, all $p$s ≥ .154.

Additionally there was no significant association between JOL accuracy (gamma

score) on the cue-alone JOL task, and performance in the mentalising or physical

condition of the strange stories task, among comparison participants and ASD

participants,  all $r$s ≤ -.36, all $p$s ≥ .109.  Concerning the cue-target JOL task there was no

significant association between JOL accuracy (gamma score) and performance in the mentalising or physical condition of the strange stories task, among comparison participants, all $r$s ≤ .24, all $p$s ≥ .290.  However, for ASD participants, JOL accuracy on the cue-target JOL task was strongly negatively association with performance on both the mentalising condition of the strange stories task, $r$ = -.63, $p$ = .002 and the physical condition of the strange stories task, $r$ = -.46, $p$ = .050.  This indicates that, for ASD participants the better their accuracy on the cue-target JOL task, the poorer their performance on the strange stories (in both conditions).

*Table 11: Means (SDs) and inferential statistics for group differences in performance on both JOL tasks, the animations task and the strange stories task.*

| | | Group | | t | p | Cohen's |
|---|---|---|---|---|---|---|
| | | ASD (n = 22) | Neurotypical (n = 21) | | | d |
| JOL Tasks: Recall performance | Cue-alone recall performance | .32 (.11) | .36 (.21) | 0.65 | .517 | 0.24 |
| | Cue-target recall performance | .50 (.21) | .47 (.21) | 0.33 | .739 | 0.14 |
| JOL Tasks: Meta-memory performance | Cue-alone gamma scores* | .89 (.12) | .92 (.12) | 0.67 | .505 | 0.25 |
| | Cue-target gamma scores* | .45 (.41) | .53 (.45) | 0.56 | .582 | 0.19 |
| Animations task | Mentalising condition | 1.27 (1.20) | 1.29 (1.06) | 0.04 | .970 | 0.02 |
| | Goal-directed condition | 3.41 (0.73) | 3.48 (0.60) | 0.33 | .745 | 0.10 |
| Strange Stories Task** | Mentalising condition | 1.62 (1.20) | 2.52 (1.29) | 2.35 | .024 | 0.72 |
| | Physical condition | 3.05 (0.71) | 3.05 (0.89) | 0.01 | .992 | 0.00 |

*Gamma scores index metamemory monitoring accuracy
**Based on 19 participants in the ASD group, and 20 participants in the neurotypical group.

**Experiment 3: Discussion**

In this experiment, children in the ASD group did demonstrate mindreading impairments on the strange stories task.  However, the results of Experiment 3 indicated that children with ASD do not demonstrate diminished metacognitive accuracy, relative to neurotypical children, even when either cue-alone or cue-target JOL assessments.  This was in keeping with the results of the previous two experiments, but not in keeping with the study's a priori predictions, nor predictions made by the one-mechanism account.

**General Discussion**

Until now, only one study has previously explored JOL accuracy in individuals with ASD (Wojcik et al., 2014).  Given methodological concerns with this paper, the primary aim of this chapter was to establish whether individuals (both children and adults) with ASD are genuinely impaired at accurately judging how well they have learnt a piece of information.

It was predicted that individuals with ASD would show significant impairments in JOL accuracy, relative to neurotypical participants.  However, there was no evidence to support this prediction; the results from all three JOL experiments clearly found no significant differences in JOL accuracy among either children or adults with ASD, relative to age- and IQ- matched controls.  Likewise there was no evidence to suggest that mindreading performance was significantly positively related to JOL accuracy in any of the three experiments.  As such, the two main prediction main in this chapter were not supported by the results.

One result was in keeping with the study predictions.  In experiment two, although individuals in the ASD group predicted they would remember more words

from the low-frequency word list than the high-frequency word list, they did not

demonstrate signs of strategy use during the learning phase of the task.  Individuals

with ASD spent an equal amount of time learning both word lists.  However,

importantly, neurotypical individuals in this study also failed to show any sign of

strategy use concerning the time spent learning different lists.  Thus the failure to find

evidence of study time strategy use among participants with ASD is instead most likely

due to methodological issues associated with the procedure.

One potential explanation, that was proposed in the introduction to explain

unimpaired accuracy on a JOL task in individuals with ASD, was that individuals may

only demonstrate metamemory impairments if they also demonstrate mindreading

impairments.  According to one-mechanism accounts of the relation between

mindreading and metacognition (e.g., Carruthers, 2009) both abilities rely on the same

underlying mechanisms.  Thus, proponents of this theory would predict impaired

metacognition only in individuals with ASD who also demonstrate clear mindreading

impairments.  However, this explanation cannot explain the results from these studies.

Despite mindreading impairments, adults and children with ASD, in both experiment

one and three, demonstrated typical JOL accuracy (gamma scores).  As such, even when

individuals with ASD demonstrate impairments in mindreading, they still do not appear

to demonstrate impairments in monitoring on JOL tasks.

Alternatively, it was also suggested that intact performance on JOL tasks in

individuals with ASD, may not be the result of typical metacognitive competence.

Instead, it was hypothesised that individuals with ASD may rely on their current

memory for target words, when making JOL assessments, to make relatively accurate

JOL predictions despite impaired monitoring on the task.  However, the results of the

final experiment are not in keeping with this suggestion.  Whilst both groups

demonstrated poorer metamemory accuracy on the cue-target JOL task, the individuals

with ASD did not demonstrate significantly greater difficulty on the cue-target task,

relative to neurotypical children (as would be expected if gamma scores on the cue-

alone task were specifically driven by relying on one's present memory for target

words, in individuals with ASD). Instead, these results again support suggestions that

the metacognitive monitoring processes required to make accurate judgments of

learning are not impaired in ASD.

All results from this chapter seem to suggest that monitoring accuracy on JOL

tasks is unimpaired in ASD. Interestingly, this is not in keeping with the results of

chapter three, which found impairments in monitoring accuracy on a FOK task. Nor is

this finding it in keeping with previous literature, that suggests awareness of one's own

mental states is impaired in ASD (Williams & Happé, 2009b, 2010). One potential

explanation for differences in metamemory accuracy in ASD, across different

metamemory paradigms, is that metamemory impairments in ASD may arise due to

impairments in basic memory ability, not meta-monitoring ability (Wojcik et al., 2014).

Wojcik suggests that recollection ability in particular may mediate different

metamemory judgements (e.g., Hicks & Marsh, 2002; Souchay, Moulin, Clarys, Taconnat,

& Isingrini, 2007), and that some judgements, particularly those made at retrieval (such

as FOK judgements) may rely heavily on the recollection of contextual information

recollected about the target word. Judgements made before retrieval (such as JOLs)

may not rely on recollection to the same extent. This suggestion is of course

speculative, and needs to be researched further, but could potentially explain

discrepancies between metamemory monitoring accuracy in ASD, across different

metamemory tasks (this issue is considered in further detail in the general discussion,

please see p.248- 252).

The results of this chapter have theoretical implications. The one-mechanism account would have predicted such associations between metamemory and mindreading, and so the current results did not support the theory in this respect. Indeed in experiment three, on the cue-target JOL task, gamma scores were strongly negatively associated with performance on both the mentalising condition and the physical condition of the strange stories task, in children with ASD. For ASD participants the *better* their monitoring accuracy on the cue-target JOL task, the *poorer* their performance on the strange stories. This is certainly not in keeping with either the study's predictions, or one-mechanism accounts. However, this association was not specifically found for performance on the mentalising condition of the strange stories task alone, but was found in both conditions (mentalising and physical) of the task. This would suggest that the association between performance on the strange stories and JOL task in children with ASD was not driven by individuals' mentalising ability, but perhaps by different cognitive processes.

As things stand two studies of metamemory accuracy have been carried out in this thesis, with contradictory results. Whilst clear impairments in FOK accuracy were found in chapter three, the results from this chapter show that both adults and children with ASD appear to make accurate JOL assessments. The following chapter also explores metamemory in children with ASD. Alongside exploring metamemory monitoring accuracy on a JOC task, chapter five also explores an aspect of metacognition not yet explored in ASD, metacognitive *control* processes, and the extent to which children monitoring judgments influence their control processes on a task.

**CHAPTER FIVE**

---

**METACOGNITIVE MONITORING AND CONTROL PROCESSES IN CHILDREN WITH**

**AUTISM SPECTRUM DISORDER: DIMINISHED JUDGEMENT OF CONFIDENCE**

**ACCURACY**

Metacognition refers to an individual's beliefs and knowledge about cognition (often referred to as metacognitive knowledge), as well as an individual's ability to monitor and control their own cognitive processes (often referred to as metacognitive skill). With regard to metacognitive skills, accurate *monitoring*/awareness of one's own cognition is thought to facilitate effective *control* over cognition and subsequent behaviour (e.g., Nelson & Narens, 1990). Certainly, both monitoring and control processes play a key role in self-regulated learning, and have been shown to influence study behaviour and test performance in neurotypical children (e.g., Hacker, Bol, Horgan, & Rakow, 2000).

The *accuracy* of one's confidence in one's states of knowledge is considered an important index of one's metacognitive monitoring ability. Studies assessing judgments of confidence (JOC) typically involve participants answering questions about recently-studied material or stored semantic knowledge, and then reporting their confidence in the answers they provided. If an individual's meta-monitoring ability is high, then their confidence judgements should discriminate accurately between correct and incorrect answers. In some studies, participants are subsequently given the opportunity to exclude some of their answers, such that those answers will not contribute to the participant's final "score". This aspect of self-monitoring is particularly important,

because confidence judgements are often used by individuals to control their behaviour (see Koriat & Goldsmith, 1996).

**Metacognition in Autism Spectrum Disorder**

Autism spectrum disorder (ASD) is a developmental disorder diagnosed on the basis of social-communication deficits, and fixated interests and repetitive behaviours (American Psychiatric Association, 2013).  It is widely acknowledged  that individuals with ASD demonstrate impairments in their ability to accurately assess others' mental states (see Yirmiya et al., 1998).  Furthermore, there is a growing body of research that suggests individuals with ASD may also demonstrate impaired awareness of their own mental states (Grainger et al., 2014; Williams & Happé, 2009b; Wojcik et al., 2013). These findings, among others, are in keeping with the view that metacognition and mindreading ability rely on the same underlying cognitive processes/mechanisms (e.g., Carruthers, 2009; Frith & Happé, 1999; Williams, 2010).

However, the evidence to suggest that metacognitive monitoring is definitively impaired in ASD is far from conclusive.  To date, four studies have assessed JOC accuracy among individuals with ASD, with mixed results (Sawyer et al., 2014; Wilkinson et al., 2010; Wojcik et al., 2011; Elmose & Happé, 2014).  In Wilkinson et al., (2010; Exp.  1), children with ASD, as well as age- and IQ-matched neurotypical comparison participants, were tested for their ability to recognise (via an old/new recognition test) recently-presented faces.  After each response, during the recognition test phase, participants made a confidence judgement about their answer, reporting whether they were "certain", "somewhat certain", or "guessing".  Wilkinson et al., (2010) found that the confidence judgments made by children with ASD were significantly less

accurate than those made by neurotypical children, implying diminished metacognitive monitoring in ASD.  However, the same procedure among adults (Exp.  2) revealed no significant between-group differences in JOC accuracy, leading Wilkinson et al., to conclude that metacognition was not diminished among adults with ASD.  Despite this conclusion, it is notable that, although not statistically significant, more than a quarter of the answers adults with ASD reported they were *certain* of were, in fact, incorrect. While neurotypical adults got 85% of the answers they reported they were certain of correct, adults with ASD only got 72% of their "certain" answers correct, and this difference was moderate in size (*Cohen's d* = 0.53).  This suggests at least a subtle diminution of meta-monitoring ability even in adults with ASD.  Additionally, a narrow (three-point) response scale was used in both experiments, which may have reduced response variability and, thus, masked group differences on the task.

In a second study, Wojcik et al., (2011) asked children to make confidence judgements about whether they had correctly performed a series of recently-observed actions.  In contrast to Wilkinson's findings (2010), this study reported no significant between-group differences in JOC accuracy, implying that meta-monitoring is undiminished in ASD.

Recently, Elmose and Happé (2014) have also investigated JOC accuracy in children with ASD.  In this study, during the study phase of each trial, children were presented with a series of six pictures (on some trials pictures of buildings, on other trials pictures of faces). After completing a short distractor task, participants then completed a recall phase in which they were asked to select the six pictures that had been presented from a selection of 12 pictures (six "old" pictures, from the study phase, and the six lure items), and place them in correct serial order. Participants' metacognitive monitoring accuracy on the task was measured in three ways:

1)    Firstly, after the learning phase (but before the recall phase) participants were asked "How many pictures do you think you *will be able to place* in the same order?". This type of judgement can be considered an aggregate JOL assessment (see chapter four, for a discussion). Metacognitive accuracy on this question was assessed by calculating the difference between an individual's prediction of how many pictures they would recall correctly and how many pictures they did in fact recall correctly.

2)    Secondly, after participants had completed the recall stage of each trial, they were asked "How many pictures do you think you *have placed* in the same order?" This type of judgement can be considered an aggregate JOC assessment. Again, metacognitive accuracy on this question was assessed by calculating the difference between an individual's assessment of how many pictures they had recalled correctly, and how many pictures they had in fact recall correctly.

3)    Finally, after participants had completed the recall stage of each trial (with either building stimuli or face stimuli) and after they made an aggregate JOC surrounding their performance on the task (see above), they were asked to judge how confident they were that *each* of the pictures they had selected was correct. This type of judgement can be considered an item-by-item JOC.  Participants were asked to judge their level of confidence as either "sure", "pretty sure", or "unsure". For the item-by-item JOC decisions, metamemory accuracy was assessed using a coding scheme. Participants scored 0 if their confidence judgement matched their recall ability (e.g., they said they were "sure" of a picture and it was the correct picture in the correct position/ they said they were "unsure" of a picture, and it was an incorrect picture). Participants scored 1 if their confidence judgement was somewhat in accordance with their recall performance (e.g., they said they were "pretty sure" for pictures they had recalled correctly, but not placed correctly). Finally participants scored 2 if their confidence judgement was not in

keeping with their recall performance at all (e.g., they said they were "sure" for a picture that was incorrect).

Overall, Elmose and Happé (2014) report that children with ASD made as accurate judgements of confidence as children without ASD (regardless of the type of stimuli used during the task or whether JOCs were aggregate/ item-by-item). Participants were also as accurate at judging overall how many stimuli they would get correct on each trial. This suggests that children with ASD do not show diminished meta-memory. However, there are three potential methodological issues with the study that should be considered when interpreting the results. These issues concern the criteria used to assess item-by-item JOC accuracy.

Firstly, one problem with the criteria used to assess accuracy on the task was that participants' object level recall responses were still considered somewhat accurate even when they recalled the correct picture in the wrong location. In contrast, participants' assessments concerning how confident they were that their recall of a given picture was correct presumably reflected how confident they were that they had provided the correct picture *and* placed it in the correct location. Given that participants were not told that previously-studied pictures recalled out of serial order would still be considered partly correct, they could not have taken this into account when making confidence judgements. It is likely that this will have reduced participants' accuracy scores during the task.

Secondly, Elmose and Happé (2014) employed a non-standard means of measuring metacognitive accuracy. Most studies in the literature on metacognitive monitoring employ gamma correlations, which take into account each participant's individual response criteria when assessing accuracy (and thus take into account participants who do not respond using the full range of JOC responses available). For

example, if an individual consistently reported they were "pretty sure" about all their correct answers, and "unsure" about all their incorrect answers, participants gamma score would be 1 (reflecting perfect monitoring accuracy on the task). However, using the scoring criteria used in Elmose and Happé, a response of "pretty sure" for a correct response was not considered accurate. Thus, a major problem with the scoring criteria Elmose and Happé (2014) used was that it did not take into account participants' individual response criteria. As such, it is possible, that group difference in accuracy may have been apparent, if gamma correlations were used to assess accuracy on the task.

One final issue with the method employed in the study is with the response options participants were given during the item-by-item JOL phase of the task. As discussed, participants were asked to judge their level of confidence as either "sure", "pretty sure", or "unsure". However, these response options gave participants no opportunity to respond that they were extremely *un*confident that an answer they had provided was correct. Intuitively, a confidence response of "unsure" indicates that one is not certain the answer is correct, but equally not certain the answer is incorrect. This represents a serious methodological problem with rating scales used in this task. Additionally, only providing three response options for participants to judge their confidence on may have reduced response variability on the task (see a similar critique of Wilkinson et al., 2010 above).

Finally, Sawyer et al., (2014) employed a JOC task that assessed both monitoring *and* control in adults with ASD. In this study, participants were asked to complete an emotion recognition task involving facial stimuli. Participants were instructed that the aim of the study was to submit as many correct responses as possible. For each emotion recognition judgement, participants rated how confident they were that they had

selected the correct response. Participants were then given the opportunity to submit each answer towards their total score (and gain a point for each correct answer), or discard the answer (and avoid losing a point for getting an answer wrong). This provided a measure of metacognitive control. In a second experiment, the same procedure was used but participants' judgements concerned their answers to general knowledge questions, rather than emotion recognition.

In both experiments, Sawyer et al., reported no significant between-group differences in JOC accuracy, implying undiminished meta-monitoring ability in ASD. However, it is important to note that the between-group difference in JOC accuracy *on the general knowledge task* was associated with a one-tailed *p* value of .06, potentially implying a subtle monitoring impairment in ASD.

In terms of metacognitive control, Sawyer et al., (2014) found no between-group differences on their key index ($d'$), implying undiminished metacognitive control in ASD. However, Sawyer et al., performed additional *post-hoc* tests, which suggested that a significantly higher proportion of ASD participants (*n* = 12) than neurotypical participants decided not to withhold *any* answers. This could imply that these 12 ASD participants were not showing any metacognitive control at all. Alternatively, it could reflect a mere failure to understand the task demands among these participants. On this basis, the extent to which metacognitive control is diminished in ASD is still not clear. What is particularly notable is that these 12 ASD participants *did* appear to show diminished monitoring ability relative to the neurotypical control group, the difference in gamma score being associated with a Cohen's *d* value of 0.62 (although, of course, we are not aware whether these participants were matched with the neurotypical controls).

Although the study by Sawyer et al., (2014) was exemplary in many respects, there are additional issues that might suggest caution should be taken when interpreting the results. Firstly, participant groups were not matched for age or performance IQ (note that this may apply to Wojcik et al., 2011 with respect to verbal IQ). Between-group differences in baseline variables could well explain the experimental findings of the study (see Mervis & Klein-Tasmin, 2004). Indeed, Dr Sawyer very kindly provided us with additional unreported data about this. In Sawyer et al., age was significantly negatively associated with $d'$ among participants with ASD, $r$ = -.37, $p$ = .04 (A. Sawyer, personal communication, August 22nd, 2014). Thus, differences in age (ASD participants were older than comparison participants) could well explain the trend towards group differences in metacognitive control in this study.

Apart from methodological differences (and potential difficulties) between existing studies, one explanation for inconsistencies in the literature could be that the samples of participants in each study differed in their mindreading ability. To our knowledge, mindreading ability was not assessed in any of these studies. Yet, as discussed above, according to one major theory (e.g., Carruthers, 2009) we should only expect to find metacognitive impairments in individuals with ASD who also have diminished mindreading. Given that some studies do not find diminished mindreading task performance in high-functioning individuals with ASD, it is conceivable that ASD participants in Wojcik et al., (2011), Elmose and Happé (2014) and Sawyer et al., (2014) did not show impairments in JOC accuracy because they would have shown undiminished mindreading task performance. To date, no study has directly examined JOC accuracy alongside mindreading ability.

**The Current Study**

The central aim of this study was to extend the current findings concerning metacognition in ASD, by examining both monitoring *and* control accuracy in children with ASD. To examine this the study employed a JOC task, during which children were asked a series of questions about recently-studied material, and were then asked to judge how certain they were that the answers they had provided were correct (providing a measure of metacognitive monitoring accuracy). Additionally, children were told that for each correct answer they submitted they would receive a point, but for each incorrect answer they would lose a point. At the end of the task children were given the opportunity to remove any of the answers they had previously provided (providing a measure of metacognitive control accuracy). The main prediction was that participants with ASD would demonstrate impairments in metacognitive monitoring ability. This was predicted on the basis of theoretical inclinations concerning the underlying mechanisms involved in mindreading/metacognition, as well as an interpretation/critical analysis of the few previous studies of this ability in ASD. Predictions concerning group differences in metacognitive control ability were less straightforward. The only study ever (by Sawyer et al., 2014) to explore this ability in ASD reported a trend toward a group difference in this ability, but there are arguably some difficulties with Sawyer et al.'s findings that prevent definitive conclusions from being drawn (see above). Thus predictions concerning this aspect of the study were non-directional. The issue of whether metacognitive monitoring and/or control is diminished in ASD is separate from the issue of whether monitoring is *used for the purpose* of control by people with ASD. For example, it is possible that monitoring ability is undiminished in ASD, but not used appropriately for the purpose of

metacognitive control. Alternatively, even if monitoring ability is diminished in ASD, residual monitoring ability might influence control processes to the same extent among individuals with ASD as among neurotypical individuals. Given these possibilities and given the fact that no previous study has explored the extent to which monitoring influences control in ASD, a non-directional prediction with respect to this aspect of the study was also made.

Two measures of mindreading ability were also included in the study. These were included to establish (a) whether the sample of ASD participants was typical in displaying diminished mindreading, and (b) the size of any association between mindreading and metacognition in neurotypical children and children with ASD. To our knowledge, no study has examined whether performance on a JOC task (either monitoring or control accuracy) relates to mindreading ability. It was predicted that mindreading ability would be associated significantly with metacognitive monitoring ability, but not with metacognitive control ability.

## Method

### Participants

Thirty-two children with ASD and 30 neurotypical children took part in this study, after their parents had given written, informed consent. Participants in the ASD group had formal diagnoses of autistic disorder or Asperger's disorder, according to established criteria (American Psychiatric Association, 2000; World Heath Organisation, 1993). To assess severity of ASD features, parents of participants with ASD completed the Social Responsiveness Scale (SRS; Constantino et al., 2003). In all but one case, participants with ASD scored above the defined cut-off for ASD on the SRS (total score ≥

60; Constantino et al., 2003). The remaining participant scored 55 on the SRS, which is just below the conventional ASD cut-off of 60.

*Table 12: Participant Characteristics (Means, Standard Deviations and Inferential Statistics).*

| | Group | | $t$ | $p$ | Cohen's $d$ |
|---|---|---|---|---|---|
| | ASD (n = 32) | Neurotypical (n = 30) | | | |
| Age (years) | 13.59 (1.36) | 13.27 (1.06) | 1.01 | .315 | 0.26 |
| VIQ | 101.28 (16.69) | 103.87 (14.92) | 0.64 | .524 | 0.16 |
| PIQ | 100.72 (13.39) | 105.67 (14.32) | 1.41 | .165 | 0.36 |
| FSIQ | 101.19 (14.85) | 105.53 (15.27) | 1.14 | .261 | 0.29 |
| SRS Total Score | 84.16 (8.79) | 45.67 (10.50) | 15.69 | < .001 | 3.98 |

SRS: Social Responsiveness Scale (Constantino et al., 2003); VIQ = verbal IQ; PIQ = performance IQ; FSIQ = full scale IQ.

Parents of the neurotypical children also completed the SRS. All but four participants in the neurotypical group scored below the defined cut-off for ASD. The remaining participants' SRS scores ranged between 60 and 73. To ensure that including these participants in the overall sample did not affect the results of the study all analyses in the paper were re-run, excluding all five participants who scored outside the expected range on the SRS. After removing these participants from analyses, none of the results (nor study conclusions) changed substantively (i.e., no $p$ value changed from significant to non-significant or vice versa, and no effect size changed category – small, moderate, large). The participant groups were closely equated for verbal and non-verbal ability (see Table 12 for participant characteristics), using the Wechsler

Abbreviated Scale of Intelligence (WASI; Wechsler, 1999).  Both groups were also

equated closely for chronological age.

**Materials and Procedures**

**Judgement-of-confidence task.**  This task was designed to assess the accuracy

of children's metamemory monitoring and control processes, and was based on tasks

used in previous studies of metacognition in neurotypical children (e.g., Krebs &

Roebers, 2010; Roebers, Schmid, & Roderer, 2009).  The task consisted of a study phase,

a test phase, a JOC phase (during which confidence in the accuracy of recall was

assessed) and a metacognitive control phase (during which the accuracy of

metacognitive control processes were assessed).  In total, the task took approximately

twenty minutes to complete.

*Study phase.*  Participants were shown a short (4 minute) video, presented to

them on a laptop computer.  This video was downloaded from a website of educational

videos, suitable for 11- to 16-year-olds.  The experimenter explained to each participant

that the video they were about to watch was about kangaroos and about how kangaroos

survive in the Australian outback.  Participants were told to pay full attention to the

video, because afterwards they would be asked some questions about the information

presented in it.

*Test phase.*  After watching the video, participants were given a short

worksheet, which consisted of 16 test questions, for which the answer had been

explicitly presented in the video, during the study phase.  Additionally, the sheet

included eight questions to ensure that participants understood the rating scale for

their subsequent confidence judgements.  Four "easy" questions were designed so that

it was almost certain participants would know the answers to them (e.g., how many

eyes do kangaroos have?) and four "impossible" questions were included so that it would be very unlikely children would know the correct answers (e.g., what is the Latin species name for kangaroo?). If participants were not able to make the blatant distinction in their judgements of confidence between answers to easy and impossible questions, then one could not be sure they understood the nature of the task at all.

For the test phase, the edge of the worksheet containing the JOC scale was folded over, so that it was not visible (please see Figure 11). Once presented with the worksheet, participants were given a pen, and were asked to write down an answer for each question. It was explained to participants that it was important they provided an answer for every question, and that if they did not know the correct answer they should take a guess. Participants were given an unlimited amount of time to answer the questions on the worksheet.
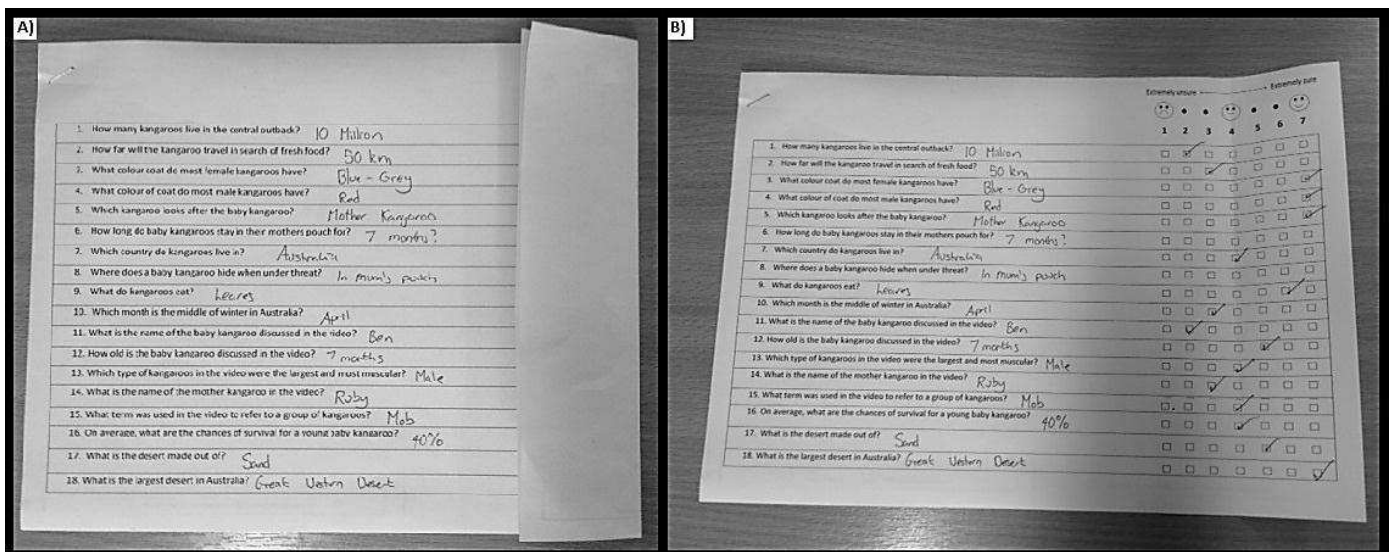


Figure 11: Pictures of the worksheet used during the JOC task. A) Picture of the JOC sheet presented to participants during the test phase of the task, with the JOC rating scales covered. B) Picture of the JOC worksheet presented during the JOC phase of the task.

***JOC Phase.*** After the experimenter checked that participants had provided an answer for all the questions on the worksheet, they turned over the section of the worksheet that had previously been folded over, making the JOC scale visible. It was explained to participants that "I would now like to know how confident you are that each of the answers you wrote down is correct". Participants were asked to judge their confidence for each answer on a 7-point likert scale, ranging from 1-7 (extremely unsure to extremely sure). The experimenter fully explained the confidence scale to participants, explaining that higher numbers on the scale indicated a higher certainty that the answer provided was correct. Participants gave confidence judgments for each answer on a scale next to that answer on the worksheet. The experimenter made sure that each participant provided a JOC for each answer.

***Metacognitive Control Phase.*** Finally, participants were told that at a later point the experimenter would mark each of the answers on the worksheet. It was explained to participants that for each correct answer they had given they would get one point, but for each incorrect answer one point would be taken away from them. Participants were then given a different coloured pen and told that they had the opportunity to improve their performance on the task. Participants were told that they were now able to cross out any of their answers. If they crossed out an answer they would not get a point for this answer if that answer was correct, but nor would they lose a point if the answer was incorrect. Participants were told that they could cross out as many or as few answers as they liked.

**Animations task.** During the Animations task, participants were required to provide a verbal description of four silent video clips, each of which displayed an interaction between a large red triangle and a small blue triangle. These clips were taken directly from Abell et al., (2000). In two of the clips (the "Coaxing" and "Tricking"

animations), an adequate explanation of the triangles' interaction required the attribution of propositional attitudes, such as beliefs, intentions, and/or desires. As in Abell et al.'s study, these clips comprised a "mentalising" condition (assessing higher-level mindreading). In the remaining two clips (the "Fighting" and "Following" animations), an adequate explanation of the triangles' interaction required the attribution of goal states, but did not necessarily require the attribution of propositional attitudes/epistemic mental states. As in Abell et al., (2000), these clips comprised a "goal-directed" condition.

Each clip was presented to participants on a laptop computer. Before undertaking the experimental trials, participants also completed two practice trials, to familiarise themselves with the task (one goal-directed and one mentalising). Experimenter feedback was given on practice trials. For each of the experimental animations, participants watched each clip twice. First, participants watched the clip through once in silence and were told to "watch the clip and see how the triangles are interacting". Participants then watched the clip again and were asked "as you watch the clip again I would like you to tell me how the triangles are interacting". Participants provided a running commentary on the triangles' interactions during this second presentation of the clip. For the experimental trials, a digital audio recording of participants' responses was made for later transcription. No feedback was given on the experimental trials.

**Strange Stories Task.** The Strange Stories task (Happé, 1994) was used as a second measure of mindreading ability. During the Strange Stories task participants were presented with four short vignettes (two mentalising stories and two physical stories). These stories were taken from Happé (1994). Correct answers to the mentalising stories required the attribution of mental states to story characters,

whereas correct answers to physical stories required an understanding of physical

causality only.  First, each story was individually presented on cards to participants.

Children either read the story aloud or, if children did not feel comfortable reading

aloud, the experimenter read each story to them.  After each story had been read the

experimenter produced a second card and presented participants with a question about

the content of the story (e.g., Why did the prisoner say that?).  The experimenter read

this question aloud, and participants provided a verbal response to each question.

Before undertaking the experimental trials, participants also completed one practice

trial, to familiarise themselves with the task (one mentalising story).  During practice

trials, participants read a story and provided an answer to the practise question.  The

experimenter then gave feedback after participants' response to the practise question.

For the experimental trials, a digital audio recording of participants' responses was

made for later transcription, and no feedback was given[2].

**Scoring**

**Judgement-of-confidence task.**

***Object-level test performance.***  A measure of participants' *object-level* memory

performance was calculated on the JOC task.  Participants recall ability was calculated as

the proportion of answers participants correctly remembered during the recall stage of

the task.

---

[2] Due to restrictions that occurred during data collection, six participants did not complete the Strange Stories task (four participants with ASD and two neurotypical participants).

***JOC rating scale use.***  The average confidence judgements given by participants for answers to the "easy" and "impossible" questions were calculated.

***Metamemory monitoring accuracy.***  Firstly, the average confidence judgement participants gave for their correct answers was calculated, as well as the average confidence judgement participants gave for their incorrect answers.  This provided a basic measure of metacognitive accuracy.  The better participants monitoring accuracy, the larger the difference should be between their confidence ratings for correct answers than incorrect answers.  For the purpose of correlational analyses a difference score was calculated, by subtracting the average confidence rating participants gave for incorrect answers, from the average confidence rating participants gave for correct answers.

Secondly, gamma scores (Goodman & Kruskal, 1954) were calculated to provide an index of JOC accuracy for each participant.  This type of analysis is recommended by Nelson (1984; Nelson, Narens, & Dunlosky, 2004) and is commonly used to analyse monitoring accuracy on JOC tasks (e.g., Roebers et al., 2009).  Gamma correlations range between + 1 and -1; a score of 0 indicates chance-level accuracy, in which confidence judgements are not associated in any way with whether an answer is correct or not.  A large positive gamma value indicates a good degree of accuracy in JOC judgments and a large negative value indicates less than chance-level performance on the task.  Gamma scores were calculated individually for each participant, based on participants' answers to the 16 experimental questions and corresponding confidence judgments.

***Metamemory control accuracy.***  *C*ontrol effectiveness was calculated using the measure of *d*-prime (*d'*).  This measure is how control accuracy was assessed in Sawyer et al., (2014), and is calculated using participants "hit-rate" (*H*) and "false-alarm rate" (*FA*).  Hit rate was calculated as the number of answers participants removed that were

incorrect (and thus correctly removed) plus the number of answers participants kept that were correct (and thus correctly kept), divided by the total number of answers. False alarm rate was calculated as the number of answers participants removed that were in fact correct divided by the total number of correct answers participants provided. *d'* was calculated using the following formula:

$$d' = Z(\text{H}) - Z(\text{FA})$$

A *d'* score of 0 indicates no difference between hit rate and false alarm rate, demonstrating ineffective control measures on the task. In contrast, the higher the *d'* value, the greater the tendency to remove incorrect answers and keep correct answers, thus the greater the effectiveness of the control strategy on the task.

***The effect of monitoring processes on control accuracy***. The average confidence judgment participants gave for the answers they removed was calculated, as was the average confidence judgement participants gave for the answers they kept. These scores were calculated to provide a measure of how participants monitoring judgements influenced their control performance. If participants control processes were strongly influenced by their JOC ratings, then they should remove answers they had the lowest confidence in, and keep answers they had highest confidence in. For the purpose of correlational analyses a difference scores was calculated, by subtracting the average confidence judgement for removed answers from the average confidence judgement for kept answers. Two participants (1 ASD, 1 neurotypical) chose not to remove any of their answers and thus difference scores for these participants could not be calculated.

**Animations task.** Voice recordings of participants' commentaries were transcribed verbatim. These transcriptions were then scored by a second rater who was blind to the diagnostic status of the participants, according to the scoring criteria outlined in Abell et al., (2000). Inter-rater reliability for scores across the four animations was excellent, *Cronbach's α* = .93

**Strange Stories Task.** Voice recordings of participants' answers to each question were transcribed verbatim. These transcriptions were also scored by a second independent rater who was blind to the diagnostic status of each participant, according to the scoring criteria outlined in White, Hill, Happé and Frith (2009). Again, inter-rater reliability for scores across the four stories was excellent, *Cronbach's α* = .92

## Results

### Mindreading performance

**Strange Stories task.** Table 13 shows the means and standard deviations for performance on the Strange Stories task. A mixed-model ANOVA revealed significant main effects of Group, $F$ (1, 54) = 6.19, $p$ = .016, $\eta_P^2$ = .103, and Condition, $F$ (1, 54) = 49.05, $p$ < .001 $\eta_P^2$ = .476. However, these were qualified by a significant Group × Story Type interaction, $F$ (1, 54) = 4.13, $p$ = .047, $\eta_P^2$ = .071. Paired samples *t*-tests indicated that both groups performed less well in the mentalising condition than in the physical condition, all *ps*< .003, all *d*s > 0.77. Independent samples *t*-tests indicated that the ASD group performed significantly less well than neurotypical participants in the mentalising condition of the task, but not the physical condition (see Table 13). As predicted, the ASD group demonstrated selective impairments in the mindreading condition of the task.

*Table 13: Means (SDs) and inferential statistics for group differences in performance on the Animations task and Strange Stories task.*

| Experimental Measure | Condition | Group | | $t$ | $p$ *(two-tailed)* | Cohen's $d$ |
|---|---|---|---|---|---|---|
| | | ASD | Neurotypical | | | |
| Animations Task | Mentalising | 1.41 (1.24) | 1.60 (1.19) | 0.63 | .534 | 0.16 |
| | Goal-directed | 3.34 (0.75) | 3.47 (0.68) | 0.68 | .501 | 0.18 |
| Strange Stories task | Mentalising | 1.54 (1.20) | 2.43 (1.17) | 2.82 | .007 | 0.75 |
| | Physical | 2.96 (0.92) | 3.21 (0.83) | 1.07 | .292 | 0.29 |

**Animations task.** Table 13 shows the means and standard deviations for performance on the Animations task. A mixed-model ANOVA revealed a significant main effect of Animation Type, indicating that, across both groups, scores were higher in the goal-directed condition than the mentalising condition, $F(1, 60) = 130.10$, $p <$ .001 $\eta_p^2 = .684$. No other effects were significant, suggesting that the ASD group did not show diminished performance on this mindreading task, all $ps > .41$.

**Judgment of confidence task**

Table 14 shows descriptive statistics and the results of independent-samples $t$-tests for all aspects of the experimental JOC task. Given the predicted group differences in meta-level performance on the JOC task, all $p$ values associated with group differences on this aspect of the task are reported one-tailed. Before reporting the main results, it is important to ensure that participants were able to use the JOC rating scale appropriately. Thus, an initial analysis of group differences in the average JOC rating for

the four "easy" and four "impossible" questions was conducted.  A mixed-model ANOVA

revealed a significant main effect of Question-type, reflecting that JOC ratings were

significantly higher for "easy" questions than for "impossible" questions, $F$ (1, 60) =

678.21, $p$ <.001 $\eta_P^2$ = 0.91.  Crucially, neither the main effect of Group nor the Group ×

Question-type interaction were significant, $F$s <0.57, $p$s > .45.  This suggests that both

groups were able to use the rating scale appropriately.

**Cognitive (object-level) test performance.**  An independent-samples $t$-test

revealed no significant between-group difference in the proportion of answers correctly

recalled (see Table 14).

**Metamemory monitoring performance.**  An independent-samples $t$-test

revealed significantly lower gamma scores among ASD than neurotypical participants.

This suggests that metamemory monitoring accuracy is impaired in children with ASD

relative to the neurotypical participants.  Table 14 shows the means and standard

deviations for the average JOC rating for correct and incorrect answers.  A mixed-model

ANOVA was also carried out on these data with Group (neurotypical/ASD) entered as

the between-subjects variable, and Answer Type (correct/incorrect) entered as the

within-subject variable.  There was a significant main effect of Group on JOC ratings,

reflecting the fact that participants with ASD made significantly higher JOC ratings than

neurotypical participants on the task overall, $F$ (1, 60) = 8.82, $p$ = .004, $\eta_P^2$ = 0.13.  There

was also a significant main effect of Answer Type, indicating that, across both groups,

JOC ratings were significant higher for correct answers than incorrect answers, $F$ (1, 60)

= 643.97, $p$ <.001 $\eta_P^2$ = 0.91 Importantly, there was a significant Group by Answer Type

interaction, $F$ (1, 60) = 12.02, $p$ = .001, $\eta_P^2$ = 0.16.  To investigate this interaction further a

series of $t$-tests were carried out.  Paired samples $t$-tests indicated that participants' JOC

ratings were significant higher for correct answers than incorrect answers, in both the

ASD and neurotypical group, all *p*s < .001, all *d*s > 2.95.  Independent samples *t*-tests

indicated that the ASD group provided significantly higher confidence ratings for the

answers they had got incorrect, relative to neurotypical participants, but did not

significantly differ in their confidence for correct answers relative to neurotypical

participants (see Table 14 for descriptive and inferential statistics).  These results

suggest that individuals with ASD demonstrated impairment in metacognitive

monitoring, and reported subjectively higher confidence in answers they had got

incorrect, relative to neurotypical participants.

**Metamemory control performance.**  An independent-samples *t*-test indicated

that *d'* was lower among participants with ASD than among neurotypical participants.

However, the between-group difference in *d'* was not significant (see Table 14)

**The influence of monitoring on control processes.**  Table 14 presents the

average JOC rating for answers participants kept, and answers participants removed, for

both the ASD and neurotypical group.  In our opinion, this contrast examines the extent

to which children's control processes were based on their monitoring judgements.  A

mixed-model ANOVA was carried out on these data with Group (neurotypical/ASD)

entered as the between-subjects variable, and Answer Type (kept/removed) entered as

the within-subject variable.  Three ASD participants and one neurotypical participant

chose not to remove any answers to the 16 experimental questions.  Thus this ANOVA

was based on 29 participants in the ASD group and 29 participants in the neurotypical

group.  Again, there was a significant main effect of Group on JOC ratings, reflecting the

fact that participants with ASD made significantly higher JOC ratings than neurotypical

participants on the task overall, $F(1, 56) = 4.24$, $p = .044$, $\eta_p^2 = 0.07$[1].  There was also a

significant main effect of Answer Type, indicating that, across both groups, JOC ratings

were significantly higher for kept answers than incorrect answers, $F(1,56) = 807.32$, *p*

<001. $\eta_p^2$ = 0.93 There was a significant Group by Answer Type interaction, $F$ (1, 56) = 5.90, $p$ = .018, $\eta_p^2$= 0.09.  Paired samples $t$-tests indicated that both the ASD and neurotypical group gave significantly higher JOC ratings for answer they subsequently kept than answers they subsequently removed, all $ps < .001$, all $ds$>3.36.  Independent samples $t$-tests indicated that the ASD group provided higher confidence ratings for the answers they subsequently removed, relative to neurotypical participants, but were equally as confident in the answers they kept.  These results suggest that, at least when judging which answers to remove, individuals in the ASD group relied on their JOC ratings less than neurotypical individuals.

*Table 14: Means (SDs) and inferential statistics for group differences in performance on the JOC task.*

| Experimental Measure | Dependent variable | Group | | $t$ | $p$ | Cohen's $d$ |
|---|---|---|---|---|---|---|
| | | ASD | Neurotypical | | | |
| Object-level performance | Proportion of answers recalled | .58 (.18) | .60 (.12) | 0.57 | .572 | 0.13 |
| Monitoring performance | Average JOC for correct answers | 6.22 (0.63) | 6.16 (0.52) | 0.41 | .341[a] | 0.10 |
| | Average JOC for incorrect answers | 3.59 (1.09) | 2.70 (0.77) | 3.70 | <.001[a] | 0.94 |
| | Gamma score* | .84 (.16) | .90 (.13) | 1.75 | .043[a] | 0.41 |
| Control performance | $d$-prime** | 4.15 (1.56) | 4.74 (1.54) | 1.48 | .145 | 0.38 |
| Influence of monitoring over control | Average JOC for kept answers | 5.80 (0.84) | 5.79 (0.67) | 0.08 | .937 | 0.01 |
| | Average JOC for removed answers | 2.55 (1.13) | 1.85 (0.72) | 2.80 | .007 | 0.74 |

Note: ASD = autism spectrum disorder; JOC = Judgment of confidence
[a]$p$ values reported one-tailed, because the direction of the effect was predicted *a priori*
*Gamma scores index metamemory monitoring accuracy
**$d$-prime scores index metamemory control accuracy

**Associations between metamemory monitoring ability, metamemory control ability and mindreading ability.**

A series of correlational analyses was carried out to explore the relations among (a) metacognitive monitoring ability (gamma score; difference between JOC for correct answers and JOC for incorrect answers), (b) metacognitive control ($d'$), (c) the extent to which monitoring influenced control (difference between JOC ratings for kept answers and JOC ratings for discarded answers), and with mindreading ability (Strange Stories mentalising/physical; Animations mentalising/goal-directed). Only one correlation reached conventional levels of statistical significance. In the ASD group only, the difference in JOC ratings for kept and removed answers was associated significantly with performance in the goal-directed condition of the Animations task, $r = .43$, $p = .019$. This association was not predicted and would not survive correction for multiple comparisons. Contrary to predictions, no other associations were statistically significant in either group, $r$s ≤ .31, $p$s ≥ .12

## Discussion

As predicted, this study found that participants with ASD showed diminished metacognitive monitoring accuracy, as reflected by significant between-group differences in gamma scores and in the difference score between JOC ratings for correct versus incorrect answers. These results are in keeping with the study's predictions and the findings of Wilkinson et al., (2010), who also found that confidence judgments made by children with ASD were less accurate than those made by neurotypical children. However, these results are not in keeping with three previous studies, that concluded JOC accuracy was unimpaired in children (Wojcik et al., 2011; Elmose & Happé, 2014) and adults with ASD (Sawyer et al., 2014). Importantly, in this study participants with ASD were (a) closely-matched with neurotypical participants in terms of age, VIQ, and PIQ, and (b) showed characteristic

mindreading impairments (on the Strange Stories task, at least). Moreover, the seven point response scale that was employed in this study provided greater opportunity for detecting variation in metacognitive monitoring ability than the narrower scales used in some previous studies. This may explain the differences in results between this study, on the one hand, and those of Wojcik et al., Elmose & Happé (2014) and Sawyer et al., on the other hand.

Although evidence that metacognitive monitoring is diminished among children with ASD was found in this study, the study found little evidence to support the idea that metacognitive control was similarly diminished; the between-group difference in $d'$ did not approach statistical significance. Having said this, it is notable that $d'$ was lower among participants with ASD than among neurotypical participants (albeit non-significantly so) and that the effect size associated with the between-group difference ($d = 0.38$) was strikingly similar to the effect size for the group difference in control ability reported by Sawyer et al., (2014; $d = 0.32$, averaged across general knowledge and emotion recognition conditions). On the one hand, this consistency across the only two studies to have explored metacognitive control ability in ASD might suggest that this ability is subtly diminished in ASD. On the other hand, one should be wary about a general tendency in the field to interpret any hint of difference between ASD and comparison groups as a sign of an impairment in ASD. Therefore, whilst this consistency across studies is noted, it is likely that if such a diminution exists, it is questionable whether it is clinically significant, given the small magnitude of the effect.

Clearer evidence of a meaningful difference between the groups was found with regard to the extent to which metacognitive monitoring influenced metacognitive control. Here, the mean difference between JOC ratings for kept and removed answers was significantly smaller among ASD participants than among comparison participants, indicating that the extent to which monitoring influenced control was diminished in participants with ASD to a significantly lesser extent than it did among comparison participants. This suggests that the means by which individuals achieved control over their behavioural choices (i.e., of

which answers to remove) was different among ASD and comparison groups. It is possible that children with ASD employed a compensatory strategy on the JOC task, allowing them to perform relatively well on the control aspect of the task despite impaired monitoring ability. For example, participants' decisions to keep or remove an answer could have been determined simply on whether they could bring to mind an answer or not. For example, when presented with the question "What animal hunts kangaroos?", participants may have chosen to keep their answer only if they could bring to mind a picture of a dingo (which had been presented in the video). This kind of all-or-nothing strategy would be successful for many aspects of control, but would likely come unstuck when the demands of the situation were complex enough to require a graded monitoring of internal states. As yet, it is not clear what information children with ASD based their control performance on and what the processing costs associated with using alternative strategies to influence control performance are. As such, future research is needed to investigate what underlying processes influence control performance in individuals with ASD.

Finally, contrary to predictions, no evidence of a significant association between mindreading and metacognition abilities was found. It is important to note that this study was suitably powered to detect at least a moderate sized ($r$ = 0.30) predicted association between these abilities if one existed (actual power = .78, which is only just short of Cohen's 1992 recommendation of ≥.80). Indeed, although group sizes of 30+ participants might be considered relatively small by the standards of many general cognitive psychology studies, a sample this size is relatively large for studies of cognition in ASD, which is strength of the study. Thus, the failure to find a significant association is unlikely to be due to insufficient power.

Overall, the results are mixed with respect to the debate regarding the connection between mindreading and metacognitive monitoring. Findings of diminished metacognitive monitoring in ASD are consistent with one mechanism accounts of the relation between these

two abilities (Carruthers, 2009, 2012; Frith & Happé, 1999; Williams, Lind, & Happé, 2009). However, they do not definitively confirm that these abilities are not underpinned by two separate mechanisms (e.g., Goldman, 2006). It may be that both systems (one underpinning mindreading and the other underpinning metacognition) are impaired in ASD. Equally, the finding that mindreading performance and metacognitive monitoring performance were not associated significantly in this sample does not rule out one system theories of the relation between these two abilities. Cognitive measures are rarely (if ever) process pure and thus an association or lack thereof should not be taken in-and-of-itself as evidence for/against the notion that the underlying cognitive abilities are/are not associated. Fortunately, the purpose of this paper was *not* to decide between these competing theories of the relation between mindreading and metacognition (for reasons just discussed, this study was not fit for that purpose). Rather, the aim was to understand the nature of metacognition in ASD.

The results of this study support the notion that metacognitive monitoring is diminished in ASD and that what residual monitoring is available to individuals with ASD is used for the purpose of metacognitive control to a lesser extent than are the monitoring resources available to neurotypical individuals. It is important future research explores metacognition in ASD further, particularly exploring both metacognitive monitoring and control processes, and how these processes relate to each other. Assessing awareness of memory performance in populations with memory difficulties is a critical matter. Studies have shown that memory awareness is linked to memory rehabilitation effectiveness (e.g., Clare et al., 2000). As such, it is essential to create a clearer picture of memory awareness impairments in individuals with ASD. Additionally, research within the typically developing literature suggests that metacognition (confidence judgments in particular) plays an important role in everyday functioning and decision making (see e.g., Yeung & Summerfield, 2012). Metacognition, control processes in particular, also play a vital role in self-regulation, and self-regulated learning. Given memory impairments in individuals with ASD (see Boucher

& Bowler, 2008) future research should aim to establish a comprehensive account of metacognition in ASD, with the aim of informing intervention efforts designed to remediate cognitive impairments in this disorder.

Chapters three, four and five all employed typical metamemory paradigms to explore psychological/ private self-awareness in ASD. However, as Neisser's framework suggests, there are several aspects of self-awareness which do not involve awareness of one's mental self. The following two chapters report the results of studies exploring aspects of *ecological* self-awareness in individuals with ASD, and investigate the extent to which individuals with ASD are aware of their own physical selves.

**CHAPTER SIX**

## ONLINE ACTION MONITORING AND MEMORY FOR SELF-PERFORMED ACTIONS IN

## AUTISM SPECTRUM DISORDER[3]

Autism spectrum disorder (ASD) is a developmental disorder diagnosed on the basis of behavioural impairments in social-communication, and by fixated interests and repetitive behaviours (e.g., American Psychiatric Association, 2013). On the cognitive level, it has been suggested that ASD is characterised by diminished self-awareness (e.g., Hobson, 1990; Russell, 1996; Williams, 2010) and, more recently, that specific aspects of self-awareness are selectively diminished in this disorder. For example, it has been suggested that individuals with ASD have diminished awareness of *psychological* aspects of the self (e.g., awareness of one's own thoughts, personality characteristics etc.), but undiminished awareness of *physical* aspects of self (e.g., awareness of one's own physical appearance; Lind, 2010; Williams, 2010). In keeping with this proposal, studies have shown that individuals with ASD demonstrate difficulties representing and reporting their own thought processes (e.g., Hurlburt et al., 1994), intentions (e.g., Williams & Happé, 2010), emotional feelings (e.g., Ben Shalom et al., 2006; Hill, Berthoz, & Frith, 2004) and beliefs (e.g., Williams & Happé, 2009b), all of which supports the suggestion that individuals with ASD show diminished psychological self-awareness. In contrast, several lines of evidence suggest that awareness of the physical self is relatively unimpaired in individuals with ASD. For example, individuals with ASD typically show undiminished performance on mirror self-recognition tasks(Ferrari & Matthews, 1983),

---

[3] This chapter has been adapted from Grainger, C., Williams, D.M. & Lind, S.E. (2013). Online action monitoring and memory for self-performed actions in autism spectrum disorder. *Journal of Autism and Developmental Disorder, 44* (5), pp. 1193-1206. It has been modified to fit the format of this thesis.

delayed video self-recognition tasks (Lind & Bowler, 2009a), and action imitation tasks (Hamilton, Brindley, & Frith, 2007).

However, some researchers have queried whether awareness of physical aspects of self is truly undiminished in ASD, citing studies that apparently show diminished "action monitoring" in ASD (Russell & Jarrold, 1998, 1999).  Russell and Hill (2001, p.317) define action monitoring as, "the mechanisms that ensure that agents know, without self-observation, (a) for which changes in perceptual input they are responsible and (b) what they are currently engaged in doing".  As such, action monitoring allows an individual to distinguish those changes in perceptual experience that are "self-caused" from those that are externally-caused.  Thus, action monitoring gives rise to the experience of agency.  If individuals with ASD do show impairments in action monitoring this contradicts theories that suggest awareness of the physical self is not impaired in ASD, despite limitations in awareness of psychological aspects of self (e.g., Lind, 2010; Williams, 2010).

Action monitoring ability is commonly assessed through tasks that examine an individuals' ability to monitor and correct their own errors.  Typically, individuals are able to correct errors so rapidly that they cannot simply be relying on visual feedback alone.  Instead correcting errors at this speed is thought to depend on monitoring so called "efference copies" of motor plans.  This enables errors to be corrected before a motor command for the particular action is initiated.  Typically, error correction problems are found in individuals with schizophrenia (e.g., Frith & Done, 1989), and are normally interpreted as reflecting diminished action monitoring .  Studies have also indicated that individuals with ASD show impairments in correcting errors (e.g., Russell & Jarrold, 1998).  As such, this is one source of evidence that suggests action monition may be impaired in ASD.  Another source of evidence, which has been taken as evidence that ASD involves diminished action monitoring ability, concerns findings from studies that have assessed relative memory for self-performed actions versus memory for observed actions.  It is well established that neurotypical individuals show

reliably superior memory for actions that they themselves have performed than actions that they have observed other people perform (e.g., Baker-Ward et al., 1990; Engelkamp, 1998). Superior memory for self-performed actions over other-performed actions is referred to as the "enactment effect" and is thought to result from additional motoric components involved in performing an action leading to those actions being more deeply encoded than observed actions (e.g., Engelkamp & Zimmer, 1989). The fact that several studies have reported reduced or absent enactment effects in ASD has led to the suggestion that ASD may be characterised by diminished action monitoring (Farrant et al., 1998; Hala et al., 2005; Millward et al., 2000; Russell & Jarrold, 1999; Wojcik et al., 2011; Zalla et al., 2010). Indeed, some studies have even reported reversed enactment effects (i.e., an "observer effect"; superior memory for observed actions than self-performed actions) in ASD, suggesting a marked atypicality in physical self-awareness in this disorder (Millward et al., 2000; Russell & Jarrold, 1999). All of these studies have led researchers to conclude "the reduced enactment effect in adults with AS reveals an impaired action monitoring system" (Zalla et al., 2010, p.6).

However, there are grounds for questioning whether the results from these studies, and thus the conclusion that action monitoring is diminished in ASD, are indeed valid and reliable. In the studies by Farrant et al., (1998) and Hala et al., (2005), participants with ASD showed a flat profile of memory for self-performed and observed actions. However, in both of these studies comparison participants did *not* show a significant enactment effect either. Rather, they too showed a flat profile of performance and there were *no significant between-group differences* in this respect in either of the studies. Therefore, the failure of individuals with ASD to show an enactment effect in these studies cannot be taken to support the view that action monitoring/physical self-awareness is diminished in this disorder. Instead, because neurotypical individuals in these studies also failed to show enactment effects, the failure to find an enactment effect among participants with ASD is likely to be due to methodological issues associated with the procedure/stimuli used in the studies.

In the study by Millward et al., (2000), comparison participants did show a significant enactment effect, whereas participants with ASD showed an atypical observer effect. This does indicate that action monitoring may be impaired in individuals with ASD. However, one major problem with this study is that Millward et al., (2000) did not match ASD and comparison participants for VIQ. Although the groups were matched for verbal mental age, the comparison group had a mean chronological age that was seven years below that of the ASD group. Thus, as Lind (2010) highlights, participants with ASD had VIQ scores that were approximately *54 points* below those of comparison participants. As such, it is simply not possible to compare meaningfully the experimental task performance of ASD and comparison participants in Millward et al.'s study. Matching for VIQ is essential in studies of cognitive function in individuals with ASD. It is possible that differences between groups in this respect can potentially entirely explain between-group differences in Millward et al.,'s experimental task (see Mervis & Klein-Tasman, 2004).

In an attempt to overcome this limitation, Millward et al., (2000) conducted a second study. This study assessed a sample of children with intellectual disability who, like the ASD group in their Study 1, had lower verbal mental ages than chronological ages. However, no ASD or neurotypical comparison groups were included, and the group with intellectual disability in Study 2 was *not* comparable to the ASD or neurotypical groups from Study 1 in terms of either verbal mental age or chronological age. Furthermore, the group of children with intellectual disability in Study 2 experienced a different set of events in different locations to those used in Study 1. As such, it is difficult to draw any meaningful conclusions from this study. In a more recent study Wojcik et al., (2011) also failed to find significant enactment effects among children with ASD. In contrast, on the same task neurotypical children showed a clear enactment effect. However, the study does not report whether groups were matched for verbal intelligence. As is the case with Millward et al., (2000) it thus

remains unclear whether group differences in VIQ could potentially account for group differences in memory performance.

It is difficult to draw conclusions from the studies by Farrant et al., (1998), Hala et al., (2005), Millward et al., (2000) and Wojcik et al., (2011) because of the methodological problems that are arguably inherent in the design of each study. In contrast, the studies by Russell and Jarrold (1999), and Zalla et al., (2010) both used sound experimental procedures. In both studies, individuals with ASD and comparison participants were closely matched for age and verbal intelligence, and comparison participants in each study did show significant enactment effects. As such, if reliable, the findings of reduced/reversed enactment effects in ASD in these studies provide a serious challenge to theories that action monitoring ability is typical in individuals with ASD. However, there is reason to question the reliability of the results reported by Russell and Jarrold. Using a slightly modified version of Russell and Jarrold's original task, Williams and Happé (2009a) found that participants with ASD showed a typical enactment effect in a source memory task. They did not observe significant differences between the (well-matched) groups of ASD and comparison participants, in this respect. This suggests that Russell & Jarrold's findings (1999) may not be replicable.

The fact that Williams & Happé (2009a) could not replicate the results of Russell and Jarrold (1999) highlights the importance of replicating methodologically rigorous studies with well-designed methods. With this in mind, the current study represents an attempt to replicate, the only other methodologically rigorous study that has failed to find an enactment effect in ASD; Zalla et al., (2010). Zalla and colleagues explored whether adults with ASD would show an enactment effect when their memory was tested for actions they themselves had performed than actions that they had observed someone else perform in a video clip. Participants' memory for performed and observed actions was tested using three memory tasks; a free recall task, a recognition task and a source memory task. Zalla et al., (2010) found that participants with ASD showed similar performance to control participants on the

recognition and source memory tests, showing better memory for actions they had performed than actions they had observed being performed by someone else.  However on the *free recall task* participants showed no significant difference in the proportion of enacted actions they recalled than the proportion of observed actions they recalled.  In contrast, control participants recalled significantly more actions that they had performed than those they had observed.

As such, in order to provide clearer evidence of whether action monitoring abilities are diminished in ASD, the first experimental task reported in this chapter attempted to replicate the findings of Zalla et al.,. In so doing, the study aimed to test whether Zalla's findings were reliable, or whether, as was the case with Russell and Jarrold's (1999) findings, they could not be replicated.  If individuals with ASD do show impairments in their ability to monitor their own actions then you would expect to find a reduced or absent enactment effect on this task.  However if, as predicted, action monitoring remains unimpaired in individuals with ASD, then performance on the task should be similar in both the neurotypical and ASD participants.

As a second experimental task, a version of the "online" action monitoring task employed by Williams and Happé (2009a; Experiment 1) was also included.  According to Russell (Russell & Jarrold, 1999), tasks which require individuals to discriminate online between their own actions and actions initiated by something/someone else provide a direct measure of action monitoring ability.  Following Russell and Hill (2001), Williams and Happé employed a task that involved participants moving a computer mouse (which was placed inside a box, obscuring it from view) and were asked to decide which, from a number of moving coloured squares displayed on a computer screen, was the stimulus being controlled by their own hand movements.  Success on the task relied on participants deciding which of the movements on the screen corresponded with their own proprioceptively experienced movements.  The study also included a second "Other" condition.  In this condition participants placed their hand on the computer mouse, but the movements of the mouse were

controlled by the experimenter.  Thus, in this condition, participants experience no motor intentions for the movements of the mouse in the Other condition, and so cannot rely on feelings of agency to determine which of the stimuli is being controlled by the mouse.  For an individual with an unimpaired sense of their own agency, this condition should be significantly more challenging than the Self condition.  In contrast, if individuals are unable to accurately monitor their own actions then it should not matter who controls the mouse, because in both cases participants cannot rely on an experience of agency to perform the task, and instead can only rely on their ability to match felt actions with the observed consequences of these actions.  Williams and Happé (2009a) did not observe any significant between-group difference in either the level or pattern of performance shown by individuals with and without ASD on the task.  Additionally, (when diagnostic groups were collapsed) Williams & Happé (2009a) found that performance on the *self* condition of the Squares task was significantly associated with source memory for *self* performed actions, independently of verbal mental age.  The better participants' action monitoring ability on the squares task the greater the enactment effect shown by participants on the memory task.  Thus, Williams and Happé found a direct link between online action monitoring and the enactment effect.  This squares task was included to provide an additional measure of action monitoring ability to the measures used by Zalla et al., (2010).  Including this test of action monitoring also allowed us to investigate whether action monitoring ability as assessed by an online measure relates to action monitoring ability assessed by the enactment effect.  It was predicted that individuals with ASD would show similar performance on the Self condition of the task as comparison participants, and that both groups would find the Self condition considerably easier than the Other condition.  It was also predicted that individuals' action monitoring ability on both experimental tasks would be related.  All predictions were in keeping with suggestions that, regardless of whether it is assessed online or via memory, action monitoring ability should be unimpaired in individuals with ASD.

## Method

**Participants**

Ethical approval for this study was obtained from Durham University ethics committee. Seventeen adults with ASD and 17 neurotypical comparison adults took part in this study, all of whom gave written informed consent before participating. Participants in the ASD group had all received formal diagnoses of autistic disorder (*n* = 4) or Asperger's disorder (*n* = 13), according to conventional criteria (American Psychiatric Association, 2000; World Heath Organisation, 1993). Participants with ASD were recruited via an advertisement on The National Autistic Society website; ASD support groups; Durham University Service for Students with Disabilities; and word of mouth. The majority of comparison participants were recruited through advertisements in local newspapers. However, a small number took part in order to receive course credits in partial fulfilment of their undergraduate psychology degrees.

Fifteen of the 17 participants in the ASD group were administered with the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000). The ADOS is an in-depth observational assessment of ASD characteristics. Two participants did not wish to complete the ADOS, because they did not feel comfortable being filmed. The mean ADOS total score for the ASD group was in the autism range (see Table 15). All participants who completed the ADOS received a total score ≥7, above the defined cut-off for ASD (Lord et al., 2000). Additionally, all participants in the ASD group completed the Autism-spectrum Quotient (AQ; Baron-Cohen et al., 2001), a self-report questionnaire that assesses ASD characteristics. Fourteen out of 17 participants scored above the defined cut-off for ASD on the AQ (total score ≥26; Woodbury-Smith et al., 2005). Three participants did not self-report a score above this cut-off. However, all three of these participants scored well above the defined ASD cut-off on the ADOS (all scored ≥12).

*Table 15: Participant Characteristics (Means, Standard Deviations and Inferential Statistics).*

| | Group | | | | |
| --- | --- | --- | --- | --- | --- |
| | ASD (*n* = 17) | Neurotypical (*n* = 17) | *t* | *p* | Cohen's *d* |
| Age | 29.11 (10.55) | 29.43 (14.95) | 0.07 | .944 | 0.02 |
| VIQ | 113.06 (13.83) | 111.00 (10.77) | 0.48 | .632 | 0.17 |
| PIQ | 110.71 (15.58) | 113.24 (11.20) | 0.54 | .591 | 0.19 |
| FSIQ | 114.53 (14.51) | 113.59 (10.94) | 0.21 | .832 | 0.07 |
| AQ Total Score | 33.71 (9.23) | 12.65 (5.22) | 8.19 | <.001 | 2.81 |
| ADOS Total Score* | 11.13 (3.58) | | | | |

AQ: Autism-spectrum Quotient; ADOS: Autism Diagnostic Observation Schedule
*Based on 15/17 participants

All comparison participants completed the AQ and all scored below the defined cut-off for ASD.  No participants, in either the ASD or neurotypical group, reported using any psychotropic medication.  Additionally, none of the participants reported a history of having a neurological or psychiatric condition (apart from ASD).  The participant groups were closely equated for verbal and non-verbal ability, as assessed using the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999).  Groups were also closely matched for chronological age (see Table 15 for participant characteristics).

**Design and Procedure**

**Enactment effect task.**  The procedure for the enactment task employed by Zalla et al., (2010) was replicated as closely as possible.  This task consisted of a study phase and test phase.  During the study phase participants were presented verbal descriptions and videos of 30 actions, 15 of which they simply observed being performed and 15 of which they observed

being performed *and* performed themselves.  Participants were informed that after the study phase they would be asked about what they had heard, seen and acted out, but were not explicitly told that their memory for the action phrases would be tested.  Eight different 30-item lists of actions were used during the task, and list presentation was counterbalanced across all participants.  Each 30 item-list consisted of actions phrases drawn from an overall set of 60 action phrases.  Four lists were created using action phrases from one half of this set, and another four lists were created using action phrases from the other half of this set.  Each list presented the action phrases in a different randomised order, and in each list 15 actions were randomly selected as the actions assigned to be performed.  All eight lists were equated for mean syllable length of items and mean spoken word frequency as indexed by Kucera and Francis (1967), all of which were reported in the MRC Psycholinguistic Database (Coltheart, 1981).  The adequacy of this matching was confirmed by a non-significant effect of List in a multivariate ANOVA (using Wilks' Lambda criterion) that included syllable length and word frequency as the dependent variables, $F$ (14, 464) = 0.31, $p$ = .993, $\eta_p^2$ = .01.

During the study phase actions were presented to participants on a computer screen which participants stood approximately 1.5 meters away from.  Participants were instructed that the beginning of each trial would be signalled by the presentation of either a green or red dot at the top of the screen.  Both green and red dots were identical in size (2x2cm) and appeared at the top of the computer screen for the entirety of a trial.  After the dot (either red or green) had been presented for 1000ms it was followed by a recording of a male voice describing an action phrase (e.g., "pour some water").  All voice clips used in the task were 2000ms long.  Immediately after auditory presentation of the action phrase participants were then presented with a video of an actor performing the appropriate action.  All video clips were 6000ms long, during which a male actor acted out the appropriate action and then adopted a neutral stance for the remainder of the clip. Figure 12 shows a representation of stimuli presentation on each trial of the task.

Participants were instructed that if a green dot appeared on the screen at the start of a trial they should listen to the action and then simultaneously mimic the action described while the video was being presented. If a red dot appeared at the start of a trial participants were instructed simply to listen to the action phrase and then watch the video clip of the actor performing the phrase. During these trials participants were asked to stand in a neutral position. At the end of each trial the experimenter clicked the mouse, after which the next trial began immediately. During green dot trials the experimenter moved onto the next trial only after the video had been fully presented and they had observed the participant adequately perform the appropriate action.



1000ms Cue
Green signals 'Perform'
Red signals 'Observe'

2000ms Voice clip

6000ms Video clip

Experimenter
ends trial
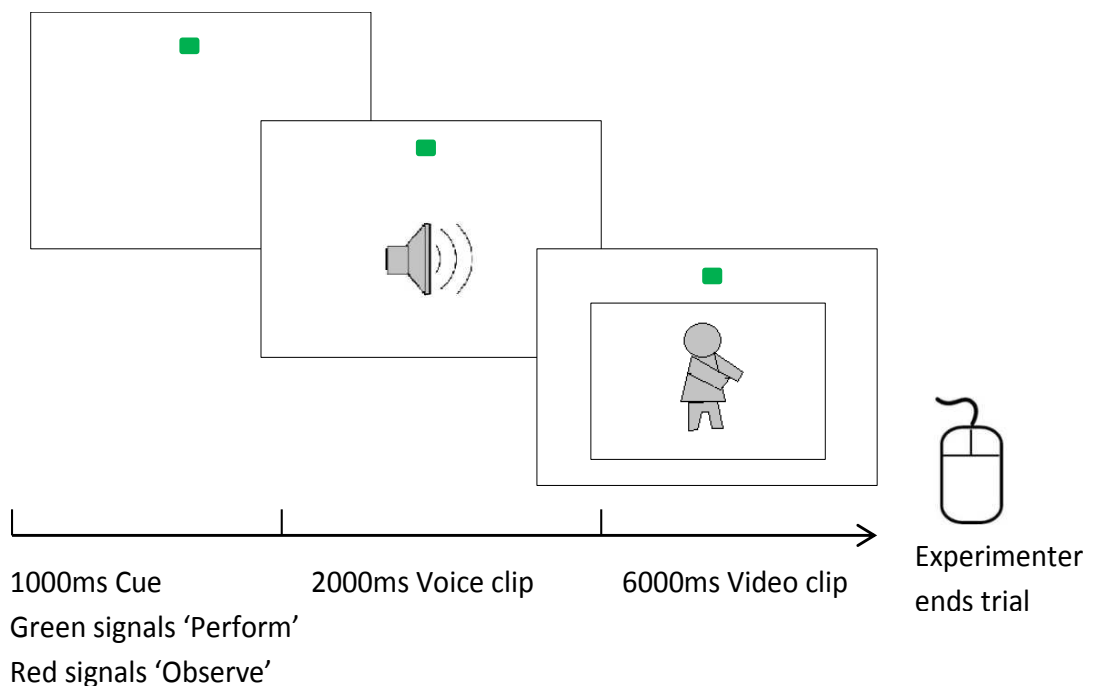
*Figure 12: Graphical representation of the procedure used during a trial of the action monitoring task.*

After a short 5 minute break participants completed the test phase. Firstly participants' memory for the action phrases was tested using a basic recall task in which participants were given five minutes to write down as many of the action phrases as they could remember. Secondly participants completed a recognition and source memory task. In

this task participants were presented with 60 action phrases (15 they had performed, 15 they had observed and 30 novel "lure" phrases). One of the remaining stimuli lists was used to make up the 30 novel actions phrases, and so novel phrases were matched to the "old" action phrases for syllable length and word frequency (as stated above). Participants were asked to judge whether each action phrase was "old" (had been presented to them previously during the study phase) or "new" (had not been presented in the study phase). If participants thought that an action phrase was old they were also asked to decide whether they thought the action was one they had performed, or one they had observed being performed but not performed themselves. During this task participants were presented with each action phrase individually on the computer screen, and an experimenter recorded their responses, and then moved onto the next trial.

**Action monitoring task.** The action monitoring task employed in this study was based on the task used in Williams and Happé (2009a). There were two conditions (Self and Other) in the task. In each condition, a series of different coloured squares moved across a computer screen. All the squares moved whenever the mouse was moved and froze whenever the mouse was not being moved. However, during each trial only one of the squares (the target square) consistently moved in accordance with the movements of the mouse. In other words, only the target square was directly controlled by the participant. The remaining squares in each trial (the distractor squares) moved in a pseudo-random fashion.

In both the Self and Other conditions of the task there were a total of 18 levels, which increased in difficulty (Table 16 summarises each level of the task). Participants completed five 30s trials at each level, and moved onto the next level only if they successfully completed more trials than would be expected by chance. For example at level 1 four squares were presented on the screen, one of which was the target square and three of which were distractor squares. Across five trials, if an individual randomly chose squares, by chance they would be expected to successfully identify the target square on one in every four trials. As

such, to pass level 1 a participant had to successfully identify the target square at least twice (in five trials) to perform above chance on this level. The number of distractor squares increased as levels got harder. At level two there were 8 distractors, at level three there were 15 distractors, at level four there were 24 distractors, at level five there were 35 distractors and at level six there were 48 distractors.

Task difficulty was also manipulated by varying the degree of similarity between the movement of the target square and the movements of the distractor squares. The vector movements of the distractor squares could be varied between 0° to 360° relative to the target square. If the movements of the distractor squares varied by 0° degrees from the target

*Table 16: Stimulus characteristics for each level of the action monitoring task.*

| Level | No. of distractor squares | Minimum no. of trials required to move onto next level (out of 5) | Distractor movement arc (°) |
|---|---|---|---|
| 1 | 3 | 2 | 360 |
| 2 | 8 | 1 | 360 |
| 3 | 15 | 1 | 360 |
| 4 | 24 | 1 | 360 |
| 5 | 35 | 1 | 360 |
| 6 | 48 | 1 | 360 |
| 7 | 3 | 2 | 180 |
| 8 | 8 | 1 | 180 |
| 9 | 15 | 1 | 180 |
| 10 | 24 | 1 | 180 |
| 11 | 35 | 1 | 180 |
| 12 | 48 | 1 | 180 |
| 13 | 3 | 2 | 90 |
| 14 | 8 | 1 | 90 |
| 15 | 15 | 1 | 90 |
| 16 | 24 | 1 | 90 |
| 17 | 35 | 1 | 90 |
| 18 | 48 | 1 | 90 |

square all the squares would move the same, and so the task would be impossible.  In

contrast, if the movement of the distractor squares varied by 360° the distractor squares

could move in any direction relative to the movement of the target square.  If participants

successfully completed levels 1 to 6 with the distractor squares set at a movement arc of 360°

they moved onto level 7, which was the same as level 1 (i.e., one target and three distractor

squares), but the movements of the distractor squares were contained to an 180° movement

arc.  Levels 8, 9, 10, 11, and 12 were the same as levels 2, 3, 4 5, and 6, respectively, expect

that the movements of the distractor squares were also contained to a 180° arc.  If

participants completed these levels successfully they moved onto level 13, which again was

the same as level 1 but with the movements of the distractor squares contained to a 90° arc.

Again levels 14, 15, 16, 17, and 18 were the same as levels 2, 3, 4 5, and 6 respectively, except

that the movement of the distractor squares was now restricted to 90°.  Table 16 summarises

each level of the task.

Participants always completed the Self condition of the task first.  In this condition

participants moved the mouse and were instructed to press the spacebar once they thought

they had identified that target square.  Once the spacebar was pressed all squares on the

screen froze and did not move when the mouse was moved, allowing participants to then click

on the square they thought was the target square.  The next trial began immediately after

participants clicked on a square.  If participants did not press the spacebar within 30s the trial

ended and participants moved onto the next trial.  This was considered an incorrect trial.

Before participants completed the experimental trials, they were shown a demonstration of

the task.  The experimenter demonstrated two trials at level 1, explaining to the participant "I

think I know which square I am controlling so I am going to press the spacebar.  Then I can

indicate which square I think I was controlling by clicking on it".  Participants then completed

two practice trials at level 1 followed by the experimental trials.  On all trials the mouse was

placed inside a cardboard box, which could be reached through openings at both ends of the

box. This obscured vision of both the mouse and the participant's hand and was used to ensure that participants did not succeed at the task simply by matching their hand movements with the movements of the square on the screen. If participants did not successfully complete enough trials for their performance to be better than chance at a particular level a "Game Over" screen appeared, signalling the end of that condition.

After completing the Self condition participants then completed the Other condition of the task. This condition was identical to the Self condition except that as well as the participant placing their hand on the mouse, the experimenter took hold of the mouse from the opposite end and gripped the top of the mouse with their index and thumb fingers. This allowed the experimenter to control the movements of the mouse. The participant was instructed to allow the experimenter to control the movements of the mouse, and not to try and move the mouse themselves. During each trial the experimenter continuously moved the mouse, first up and down and then left and right. The same mouse movements were standardised across all participants and trials. Once the participant thought they had identified the target square they were instructed to press the spacebar. Like before, all squares on the screen froze and the participant were then able to control the movements of the mouse, and click on the square they thought was the target square. As before, the experimenter demonstrated the task and participants then completed two practise trials at level one, before beginning the experimental trials.

**Scoring**

For the enactment effect memory task participants' recall performance was calculated as the proportion of actions individuals correctly recalled, both for enacted actions and observed actions. As a measure of recognition performance two separate corrected hit rates were calculated[1]. Corrected hit rates were calculated using the formula H-FA, where H represents hit rate (the proportion of old items participants *correctly* identifying as "old") and

FA represents false alarm rate (the proportion of new actions participants *incorrectly* identifying "old"). Two corrected hit rate scores were calculated, using separate hit rates for enacted and observed actions. A single false alarm rate was used to calculate both corrected hit rates, since false alarm rates were derived from performance on distractor items which by definition were neither enacted nor observed. Source monitoring performance was calculated as the proportion of action phrases participants made correct source attributions for (i.e. correctly identified performed actions as "performed", and observed actions as "observed"), for both enacted and observed action phrases.

During both conditions of the action monitoring task the computer automatically recorded which squares participants clicked on during each trial, and whether this was the correct target square. For each participant the total number of successfully completed trials (from a maximum of 90 trials) and the total number of successfully completed levels (from a maximum of 18 levels) in both the Self and Other condition was calculated.

## Results

### Enactment Task

**Free Recall.** Table 17 shows the average proportion of actions that individuals in the ASD and neurotypical group recalled correctly, for both enacted and observed actions. A 2 (Group: ASD/neurotypical) × 2 (Condition: Enacted/Observed) ANOVA was conducted on the proportion of actions correctly recalled. A significant main effect of Condition was found, $F (1, 32) = 28.42$, $p < .001$, $\eta_p^2 = .47$, reflecting the fact that across groups individuals recalled significantly more actions that they had performed than actions they had observed. There was no significant main effect of Group, $F (1, 32) = 1.31$, $p = .261$, $\eta_p^2 = .04$, and no significant interaction between Group and Condition $F (1, 32) = .01$, $p = .906$, $\eta_p^2 < .01$. This reflected the fact that participants in both groups showed the same pattern of performance on the recall

task, and recalled significantly more enacted actions than observed actions (showing a

statistically significant enactment effect).

*Table 17: Memory performance on the recall, recognition, and source monitoring tasks*

*for enacted and observed action phrases in the ASD and neurotypical groups.*

|  |  | Group | |
|---|---|---|---|
|  |  | ASD | Neurotypical |
| Recall performance | Enacted | .63 (.13) | .62 (.14) |
|  | Observed | .37 (.13) | .38 (.14) |
| Recognition performance | Enacted | .95 (.09) | .97 (.05) |
|  | Observed | .83 (.14) | .81 (.13) |
| Source monitoring performance | Enacted | .95 (.06) | .96 (.07) |
|  | Observed | .80 (.20) | .81 (.11) |

**Recognition.** Corrected hit rates for enacted and observed actions among ASD and

comparison participants are reported in Table 17[4].  A 2 (Group: ASD/neurotypical) × 2

(Condition: Enacted/Observed) mixed ANOVA was conducted on these data.  A significant

main effect of Condition was found, $F (1, 32) = 58.96$, $p < .001$, $\eta_p^2 = .65$, reflecting superior

recognition of enacted items than observed items.  There was no significant main effect of

---

[4] It should be highlighted that Zalla et al., (2010) used the nonparametric measures of $A'$ and $B''_D$ to assess participants recognition discrimination. However, when $A'$ scores were calculated for recognition performance in this study, one sample $t$-tests indicated that scores did not significantly differ from ceiling level accuracy (100% discrimination accuracy) for enacted actions in both in the TD group, $t (16) = 1.97$, $p = .06$, and ASD group, $t (16) = 1.87$, $p = .08$.  As such, to maximise the rigour of our statistical analysis, corrected hit rates were used as an alternative measure of recognition performance on the task.

Group, $F$ (1, 32) = 0.004, $p$ = .952, $\eta^2_p$< .01, and no significant interaction between Group and condition $F$ (1, 32) = 2.09, $p$ = .158, $\eta^2_p$ = .06.  This reflected the fact that both groups showed a similar pattern of recognition performance, demonstrating better recognition discrimination for enacted actions than actions they had observed.

**Source Monitoring.**  Table 17 also shows the average proportion of actions that participants in the ASD and neurotypical group made correct source memory judgements for, for both enacted and observed actions.  A 2 (Group: ASD/neurotypical) × 2 (Condition enacted/observed) mixed ANOVA was conducted.  A significant main effect of Condition was found, $F$ (1,32) = 25.9, $p$ <.001, $\eta^2_p$ = .45, reflecting the fact that, across groups, individuals made more correct source judgements for enacted actions than actions they had observed.  There was no significant main effect of Group, $F$ (1, 32) = 0.04, $p$ = .844, $\eta^2_p$< .01, and no significant interaction between Group and condition $F$ (1, 32) = .01, $p$ = .944, $\eta^2_p$< .01.  This reflected the fact that both groups made more correct source monitoring judgments for enacted actions than observed action.  A series of one sample $t$-tests was carried out, to establish whether performance on the enactment task was at floor or ceiling level, for any of the memory measures.  These $t$-tests indicated that, in both the ASD and neurotypical groups, performance on the recall, recognition, and source monitoring tasks significantly differed from floor or ceiling level performance, all $t$s > 2.38, all $p$s < .030.

To summarise, on all three tests of memory (recall, recognition, and source monitoring) participants in *both* the ASD and neurotypical groups showed better memory for actions that they had enacted than actions they had observed.  This pattern of memory performance did not differ between ASD participants and neurotypical participants on any measure, as indicated by *no* significant interactions between participants' diagnostic group and their memory for enacted/observed actions.

**Squares Task**

Table 18 shows the mean number of levels and trials completed in both the Self and Other conditions of the task, for both ASD and neurotypical participants. Firstly a 2 × 2 mixed ANOVA was carried out, with the number of *trials* completed in each condition (Self/Other) entered as a within-subjects variable and diagnostic group (ASD/neurotypical) entered as a between subject variable. There was a significant main effect of condition on the number of trials completed, $F(1, 32) = 75.66$, $p < .001$, $\eta_p^2 = .70$. This reflected superior performance in the Self condition than the Other condition. There was no significant main effect of group, $F(1, 32) = 0.46$, $p = .503$, $\eta_p^2 = .01$, indicating, across both conditions, that participants in the ASD group showed similar performance to comparisons participants. There was also no significant interaction between group and condition, $F(1, 32) = 1.45$, $p = .237$, $\eta_p^2 = .04$.

Another 2 × 2 mixed design ANOVA was carried out, with the number of *levels* participants successfully completed in each Condition (Self/Other) entered as a within-subjects variable, and Group (ASD/neurotypical) entered as a between subjects variable. Again, there was a significant main effect of condition on the number of levels successfully completed, $F(1, 32) = 87.23$, $p < .001$, $\eta_p^2 = .73$. There was no main effect of Group, $F(1, 32) = 1.54$, $p = .223$, $\eta_p^2 = .05$. However, the interaction between Group and Condition was marginally significant, $F(1, 32) = 3.95$, $p = .056$, $\eta_p^2 = .11$. To investigate this interaction further a series of independent sample *t*-tests was carried out. These indicated that within the Self condition there was no significant group difference in the number of levels successfully completed, $t(32) = 1.67$, $p = .105$, $d = 0.57$. This was also the case in the Other condition, $t(32) = 1.63$, $p = .112$, $d = 0.56$. Paired-samples *t*-tests showed there was a significant difference between the number of levels completed in the Self condition, relative to the Other condition, in both the ASD group, $t(16) = 4.83$, $p < .001$, $d = 2.87$, and the neurotypical group, $t(16) = 8.73$, $p < .001$, $d = 1.75$. The significant interaction between Group and Condition

appeared to be driven by the relatively larger difference between performance on the Self and Other conditions of the task shown by ASD participants relative to neurotypical participants. On average individual in the neurotypical group completed 7.29 (6.22) more levels on the Self condition of the task relative to the Other condition, whereas individuals in the ASD group completed on average 11.24 (5.31) more levels on the Self condition relative to the Other condition. Figure 13 and Figure 14 show the number of participants in each group that successfully completed each level of the task, in both the Self and Other conditions.

*Table 18: Mean (standard deviation) number of levels and trials completed in the Self and Other condition, by both the ASD and neurotypical group.*

|  |  | Group | |
|---|---|---|---|
|  |  | ASD | Neurotypical |
| Trials Completed | Self | 40.59 (20.77) | 33.53 (22.43) |
|  | Other | 3.59 (2.98) | 5.53 (4.00) |
| Levels Completed | Self | 12.59 (5.42) | 9.41 (5.69) |
|  | Other | 1.35 (1.11) | 2.12 (1.58) |
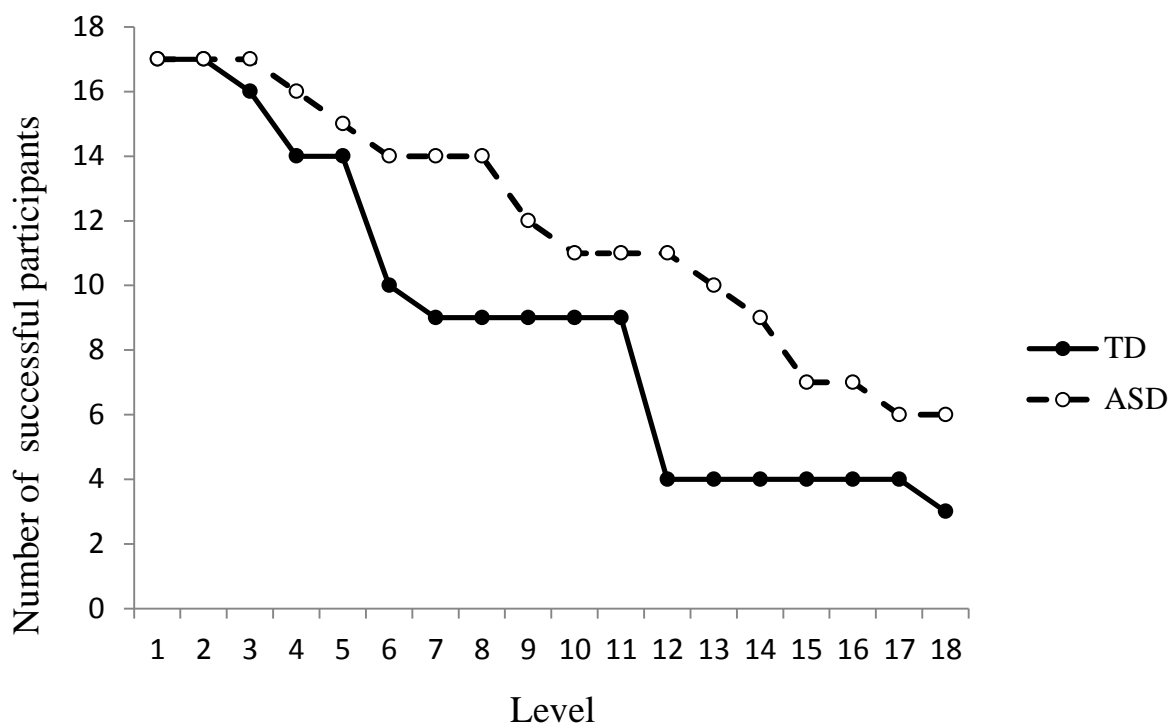
F*igure 13: Number of participants in each group who successfully completed each level in the "Self" condition of the action monitoring task.*
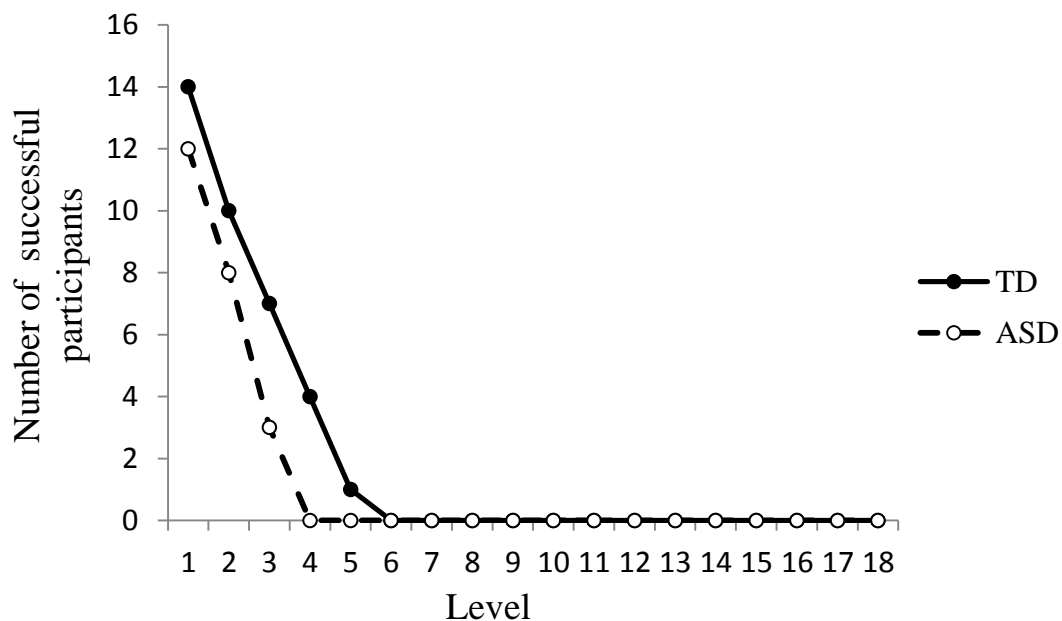


Figure 14: Number of participants in each group who successfully completed each level of the "Other" condition of the action monitoring task.

**Relations between action monitoring ability on the Enactment task and Squares task**

On the Squares task, the size of the effect of action monitoring was calculated by subtracting the number of successful trials participants made in the Other condition from the number of successful trials participants made in the Self condition of the task. This difference score represented the size of the advantage of action monitoring on the Squares task. The greater the score, the more sensitive action monitoring ability was on the task.

On the Enactment task, the size of the enactment effect participants showed was calculated by subtracting participants' memory scores for observed actions from their memory scores for enacted actions. Three separate difference scores were calculated for 1) recall memory, 2) recognition memory and 3), source monitoring memory. For all three scores the greater the score, the greater the memory advantage for enacting actions than observing them.

To investigate the relation between the effect of action monitoring on the Squares task and extent to which participants showed an enactment effect on the Enactment task, a series of Pearson's correlations was conducted. There was no significant relation between the size of the effect of action monitoring (on the Squares task) and the extent to which participants showed enactment effects on the recall, recognition and source monitoring tests (on the Enactment task), among either ASD or comparison participants, all $r$s < .28, all $p$s > .279.

### Discussion

Individuals with ASD showed no evidence of action monitoring impairments in this study. Results from the Squares task indicated that the ASD group were *as* able as comparison individuals to detect which square was controlled by their own actions. Importantly, both groups of participants also found it significantly easier to identify the target square when it was controlled by their own intentional movements than when the movement of the target square was controlled by the experimenter. Thus, individuals with ASD were able to monitor

their own motor commands and benefit from the feelings of agency that were unique to the

Self condition of the task. Indeed, there was some evidence that individuals with ASD were

somewhat more sensitive to their agency than comparison participants; when the number of

successful *levels* completed on the Squares task was taken as the dependent measure of

performance. These results are in keeping with previous suggestions that individuals with

ASD may in fact make *more* efficient use of non-visual, motor cues than neurotypical

individuals, and rely relatively less on visual cues (Frith & Hermelin, 1969). This suggests

that, far from being impaired, individuals with ASD might show heightened physical self-

awareness. Regardless of whether this is the case, the results from this study certainly do not

suggest that individuals with ASD are impaired at online action monitoring.

During the enactment task, as predicted, adults with ASD showed better memory for

actions they had enacted than actions they had observed, on *all three* memory tests. On the

recall, recognition, and source monitoring tests, memory performance was both qualitatively

and quantitatively similar in the ASD and neurotypical groups; individuals with ASD showed

enactment effects of a closely similar magnitude to those shown by neurotypical individuals.

This is in keeping with many previous studies that have reported finding typical enactment

effects in individuals with ASD (e.g., Hare et al., 2007; Lind & Bowler, 2009c; Maras et al.,

2012; Summers & Craik, 1994; Williams & Happé, 2009a), and also in keeping with many

studies in the broader action monitoring literature (e.g., Blakemore et al., 2006; David et al.,

2008). Indeed, with respect to the recognition and source memory tasks, these results

replicate those of Zalla et al., (2010); in both this study and in Zalla et al.'s study, typical

enactment effects for recognition and source memory were observed among individuals with

ASD. Nonetheless, in their discussion of their data, Zalla et al., mainly focussed on the only

difference they observed between the groups, which was in free recall only. It was in this

respect the results of the study did *not* replicate Zalla et al.'s results.

In the introduction it was argued that only two studies (Russell & Jarrold, 1999; Zalla et al., 2010) have reported finding that individuals with ASD do not show a typical enactment effect, using samples we know are well matched. Russell and Jarrold's study could not be replicated by Williams and Happé (2009a). Similarly, in this paper, the only finding that the study failed to replicate was atypical memory performance in individuals with ASD on the recall test (the only results in the paper that in any way indicated action monitoring is impaired in ASD). As such, it is argued that this study, and the enactment effect literature in general, provides support for theories that suggest action monitoring should remain unimpaired in individuals with ASD (e.g., Lind, 2010; Williams, 2010).

A number of studies have now explored whether individuals with ASD demonstrate typical enactment effects. A summary of the results from these studies is shown in Table 19. This table shows that, in fact, the majority of these studies are in keeping with the results of this study, and report similar performance in ASD and comparison participants (with both groups either showing an enactment effect of similar magnitude, or both groups showing no enactment effect). Across these studies, when you look at the average size of the memory advantage individuals demonstrate for self-performed items than other-performed items, there is no significant difference between ASD and neurotypical participants, $t$ (38) = .40, $p$ = .692, $d$ =0.11. On average, across all studies looking at the enactment effect, ASD participants remembered 10% more actions that they performed than actions they observed. Comparably, on average, neurotypical participants remembered 11% more actions when they enacted them. This lends more support to the view that action monitoring (and physical self-awareness, more generally) is undiminished in ASD.

Potential reasons for why both Russell & Jarrold (1999) and Zalla et al., (2010) found discrepant results, compared to other studies of the enactment effect in individuals with ASD, should be considered. One potential reason for discrepancies between studies could be differences in the developmental ability of participants in different studies. It is possible that

problems with action monitoring may be evident in individuals with ASD who have low verbal

mental ages.  The average verbal mental age (VMA) of the children with ASD in Russell &

Jarrold's study (VMA= 7.13 years) was lower than the level among children with ASD in

Williams & Happé (2009a) study (VMA= 8.44 years), potentially explaining why only Williams

and Happé observed an enactment effect among their sample of ASD participants.  However,

the developmental level (i.e., VMA) of participants is unlikely to be the sole explanation for

differences across studies, because other studies have reported a typical enactment effect

among children with ASD whose VMAs were on average *lower* than participants in Russell &

Jarrold's study.  For example, Lind and Bowler (2009b) found that a sample of children with

ASD with an average VMA of 6.66 years showed typical enactment effects than comparison

participants.  An alternative explanation for Russell and Jarrold's failure to observe an

enactment effect among individuals with ASD may be to do with the verbal intelligence, rather

than developmental level, of participants.  That is, problems with action monitoring may be

evident only among intellectually low-functioning individuals with ASD.  In Russell and

Jarrold's study, participants with ASD had a mean VIQ of only approximately 53.89[5], which is

notably low, relative to other relevant studies.  However, it is also notable that a large

proportion of the individuals with ASD in Williams and Happé's sample would still be

considered relatively intellectually low-functioning (the mean VIQ among participants with

ASD was 73.50), but nonetheless showed typical enactment effects.  Moreover, reference to

VIQ does not explain why adults in Zalla et al's study (mean VIQ =114.2) did not show an

---

[5] The average VIQ for participants in the ASD group is not reported in Russell and Jarrold (1999). However, Russell and Jarrold (1999) reports the average VMA (7.13 years) and the average chronological age (CA; 13.23) for participants in the ASD group. These were used to estimate the average VIQ of the ASD group, using the formula VIQ = VMA/CA × 100.

enactment effect on the recall task.  As such, developmental differences in VMA/ VIQ cannot fully explain discrepancies between each study's findings.

Within the literature, three studies (the current study; Williams and Happé 2009a; Russell & Hill, 2001) have now used the Squares task (or a variation of it) to investigate action monitoring ability in individuals with ASD.  All three of these studies find convergent results, indicating undiminished task performance among adults and children with ASD.  That being said, the current study did not find a significant correlation between the effect of action monitoring on the Squares task and the extent to which participants showed an enactment effect on the action memory task , in either the ASD or the neurotypical group.  In this respect, this study did not replicate Williams and Happé's findings (2009a).  The enactment tasks that was used was based as closely as possible on the method used in Zalla et al., (2010) and thus was not the same task as the task used by Williams & Happé.  As such, it may be the case that the enactment task used in the current study measured distinct aspects of action monitoring to that used in Williams and Happé.  Furthermore, this study assessed action monitoring ability in adults, whereas Williams & Happé assessed action monitoring ability in children. Any of these differences between the study designs/methods could explain the discrepancy between the results in terms of this specific finding.  However, what is more notable is the high consistency of findings across these two studies.  Despite the differences in methods used, the results from the two studies converge in most respects.

In terms of the broader literature concerning self-awareness in ASD, the majority of evidence (including the results from both tasks in this study) suggests that physical self-awareness is undiminished in ASD.  In contrast, studies have indicated that autobiographical (episodic) memory (Crane & Goddard, 2008; Crane, Pring, Jukes, & Goddard, 2012), and episodic future thinking, is impaired in ASD ( Lind, Williams, Bowler, & Peel, Lind & Bowler, 2010; 2014), and arguments have been put forward to suggest that such impairments may (partially) result from impairments in self-awareness (see Lind, 2010).  However, diminished

memory for personally experienced events, and the diminished ability to imagine events in the future, is likely to rely on an awareness of a psychological, temporally-extended self. Thus, it appears that individuals with ASD may demonstrate selective impairments only in psychological self-awareness.

Theoretically, the results of this study can also inform cognitive theories surrounding the sense of agency. Within the literature debate exists concerning how different aspects of social cognition, such as agency, imitation and mentalising relate to one and other. Arguably, if a sense of agency acts as a precursor to mentalising ability (Russell, 1996), then stereotypical mentalising deficits in individuals with ASD (e.g., Baron-Cohen et al., 1985; Yirmiya et al., 1998) should be associated with similar deficits in action monitoring. These results, alongside results from several other studies (e.g., David et al., 2008; Hamilton et al., 2007; Sebanz, Knoblich, Stumpf, & Prinz, 2005) suggests that a dissociation may exist between the sense of agency and mentalising ability, which is stereotypically impaired in individuals with ASD. Instead, these results support suggestions (e.g., David et al., 2008) that social-cognitive deficits associated ASD occur at a higher-level than that needed for a sense of agency.

As well as having theoretical implications, establishing the extent of action monitoring abilities in ASD has practical importance. It is well established that neurotypical individuals show better memory for information they have enacted. On the basis that individuals with ASD also benefit from self-enactment then strategies aimed at improving learning and memory in ASD should focus on encouraging approaches that capitalise on this particular memory strength (i.e., encouraging motor participation in the learning process). More generally, it may be possible to improve everyday functioning in individuals with ASD if memory can be enhanced by physical self-enactment. Additionally, establishing whether individuals with ASD show an intact ability to accurately determine the source of actions gives rise to potential forensic implications. For example, being able to accurately recall an event

one was involved in, and distinguish between actions carried out by oneself and others, influences the reliability of eyewitness testimonies (see Maras et al., 2012).  This study indicates that action monitoring is a relative strength in ASD, something that should be taken into account in future research.

*Table 19: Summary of studies reporting memory for self-performed items in individuals with ASD and neurotypical comparison participants.  This table reports the group size for each study, and the average difference in memory performance between the proportion of self-performed items remembered compared to the proportion of other-performed items remembered.*

| | Memory Test | Report a significant difference in the size of the enactment effect shown by ASD and neurotypical participants | Proportion change across conditions (Self-Other) | | *n* | |
| --- | --- | --- | --- | --- | --- | --- |
| | | | ASD | Neurotypical | ASD | Neurotypical |
| Summers & Craik (1994)[a] | Free Recall | No | .20 | .05 | 8 | 8 |
| | Recognition | No | .17 | .24 | 8 | 8 |
| Farrant, Boucher & Blades (1998) | Source memory | No | -.18 | -.06 | 15 | 15 |
| Russell & Jarrold (1999)[b] | Source memory | Yes | -.03 | .02 | 22 | 22 |
| Millward, Powell, Messer & Jordan (2000)[c] | Recall | Yes | / | / | 12 | 12 |
| Hill & Russell (2002) | Source memory | No | .06 | .01 | 20 | 20 |
| Hala, Rasmussen & Henderson (2005) | Source memory | No | .16 | .05 | 13 | 13 |
| Hare, Mellor & Azmi (2007) | Free recall | No | .06 | .12 | 12 | 14 |
| | Cued Recall | No | .17 | .23 | 12 | 14 |
| Williams & Happé (2009a)[b] | Source memory | No | .04 | -.01 | 16 | 16 |
| Lind & Bowler (2009b) | Recognition | No | .19 | .22 | 53 | 50 |
| | Source memory | No | .12 | .07 | 53 | 50 |
| | Free Recall | Yes | .08 | .24 | 18 | 18 |
| Zalla, Daprati, Sav, Chaste, Nico & Leboyer (2010) | Recognition | No | .10 | .14 | 18 | 18 |
| | Source memory | No | .09 | .13 | 18 | 18 |
| Wojcik, Allen, Brown & Souchay (2011) | Free Recall | Yes | .00 | .03 | 16 | 16 |
| Maras, Memon, Lambrects & Bowler (2012) | Free recall | No | .25 | .20 | 18 | 18 |
| | Cued Recall | No | .02 | .05 | 18 | 18 |
| Grainger, Williams & Lind (current study) | Free Recall | No | .26 | .24 | 17 | 17 |
| | Recognition | No | .12 | .16 | 17 | 17 |
| | Source memory | No | .15 | .15 | 17 | 17 |
| | Average across all studies (SD): | | .10 (.10) | .11(.09) | 18.46 (11.02) | 18.38 (10.15) |
| | Total no.  of participants in all studies: | | | | 240 | 239 |

[a] In this study there was no other-performed condition, but instead self-performed items were compared to verbally presented items.

[b] For both Russell & Jarrold (1999) and Williams & Happé (2009a) the statistics reported here refer to differences in memory for cards turned over by participants themselves (on behalf of themself and their doll) compared to cards turned by another (on behalf of the experimenter and the experimenter's doll).  In the case of Russell & Jarrold (1999) the statistics refer to proportion change across both conditions (expected/unexpected) of the task.

[c] Millward et al., (2000) do not report means and standard deviations for their results, making it not possible to calculate effect sizes and proportion differences between the Self and Other condition.

**CHAPTER SEVEN**

## THE INTENTION-SUPERIORITY EFFECT IN CHILDREN WITH AUTISM SPECTRUM

## DISORDER

In chapter six, evidence was presented that suggested adults with ASD demonstrate typical enactment effects, demonstrating superior memory for actions they performed themselves than for actions they viewed someone else perform.  However, within the typically developing literature studies have shown that it is not just memories for actions one *has* performed in the past that hold a privileged status in memory, but also memory for actions one *intends* to perform in the future (e.g., Goschke & Kuhl, 1993; Marsh, Hicks, & Bink, 1998).  For example, in a seminal study, Goschke and Kuhl (1993) found that across four experiments, participants were significantly faster and more accurate at recognising words from descriptions of actions participants intended to perform at a later point, relative to words from descriptions of actions participants did not intend to perform.  This "intention-superiority effect" appears to be robust, and a series of studies have now shown that individuals tend to demonstrate superior memory for content associated with uncompleted intentions, evidenced by better recall (Jahn & Engelkamp, 2003; Koriat, Ben-Zur, & Nussbaum, 1990; Maylor, Chater, & Brown, 2001), better recognition (Jahn & Engelkamp, 2003), faster recognition latencies (Goschke & Kuhl, 1993) and faster lexical decision latencies (Marsh et al., 1998) for content associated with actions individuals intend to perform than actions they do not intend to perform.

The ISE directly involves processing one's future intentions.  However, whilst this effect certain involves some aspect of self-awareness, it does not necessary rely on

forming meta-representations of one's own intention as such. Instead, it is possible that this effect relies on imagining oneself completing the action at a future point. If this is the case, then it is probable that intention superiority relies on intact extended self-awareness.

**The intention superiority effect and extended self-awareness**

Extended self-awareness involves an awareness of the self that encompasses one's present self, past self and future self. As such extended self-awareness involves the understanding that several alternative representations of the self can reflect different representation of the same enduring self (across time). One explanation of the ISE it that it relies on extended self-awareness and occurs due to prospection, sometimes also referred to as episodic future thinking. Episodic future thinking involves imagining anticipated future events (Atance & O'Neill, 2005; Buckner & Carroll, 2007). Through the simulation of possible future scenarios, one can hypothetically test alternative plans of action which may considerably improve behavioural flexibility and self-control (Suddendorf & Corballis, 2007). It is distinctly possible the future thinking system plays an important role in forming/encoding intentions. Whilst encoding a prospective intention it may be necessary to mentally imagine the future activity that one is going to perform. Future thinking might increase the salience of one's intention, explaining the ISE. Theoretically, it has been suggested (see Gilbert, Armbruster, & Panagiotidi, 2012) that the chances of a delayed intention being carried out increase when the context an intention is encoded in is similar to the context it will be retrieved in (Gilbert et al., 2012). Studies have suggested that imagining a future activity produces similar brain activity to actually performing that activity (see e.g., Stokes, Thompson, Cusack, & Duncan, 2009). As such, whilst encoding a prospective intention,

mentally imagining performing that intended action in the future may increase the chances of the intention being carried out, as imagining performing the future intention should increases the similarity between encoding and retrieval contexts.

Intuitively, it makes sense that the same cognitive processes are engaged during the ISE and episodic future thinking.  However, surprisingly this relation has never been explicitly examined among neurotypical adults or adults with ASD.  As such, the question of whether the episodic future thinking system is responsible for successful intention-superiority effects remains hypothetical, and still needs to be established.

**Can the ISE be considered an *enactment superiority* effect?**

Whilst parallels can be drawn between the cognitive processes engaged during the ISE and episodic future thinking several other theories have been proposed to explain the intention superiority effect.  Another plausible and very different explanation for the ISE has suggested that this effect occurs due to encoding additional motor information associated with the to-be-performed action (Freeman & Ellis, 2003; Koriat et al., 1990).  More specifically, it has been suggested that when an individual encodes an intention to perform an action, they activate motor information associated with performing that action in the future.  Thus, the future action is encoded in multiple formats, including an action-based format.

As discussed in chapter six, it is well established that neurotypical individuals show significantly better memory for actions that they have performed than actions that they have simply read or observed (Baker-Ward et al., 1990; Engelkamp, 1998).  Such superior memory for self-performed actions is referred to as the "enactment effect" and is thought to result from additional motoric components involved in performing an action leading to that action being more deeply encoded than actions simply studied

(Engelkamp & Zimmer, 1989).  In a similar fashion, it has been suggested that the ISE could be considered an *intention enactment* effect, and superior memory for intended actions is the result of sensorimotor information encoded when an intention is formed (Freeman & Ellis, 2003).

An action superiority interpretation of the ISE gains support from a study conducted by Freeman and Ellis (2003, experiment 3).  In this experiment Freeman and Ellis manipulated whether participants simply read actions or performed them during the encoding stage (to assess the ISE), but also manipulated whether participants expected to perform these actions at a later point or whether they expected to simply verbally recall the actions.  Freeman and Ellis reasoned that if the ISE occurred due to sensorimotor information being encoded, participants should demonstrate no difference in their memory for actions they intended to perform but had not enacted than actions they intended to perform but had enacted.  Additionally, they predicted that participants would demonstrate general "action superiority" effects and would show similar memory advantages for actions whenever the actions were to be enacted (either during enactment at encoding or enactment at a later point).  Interestingly, Freeman and Ellis found exactly this pattern of memory in neurotypical adults lending support to their suggestion that the ISE is underpinned by enactment superiority.

However, whether "action superiority" can adequately account for the ISE is not yet certain.  It has been argued that often the specific motor components associated with a future intention are not known to an individual when they encode the intention, and some intentions may be too rich/complex to be represented through sensorimotor encoding (McDaniel & Einstein, 2007).  McDaniel and Einstein (2007) question whether individuals engage in sensorimotor encoding of complex intention such as the intention

to take a trip (which itself encompasses the intention to perform several actions e.g., to pack your bag, check in to your flight, cancel your mail etc.).

**Intention superiority in individuals with ASD**

To date, no study has investigated the ISE directly among people with ASD. However, there are reasons to predict that individuals with ASD will demonstrate diminished intention superiority effects.  A growing body of evidence suggests individuals with ASD demonstrate impairments in extended self-awareness.  Evidence of this comes from studies that report individuals with ASD demonstrate impairments in prospective memory tasks (i.e., tasks that require participants to actually carry out an intended action at specific point in time; see e.g., Altgassen, Koban, & Kliegel, 2012; Williams, Boucher, Lind, & Jarrold, 2013; Williams, Jarrold, Grainger, & Lind, 2014), as well as studies that find individuals demonstrate impairments on measures of EFT ability (see e.g., Lind & Bowler, 2010; Lind, Bowler, & Raber, 2014; Lind, Williams, et al., 2014; Terrett et al., 2013).  One potential explanation for such these findings is the suggestion that individuals with ASD are impaired at mentally projecting the self into a different, future situation (reflecting diminished extended self-awareness).  Given the earlier suggestion that the ISE may directly rely on EFT/extended self-awareness, there are thus reasons to believe that individuals with ASD will demonstrate reduced or absent intention superiority.

However, the prediction that individuals with ASD will show a diminished ISE is based on the theory that this effect relies on EFT/ extended self-awareness.  If, instead, the ISE is merely a variant of the enactment effect (see Freeman & Ellis, 2003), then there is no reason to suppose that it will be diminished in ASD (given the robust evidence of undiminished enactment effects in this disorder; see chapter six).  Thus, the

study of the ISE in ASD has the potential to inform general theory concerning the cognitive underpinnings of this effect.

**The Current Study**

To date, no study has explored the ISE in individuals with ASD. As such the primary aim of this study was to assess whether individuals with ASD demonstrate superior memory for actions they intend to perform. To explore the ISE, children with ASD were presented with a series of written action phrases (e.g., "turn on the spot"). Participants were asked either to read the action statement aloud (Read condition), perform the action stated (Perform condition), or to plan to perform the stated action at the end of the task (Plan condition), thus forming an intention to perform that action. After being presented with these actions, participants' memory for actions in all three conditions was assessed. As such, the study employed a task that assessed both the ISE and the enactment effect in children with ASD and neurotypical children. Thus, the current study aimed to explore the ISE in ASD, but also aimed to extend the results of research conducted in chapter six, which explored the enactment effect in adults with ASD. Whilst the results of chapter six indicated that adults with ASD demonstrate typical enactment effect, it is possible (though no predicted) that children with ASD may demonstrate diminished enactment effects, and impairments in enactment effects are developmental in nature. By exploring the enactment effect alongside the ISE effect, this study addressed this issue.

In keeping with the results of chapter six, it was predicted that children with ASD would also demonstrate typical enactment effects on the task. However, given research that suggests extended self-awareness may be impaired in ASD it was predicted that the ISE would be diminished among children with ASD, relative to the ISE

among age and IQ-matched neurotypical participants.  This prediction was however

tentative, given the debate concerning the underlying processes engaged during

intention superiority.

Additionally, by exploring the ISE alongside the enactment effect in individuals

with ASD the current study provides the potential to inform the theoretical debate

surrounding the underlying processes involved in the ISE.  As discussed, individuals

with ASD appear to demonstrate intact physical self-awareness and motor planning

skills (see chapter six) but diminished extended self-awareness.  If individuals with ASD

do demonstrate typical intention superiority, this would support motor planning

theories of the ISE.  In contrast, diminished intention superiority effects in ASD, despite

intact enactment effects, would support suggestions the ISE is not simply an extension

of the enactment effect.  Instead this pattern of results would support the suggestion

that the ISE involves self-prospection, and episodic future thinking.  It is of course

possible that children with ASD may demonstrate typical intention superiority effects,

yet such effects will rely on atypical underlying processes.  It was predicted that, if

children with ASD did demonstrate typical intention superiority effect, these effect

would relate differently to motor planning skills, relative to neurotypical comparison

children.

**Method**

**Participant**

Twenty two children with ASD and 20 neurotypical comparison children took

part in this experiment, after their parents had given written, informed consent.

Participants in the ASD group had formal diagnoses of Autistic Disorder or Asperger's

disorder, according to established criteria (American Psychiatric Association, 2000;

World Health Organisation, 1993).  To assess severity of ASD features, parents of

participants with ASD completed the Social Responsiveness Scale (SRS; Constantino et

al., 2003).  In all but one case, participants with ASD scored above the defined cut-off for

ASD on the SRS (total score ≥ 60; Constantino et al., 2003).  The remaining participant

scored 55 on the SRS, which is just below the conventional ASD cut-off of 60.  This

participant had a formal diagnosis of Autistic Disorder.  Parents of neurotypical children

also completed the SRS.  All but one participant in the neurotypical group scored below

the defined cut-off for ASD, with one participant scoring just above the cut off (66).

After removing these participants from analyses, none of the results (or the study

conclusions) changed substantively, and thus these participants were included in

analysis.  Using the Wechsler Abbreviated Scale of Intelligence (WASI; Psychological

Corporation, 1999), the groups were equated closely for verbal IQ (VIQ), performance

IQ (PIQ), and full-scale IQ (FSIQ).  Both groups were also equated closely for

chronological age.  Participant characteristics are presented in Table 20.

*Table 20: Participant Characteristics (Means, Standard Deviations and Inferential Statistics).*

| | Group | | $t$ | $p$ | Cohen's $d$ |
|---|---|---|---|---|---|
| | ASD ($n$ = 22 ) | Neurotypical ($n$ = 20) | | | |
| Age (years) | 13.42 (1.12) | 13.22 (1.01) | 0.62 | .539 | 0.19 |
| VIQ | 106.00 (19.34) | 106.90 (14.43) | 0.21 | .838 | 0.05 |
| PIQ | 106.05 (12.90) | 109.80 (14.48) | 0.89 | .379 | 0.27 |
| FSIQ | 106.73 (11.84) | 109.50 (15.00) | 0.67 | .508 | 0.20 |
| SRS Total Score | 83.59 (9.87) | 43.25 (7.86) | 14.45 | <.001 | 4.52 |

SRS: Social Responsiveness Scale (Constantino et al., 2003); VIQ = verbal IQ; PIQ = performance IQ; FSIQ = full scale IQ.

**Materials and Procedures**

    **Intention Superiority Task.** The Intention Superiority Task consisted of a study phase and a test phase. During the study phase participants were presented with 45 action phrases, 15 of which they read (Read condition), 15 of which they performed (Enactment condition), and 15 of which they made the intention to perform at the end of the task (Intend condition). Three different 15-item lists of action phrases (e.g., "rub your stomach") were used as stimuli during the study phase of the task. A set of 45 novel actions was also compiled, which was used for the purpose of providing "lure" items during the test phase of the task. All four lists were equated for mean syllable length and mean spoken word frequency of action phrases, as indexed by Kucera and Francis (1967) and reported in the MRC Psycholinguistic Database (Coltheart, 1981). The adequacy of this matching was confirmed by a non-significant effect of List in a multivariate ANOVA (using Wilks' Lambda criterion) that included syllable length and word frequency as the dependent variables, $F$ (3, 86) = .20, $p$ = .894, $\eta_p^2$ = .007. During the study phase, each of the 15 item lists was assigned to a different condition (Read, Intend, or Enactment), and the order in which lists were assigned to each condition was counterbalanced across all participants. This created six different conditions of the task. During each condition, action phrases were presented to participants in a different, pseudo-randomised order, in which no more than two action phrases from any one condition appeared on successive trials (i.e., participants were never presented with more than two action phrases from the same condition one after another).

    Figure 15 shows a graphical representation of stimuli presentation on each trial of the task. During the study phase, actions were presented to participants on a computer screen which participants stood approximately 1 meter away from.

Participants were instructed that the beginning of each trial would be signalled by the presentation of one of three instructions (Read/Plan/Perform), which would appear individually at the top of the screen. Participants were told that after the instruction word had been presented, it would be followed by an action phrase, presented directly below the instruction. Participants were told that if the instruction "Read" appeared on the screen they should read the action phrase aloud. If the instruction "Perform" appeared on the screen participants were told that they should mime (act out) the action phrase. Finally, participants were instructed that if the instruction "Plan" appeared on the screen they should make a plan to perform the action at the end of the task. During each trial of the task the instruction word was presented individually on the screen for 1500ms, followed by the action phrase, which was presented below the instruction word for another 5000ms.

Before completing the study phase of the task, participants completed a practice task, which consisted of six trials (two "Read" trials, two "Intend" trials, and two "Enactment" trials). None of the action phrases that appeared in the practice task appeared during the experimental trials. After participants had completed the practice task they completed the experimental trials. During the experimental trials the experimenter observed the participant to check that they performed the appropriate action on Enactment trials of the task. If a participant did not perform an action during an Enactment trial the experimenter took note of this trial and the trial was subsequently removed during data analysis. Similarly, if participants accidentally performed an action during a "Read" or "Intend" trial the experimenter took note of this, and this action phrase was also removed during data analysis. Errors in following the condition instructions correctly were very rare (and only occurred on three trials). Participants were informed before starting the experimental trials that after the study

phase had been completed they would be asked some questions about what they had read, planned and performed, but were not explicitly told that their memory for the action phrases would be tested.

After the study phase, participants completed the test phase. Participants' recognition and source memory for the action phrases was tested during this phase. During the test phase, participants were shown an action phrase individually on the computer screen and were asked to judge whether each action phrase was "old" (had been presented to them previously during the study phase) or "new" (had not been presented in the study phase). If participants responded that an action phrase was old, they were asked to decide whether they thought the action was one that they had read, one they had planned to perform, or one they had performed during the task. The experimenter recorded participants' responses.
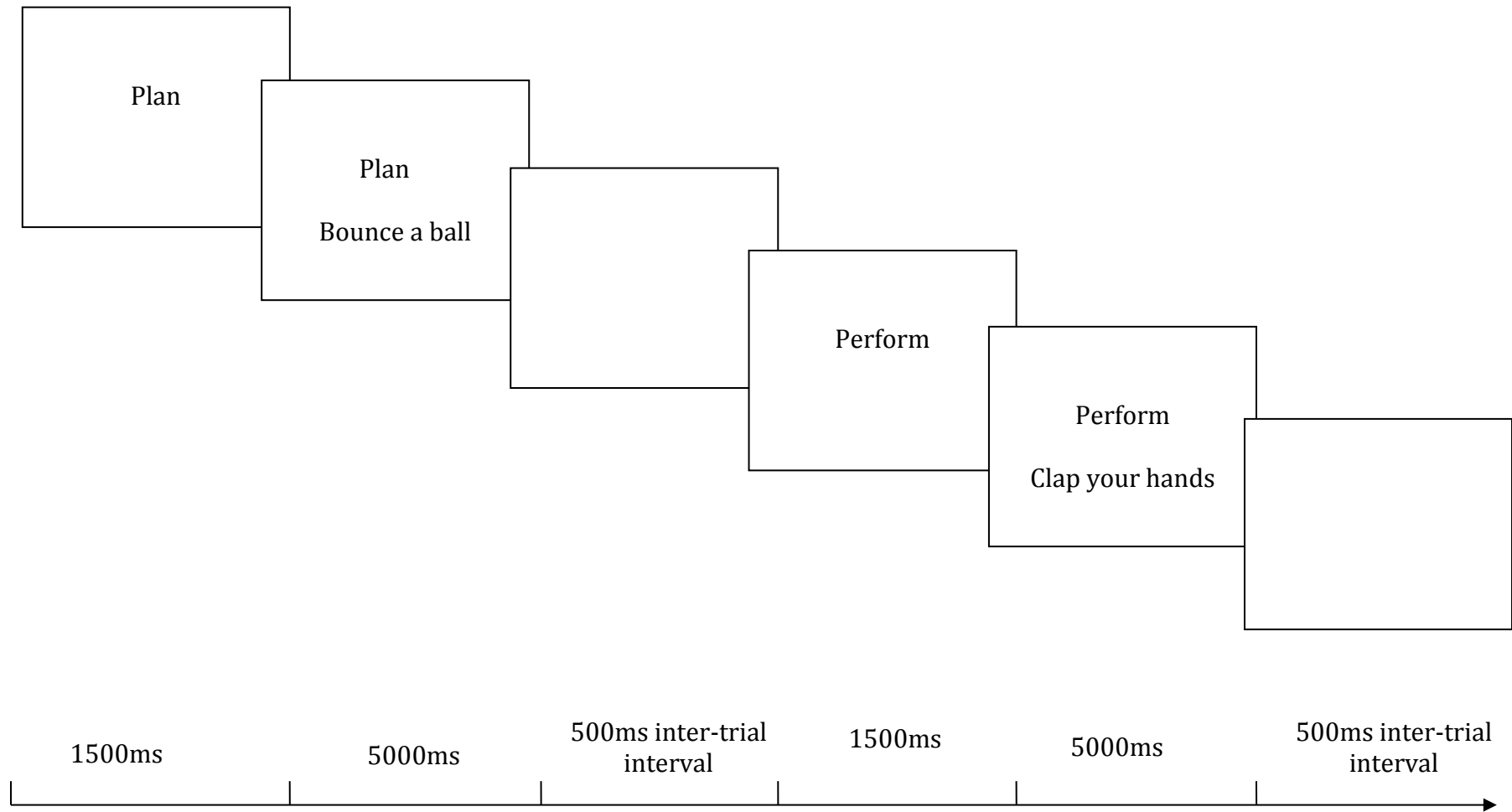
*Figure 15: Graphical representation of the procedure used during the study phase of the Intention Superiority Task (providing an example of two trials).*

**Scoring**

**Intention Superiority Task.** Performance on the Intention Superiority Task was analysed using the measure of $d$-prime ($d'$), a parametric measure of item discrimination. As a measure of recognition performance three separate $d$-prime scores were calculated using the formula below. In this equation H represents a hit rate (the proportion of old items participants *correctly* identifying as "old") and FA represents false alarm rate (the proportion of new actions participants *incorrectly* identifying "old").

$$d' = z(H) - z(FA)$$

Three separate $d'$ scores were calculated, using separate hit rates based on the proportion of performed actions correctly recognised, the proportion of planned actions recognised, and the proportion of read actions recognised. A single false alarm rate was used to calculate all three $d'$ scores, since false alarm rates were derived from performance on distractor items, which by definition were neither read, planned nor performed during the task. Source monitoring performance was also assessed using $d'$ and scores were calculated separately based on hit rates for the proportion of read actions participants made correct source monitoring judgements for, the proportion of planned actions participants made correct source monitoring judgments for and the proportion of performed actions participants made correct monitoring judgements for.

For the purpose of correlation analyses, two difference scores were calculated; one to indicate the size of the enactment effect and one to indicate the size of the intention superiority effect. The difference between $d'$ for actions participants performed (enactment condition) minus $d'$ for actions they had simply read (read condition) was used to calculate the size of participants enactment effect. The

difference between $d'$ for actions participants had planned to perform at the end of the task (intend condition) minus $d'$ for actions participants read (read condition) represented the size of participants intention superiority effect.  These difference scores were calculated for both recognition and source memory performance.

## Results

**Intention Superiority Task**

*Table 21: Means (SDs) and inferential statistics for group differences in performance on the intention superiority task (d').*

| Memory Type | Action condition | Group | | | | |
|---|---|---|---|---|---|---|
| | | ASD (*n*=22) | Neurotypical (*n*=20) | *t* | *p* | Cohen's *d* |
| Recognition | Read | 1.75 (0.73) | 2.02 (0.65) | 1.25 | .220 | 0.39 |
| | Intend | 2.06 (0.60) | 2.26 (0.65) | 1.04 | .304 | 0.32 |
| | Enactment | 3.03 (0.83) | 3.38 (0.74) | 1.48 | .147 | 0.45 |
| Source Monitoring | Read | 0.84 (0.83) | 1.13 (0.94) | 1.05 | .300 | 0.36 |
| | Intend | 1.76 (0.83) | 1.94 (0.79) | 0.70 | .488 | 0.19 |
| | Enactment | 2.67 (0.84) | 2.94 (0.88) | 1.00 | .322 | 0.31 |

Performance in each condition of the intention superiority task in the ASD and neurotypical group can be seen in Table 21.  Before analysing group difference in performance a series of one-sample *t*-tests was carried out, to establish whether performance on the intention superiority task was at floor or ceiling level, for any of the

memory measures.  These *t*-tests indicated that, in both the ASD and neurotypical

groups, the proportion of actions correctly recognised in each condition significantly

differed from floor or ceiling level performance, all *t*s ≥3.47 , all *p*s ≤.003, as did the

proportion of actions participants made correct source monitoring judgements for all *t*s

≥ 4.55, all *p*s < .001.

**Recognition memory.**  D-prime scores for recognition performance in each

condition of the task (Read/Plan/Perform) are reported in Table 21, among both the

ASD and neurotypical group.  A 2 (Group: ASD/neurotypical) × 3 (Condition:

Read/Intend/Enactment) mixed ANOVA was conducted on these data.  A significant

main effect of Condition was found, $F$ (2, 80) =118.50, $p$ < .001, $\eta_P^2$ = .75.  This reflected

the fact that across both groups, recognition memory was significantly greater for

actions participants had enacted than for actions participants had planned, and

recognition memory for actions participants had planned was significantly greater than

actions participants had read, all *t*s ≥ 3.13, all *p*s ≤ .003, and all *d*s ≥ 0.42 (see Figure 16

for a graphical representation of these results).

However, there was no significant main effect of Group, $F$ (1, 40) =2.09, $p$ = .156,

$\eta_P^2$= .05, and no significant interaction between Group and Condition $F$ (2, 60) = 0.38, $p$ =

.685, $\eta_P^2$ = .01.  These results are not in keeping with predictions, and suggest that the

ASD group demonstrated a similar pattern of recognition memory to neurotypical

children on the task (demonstrating typical enactment effects and typical intention

superiority effects).

**Source Monitoring.**  D-prime scores for source monitoring performance in each

condition of the task (Read/ Intend/Enactment) are also reported among both the ASD

and neurotypical group in Table 21.  A 2 (Group: ASD/neurotypical) × 3 (Condition:

Read/Plan/Perform) mixed ANOVA was conducted on these data.  A significant main

effect of Condition was found, $F$ (2, 80) = 122.82, $p$ <.001, $\eta^2_p$ = .75.  This reflected the fact

that participants, across both groups, demonstrated significantly better source

monitoring performance for actions they had performed than for actions they had

planned to perform, and significantly greater source monitoring performance for

actions participants had planned to perform, than actions they had read, all $ts$ ≥7.39, all

$p$s < .001, and all $d$s ≥ 1.03 (see Figure 16 for a graphical representation of these

results).

However, again there was no significant main effect of Group, $F$ (1, 40) = 1.15, $p$ =

.290, $\eta^2_p$= .03, and no significant interaction between Group and Condition $F$ (2, 80) =

0.13, $p$ = .879, $\eta^2_p$ <.01.  As such, there were no significant differences between the groups

in either overall levels of performance or patterns of performance across conditions.
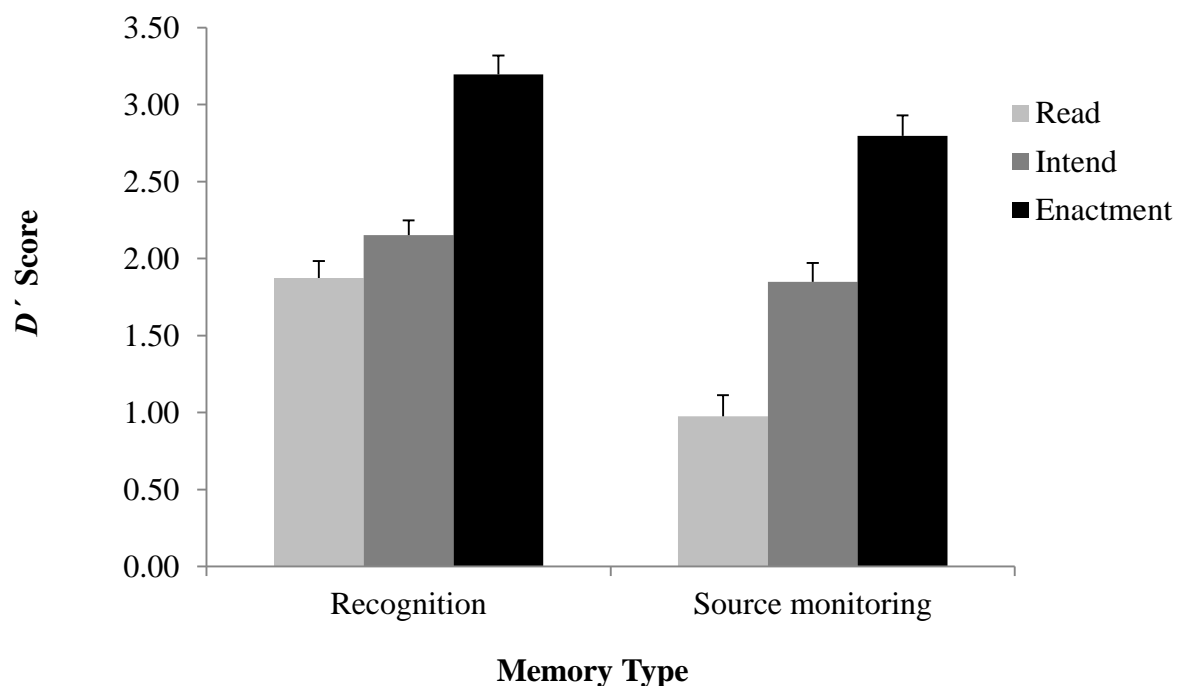


*Figure 16: D′ scores (across both groups) for performance in each condition of the*

*task, for both recognition and source monitoring performance.  Error bars represent*

*standard error of the mean.*

To summarise, on all both tests of memory (recognition and source monitoring) participants in *both* the ASD and neurotypical groups showed better memory for actions that they had enacted than actions they had read or actions they planned to perform (thus demonstrating clear enactment effects). Additionally, on all both tests of memory participants in both groups showed better memory for actions that they had planned to perform than actions they had simply read (demonstrating intention-superiority effects). This pattern of memory performance did not differ between ASD participants and neurotypical participants on any measure, as indicated by *no* significant interactions between participants' diagnostic group and their memory for read/planned/performed action for either recognition or source monitoring performance.

**Associations between the enactment effect and the intention-superiority effect**

A series of correlations was carried out to assess the extent to which the size of the enactment effect was associated with the size of the ISE among either group of participants (see Table 22). Among ASD participants, the size of the enactment effect was strongly positively correlated with the size of the ISE, in both recognition memory, $r = .646$, $p = .001$, and source memory $r = .858$, $p < .001$. For neurotypical participants, the size of participants enactment effect was also strongly positively correlated with the size of the ISE, in both recognition memory, $r = .593$, $p = .006$, and source memory $r = .483$, $p = .031$. A Fisher's $Z$ test indicated that the size of this correlation in recognition performance did not differ between the ASD group and neurotypical group, $z = 0.26$, $p = .79$. However, the relationship between the size of participants enactment effect and their ISE was significantly greater in the ASD group than the neurotypical group, when looked at in source monitoring $z = 2.27$ $p = .02$.

*Table 22: Correlations matrix showing the relationship between the size of the enactment effect and size of the intention superiority effects demonstrated by ASD and neurotypical participants on the intention-superiority task.*

| Group | Experimental Measure | ISE[a]- Source monitoring | ISE- Recognition | EE[b] – Source monitoring | EE- Recognition |
|---|---|---|---|---|---|
| ASD | ISE- Source monitoring | 1 | | | |
| | ISE- Recognition | .76** | 1 | | |
| | EE -Source monitoring | .86** | .69** | 1 | |
| | EE – Recognition | .50* | .65** | .69** | 1 |
| Neurotypical | ISE- Source monitoring | 1 | | | |
| | ISE- Recognition | .39 | 1 | | |
| | EE -Source monitoring | .48* | .12 | 1 | |
| | EE – Recognition | .10 | .59** | .35 | 1 |

*Correlation is significant at the .05 level
**Correlation is significant at the .001 level.
[a] ISE = Size of the intention superiority effect
[b] EE= Size of the enactment effect

## Discussion

Until now, no study has explored the ISE in ASD. As such, the primary aim of this study was to explore the extent to which individuals with ASD demonstrate typical intention superiority effects. In terms of the central experimental finding, the study found that children with ASD demonstrated no evidence of a diminished ISE, which was not in keeping with the study's tentative predictions. In contrast, in source monitoring and recognition memory, individuals with ASD clearly demonstrated superior memory

for actions they intended to perform at the end of the task than actions they had simply read (demonstrating clear intention superiority effects).

In the introduction it was also predicted that children with ASD would demonstrate typical enactment effects.  In keeping with this prediction, this study also found no impairments in the size of the enactment effect demonstrated by children with ASD.  This is in keeping with the results of chapter six, which demonstrated adults with ASD show typical enactment effects, and is also in keeping with the results of several studies from the literature (as reviewed in chapter six).  These results are also in keeping the idea that individuals with ASD do not demonstrate impairments in physical/ecological self-awareness

A second aim of the study was to explore the extent to which children's intention superiority effects related to the size of the enactment effects they demonstrated.  In terms of the correlational results the study found a strong relationship between the size of the enactment effect participants demonstrated on the task and the size of the ISE they demonstrated, in both ASD and neurotypical participants.  These findings provide strong support for the suggestion that the ISE may occur due to encoding additionally motor information associated with the to-be-performed action (Freeman & Ellis, 2003).  Finding a strong positive correlation between the ISE and enactment effect supports an action superiority interpretation of the ISE.  Indeed, these results concur with the results of other studies within the literature, which also suggests that the ISE relies on motor encoding (see e.g., Freeman & Ellis, 2003).  To explore the extent to which the ISE relied on motor encoding Freeman and Ellis carried out an innovative experiment.  During this experiment, participants were presented with two lists of actions, which they were asked to memorise.  Participants were told that their memory for one of these lists would be tested by a recognition test, whereas their memory for the other list

would be tested by a recognition test and a recall test, in which participants would be asked to verbally recall the actions on the list (thus participants formed intentions to recall the actions on this list, but not the other).  However, this study was novel in that immediately after participants studied the list of actions they were given one of two distractor tasks.  Participants were asked either to continuously count backwards from ten (a verbal distractor task) or to draw imaginary circles in the air (a motor distractor task).  Whilst the aim of the motor distractor task was to impair encoding motor information associated with the actions on the list, the verbal distraction task aimed to impair verbal encoding of information associated with the actions on the list.  Freeman and Ellis (2003) found that participants who had performed the motor distraction task after learning (but not the verbal distraction task) failed to demonstrate intention superiority effects.  Alongside the results of this study, the findings of the current study support the idea that the ISE is the result of sensorimotor encoding of one's intended actions.

Importantly, the results indicated that, at least with respect to source memory, the ISE demonstrated by children with ASD related more strongly to the size of the enactment effect they demonstrated, relative to neurotypical children.  This suggests that in the ASD group this effect relied *more* on motor information associated with to-be-performed actions, relative to the neurotypical group.  It is possible that children with ASD engaged in more motor encoding relative to neurotypical participants during the task, given relative strengths in motor processing in ASD and potential impairments in EFT ability (although this suggestion is not empirically supported, given no direct measure of self-prospection/EFT was employed).

The prediction that individuals with ASD would show a diminished ISE was based on the theory that this effect relies on EFT, and mentally imaging oneself perform

an intended action.  Individuals with ASD typically demonstrate impairments on

measures of EFT ability (see e.g., Lind & Bowler, 2010; Lind, Bowler, et al., 2014; Lind,

Williams, et al., 2014; Terrett et al., 2013).  As such, findings that individuals with ASD

demonstrate entirely typical intention superiority effects support the suggestion that

this effect does not solely rely on EFT.  However, one problem with this suggestion is

that studies of EFT in ASD have not always found consistent evidence of impairments in

this disorder.  In one study Lind & Bowler (2010) asked participants with ASD to

provide verbal descriptions of a series of past events, as well as a series of potential

future events.  Lind and Bowler found that descriptions of both past and future events

provided by individuals with ASD were significantly less detailed and specific than those

provided by neurotypical participants.  However in contrast a more recent study by

Crane, Lind & Bowler (2013) found no evidence of EFT (or episodic memory)

impairments.  In this study Crane and colleagues presented participants with the

beginning of sentences (e.g., "Next year I will...") and asked participants to complete the

sentence.  Using this method Crane et al., (2013) found that individuals with ASD

performed similarly to neurotypical individuals.  Given that studies have not

consistently found evidence of EFT impairments in ASD, it is of course possible that EFT

abilities were not impaired in the ASD group, relative to the neurotypical group, tested

in this study.  As such, without a measure of EFT ability in the ASD group, it is not

possible to firmly conclude that the ISE does not rely on EFT.

As well as having theoretical implications concerning the underlying bases of the

ISE, the study informs the literature on self-awareness in ASD.  In terms of the broader

literature concerning self-awareness, the majority of evidence (including the results

from this study and the results of chapter six) suggests that physical self-awareness is

undiminished in ASD.  In contrast, as discussed in the introduction, other studies have

indicated that autobiographical episodic memory (memory for events one has experienced in the past) and episodic future thinking (the ability to imagine event that might plausibly happen at a future time) are likely impaired in ASD. Diminished memory for personally experienced events, and the diminished ability to imagine events in the future, is likely to rely on an awareness of the extended self. The result of this study support suggestions that physical self-awareness is intact in ASD, and that impairments in self-awareness in ASD may be selective in nature. This idea is explored further in the following chapter. The final chapter of this thesis explores self-conception in adults with ASD, and whether individuals with ASD hold both typical psychological self-concepts and typical physical self-concepts.

**CHAPTER EIGHT**

---

## "WHO AM I?" A STUDY OF CONCEPTUAL SELF-AWARENESS IN ADULTS WITH

## AUTISM SPECTRUM DISORDER

Within the field of psychology, the concept of "the self" has been widely discussed. This term is often used to refer to multiple different phenomena and thus a single definition of "the self" is not possible. Indeed, in early conceptualisations James (1890) proposed that the self could not be considered a single entity; instead the self is multifaceted, and consists of several different dimensions including the physical self, mental self, spiritual self, and the ego. Similarly, Neisser (1988) defined five different aspects of the self and self-awareness. One important aspect of the self defined by Neisser is that of the conceptual self. The conceptual self can be considered the self as defined in terms of the theories and assumptions an individual holds about themselves. Just as it is difficult to provide one single definition for the term "the self", so individuals do not hold one single self-concept. Instead, individuals tend to conceptualise themselves in multiple different ways, and define themselves in terms of physical, social, psychological, and emotional characteristics.

Within the typically developing literature, research exploring the development of self-conception is vast. Damon & Hart (1982) provide perhaps the most comprehensive account of the development of self-understanding. These researchers suggest that children as young as two years of age demonstrate behavioural signs that they attribute certain characteristics to themselves and hold some form of self-concept (Kagan, 1982). Studies of self-conception have mostly employed methods that ask individuals to provide verbal descriptions of themselves (Damon & Hart, 1988). Such,

studies have shown that young neurotypical children tend to primarily describe themselves in terms of physical self-concepts, shifting towards describing themselves more in terms of psychological and social self-concepts over the course of the early school years (see e.g., Selman, 1980). By adulthood individuals have typically formed robust and complex self-concepts about themselves.

**Self-conception in Autism Spectrum Disorder**

There are several reasons to predict that individuals with ASD will demonstrate diminished conceptual self-awareness. In early descriptions of childhood ASD both Kanner (1943) and Bosch (1970) noted a lack of self-awareness in children with this disorder, and such observations have been supported by a number of empirical studies that suggest that individuals with ASD lack typical conceptual self-awareness. Early manifestations of conceptual self-awareness can be seen in the use of personal pronouns, which demonstrates an explicit understanding of the distinction between self and other. Several studies have shown that personal pronoun use is atypical in ASD (Hobson et al., 2010; Jordan, 1989; Lee et al., 1994; Lind & Bowler, 2009a; Loveland & Landry, 1986). Additionally, studies have found that individuals with ASD typically show reduced or absent self-reference effects (Henderson et al., 2009; Lombardo et al., 2007; Toichi et al., 2002), another indication that the concept of the self may be diminished in ASD (see chapter 1; p.23).

On a theoretical level, there are also reasons to predict that individuals with ASD will demonstrate impairment in self-knowledge. Several theories (e.g., Williams, 2010; Carruthers, 2009; Frith & Happé, 1999) suggest that individuals with ASD are as impaired at understanding their own mental states as they are at understanding mental states in others. If individuals with ASD are impaired at understanding their own

mental states, it follows they will hold poorly established psychological (but not necessarily physical) self-concepts.  However, whilst previous studies suggest that individuals with ASD demonstrate a diminished awareness of their own knowledge and mental states (Perner et al., 1989; Williams & Happé, 2009b), surprisingly few studies have explored whether individuals with ASD demonstrate diminished and/or qualitatively atypical self-concepts (Kristen et al., 2014; Lee & Hobson, 1998; Tanweer et al., 2010).  Additionally, no studies have explored whether different aspects of conceptual self-awareness in ASD are related to the extent to which individuals can understand mental states (in themselves or other).

Another index of conceptual self-awareness is that of autobiographical semantic memory.  Autobiographical semantic memory refers to an individual's memory for information about themselves.  If individuals with ASD lack self-knowledge, they should demonstrate impairments in autobiographical semantic memory.  Crane and Goddard (2008) assessed both autobiographical semantic memory and autobiographical episodic memory (an individual's memory for personally experienced events) in adults with ASD by asking them series of questions.  Whilst autobiographical episodic questions asked participants to recall specific past events (e.g., Can you tell me all the things you did before you went to bed last night?) autobiographical semantic questions asked participants to remember information about themselves (e.g., Can you tell me where your secondary school was?).  Crane and Goddard's study found that whilst adults with ASD demonstrated clear episodic memory impairments, semantic knowledge remained intact in high-functioning adults with ASD.  Such findings suggest that *conceptua*l self-awareness is not impaired in ASD.  However, a study by Bruck and colleagues (Bruck et al., 2007) assessed children's knowledge of personal facts, and found that children with ASD were significantly poorer than neurotypical children at answering such questions.

This indicates that children with ASD are impaired at autobiographical semantic memory. As such, the results of studies of episodic semantic memory suggest that self-concept development may be *delayed* in ASD, yet by adulthood impairments in semantic memory abilities appear to have recovered (see Lind, 2010). Alternatively, contradictory results may reflect methodological issue with one or more of the studies discussed.

To date, three studies have also explored conceptual self-awareness in ASD by asking individuals (both children and adults) to provide verbal descriptions of themselves (Kristen et al., 2014; Lee & Hobson, 1998; Tanweer et al., 2010). In an early study Lee and Hobson (1998) used Damon and Hart's (1988) self-understanding interview to explore self-understanding in children with ASD. The self-understanding interview is a semi-structured interview that consists of a series of questions participants are asked about themselves (e.g., "What are you like?"; "How did you get to be the way you are?"). During this interview, it was found that children with ASD described themselves significantly more using physical characteristics, relative to neurotypical children. Additionally, whilst children with ASD did not produce fewer psychological descriptions of themselves, those they did provided differed qualitatively from those provided by neurotypical children. Namely, children with ASD produced far fewer psychological descriptions of themselves that referred to social categories (e.g., I like being with friends). Lee & Hobson's study thus provides some evidence for the suggestion that children with ASD demonstrate impairments in aspects of conceptual self-awareness (at least concerning psychological aspects of the self-concept).

In a more recent study Tanweer and colleagues (Tanweer et al., 2010) also explored whether individuals with ASD define themselves in a similar way to neurotypical individuals. This study used the twenty statements task (TST; Kuhn &

McPartland, 1954), an open-ended task in which individuals are asked to respond to the question, "Who am I?", by providing twenty "I am" statements, which they believe provide an accurate description of their long-term self.  These statements were then coded into eight distinct categories; traits, social identities, evaluations, physical descriptions, emotional descriptions, peripheral information, and global descriptions. Additionally, as well as coding statements for specific categories, Tanweer and colleagues explored whether statements provided by individuals were abstract statements, alongside whether statements were autonomous or social (autonomous statements were those that did not make references to others, or specific social contexts).  Tanweer and colleagues found that adults with ASD provided significantly more abstract statements relative to neurotypical adults, and provided statements that were significantly more autonomous.  Findings that adults with ASD provided significantly more autonomous statements than neurotypical adults are in keeping with the results of Lee and Hobson (1998) who reported that children with ASD tended to provide fewer social descriptions relative to neurotypical children.

Additionally, the study found that the statements provided by adults with ASD did not cover as wide a range of categories as those provided by the neurotypical group, suggesting that adults with ASD provide diminished identity complexity, relative to comparison adults.  However, unfortunately the paper does not report whether significant group differences in identity complexity (the number of categories participants' statements fell into) were driven by group differences in specific categories or not.  It is possible that participants with ASD did not produced statements that fell into as many categories as neurotypical participants because they consistently failed to provide descriptions of their own psychological traits/ social identities.  As such, it is not clear from this study (Tanweer et al., 2010) whether adults with ASD

show a similar pattern in the *type* of self-concepts they use to describe themselves, as Lee and Hobson (1988) found children with ASD did.

Finally, Kristen and colleagues (Kristen et al., 2014) assessed self-conception in adults with ASD using a version of a mind-mindedness task (Meins & Fernyhough, 2010). During this task adults were asked to answer the question "Can you describe yourself for me?". Adults' descriptions were then either coded as mental descriptions, behaviours, physical descriptions, or general descriptions. Kristen et al., (2014) found that adults with ASD described themselves significantly less in terms of mental state terms than neurotypical adults. However, adults with ASD provided a similar number of physical descriptions of themselves, relative to neurotypical control participants. These findings are relatively in keeping with the results of Lee & Hobson (1998), who found that children with ASD tended to provide more physical self-descriptions than neurotypical children (both results indicate that physical self-concepts remain intact in ASD). However, this study used a relative sparse coding scheme, and did not code statements on whether they referred to social categories. As such, it is not possible to tell whether the impairments in mental self-conception demonstrated by adults with ASD in this study are similar to the pattern of results demonstrated by children with ASD in Lee and Hobson's study (1988).

**The Current Study**

It is clear that relatively few studies have explored self-descriptions in individuals with ASD, and those that have report mixed results. Additionally, to date, no study has explored the relationship between the ability to understand mental states in others (mindreading/theory of mind) and the extent to which an individual's defines themselves in terms of mental states or psychological terms. As discussed, theories that

suggest individuals with ASD are impaired at understanding their own mental states (e.g., Williams, 2010; Carruthers, 2009) predict that individuals with ASD should demonstrate less elaborate psychological self-concepts. However, such theories do not predict poorer physical self-concepts in ASD. Although Kristen et al., (2014) measured participants mindreading ability, unfortunately they do not report whether mindreading abilities related to the extent participants conceptualised themselves in terms of mental terms.

To address these questions the current study explored conceptual self-awareness in adults with ASD using the twenty statements task (Kuhn & McPartland, 1954). This task benefited from being open-ended and thus allowed participants to describe themselves using any characteristics they believed to represent their long-term identity well. Participants' statements were coded using a comprehensive coding scheme, similar to that employed in Tanweer et al., (2010). However, statements were also coded based on whether they were psychological or physical in nature, allowing for a direct comparison of psychological and physical self-concepts in ASD.

It was predicted that adults with ASD would demonstrate poorly established psychological self-concepts, and would describe themselves less in terms of psychological characteristics relative to neurotypical adults. However, it was expected that there would be no difference in terms of the extent to which adults with ASD described themselves using physical descriptions. Following the results of Lee and Hobson (1998), and Tanweer et al., (2010), alongside characteristic impairments in social understanding in ASD, it was also predicted that participants with ASD would not describe themselves as much in terms of social categories relative to neurotypical participants.

To date, the relationship between mindreading ability and conceptual self-awareness has not been explored in ASD.  As such, to explore this relationship, the study also included a measure of mindreading ability (the animations task).  In terms of the relationship between mindreading ability and conceptual self-awareness it was predicted that performance on the animations task would significantly relate the extent to which participants described themselves in terms of psychological characteristics, but not to the extent participants described themselves in terms of physical characteristics.

**Method**

**Participants**

Fifteen adults with ASD and 20 neurotypical comparison adults took part, all of whom gave written, informed consent before participating.  Participants in the ASD group had formal diagnoses of autistic disorder or Asperger's disorder, according to established criteria (American Psychiatric Association, 2000; World Heath Organisation, 1993).  In order to assess current ASD features, 13 of the 15 participants in the ASD group completed Autism Diagnostic Observation Schedule-Generic (ADOS; Lord et al., 2000) assessments.  The remaining two participants declined to complete the ADOS, as they did not feel comfortable being filmed.  The two participants who did not complete the ADOS had rigorous diagnoses and scored above the cut-off on the Autism-spectrum Quotient (see immediately below).  All participants who completed the ADOS received a total score ≥7, the defined cut-off for ASD (Lord et al., 2000).  All participants completed the Autism-spectrum Quotient (AQ; Baron-Cohen et al., 2001), a self-report questionnaire that assesses ASD/ASD-like features.  Thirteen out of 15 participants with ASD scored above the defined cut-off for ASD on the AQ (total score ≥26; Woodbury-

Smith et al., 2005).  Only two participants missed this cut-off.  However, both of these participants scored well above the defined ASD cut-off on the ADOS (ADOS scores for both participants were 12).  All comparison participants scored below the defined cut-off for ASD.

*Table 23: Participant characteristics (means, standard deviations and inferential statistics).*

| | Group | | *t* | *p* | Cohen's |
| | ASD | Neurotypical | | | |
| | (*n* = 15) | (*n* = 20) | | | *d* |
|---|---|---|---|---|---|
| Age (years) | 29.88 (10.44) | 29.43 (14.14) | 0.10 | .918 | 0.04 |
| VIQ | 112.47 (14.99) | 111.30 (10.77) | 0.27 | .790 | 0.09 |
| PIQ | 110.20 (15.56) | 113.85 (10.57) | 0.83 | 414 | 0.27 |
| FSIQ | 112.87 (15.34) | 114.10 (10.26) | 0.29 | .777 | 0.09 |
| AQ Total Score | 33.40 (9.01) | 13.50 (6.12) | 7.78 | <.001 | 2.58 |
| ADOS Social + Communication Score* | 11.62 (2.02) | | | | |

AQ: Autism-spectrum Quotient; ADOS: Autism Diagnostic Observation Schedule; PIQ = performance IQ; FSIQ = full scale IQ; VIQ = verbal IQ
*Based on 13/15 participants

No participants, in either group, reported using any psychotropic medication or any history of neurological or psychiatric disorders (apart from ASD).  The participant groups were closely equated for verbal and non-verbal ability (see Table 23 for participant characteristics).  Verbal IQ (VIQ), performance IQ (PIQ), and full-scale IQ

(FSIQ) were assessed using the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999). Groups were also closely equated for chronological age.

**Materials and Procedures**

**Twenty statements task.** Following the method employed in Tanweer et al., (2010) the Twenty Statements Task (Kuhn & McPartland, 1954) was administered, to assess participants' sense of identity. The twenty statements task is a relatively simple, open-ended task that encompasses the idea that the self is multifaceted. During the task participants were simply asked to respond to the question "Who am I?" by writing down twenty statements about themselves, beginning with the phrase "I am". Participants were asked to write statements that best defined their long-term identify and to avoid writing statements that described non-permanent, short-term/temporary aspects of their current self (e.g., I am tired, I am hungry). There was no time limit to how long participants could take to complete the task.

**Animations task.** As a measure of mindreading ability the animations task (Abell et al., 2000) was administered. A full description of this task is provided in chapter two (please see p.59).

**Scoring**

**Twenty statements task.** Participants' statements were coded using the modified coding system outlined in Rhee et al., (Rhee, Uleman, Roman, & Lee, 1995; please see appendix three). This coding system was the same as the one used in Tanweer et al., (2010). Statements were coded into one of eight categories; 1) traits; 2) social identities; 3) specific attributes; 4) evaluative descriptions; 5) physical description; 6) emotional states; 7) peripheral information and 8) global descriptions.

Each category was also divided into several distinct subcategories that statements were categorised into (please see p.275 for details of these specific subcategories). An additional subcategory of sexual orientation, which not included in previous coding schemes, was added as a social identity subcategory, as statements concerning sexual orientation did not easily fit into the other predefined subcategories used in previous coding systems. Following Tanweer et al., (2010) and Rhee et al., (1995) statements were also either classed as autonomous or social (social statements referenced other people or social context), and either abstract or specific.

Not all participants were able to produce 20 statements. As such, proportion scores were calculated that represented the proportion of statements participants produced for each category (e.g., if a participant provided 16 statements overall, 8 of which were trait descriptions, their overall trait description score would be 0.5). The proportion of autonomous statements participants provided was also calculated. Additionally, following Tanweer et al., (2010) and Rhee et al., (1995), an "identity strength" score was calculated as the total number of statements generated overall, and an "identity complexity" score was calculated as the total number of different categories and subcategories participants' statements covered. An "identity quality" score was also calculated as the proportion of statements participants generated that were abstract. Unlike previous coding systems the proportion of overall descriptions that were physical was calculated as well as the proportion of overall descriptions that were psychological in nature (please see appendix three, p.275 for a list of the categories considered psychological and physical). Finally, following Rhee et al., (1995) responses that did not make sense were coded as "nonsense", and statements that participants repeated were only coded once. Additionally, on the rare occasion that participants provided more than one meaning in a statement (e.g., "I am a good person, as I usually

intend well") the statement was classified based on the first meaning provided. 25% of participants were randomly selected and their statements were also coded by a second rater who was blind to the diagnostic status of the participants. Inter-rater reliability was good, *Cronbach's α* = .80.

       **Animations task.** Voice recordings of participants' commentaries were transcribed verbatim. These transcriptions were then scored by a second rater who was blind to the diagnostic status of the participants, according to the scoring criteria outlined in Abell et al., (2000). Inter-rater reliability for scores across the four animations was excellent, *Cronbach's α* = .93

## Results

### Twenty statements task

Table 24 shows the means and standard deviations for overall performance on the twenty statements task. Independent *t*-tests indicated that adults with ASD produced significantly fewer statements during the task, relative to neurotypical participants, as evidenced by significantly lower identity strength scores on the task. However, there was no difference in the mean identity complexity scores in the ASD and neurotypical groups. This indicates that the statements adults with ASD provided ranged across as many categories as those provided by neurotypical participants. Additionally, the proportion of abstract statements (measured by identity quality scores) and autonomous statements provided by adults with ASD was similar to the proportion of abstract and autonomous statements provided by neurotypical adults.

*Table 24: Means (SDs) and inferential statistics for group differences in performance on the twenty statements task.*

| Experimental Measure | Group | | t | p | Cohen's d |
|---|---|---|---|---|---|
| | ASD (*n*= 15) | Neurotypical (*n*=20) | | | |
| Identity strength | 12.07 (4.48) | 15.25 (4.31) | 2.13 | .041 | 0.72 |
| Identity complexity | 6.20 (3.14) | 6.60 (1.90) | 0.47 | .643 | 0.15 |
| Identity quality | .35 (.31) | .43 (.25) | 0.81 | .423 | 0.28 |
| Prop. autonomous statements | .53 (.24) | .59 (.25) | 0.72 | .480 | 0.24 |

Figure 17 shows the mean proportion of statements ASD and neurotypical participants produced in each of the eight statement categories. To explore group differences in the types of statements provided by ASD and neurotypical participants a multivariate ANOVA (using Wilks' Lambda criterion) was carried out, which included the eight overall statement categories (traits, social identities, specific attributes, evaluative descriptions, physical descriptions, emotional descriptions, peripheral information and global descriptions) as dependent variables. This indicated that there was no significant effect of group (ASD/neurotypical), $F$ (8, 26) = 1.34, $p$ = .317, $\eta_p^2$ =.28. This suggests that the types of statements produced by participants in the ASD group did not significantly differ from the types of statements produced by neurotypical participants.
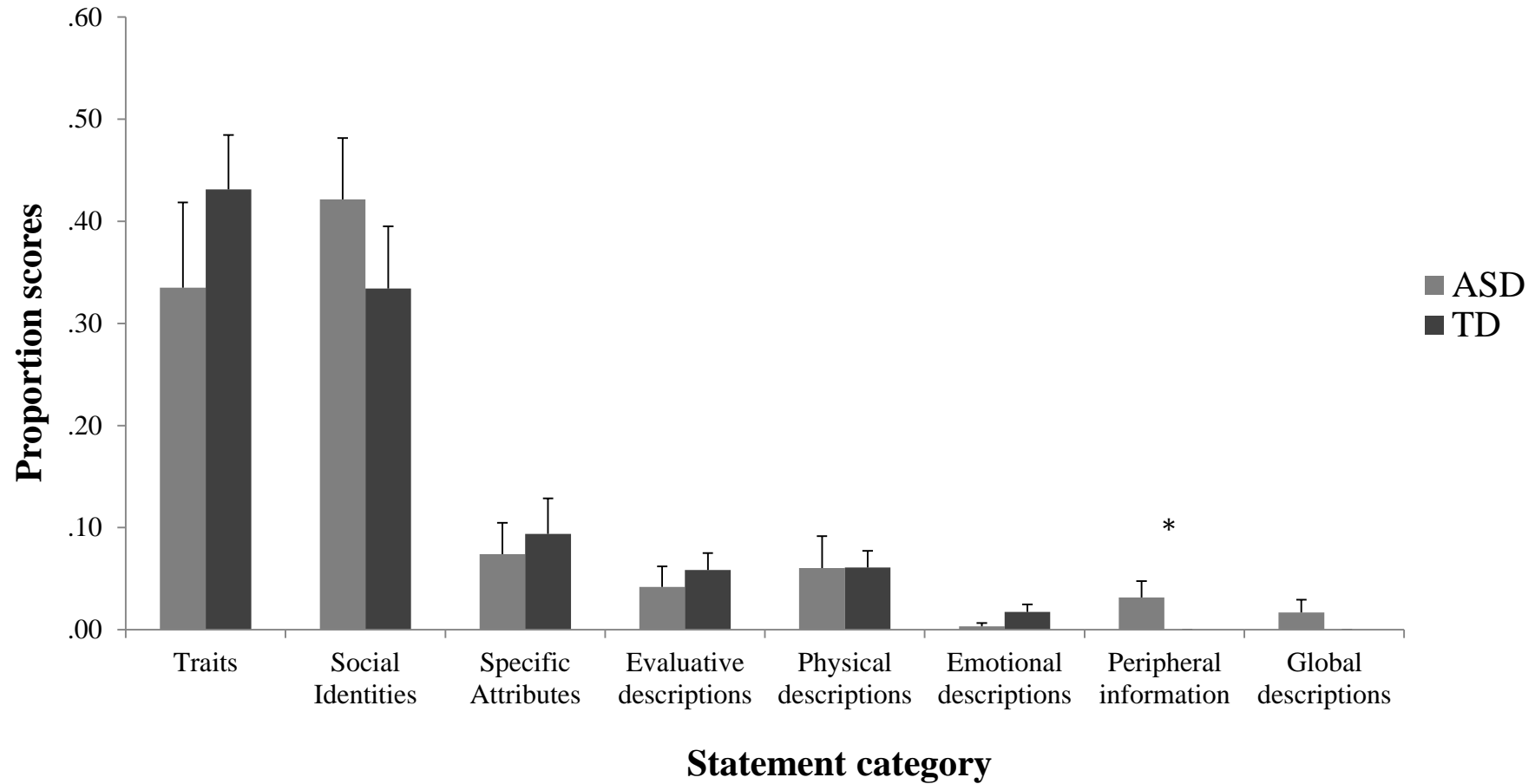
*Figure 17: The proportion of statements ASD and neurotypical participants produced, in each of the overall statement categories. * = p < .05.*

Further analysis was carried out to directly explore predictions that individuals with ASD would demonstrate specific impairments in their awareness of psychological self-concepts, but not physical self-concepts.

Figure 18 shows the mean proportion of statements ASD and neurotypical participants produced that were either physical or psychological in nature. A mixed-model ANOVA was carried out on these data with group (neurotypical /ASD) entered as the between-subjects variable, and statement type (psychological/physical) entered as the within-subject variable. There was a significant main effect of Statement Type, $F$ (1, 33) = 43.80, $p < .001$, $\eta_p^2 = .570$, reflecting the fact that across both groups, participants provided a significantly higher proportion of psychological statements than physical statements. However, there was no significant main effect of group, $F$ (1, 33) = 2.59, $p = .117$, $\eta_p^2 = 0.073$, and no significant Group × Statement Type interaction, $F$ (1, 34) = 1.78, $p = .191$, $\eta_p^2 = .051$. These results are not in keeping with predictions, and suggest that ASD participants did not define themselves less in terms of psychological self-concepts than neurotypical participants.
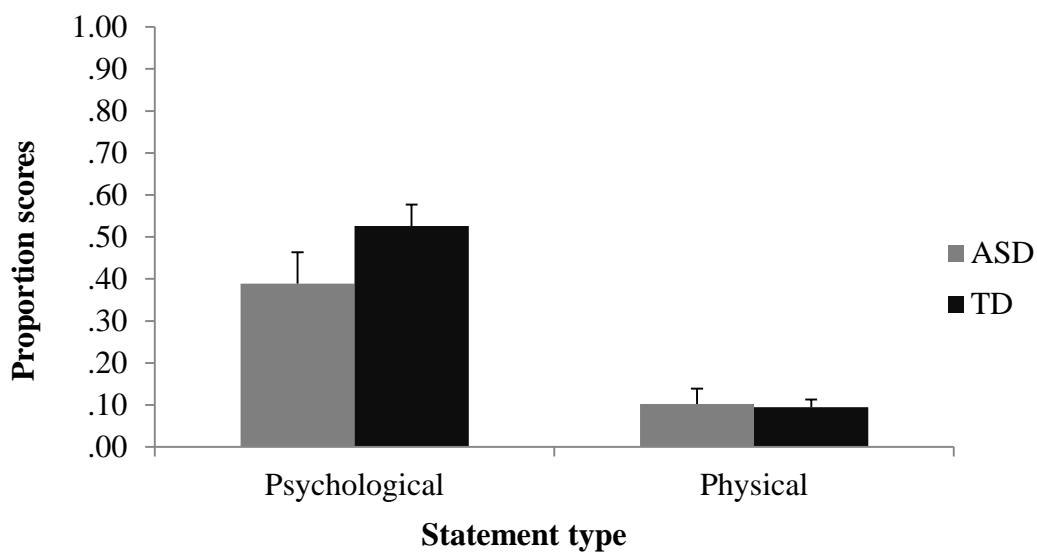


*Figure 18: The proportion of statements ASD and neurotypical participants produced that that referred to psychological and physical characteristics.*

**Animations task**

Table 25 shows the means and standard deviations for performance on the animations task. A mixed-model ANOVA was carried out on these data with Group (neurotypical/ASD) entered as the between-subjects variable, and Animation Type (mentalising/goal-directed) entered as the within-subject variable. There was a significant main effect of Group on animations scores, reflecting the fact that participants with ASD performed significantly less well than comparison participants on the task overall, $F$ (1, 33) = 16.56, $p$ = < .001, $\eta_p^2$ = .33. There was also a significant main effect of Animation Type, indicating that, across both groups, scores were higher in the goal-directed condition than the mentalising condition, $F$ (1, 33) = 78.22, $p$ < .001. $\eta_p^2$ = .70. There was no significant Group by Animation Type interaction, $F$ (1, 33) = 0.06, $p$ = .815, $\eta_p^2$ = < .01, suggesting that individuals in the ASD group were impaired at both higher- and lower-level mindreading, relative to individuals in the neurotypical group.

*Table 25: Means (SDs) and inferential statistics for group differences in performance on the animations task.*

| Animations task | Group | | | $t$ | $P$ | Cohen's $d$ |
|---|---|---|---|---|---|---|
| | ASD | Neurotypical | | | | |
| | ($n$= 15) | ($n$=20) | | | | |
| Overall | 8.93 (2.63) | 12.05 (1.90) | | 4.07 | <.001 | 1.36 |
| Mentalising condition | 3.40 (1.35) | 4.90 (1.65) | | 2.87 | .007 | 1.00 |
| Goal-directed condition | 5.53 (1.46) | 7.15 (0.75) | | 4.28 | <.001 | 1.40 |

**Associations between mindreading ability and performance on the twenty**

**statements task**

A series of correlational analyses was carried out to explore the relation between performance in each condition of the animations (mindreading) task and the proportion of physical/psychological self-related statements participants produced. Neither performance in the mentalising or goal-directed condition of the animations task correlated with the proportion of psychological or physical self-concepts participants with ASD produced, all $rs \leq .35$, all $ps \geq .201$. In the neurotypical group performance in the mentalising condition of the animations tasks did not correlate with either the proportion of physical or psychological self-concept participants produced, $rs \leq .29$, all $ps \geq .222$, nor did performance in the goal-directed condition correlate with the proportion of physical self-concepts participants produced, $r = -.017$, $p = .944$. However, performance on the goal-directed condition of the task was significantly negatively correlated with the proportion of psychological statements neurotypical participants produced, $r = -.525$, $p = .017$. This suggests that unexpectedly, the poorer neurotypical participants performed on the goal-directed condition of the animation task, the more psychological self-concepts they produced.

## Discussion

The results of this study suggest that individuals with ASD demonstrated undiminished self-concepts relative to neurotypical participants. Unexpectedly, the self-concepts provided by the ASD group were in general *qualitatively* similar to those provided in the neurotypical group and ASD participants provided statements that were distributed across similar categories to those provided by neurotypical adults. The results also indicated that individuals with ASD reported self-descriptions that were as

diverse and abstract as those provided by neurotypical adults.  These results are not in

keeping with the findings from previous studies that suggest individuals with ASD tend

to define themselves more using abstract descriptions (e.g., Tanweer et al., 2010).  Nor

are they in keeping with other findings that suggest both children and adults with ASD

tend to describe themselves less in terms of social categories than neurotypical

individuals (e.g., Lee & Hobson, 1998; Tanweer et al., 2010).

Additionally, the study found no evidence for the suggestion that psychological

self-concepts are diminished in individuals with ASD.  These results are not in keeping

with the study predictions, nor previous findings within the literature.  Both Kristen et

al., (2014), and Lee and Hobson's (1998) studies suggested that psychological self-

knowledge may be diminished in ASD, whilst physical self-knowledge remains intact.

This pattern of results is also mirrored in the results of other measures of conceptual

self-awareness, that find individuals with ASD demonstrate diminished self-reference

effects (Henderson et al., 2009; Lombardo et al., 2007; Toichi et al., 2002)and pronoun

use (Hobson et al., 2010; Jordan, 1989; Lee et al., 1994; Lind & Bowler, 2009a; Loveland

& Landry, 1986) yet unimpaired mirror self-recognition (Dawson & McKissick, 1984;

Ferrari & Matthews, 1983; Neuman & Hill, 1978; Spiker & Ricks, 1984).

One possible explanation for these results is the suggestions that perhaps impairments

in acquiring typical psychological self-concepts are developmental in nature.  Indeed,

this suggestion is in keeping with findings from studies exploring semantic

autobiographical memory in ASD (a different index of self-knowledge), which suggest

semantic knowledge about oneself is impaired in children with ASD, yet appears to

resolve by adulthood (see Lind, 2010).  As stated by Williams and Bowler:

> We should never forget that the clinical picture we see among individuals with a
>
> diagnosis of ASD represents a particular point in an atypical developmental

trajectory, in which both the clinical features and any putative underlying factors may be in a process of change. The challenge to understand how this process of change operates will no doubt add an additional layer of complexity to the picture, but we have little doubt that this layer will be necessary. (2014, p.5)

As such, it is possible that the development of psychological self-concepts is delayed/ atypical in children with ASD (e.g., Lee & Hobson, 1998), but by adulthood individuals with ASD have formed at least relatively robust self-concepts.

Indeed, whilst the adults with ASD may possess typical psychological self-concepts, these concepts themselves may have been acquired through atypical processes. For example, an individual may know that they are a kind person because they are aware of psychological aspects of themselves. However, this knowledge may also be learnt through indirect means (e.g., being told by other individuals that they are kind). It is not clear from the results of this study what processes individuals used to generate self-concepts during the task.

It is of course possible that the ASD group employed an atypical strategy on the task (thus producing statements that were qualitatively similar to those provided by neurotypical participants), despite impaired conceptual self-awareness. For example, when asked to provide descriptive statements about themselves participants may simply have provided generic statements (that did not specifically apply to themselves), or provided statements based on stereotypes. As such, although the results of this study inform us about the types of self-concepts individuals with ASD spontaneously attribute to themselves, which is of course important, we have no indication of how *accurate* these self-concepts are. Indeed, in chapter three we saw that individuals with ASD reported they were very aware of their own minds. However, the empirical evidence

reported in this chapter suggested that individuals in the ASD group were in fact impaired in metacognition.

Despite finding no evidence of impairments in identity complexity in the ASD group, the study did find that individuals with ASD demonstrated significant impairments in their *identity strength* scores, relative to neurotypical individuals. This reflected the fact that overall adults with ASD provided significantly fewer descriptions of themselves relative to neurotypical participants. One explanation for this finding is the suggestion that overall adults with ASD hold self-concepts that are less detailed than neurotypical adults. However, an alternative explanation for these results is the suggestion that poorer identity strength scores in the ASD group may be the result of impairments in generativity abilities in ASD. Generativity ability refers to an individual's ability to access stored information in memory and generate spontaneous, novel responses during a task (e.g., the ability to generate word beginning with a particular letter, etc.). Several previous studies have suggested that individuals with ASD demonstrate diminished generativity abilities (e.g., Ambery et al., 2006; Dichter, Lam, Turner-Brown, Holtzclaw, & Bodfish, 2009). For example, Dichter and colleagues (Dichter et al., 2009) found that, given the same amount of time, individuals with ASD were able to generate fewer examples of animal names and fewer uses of an object, relative to neurotypical individuals. It is possible that individuals with ASD provided fewer self-concepts during the task than neurotypical participants, due to impaired generativity abilities. As such, impairments in identity strength scores in the ASD group may not necessarily be indicative of impairments in conceptual self-awareness in ASD. However, one problem with this suggestion is that not all studies of generativity ability in ASD find evidence of impairments (see e.g., Lind & Bowler, 2010; Crane & Goddard, 2008). For example, Lind and Bowler (2010) found no evidence of generativity

impairments in individuals with ASD.  As such, evidence of generativity impairments in

ASD is mixed.  Given this study did not asses generativity abilities in ASD it is thus

unclear whether such impairments (rather than impairments in conceptual self-

awareness) explain impairments in identity strength scores on this task.

Regarding the correlational analyses carried out in this study, the study found no

evidence of a significant relationship between participants' mindreading abilities and

the extent to which participants described themselves using psychological

characteristics.  One mechanism accounts of mentalising predict that individuals who

demonstrate mindreading impairments (such as the ASD group in this study) should

also demonstrate impairments understanding their own mental states, and thus you

would predict that mindreading impairments would be associated with impaired

psychological self-concepts.  As such, the study did not provide direct support for the

suggestion that mindreading is related to reports about one's own psychological self.

However, some caution should be taken interpreting these findings.  As stated above, it

is not clear from the results of this study how accurate the psychological self-concepts

reported by participants were (and thus we do not know the extent to which the self-

concepts participants provided were based on introspective awareness of their own

mental states).

**CHAPTER NINE**

<div align="right">

**GENERAL DISCUSSION**

</div>

Historically, "the self" has been central to theories of autism spectrum disorder (ASD). Indeed the term autism itself derives from the word "autos", the Ancient Greek word for "self". However, paradoxically, whilst some researchers have suggested that classic autism involves a total focus on the self (Baron-Cohen, 2005) individuals with ASD have also been characterised as having an "absent self" (see e.g., Baron-Cohen, 2005; Frith, 2003; Frith & Happé, 1999). One resolution for this paradox is the suggestion that the self is itself not a single entity, but is instead multifaceted. Whilst ASD may involve a total focus on some aspect of the self, it may be characterised by the absence of a different aspect of self. The aim of this thesis was to explore self-awareness in ASD, investigating whether ASD is defined by absolute impairments in self-awareness, or whether self-awareness is selectively impaired in children and/or adults with this disorder.

It is evident from several of the studies reported in this thesis that individuals with ASD clearly do not demonstrate a completely "absent self" (Frith, 2003). For example, the studies reported in chapter six and seven indicate that both children and adults with ASD demonstrate entirely typical action monitoring abilities. Such findings are in keeping with the literature on action monitoring in ASD, and suggest that physical self-awareness appears to be intact in this disorder. Additionally, the results reported in chapter eight suggest that self-concepts reported by individuals with ASD are similar to those provided by neurotypical individuals. However, in contrast, the results reported in chapter three and five indicate that metacognitive monitoring is impaired in ASD, at least on some measures of metamemory ability. This appears to suggest that

children and adults with ASD demonstrate impaired understanding of their own mental states. It is widely acknowledged that the self is multifaceted (e.g., Neisser, 1988, Rochat, 2003; Zahavi, 2010), and thus an individual can be aware of several aspects of the self at any one time. As such, it follows that different aspects of self-awareness can be selectively impaired (see Zahavi, 2010). The idea that ASD involves selective impairments in self-awareness is explored in this discussion. More specifically it is suggested that, whilst individuals with ASD appear to be fully aware of their own physical selves, they demonstrate impairments in private/psychological self-awareness, and understanding their own minds (see Williams, 2010). The following section considers whether the results reported in this thesis support this suggestion and whether these results are in keeping with the broader literature.

**Physical self-awareness in ASD**

Chapters six, seven, and eight all assessed aspects of the physical self in individuals with ASD and presented a consistent pattern of findings. As predicted, throughout the studies reported in this thesis there was no hint of impairments in physical/ecological self-awareness, in either children or adults with ASD.

One important aspect of physical self-awareness is the ability to monitor one's own actions. This ability was explored in chapter six and seven, which employed measures of action monitoring in adults and children with ASD. Results from the squares task employed in chapter six suggest that adults with ASD are able to accurately monitor their online actions. Additionally, the studies of the enactment effect reported in chapters six and seven suggest that both children and adults with ASD demonstrate typical action monitoring ability. Superior memory for self-performed actions over other-performed actions (the enactment effect) is thought to result from additional

motoric components involved in performing an action leading to those actions being more deeply encoded than observed actions (e.g., Engelkamp & Zimmer, 1989). As such the presence of typical enactment effects in individuals with ASD also suggest that action monitoring is unimpaired in this disorder. Findings that suggest individuals with ASD appear to demonstrate typical action monitoring ability are in keeping with other studies of online action monitoring ability (e.g., Russell & Hill, 2001; Williams & Happé, 2009a) alongside studies in the broader action monitoring literature in ASD (e.g., Blakemore et al., 2006; David et al., 2006). As such, it appears that the majority of the literature surrounding action monitoring in ASD support the suggestion that physical self-awareness is a relative strength in ASD.

Chapter eight explored a different aspect of self-awareness in ASD, and investigated whether individuals with ASD conceptualise themselves in a similar way to neurotypical individuals. The results of this study found that adults with ASD described themselves using physical concepts as often as neurotypical individuals. These results are in keeping with studies that have fairly consistently found that a large proportion of the children with ASD tested successfully recognise their own image in the mirror (Dawson & McKissick, 1984; Ferrari & Matthews, 1983; Neuman & Hill, 1978; Spiker & Ricks, 1984). Mirror self-recognition can be considered another indication that a child holds a physical self-concept, and appears to be a relative strength in children with ASD.

Altogether, the results reported in this thesis support suggestions that physical self-awareness is intact in ASD. Indeed, there was some tentative evidence to suggest individuals with ASD were somewhat more sensitive to their agency than comparison participants. When looking at performance on the Squares task, employed in chapter six, participants with ASD performed better than neurotypical participants, when the number of successful *levels* completed on the Squares task was taken as the dependent

measure of performance. These results are also in keeping with those from an early study by Frith and Hermelin (Frith & Hermelin, 1969), which also suggested children with ASD appeared to be *better* at monitoring their own efference copy, relative to comparison participants. In this study participants were required to move a stylus along a track that had been cut into a piece of Perspex. Participants were then asked to complete the task again, this time without the aid of visual cues. Frith and Hermelin (1969) found that participants with ASD completed the task significantly faster than comparison participants, and concluded that these findings were consistent with *enhanced*, rather that impaired, action monitoring. This suggests that, far from being impaired, individuals with ASD might show heightened physical self-awareness. Regardless of whether awareness of one's physical self is heightened in ASD, the studies reported in this thesis certainly suggest that this aspect of self-awareness is not impaired.

**Psychological self-awareness in ASD**

Recently, Williams (2010) suggested that individuals are impaired at "theory of own mind" and are as impaired at recognising their own mental states as they are at recognising mental states in others (Williams, 2010). One of the central aims of this thesis was to explore psychological self-awareness in both children and adults with ASD in more detail than in previous studies, using classic tests of metacognition. Chapters three, four, and five assessed metacognition in children and adults with ASD. However, the results reported in these chapters are not as clear cut as those reported in the studies of physical self-awareness. Chapter three employed a classic feeling-of-knowing (FOK) paradigm to explore metamemory monitoring ability in adults with ASD. This study found that participants with ASD showed significantly diminished FOK accuracy.

This diminution was associated with a large effect size ($d$ = 0.97), indicating a substantial difficulty with metamemory monitoring. As such, this suggests that adults with ASD *are* impaired at understanding their own mental states. In keeping with this finding were the results reported in chapter five, which found children with ASD also demonstrated impairments in meta-monitoring. During a judgment of confidence (JOC) task children with ASD found it significantly harder to judge whether the answers they had provided during the task were correct or incorrect. Again, this suggests that metamemory monitoring is impaired in ASD. As such, together these studies provide evidence of monitoring impairments in both children *and* adults with ASD.

### Metacognitive bias in individuals with ASD

Whilst impairments in both FOK and JOC accuracy are in keeping with the suggestion that psychological self-awareness is impaired in ASD, there are of course several potential explanations for diminished judgements accuracy in the ASD group. One possibility is that poor judgement accuracy on both tasks was the result of a particular responding bias in ASD. For example, as discussed in chapter three, some studies have indicated that individuals with ASD may demonstrate a "positive illusory bias", showing a tendency to self-assess their perceived competence as greater than their actual ability. If people with ASD were consistently over-confident because of a tendency to be "arrogant" (for example), or indeed consistently under-confident because of a tendency to be "modest" (for example), then they may well demonstrate impaired performance on metamemory tasks (and such impairments would not necessarily suggest monitoring is impaired in ASD).

Concerning the JOC task, it appears that children with ASD demonstrated monitoring impairments that were driven by *over*confidence on the task. The children in the ASD group reported significantly higher confidence in their incorrect answers,

relative to neurotypical children. Children with ASD were also more confident in the answers they then removed, relative to the neurotypical group. This is in keeping with previous findings that suggest individuals with ASD tend to overestimate their cognitive abilities. Some studies have indicated that individuals with ASD tend to self-report their own social functioning more positively than parents will report (e.g., Lerner et al., 2012), and self-report their own ASD traits as less severe than parents will report (e.g., Johnson et al., 2009). Indeed, the results from the meta-cognitions questionnaire reported in chapter three suggested that adults with ASD self-report *greater* awareness of their own mental states than neurotypical participants report. These studies all suggest that, at least when providing self-reports, individuals with ASD demonstrate a tendency to report their perceived competence as greater than their actual ability.

However, the results reported in chapter three found that during a FOK task, monitoring impairments in individuals with ASD were not driven by overconfidence. Instead, among adults with ASD, diminished performance on the FOK task was driven by a relative *under*confidence existing knowledge, rather than by overconfidence in judgements during the task. Thus it appears that the type of errors individuals with ASD made differed depending on the type of metamemory judgement. The fact that individuals with ASD did not demonstrate a consistent type of bias across the FOK and JOC tasks suggests that one particular response bias in ASD does not best explain impaired performance on these tasks. Instead, this pattern of findings is in keeping with the idea that metacognition is impaired in ASD. If individuals with ASD are impaired at understanding their own mental states, one would predict diminished accuracy on metamemory tasks. However, you would not expect individuals with ASD to make metamemory monitoring mistakes that were exclusively under-confident or over-confident.

**Judgment of learning (JOL) accuracy**

Together, diminished performance on both the FOK and JOC task suggests that individuals with ASD are impaired at understanding their own mental states. However, there was no hint of metamemory impairments in the experiments reported in chapter four. The results from all three judgement of learning (JOL) experiments reported no significant differences in JOL accuracy among either children or adults with ASD. These results are clearly not in keeping with the results reported in chapter three or five, nor previous literature, that suggests awareness of one's own mental states is impaired in ASD (Williams & Happé, 2009b, 2010). However, they are in keeping with the only other study to explore JOL accuracy in ASD (Wojcik et al., 2014). Wojcik and colleagues also found that adolescents with ASD can make accurate JOL assessments. However, there were several methodological issues associated with this study. Firstly, the ASD group was not matched to the neurotypical group for VIQ. Additionally, the study did not assess whether mindreading ability was impaired in the sample of ASD participants they assessed. This is important, as several theories only ever predict impaired JOL accuracy in individuals who also demonstrate mindreading impairments. Given these considerations, the results of the JOL tasks reported in chapter four provide more definitive evidence that JOL accuracy is impaired in ASD.

If metacognitive monitoring ability (and psychological self-awareness) truly is impaired in individuals with ASD this begs the question why do individuals with ASD not demonstrate impairments on JOL tasks (neither cue-alone, cue-target, nor aggregate JOL task).? There are at least three possible explanations for the discrepancy in the results across chapters three, four, and five.

**1) Different underlying bases of metamemory judgements**. One possible explanation of this inconsistent pattern of results is the suggestion that different

metamemory judgements rely on different sources of information.  Although making

JOL, JOC and FOK judgements clearly involves monitoring one's own memory, it has

been argued that these types of judgements rely on different memory heuristics (see

Kelemen et al., 2000; Schwartz, 1994; Souchay & Isingrini, 2012).  Within the memory

literature the distinction has been made between the processes of "familiarity" and

"recollection".  Familiarity in this sense refers to a general feeling of memory that varies

in strength, but does not contain any information about the context in which that

knowledge was acquired.  This contrasts with "recollection", which involves the

additional incorporation of contextual information in memory.  Evidence suggests that

some judgments are heavily influenced by the *familiarity* an individual has towards a

cue word during a metamemory task (see Metcalfe et al., 1993; Reder, 1987)  In

contrast, whilst making metamemory judgments individuals may alternatively rely on

the extent to which they can *recollect* partial or related information about the target

word (see Koriat, 1993).

One potential explanation for differences in metamemory accuracy in ASD is the

suggestion that impairments in memory explain metamemory impairments in this

disorder.  More specifically, Wojcik and colleagues suggest that metamemory

impairments in ASD may arise due to impairments in recollection ability in this disorder

(see Wojcik et al., 2014).  These authors suggest that some judgements, particularly

those made at retrieval (such as FOK and JOC judgements), rely heavily on the

recollection of partial or related information about the target word.  In contrast

judgements made *before* retrieval (such as JOLs) may not rely on recollection to the

same extent.  Instead, Wojcik and colleagues (see Wojcik et al., 2014)  suggest that

individuals may base judgements made at encoding (including JOLs) more on cues

related to the type of task or material they are judging, as well as basing such judgements on general knowledge of factors that affect memory performance.

In support of this suggestion, studies have found that FOK judgments in particular appear to rely heavily on recollection. For example, Souchay (Souchay et al., 2007) explored whether FOK accuracy in older adults was associated with performance on a remember/know task. A remember/know task is a variation of an old/new recognition memory task. During this task participants are presented with stimuli and asked to assess whether each item is "old" (and appeared previously during study) or "new". For memory items participants identify as "old" participants are the asked to identify whether they think they can explicitly "remember" the occurrence of that memory item during study, or whether they simply "know" it is old. Remember responses are typically associated with the process of recollection, whereas know responses are typically associated with the process of familiarity. Souchay found that FOK judgments accuracy (gamma scores) in older adults was positively associated with the number of "remember" responses participants make during a standard remember/know recognition task (Souchay et al., 2007). These results were interpreted as suggesting that individuals who report they "remember" more during recognition are better at making accurate FOK judgments.

Additionally, studies exploring the effect of divided attention at encoding suggest that FOK judgements rely heavily on recollection. Divided attention tasks appear to impair individuals' recollection ability but not familiarity ability (see Yonelinas, 2002). Sacher and colleagues found that when participants engage in divided attention at encoding, FOK gamma accuracy is much poorer (see Sacher, Taconnat, Souchay, & Isingrini, 2009). Interestingly, divided attention does not appear to affect JOL accuracy (Barnes & Dougherty, 2007)

As Wojcik and Souchay suggest, individuals with ASD may not demonstrate impairments on JOL tasks because such task rely less on recollection than other metamemory tasks.  However, there are several potential problems with this explanation.  Firstly, the results reported in this thesis are not entirely in keeping with this suggestion.  Whilst individuals with ASD demonstrated significant impairments in metacognitive accuracy during a FOK and JOC task, these impairments were apparent despite *typical* object-level performance on the task.  In both experiments object level ability was assessed using a cued-recall memory test, and accurate performance on such recall tasks is typically thought to rely on recollection.  As such, this suggests that the metacognitive impairments demonstrated by individuals with ASD were not solely driven by impairments in recollective ability.

Additionally, within the broader literature the evidence to support the suggestion that recollection abilities are impaired in ASD is mixed.  Whilst some studies suggest recollection is impaired in ASD Wojcik and Souchay themselves report evidence that suggests recollection is intact in individuals with this disorder.  Across a series of experiments (Souchay, Wojcik, Williams, Crathern, & Clarke, 2013) Souchay and colleagues assessed recollection in adolescents with ASD, using both subjective measures of recollection (remember/know tasks) and objective measured of recollection (source monitoring tasks).  Using remember/know paradigms, the study assessed recollection abilities during a memory task by asking participants to explicitly report whether they remembered a piece of information, or whether they simply knew it.  To assess the quality of these subjective judgements, participants were also then asked to provide additional contextual information to justify their "remember" responses.  In one of the experiments reported in their paper (Souchay et al., 2013; Experiment 1) individuals with ASD reported significantly fewer "remember" responses

relative to neurotypical participants. This finding is in keeping with other studies in the literature (see e.g., Bowler et al., 2000), and suggests that, if participants were making accurate remember/ know responses, ASD participants demonstrate impairments in recollection. However, the evidence to support the suggestion that these "remember" responses were accurate (in both the ASD and neurotypical groups) was fairly limited. Whilst ASD participants were able to provide correct source information for 59% of the items they reported they remembered, neurotypical participants provided correct source information for 48% of their remember response (this difference was not significant). Given that chance performance on this task would predict participants would be able to provide contextual information for 50% of the items they say they can "remember", these findings suggest that in fact, participants (in both groups) were not very good at determining whether they had in fact recollected information or not. As such, it is difficult to draw any firm conclusions from the finding that individuals with ASD report recollecting memory items less often than neurotypical participants.

Additionally, the other two remember/know experiments reported in Souchay et al., (2013) paper found that the proportion of items participants reported they remembered did not differ between the ASD and neurotypical groups. Furthermore, across three source monitoring experiments, the study found that adolescents with ASD were able to recollect contextual information about previously studied items (e.g., what colour a picture had been presented in) as accurately as neurotypical adolescents. This lead the authors to conclude that the results supported "preserved recollection [in individuals with ASD], at least as measured objectively by the source memory performance" (Souchay et al., 2013, p.1605). As such, these findings appear to contradict the suggestion that impairments in metamemory in ASD are the results of impaired recollection in this disorder.

Neuroimaging studies also suggest that individuals do rely on recollection whilst making accurate judgements of learning. For example, Do Lam and colleagues suggest that JOL are based (at least partially) on recollection (Do Lam et al., 2012). In support of this suggestion they found that the brain activation associated with making judgements of learning shared common neural correlates with successful memory retrieval, in particular the medial prefrontal cortex (mPFC). They suggests that the mPFC may be engaged in monitoring retrieval outcomes whilst making JOL assessments (although these findings might also imply that instead/as well meta-mnemonic processes are involved during retrieval).

Whilst it is unlikely that impairment in recollection in ASD can entirely explain impaired performance on FOK and JOC tasks, there are several ways this theory could be assessed. For example within the typically developing literature, neural signatures for specific mnemonic strategies (such as familiarity and recollection) have been identified (Paynter, Reder, & Kieffaber, 2009), and ERPs have been used extensively to investigate memory (Rugg & Curran, 2007) and metamemory (Paynter et al., 2009; Skavhaug, Wilding, & Donaldson, 2010; Skavhaug, Wilding, & Donaldson, 2013) among neurotypical individuals. For example, there is evidence that suggests specific early-onset ERP components are associated with feelings of familiarity during a FOK task (Paynter et al., 2009). As such, future EEG research exploring the neural underpinnings of metacognition in ASD would help to answer the question of whether impairments in metamemory accuracy in ASD are in fact driven by a relative lack of mnemonic cues (as Wojcik and colleagues suggest), or by monitoring impairments (as one-mechanism theories suggest).

**2) Individuals with ASD used an alternative strategy on the JOL tasks**.

Another potential explanation for the pattern of results found across metamemory tasks

is the suggestion that individuals with ASD were able to perform relatively well on the

JOL task despite impaired metamemory, by employing an atypical strategy during the

task. The cue-target JOL task employed in chapter four (experiment 3) directly

addressed this issue. The results from this experiment suggested that performance in

children with ASD was not driven by simply basing each JOL on whether participants

could bring to mind a missing cue word. However, studies within the typically

developing literature have shown that performance on JOL tasks can be accurate when

other strategies are used to predict future memory performance. In addition to factors

that influence encoding and retrieval processes, studies have shown that inferential

factors can also influence participants' judgments of learning decisions (see Chua,

Pergolizzi, & Weintraub, 2014). For example, Sommer and colleagues presented

participants with a series of faces during a facial recognition task (Sommer, Heinz,

Leuthold, Matt, & Schweinberger, 1995). Half of the individuals who participated in this

study were asked to make a JOL for each face, and were asked to judge whether they

thought they would recognise each face in a future recognition task. However, the other

half of the participants were simply asked to rate each face on how distinctive they

found them. Sommer and colleagues found that distinctiveness ratings were as

predictive of future memory performance as JOL rating. This suggests that it is possible

to make perfectly accurate judgments of learning by basing each judgment purely on

inferential information (a strategy that does not rely on metamemory monitoring

ability). If such a strategy was used, this might explain why individuals performed

typically on the JOL task, despite underlying metamemory impairments. This

suggestion is of course tentative, as it is not clear from the studies reported in this thesis

what information individuals with ASD used to inform their JOL decisions (or indeed

their FOK/JOC assessments). Additionally, if it is the case that individuals with ASD

"hacked out" typical performance on the JOL despite poor metacognitive skills, the question of why compensatory strategies were not used (or why compensatory strategies were not as effective) on the FOK and JOC tasks needs to be addressed. As discussed, different metamemory tasks rely on different sources of information. One possibility could be that FOK and JOC tasks require greater metacognitive demands than JOL tasks, and thus it is harder to perform as well on such task by using an alternative strategy. Again this suggestion is tentative, and future research is needed to help answer existing questions surrounding the underlying processes involved during metamemory tasks in individuals with ASD.

**3). Individuals do not possess *general* metamemory abilities.** Interestingly several studies have found that an individual's accuracy on one metamemory task does not tend to correlate with their accuracy on a different task (e.g., Souchay & Isingrini, 2012; Souchay, Isingrini, Clarys, Taconnat, & Eustache, 2004). Indeed, the evidence that individuals have stable metacognitive abilities that are employed across several time and different types of tasks is relatively weak (see Kelemen et al., 2000). This has led some researchers to argue again a general metacognitive ability. If individuals do not hold a general metamemory ability, it is possible that individuals with ASD are only impaired in some aspects of metamemory and not others. In terms of one mechanism theories of mindreading, perhaps only some aspects of metacognition rely on the same underlying mechanisms that mindreading rely on.

However, to counter this suggestion, Keleman suggests that a general metamemory ability (across tasks) is apparent when studies employ a large number of trials (Kelemen et al., 2000). Fleming has provided support for this suggestion by demonstrating participants show good reliability on a perceptual decision making task, when the task employs large numbers of trials (Fleming, Weil, Nagy, Dolan, & Rees,

2010) . Additionally, Song and colleagues found that individual differences in metacognitive accuracy on two perceptual tasks, each of which employed large trial numbers, were strongly correlated (Song et al., 2011). These researchers support the idea that metamemory paradigms measure the same underlying processes (metamemory accuracy), but reliability in performance between tasks is only apparent when large trial numbers are used.

Evidence from neuroimaging studies also supports the suggestion that metamemory tasks share the same underlying processes. For example, studies suggest that the anterior prefrontal cortex (aPFC) is particularly involved in making accurate metamemory judgements, across tasks. Whilst medial regions of the aPFC appear to be involved in making accurate JOL (Kao, Davis, & Gabrieli, 2005) and FOK assessments (Schnyer, Nicholls, & Verfaellie, 2005; Schnyer et al., 2004), more lateral areas of the aPFC are involved in making accurate JOC assessments (Yokoyama et al., 2010). This dissociation is most likely due to the fact that JOL and FOK judgements are prospective in nature, whilst JOC judgments are retrospective in nature (Chua et al., 2014) . This is in keeping with other studies that find *medial* parts of the aPFC are particularly involved in making predictions (see e.g., Bar, 2009). Of course, these studies do not tell us what areas are recruited by individuals with ASD whilst making accurate metamemory judgements, but do support the suggestion that in typical development, accuracy on metamemory tasks is driven by common underlying mechanisms.

Chua and colleagues (Chua, Schacter, & Sperling, 2009) also found that both FOK and JOC tasks appear to produce differential activation (when than in a non-metamemory task) in several overlapping brain regions, including both the left and the right temporo-parietal junctions (TPJ). Interestingly, TPJ also appears to be involved in mental state understanding in others, and reasoning about thoughts, including false

beliefs (see Perner et al., 2006; Samson, Apperly, Chiavarino, & Humphreys, 2004; Saxe & Powell, 2006). This provides some support for the one mechanism account that suggest metacognition and mindreading rely on the same underlying processes. However, to date the only two fMRI studies that have explored the brain regions associated with making judgments of learning, found no indication that TPJ was activated during JOL tasks (Do Lam et al., 2012; Kao et al., 2005). This suggest that perhaps, JOL tasks do not rely on the same processes as those engaged during mindreading, whereas FOK and JOC do. This is another alternative suggestion for why JOL accuracy appears intact in ASD whilst FOK and JOC accuracy appears diminished.

Currently, there is not enough evidence to determine which, if any of these three possibilities, if any, is the best one to explain the pattern of results found in this thesis (impaired FOK and JOC accuracy in ASD, alongside intact JOL accuracy). As such, it will be important that future research investigates the nature of the cues that people with ASD use to inform their JOL assessment (as well as other metamemory judgments).

### Metacognitive control processes in individuals with ASD

When considering the question of whether metacognition is impaired in ASD it is also important to look at the evidence from studies of metacognitive control processes. The JOC task employed in chapter five explored both monitoring and control processes in children with ASD. Although this study found clear evidence that metacognitive monitoring is diminished among children with ASD, it found little evidence to support the idea that metacognitive control was similarly diminished. Having said this, it is notable that control processes were less accurate amongst ASD participants than neurotypical participants (albeit non-significantly so) and that the effect size associated with the between-group difference ($d = 0.38$) was strikingly similar to the effect size for

the group difference in control ability reported by Sawyer et al.,(2014; *d* = 0.32, averaged across general knowledge and emotion recognition conditions). This consistency across the only two studies to have explored metacognitive control ability in ASD might suggest that this ability is subtly diminished in ASD. However, it is important to note that if such a diminution exists, it is questionable whether it is clinically significant, given the small magnitude of the effect.

**Are significant group differences in any aspect of self-awareness simply chance findings?**

Looking at the thesis as a whole, it is striking that only two of the chapters indicate self-awareness is significantly impaired in ASD. Indeed, almost all of the other analysis carried out to explore group difference in self-awareness suggests that there were no significant difference between ASD participants and neurotypical participants. As such, it is worth considering whether the significant between group differences that were found (in FOK gamma scores and JOC accuracy) were simply chance findings? It could be suggested that these differences may be simply the result of carrying out multiple statistical tests rather than indicative of underlying impairments in metacognition. Whilst this is of course possible there are several arguments against this suggestion. Firstly, there are several theoretical reasons to predict that metacognition should be impaired in ASD, and to directly predict group differences in FOK and JOC accuracy. For example, several theories suggest that individuals with ASD should be as impaired at understanding their own mental states as they are at understanding others' mental states (see e.g., Carruthers, 2009; Williams, 2010). As such, impaired accuracy on the FOK and JOC tasks was directly in keeping with *a priori* hypotheses that predicted metacognitive monitoring accuracy should be impaired in individuals with

ASD. Thus, group differences on these measures were expected, in the direction that significant group differences were found. Additionally, impairments in FOK accuracy and JOC accuracy are in keeping with the result of other studies within the literature that report diminished monitoring accuracy in individuals with ASD (Wilkinson et al., 2010; Wojcik et al., 2013). Given these factors, it is highly unlikely that these results are chance findings.

**Implications for theories of self-awareness**

From the results reported in this thesis it is not possible to conclude that metacognition is definitively impaired in ASD. However the results suggest that at least some aspects of private self-awareness are diminished in ASD, yet individuals will demonstrate no obvious impairments in physical self-awareness. As such, overall, the pattern of findings reported in this thesis provides most support for theories of self-awareness that distinguish between physical and psychological aspects of self (Gillihan & Farah, 2005; Neisser, 1988; Williams, 2010). In contrast, the evidence from ASD does not support theories that suggest physical self-awareness is dependent on having psychological self-awareness (Lewis & Ramsay, 2004). Additionally, the case of ASD does not support suggestions that awareness of the physical self is sufficient to provide awareness of the psychological self (Russell, 1996).

Alongside results from the behavioural studies summarised in this thesis, support for this suggestion also comes from neuroimaging studies. During an fMRI study, Kennedy and Courchesne (2008) presented adults with ASD, alongside a group of age and IQ matched neurotypical individuals, with various statements which either referred to psychological personality traits (e.g.., I am polite) or observable, external characteristics/behaviours (e.g., I have a dog). Individuals were asked to judge

(true/false) whether each statement described themselves (Self condition) or their mother (Other condition). One interesting result from this study was that group differences *were* observed in activation in the dorsal medial PFC and the PCC. In both these brain regions, individuals with ASD showed reduced activation relative to the neurotypical group, when judging statements that concerned psychological traits (in both the self and other condition). However they showed similar activation to neurotypical individuals when judging statements that concerned observable traits. This suggests that, perhaps, rather than showing general patterns of atypical activation during self- and other-processing, individuals with ASD may demonstrate a specific deficit in making judgements that rely on inferring mental states (in both oneself and others), but show no impairments making judgements that rely on observable phenomenon (in both oneself and others). Interestingly, Lind (2010) highlights the fact that all existing studies of self-referencing in individuals with ASD have explored self-referencing in relation to *psychological* trait adjectives. As such, they examine whether individuals encode information in relation to psychological aspects of an individual's self-concept. Following the suggestion that that awareness of the physical self may not be impaired in ASD Lind (2010) directly predicts that individuals with ASD *would* show self-reference effects when encoding information in relation to the physical self, as physical aspect of the conceptual self are not diminished in ASD. However, these predictions have never been empirically tested.

**The relation between mindreading and metacognition**

The studies reported in this thesis can also inform the theoretical debate concerning the relation between mindreading and metacognition. Findings of diminished FOK accuracy and JOC accuracy alongside diminished mindreading ability

are in keeping with the predictions of the one mechanism theory of the relation between metacognition and mindreading, which directly predict that both metacognition and mindreading should be impaired in ASD. Certainly, the results of this thesis are not in keeping with key predictions made by either the simulation theory or two-mechanism theories, which both predict that metacognition is unimpaired in ASD. Of course, the results of these studies do not definitively confirm that these abilities are not underpinned by two separate mechanisms (e.g., Goldman, 2006), because it may be that both systems (one underpinning mindreading and the other underpinning metacognition) are impaired in ASD. They are simply not in keeping with predictions made by either Goldman (2006) or Nichols and Stich (2003).

However, the results from this thesis are not entirely in keeping with the one-mechanism accounts. Firstly, one-mechanism accounts would predict global impairments across metamemory tasks. As such, findings that JOL accuracy was typical in individuals with ASD who also demonstrate mindreading impairments do not support one-mechanism accounts of mentalising. Secondly, none of the studies found evidence of a significant positive association between metacognitive accuracy and performance on either the animations task or strange stories task. The one-mechanism account would directly predict such associations between metamemory and mindreading, so the current results do not support the theory in this respect. Whilst the FOK and JOC were underpowered to detect significant associations between mindreading performance and metacognitive accuracy, the JOC study (reported in chapter five) was suitably powered to detect at least a moderate sized ($r = 0.30$) predicted association between these abilities if one existed. Thus, the failure to find a significant association is unlikely to be due to insufficient power. As such, the correlations reported in this thesis do not support one-mechanism accounts.

However, some caution should be taken when interpreting the results of the correlation analyses. Finding that mindreading performance and metacognitive monitoring performance were not associated does not rule out one system theories of the relation between these two abilities. Cognitive measures are rarely (if ever) process pure and thus an association or lack thereof should not be taken in-and-of-itself as evidence for/against the notion that the underlying cognitive abilities are/are not associated.

To summarise, whilst the results from this thesis probably provide most support for one-mechanism accounts of mentalising, they do not definitively support one-mechanism theories, two-mechanism theories or simulation theory. Indeed, several of the studies reported in this thesis found significant correlations not predicted by any theories of mentalising. For example, in experiment 3 of chapter four, JOL accuracy on the cue-target JOL task was strongly negatively association with performance on both conditions of the strange stories task. This indicates that, for ASD participants the better their accuracy on the cue-target JOL task, the poorer their performance on the strange stories (in both conditions). However, given several of the studies reported in the thesis conducted large numbers of correlations it is unsurprising that some unpredicted associations occurred. It is likely that these correlations are chance finding, and it is notable that they would not survive corrections for multiple comparisons.

The main aim of this thesis was not to systematically decide between competing theories of the relation between mindreading and metacognition ,but was rather to better understand the nature of metacognition and self-awareness in ASD. As such, the thesis was fairly limited in the methodology it employed to assess the relationship between mindreading and metacognition (in neurotypical individuals and individuals

with ASD). As discussed, cognitive measures are rarely (if ever) process pure. Thus finding an association or lack of association between two measures does not provide the best support for or against competing theories. Instead of using correlational analyses to explore this question, future research could employ dual-task paradigms to better understand the relationship between mindreading and metacognition in ASD. For example, during a dual task paradigm individuals could engage in a primary task that involved metacognition, thus drawing upon the resources of the mindreading/metacognition faculty (if one-mechanism theories are to be believed). Alongside engaging in this primary task, the task would also involve either a) a secondary mindreading task (e.g., making a key-board response when a mental state term is presented) or b) a secondary task that was equally as demanding, but did not involve mindreading (e.g., making a keyboard response when an animal name is presented). If metacognition does rely on the same underlying processes as those employed in mindreading, you would expect to see impairments on a secondary task that involves mindreading, but not on a secondary task that did not involve mindreading.

**Mindreading performance in individuals with ASD**

Another potential explanation for why the correlational results reported in chapters three, four, and five failed to support one-mechanism account of mentalising, could be the fact that individuals with ASD did not consistently demonstrate predicted patterns of performance on the animations task(a well-established measure of mindreading ability). Interestingly throughout the studies reported in the thesis, performance on the animations task was not always impaired specifically in the *mentalising* condition. In chapters three, four (experiment 1), six and eight performance

was impaired in ASD individuals in both conditions of the task. This raises the question of whether mindreading ability itself specifically was driving diminished performance on the animations task in the ASD participants tested.

Several papers have also used accuracy/appropriateness scores from both the goal-directed and mentalising condition of the animation task to assess mindreading ability in individuals with ASD (Abell et al., 2000; Campbell et al., 2006; Castelli et al., 2002; Jones et al., 2011; Lind, Bowler, et al., 2014; Lind, Williams, et al., 2014; Salter et al., 2008; White et al., 2011; Zwickel et al., 2011). Four of these papers (Lind, Bowler, et al., 2014; Lind, Williams, et al., 2014; Salter et al., 2008; Zwickel et al., 2011) explicitly report a failure to observe a significant group by condition interaction effect. Whilst Lind, Bowler et al., (2014) also failed to find a significant main effect of group, the other three studies failed to find a significant interaction effect, despite observing a significant main effect of group. These findings are in keeping with the findings from chapters three, four, six, and eight of this thesis. Out of the remaining (five) papers, three papers (Campbell et al., 2006; Castelli et al.,2002; Abell et al., 2000) do not report any inferential statistics for a group by condition interaction effect. Thus, it is not clear whether they found a significant group by condition interaction or not. The other two papers (White et al., 2011; Jones et al., 2011) analysed group differences separately in each condition of the animation task (using Mann-Whitney comparisons), and do not discuss interaction effects. As such, of these papers, the only five that explicitly report the group (ASD, neurotypical) by condition (goal-directed/mentalising) interaction effect do not find a significant effect.

One suggestion that explains this pattern of results is the suggestion that both conditions of the task involve some level of mindreading. Behavioural and neuroimaging studies support the view that the goal-directed condition of the task

involves an aspect/level of mindreading, albeit a lower level/different aspect than is tapped by the mentalising condition of the task. Behavioural studies have shown that neurotypical individuals do reliably employ mental-state language when describing the interactions of the triangles in the goal-directed condition, albeit less frequently than when describing the interactions of the triangles in the mentalising condition (e.g., Castelli et al., 2002; Castelli et al., 2000; Klein, Zwickel, Prinz, & Frith, 2009; Zwickel et al., 2011).

In terms of neuroimaging studies, there is strong evidence that mentalising is underpinned by a network of brain regions that consists of the medial prefrontal cortex (mPFC), the posterior cingulate cortex (PCC), superior temporal sulcus (STS), the temporal pole, and (importantly) the temporo-parietal junction (TPJ). The importance of the TPJ is that it is *specifically* involved in mental state understanding and reasoning about thoughts, including false beliefs (see Perner et al., 2006; Samson, Apperly, Chiavarino, & Humphreys, 2004; Saxe & Powell, 2006). Neuroimaging studies have shown that the TPJ (as well as the mPFC, STS, and TPJ) is activated when neurotypical adults watch the clips that comprise the mentalising *and* goal-directed conditions of the task, albeit that this activation is significantly greater in the mentalising condition than the goal-directed condition (see, for example, Castelli et al., 2002; Castelli et al., 2000). Thus, it is highly unlikely that the goal-directed condition has nothing to do with mindreading.

The difference between the mentalising and goal-directed conditions may lie in the fact that success in the goal-directed condition can be achieved by attributing non-propositional attitudes. Leslie (e.g., Leslie, 1994) proposed that the "ToM mechanism" consists of two distinct sub-systems, one subsystem (system$_1$) that processes the goal states that underpin agents' actions (henceforth referred to as "lower-level"

mindreading), and another subsystem (system2) that processes the propositional

attitudes that underpin agents in relation to their propositional attitudes (henceforth

referred to as "higher-level" mindreading). Although the goal-directed condition of the

animations task does not require the attribution of propositional attitudes (and thus

will not rely on system2 of the mindreading mechanism), arguably this condition still

relies on the goal-directed subsystem (system1) of the mindreading mechanism.

This suggestion explains why impairments were seen in both conditions of the

animations task in several of the studies reported in this thesis. However, it cannot

explain why children with ASD who participated in the experiments reported in

chapters five, six (experiment three), and seven failed to demonstrate impairments in

*either condition* of the animations task. In all of the studies reported in this thesis

children with ASD failed to show *any impairment* on the animations task, in either the

mentalising or goal-directed conditions (yet the same samples of participants

consistently demonstrated mindreading impairments on the strange stories task). This

could be due to children only completing four trials on the task (two mentalising trials

and two goal-directed trials). Instead adults who participated in the animations task

completed eight trials (four mentalising trials and four goal-directed trials). This

suggests that perhaps the animations task is only a sensitive measure of mindreading

when participants complete all trials in the task. Alternatively, it may be that only some

of the animation trials are a good measure of mindreading ability, and the four

animations chosen as the trials for the child studies did not measure mindreading

ability as well as other (excluded) trials from the task.

**Questions for future research**

Recently, there has been a growing interest in the study of self-awareness in ASD. Upon the onset of my PhD only two studies (Wilkinson et al., 2010; Wojcik et al., 2011) had employed metamemory paradigms to explore psychological self-awareness in ASD. Since then there have been five more published studies of metacognition in ASD (Grainger et al., 2014; Sawyer et al., 2014; Wojcik et al., 2013; Wojcik et al., 2014; Elmose & Happé, 2014). These studies have clearly extended our understanding of (psychological) self-awareness in ASD. However, it is evident that several questions concerning self-awareness in ASD still remain.

Within the typically developing literature there is a growing body of literature that had explored the underlying processes involves in metacognition using both fMRI and EEG techniques. Future research employing such techniques in individuals with ASD will be crucial to help answer questions surrounding the underlying processes engaged during tasks that assess self-awareness in ASD (see also p.252 above). Additionally, future research should also explore whether aspects of self-awareness can be fostered in individuals with this disorder. Previous studies have indicated that it is possible to improve individuals' mindreading performance (Begeer et al., 2011; Fisher et al., 2005). One-mechanism theorists should predict that, after mindreading training, individuals with ASD should also demonstrate improvements in metacognitive accuracy. In contrast, two-mechanism theories predict that mindreading training will not improve metacognitive.

Additionally, studies have shown that transcranial direct current stimulation (tDCS) to the right temporal parietal junction can improve neurotypical participants' ability to make judgments about both the self and the other (Santiesteban, Banissy, Catmur, & Bird, 2012). As the authors suggest these "findings demonstrate the efficacy of tDCS to improve social cognition, and highlight the potential for tDCS to be used as a

tool to aid self other processing in clinical populations" (Santiesteban et al., 2012, p.2274). Given studies have indicated that TPJ appears to be engaged when making FOK and JOC judgements, it would be interesting to investigate whether tDCS stimulation to TPJ could also enhance metamemory accuracy in both neurotypical individuals and individuals with ASD.

Furthermore, within the typically developing literature it is still not clear whether metacognitive accuracy is stable across domains. For example, Fleming has shown that individuals with anterior prefrontal lesions demonstrate domain-specific impairments in metacognitive accuracy, showing impairments in perceptual decision making judgements, but intact metamemory accuracy (Fleming, Ryu, Golfinos, & Blackmon, 2014). Whilst the theories proposed in this thesis predict that impairments in metacognition in ASD should be domain-general, to date no study has explore metacognitive accuracy in individuals with ASD on a perceptual decision making task. As such, it will be important to establish whether predicted impairments in metacognition are global or domain-specific in the future.

**Implications and final comments**

Alongside theoretical implications, the results of this thesis also have several potential practical and clinical implications. Ultimately, if an individual has a reduced ability to accurately assess what information they know, and what they do not know, this may have several consequences. From an educational perspective, studies have shown that several outcomes (such as exam performance) can be predicted by metacognitive monitoring accuracy (e.g., Hartwig et al., 2012; Thiede et al., 2003). Other studies have shown that metacognitive training has been shown to remediate difficulties in reading, writing and mathematical reasoning (see Brown & Campione,

1996; Fuchs et al., 2003; Sitko, 1998) in typical development. Such studies suggest that impairments in psychological self-awareness and metacognition will significantly impact an individual's educational achievements and learning abilities. This may be particularly relevant for intellectually high-functioning people with ASD, given that many of these individuals show significantly lower academic achievement than would be expected on the basis of their intelligence, which in turn impacts negatively on their life chances (see Estes et al., 2011)

It is important to establish a comprehensive account of self-awareness in ASD, given the consequences that impaired self-awareness (particularly impairments in metacognition) may have on regulation and an individual's cognitive performance. Future research should explore the educational and social consequences of specific impairments in self-awareness. It is hoped that alongside future research, the findings from this thesis will help to establish a more definitive account of self-awareness in ASD, and that a greater understanding of this area will eventually contribute to successful remediation of cognitive and behavioural impairments in this disorder.

**APPENDICES**

**Appendix 1:**   Details of the coding scheme used to score the Animations Task, based on the coding scheme Abell, Happé and Frith (2000).

**Appendix 2:**   Details of the coding scheme used to score the Strange Stories task based on the scoring criteria outlined in White, Hill, Happé and Frith (2009)

**Appendix 3:**   Details of the coding scheme used to code the Twenty Statements Task, based on the coding scheme outlined in Rhee et al., (1995).

**Appendix 1:  Details of the coding scheme used to score the Animations Task, based on the coding scheme Abell, Happé and Frith (2000).**

**General Rules**

Each description is scored 2, 1, or 0 according to how accurately it reflects the sequence.

>   2 = spot-on description of the story or the actions represented; can be concise just capturing gist, or can be discursive.

>   1 = partial description of the sequence; description is related to the sequence, but imprecise or incomplete.

>   0 = bizarre descriptions, plainly wrong descriptions, and responses that focus solely on a minor unimportant aspect of the sequence.

**Specific Rules - Random movement sequences**

**Floating/Bouncing**: Character roles: just triangles (both sequences without enclosure)

>   2 = anything implying random or purposeless movement including moving, bouncing, just dancing.

>   1 = purposeful movement without interaction, including turning round and getting dizzy, dancing in a circle.

>   0 = purposeful movement implying interaction between the triangles including copying each other; avoiding each other.

**Specific Rules - Goal directed movement sequences**

**Fighting:** Character roles: two deer.  No enclosure.

>   2 = action implying physical fight, e.g., bashing each other.

>   1 = action that conveys the idea of a conflict, but is too specific or too vague, e.g., biting; pushing.

>   0 = action that does not relate to conflict, e.g., following each other.

**Following:** Character roles: mother duck and duckling.  Enclosure.

>   2 = description which conveys following each other.

>   1 = description that is related to but somewhat remote from following, e.g., copying; chasing.

0 = action that does not relate to following each other, e.g., jumping.

**Chasing:** Character roles: two cats.  Enclosure.

2 = description that conveys the idea of a chase.

1 = description that is related to but somewhat remote from chasing.

0 = action that does not relate to chasing, e.g., going up and down.

**Dancing:** Character roles: two ponies.  No enclosure.

2 = description that conveys the idea of moving in formation, e.g., dancing; making a pattern.

1 = description that is partially correct or related to dancing, e.g., doing different things e.g., one went one way, the other went the other way.

0 = action that is not related to moving in formation, e.g., galloping along.


## Specific Rules - ToM movement sequences


**Surprising:** Character roles: grandma and grandson.  Enclosure.

2 = any mention of boy tricking, surprising his grandma; hiding, hide and seek.

1 = description which gives part of the story but misses the critical point (see above).

0 = description which gives only minor part of action e.g., knocking on the door, or does not relate to any of the events in the sequence.

**Coaxing:** Character roles: mother and child.  Enclosure.

2 = description that conveys child's reluctance to go out and mother's attempts to get child out, e.g., persuading.

1 = partially correct description focussing on one aspect of the story or one character only, e.g., child does not want to go out; or, mother is pushing child to go out.

0= actions that do not relate to the events or relate to a minor aspect of the sequence only, e.g., dancing together, or unrelated description..

**Mocking:** Character roles: teacher and boy.  No enclosure.

2 = description that conveys that boy is copying teacher without the teacher noticing, including pretending, hiding, being naughty.

1 = partially correct, e.g., following, copying.

0 = focus on a single unimportant event, e.g., boy ran away, or unrelated description.

**Seducing:** Character roles: girl prisoner and guard.  Enclosure.

2 = description that conveys the girl prisoner luring, persuading or tricking the guard.

1 = partial story with minimal action for each character, e.g., girl trying to escape; guard blocking.

0 = description which focusses on unimportant event or is extremely minimal, e.g., she got out, or unrelated description.

**Appendix 2:  Details of the coding scheme used to score the Strange Stories task based on the scoring criteria outlined in White, Hill, Happé and Frith (2009).**

**General Rules**

> 2= Full and complete aswer.
>
> 1= incomplete or partially correct answer.
>
> 0= incorrect answer.

**Specific Rules- Theory of Mind Stories**

**Prisoner ("Why did the prisoner say that?"):**

> 2 = reference to fact that other army will not believe and hence look in other place, reference to prisoner's realization that that's what they'll do, or reference to double bluff.
> 1 = reference to outcome (to save his army's tanks) or to mislead them.
> 0 = reference to motivation that misses the point of double bluff (he was scared).

**Kittens ("Why did Mrs Smith say that?"):**

> 2 = reference to persuasion, manipulating feelings, trying to induce guilt / pity.
> 1 = reference to outcome (to sell them or get rid of them in a way which implies not drowning) or simple motivation (to make Jill sad).
> 0 = reference to general knowledge or dilemma without realization that the statement was.
> not true (she's a horrible woman).

**Specific rules- Phsyical Stoires**

**Armies ("Why did the blue army win?"):**

> 2 = reference to both weather conditions and either relative ground superiority or inability of other army's planes to be useful in fog (names of armies unimportant).

1 = reference either to weather or relative superiority on ground versus air (because it was foggy); nothing about why weather makes it especially difficult for planes or nothing about planes being affected more than tanks; reference to fog to justify incorrect response (the aeroplanes won because the fog meant they could hide from the tanks).

0 = reference to irrelevant or incorrect information (they won because they had better planes); justifications for why tanks are better than planes.

**Burglar ("Why did the alarm go off?") :**

2 = reference to animal which the burglar disturbed setting off alarm by crossing beam (type of animal unimportant).

1 = reference to burglar setting off alarm (he was startled by the animal so crossed the beam); reference to animal setting off alarm without explaining it crossed the beam (he trod on a cat and it set off the alarm).

0 = reference to irrelevant or incorrect factors (the animal's screech set off the alarm); alternative reasons for alarm going off (a security camera saw him and set the alarm off)

**Appendix 3: Coding system used to code the Twenty Statements Task, based on the coding scheme outlined in Rhee et al., (1995).**

| Category and subcategory | Psychological or physical | Abstract or specific | Autonomous or social |
|---|---|---|---|
| **1. Trait** | | | |
| Pure (e.g., kind, friendly) | Psychological | Abstract | Autonomous |
| Qualified | | | |
| Contextualised (e.g., at home) | Psychological | Specific | Social |
| Temporal (e.g., sometimes) | Psychological | Specific | Social |
| **2. Social Identity** | | | |
| Role Status (e.g., student) | / | Specific | Social |
| Family Info (e.g., sister, dad) | / | Specific | Social |
| Ethnicity/race/nationality | / | Specific | Social |
| Gender (e.g., boy, female) | Physical | Specific | Social |
| Self-ascribed identities (e.g., dancer) | / | Specific | Social |
| Origin (e.g., from Scotland) | / | Specific | Social |
| Religion (e.g., a Buddhist) | / | Specific | Social |
| Occupation (e.g., a lecturer) | / | Specific | Social |
| Negation (e.g., not a Christian) | / | Specific | Social |
| Name (e.g., I am Josh) | / | Specific | Autonomous |
| Sexual Orientation (e.g., homosexual) | / | Specific | Social |
| **3. Specific Attribute** | | | |
| Preferences (e.g., interests, likes) | | | |
| Autonomous | Psychological | Specific | Autonomous |
| Social | Psychological | Specific | Social |
| Aspirations (e.g., hopes wishes) | | | |
| Autonomous | Psychological | Specific | Autonomous |
| Social | Psychological | Specific | Social |
| Activities (e.g., habits, activities) | | | |
| Autonomous | / | Specific | Autonomous |
| Social | / | Specific | Social |
| **4. Evaluative Description** | | | |
| Autonomous (e.g., good at maths) | / | Specific | Autonomous |
| Social (a good friend) | / | Specific | Social |
| **5. Physical descriptions** | | | |
| Descriptive (e.g., tall, blonde) | Physical | Specific | Autonomous |
| Age | Physical | Specific | Autonomous |
| Factual (e.g., height, weight) | Physical | Specific | Autonomous |
| Physical condition (e.g., long-sighted) | Physical | Specific | Autonomous |
| **6. Emotional State** | | | |
| Autonomous (e.g., worried) | Psychological | Abstract | Autonomous |
| Social (e.g., in love) | Psychological | Specific | Social |
| **7. Peripheral Information** | / | | |
| Immediate situations (e.g., tired) | / | Specific | Autonomous |
| Present residence (e.g., live at home) | / | Specific | Social |
| Other's descriptions | / | Specific | Social |
| Possessions (e.g., clothes) | / | Specific | Autonomous |
| **8. Global Descriptions** | | | |
| Universal (e.g., human, earthling) | / | Abstract | Social |
| Existential (e.g., I am me) | / | Abstract | Autonomous |

*Note:* Responses such as "I am Gryffindor" were coded as nonsense. Repeated descriptions were not coded.

**REFERENCES**

Abell, F., Happé, F., & Frith, U. (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. Cognitive Development, 15(1), 1-16. doi:10.1016/S0885-2014(00)00014-9

Altgassen, M., Koban, N., & Kliegel, M. (2012). Do adults with autism spectrum disorders compensate in naturalistic prospective memory tasks? Journal of autism and developmental disorders, 42(10), 2141-2151. doi:10.1007/s10803-012-1466-3

Ambery, F. Z., Russell, A. J., Perry, K., Morris, R., & Murphy, D. G. (2006). Neuropsychological functioning in adults with Asperger syndrome. Autism, 10(6), 551-564. doi:10.1177/1362361306068507

American Psychiatric Association. (2000). Diagnostic and statistical manual of mental disorders (4th edition, text revised) (DSM-IV-TR). Washington DC: American Psychiatric Association.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th edition). Washington DC: American Psychiatric Association.

Amsterdam, B. (1972). Mirror self-image reactions before age two. Dev Psychobiol, 5(4), 297-305. doi:10.1002/dev.420050403

Arbuckle, T. Y., & Cuddy, L. L. (1969). Discrimination of item strength at time of presentation. Journal of Experimental Psychology, 81(1), 126-131. doi:10.1037/h0027455

Asperger, H. (1944). Die 'Autistichen Psychopathen' im Kindersalter. Archiv fur Psychiatrie und Nervenkrankheiten, 117, 76-136.

Atance, C. M., & O'Neill, D. K. (2005). The emergence of episodic future thinking in humans. Learning and Motivation, 36(2), 126-144. doi:10.1016/j.lmot.2005.02.003

Avikainen, S., Kulomäki, T., & Hari, R. (1999). Normal movement reading in Asperger subjects. Neuroreport, 10(17), 3467-3470. doi:10.1097/00001756-199911260-00001

Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., & Charman, T. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). Lancet, 368(9531), 210-215. doi:10.1016/S0140-6736(06)69041-7

Baker-Ward, L., Hess, T. M., & Flannagan, D. A. (1990). The effects of involvement on children's memory for events. Cognitive Development, 5(1), 55-69. doi:10.1016/0885-2014(90)90012-I

Bar, M. (2009). The proactive brain: memory for predictions. Philosophical Transactions of the Royal Society B: Biological Sciences, 364(1521), 1235-1243. doi:10.1098/rstb.2008.0310

Barnes, K. A., & Dougherty, M. R. (2007). The effect of divided attention on global judgment of learning accuracy. The American journal of psychology, 120(3), 347-359. doi:10.2307/20445409

Baron-Cohen, S. (1992). Out of sight or out of mind? Another look at deception in autism. J Child Psychol Psychiatry, 33(7), 1141-1155. doi:10.1111/j.1469-7610.1992.tb00934.x

Baron-Cohen, S. (2005). Autism—"Autos": Literally, a Total Focus on the Self? In T. Feinberg & J. Keenan (Eds.), The Lost Self. Oxford: Oxford University Press.

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

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger syndrome/high-

functioning autism, males and females, scientists and mathematicians. Journal of autism and developmental disorders, 31(1), 5-17. doi:10.1023/A:1005653411471

Bastin, C., Feyers, D., Souchay, C., Guillaume, B., Pepin, J. L., Lemaire, C., . . . Salmon, E. (2012). Frontal and posterior cingulate metabolic impairment in the behavioral variant of frontotemporal dementia with impaired autonoetic consciousness. Human Brain Mapping, 33(6), 1268-1278. doi:10.1002/hbm.21282

Begeer, S., Gevers, C., Clifford, P., Verhoeve, M., Kat, K., Hoddenbach, E., & Boer, F. (2011). Theory of mind training in children with autism: A randomized controlled trial. Journal of autism and developmental disorders, 41(8), 997-1006. doi:10.1007/s10803-010-1121-9

Ben Shalom, D., Mostofsky, S. H., Hazlett, R. L., Goldberg, M. C., Landa, R. J., Faran, Y., . . . Hoehn-Saric, R. (2006). Normal physiological emotions but differences in expression of conscious feelings in children with high-functioning autism. Journal of autism and developmental disorders, 36(3), 395-400. doi:10.1007/s10803-006-0077-2

Blakemore, S. J., Tavassoli, T., Calo, S., Thomas, R. M., Catmur, C., Frith, U., & Haggard, P. (2006). Tactile sensitivity in Asperger syndrome. Brain Cogn, 61(1), 5-13. doi:10.1016/j.bandc.2005.12.013

Boria, S., Fabbri-Destro, M., Cattaneo, L., Sparaci, L., Sinigaglia, C., Santelli, E., . . . Rizzolatti, G. (2009). Intention understanding in autism. Plos One, 4(5), e5596. doi:10.1371/journal.pone.0005596

Bosch, G. (1970). Infantile autism: Springer.

Boucher, J., & Bowler, D. M. (2008). Memory in Autism. Cambridge: Cambridge University Press.

Bowler, D. M. (1992). Theory of mind in aspergers syndrome. Journal of Child Psychology and Psychiatry and Allied Disciplines, 33(5), 877-893. doi:10.1111/j.1469-7610.1992.tb01962.x

Bowler, D. M., Gaigg, S. B., & Gardiner, J. M. (2008). Subjective organisation in the free recall learning of adults with Asperger's syndrome. Journal of autism and developmental disorders, 38(1), 104-113. doi:10.1007/s10803-007-0366-4

Bowler, D. M., Gardiner, J. M., & Grice, S. J. (2000). Episodic memory and remembering in adults with Asperger syndrome. Journal of autism and developmental disorders, 30(4), 295-304. doi:10.1023/a:1005575216176

Boyd, B. A., McBee, M., Holtzclaw, T., Baranek, G. T., & Bodfish, J. W. (2009). Relationships among repetitive behaviors, sensory features, and executive functions in high functioning autism. Research in Autism Spectrum Disorders, 3(4), 959-966. doi:10.1016/j.rasd.2009.05.003

Brown, A. L., & Campione, J. C. (1996). Psychological theory and the design of innovative learning environments: On procedures, principles, and systems. In L. Schauble & R. Glaser (Eds.), Innovations in learning: New environments for education. Hillsdale, NJ, England: Lawrence Erlbaum Associates.

Bruck, M., London, K., Landa, R., & Goodman, J. (2007). Autobiographical memory and suggestibility in children with autism spectrum disorder. Development and Psychopathology, 19(1), 73-95. doi:10.1017/s0954579407070058

Brugha, T. S., McManus, S., Bankart, J., Scott, F., Purdon, S., Smith, J., . . . Meltzer, H. (2011). Epidemiology of autism spectrum disorders in adults in the community in England. Arch Gen Psychiatry, 68(5), 459-465. doi:10.1001/archgenpsychiatry.2011.38

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. Trends Cogn Sci, 11(2), 49-57. doi:10.1016/j.tics.2006.11.004

Campbell, R., Lawrence, K., Mandy, W., Mitra, C., Jeyakuma, L., & Skuse, D. (2006). Meanings in motion and faces: Developmental associations between the processing of intention from geometrical animations and gaze detection accuracy. Development and Psychopathology, 18(1), 99-118. doi:10.1017/S0954579406060068

Carruthers, P. (2009). How we know our own minds: The relationship between mindreading and metacognition. Behavioral and Brain Sciences, 32(2), 121-182. doi:10.1017/S0140525x09000545

Cartwright-Hatton, S., & Wells, A. (1997). Beliefs about Worry and Intrusions: The Meta-Cognitions Questionnaire and its Correlates. Journal of Anxiety Disorders, 11(3), 279-296. doi:10.1016/S0887-6185(97)00011-X

Castelli, F., Frith, C., Happé, F., & Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. Brain, 125(Pt 8), 1839-1849. doi:10.1093/brain/awf189

Castelli, F., Happé, F., Frith, U., & Frith, C. (2000). Movement and mind: A functional imaging study of perception and interpretation of complex intentional movement patterns. Neuroimage, 12(3), 314-325. doi:10.1006/nimg.2000.0612

Chua, E., Pergolizzi, D., & Weintraub, R. (2014). The cognitive neuroscience of metamemory monitoring: Understanding metamemory processes, subjective levels expressed and metacognitive accuracy. In S. Fleming & C. Frith (Eds.), The Cognitive Neuroscience of Metacognition. New York: Springer.

Chua, E. F., Schacter, D. L., & Sperling, R. A. (2009). Neural Correlates of Metamemory: A Comparison of Feeling-of-Knowing and Retrospective Confidence Judgments. Journal of cognitive neuroscience, 21(9), 1751-1765. doi:10.1162/jocn.2009.21123

Clare, L., Wilson, B. A., Carter, G., Breen, K., Gosses, A., & Hodges, J. R. (2000). Intervening with everyday memory problems in dementia of Alzheimer type: An errorless learning approach. Journal of Clinical and Experimental Neuropsychology, 22(1), 132-146. doi:10.1076/1380-3395(200002)22:1;1-8;FT132

Cohen, J. (1969). Statistical power analysis for the behavioural sciences. San Diego, CA: Academic Press.

Cohen, J. (1992). A Power Primer. Psychological Bulletin, 112(1), 155-159. doi:10.1037//0033-2909.112.1.155

Coltheart, M. (1981). The MRC psycholinguistic database. . Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 33(33), 497-505. doi:10.1080/14640748108400805

Constantino, J. N., Davis, S. A., Todd, R. D., Schindler, M. K., Gross, M. M., Brophy, S. L., . . . Reich, W. (2003). Validation of a brief quantitative measure of autistic traits: Comparison of the social responsiveness scale with the autism diagnostic interview-revised. Journal of autism and developmental disorders, 33(4), 427-433. doi:10.1023/A:1025014929212

Conway, M. A. (2005). Memory and the self. Journal of memory and language, 53(4), 594-628. doi:10.1016/j.jml.2005.08.005

Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. Psychol Rev, 107(2), 261-288. doi:10.1037//0033-295X.107.2.261

Courage, M. L., Edison, S. C., & Howe, M. L. (2004). Variability in the early development of visual self-recognition. Infant Behavior and Development, 27(4), 509-532. doi:10.1016/j.infbeh.2004.06.001

Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. Journal of experimental Psychology: general, 104(3), 268. doi:10.1037//0096-3445.104.3.268

Crane, L., & Goddard, L. (2008). Episodic and semantic autobiographical memory in adults with autism spectrum disorders. Journal of autism and developmental disorders, 38(3), 498-506. doi:10.1007/s10803-007-0420-2

Crane, L., Lind, S. E., & Bowler, D. M. (2013). Remembering the past and imagining the future in autism spectrum disorder. Memory, 21(2), 157-166. doi:10.1080/09658211.2012.712976

Crane, L., Pring, L., Jukes, K., & Goddard, L. (2012). Patterns of Autobiographical Memory in Adults with Autism Spectrum Disorder. Journal of autism and developmental disorders, 42(10), 2100-2112. doi:10.1007/s10803-012-1459-2

Damon, W., & Hart, D. (1982). The development of self-understanding from infancy through adolescence. Child development, 53(4), 841-864. doi:10.1111/j.1467-8624.1982.tb01347.x

Damon, W., & Hart, D. (1988). Self-understanding in childhood and adolescence. New York: Cambridge University Press.

Dapretto, M., Davies, M. S., Pfeifer, J. H., Scott, A. A., Sigman, M., Bookheimer, S. Y., & Iacoboni, M. (2006). Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. Nature Neuroscience, 9(1), 28-30. doi:10.1038/nn1611

David, N., Bewernick, B. H., Cohen, M. X., Newen, A., Lux, S., Fink, G. R., . . . Vogeley, K. (2006). Neural representations of self versus other: Visual-spatial perspective taking and agency in a virtual ball-tossing game. Journal of cognitive neuroscience, 18(6), 898-910. doi:10.1162/jocn.2006.18.6.898

David, N., Gawronski, A., Santos, N. S., Huff, W. G., Lehnhardt, F. G., Newen, A., & Vogeley, K. (2008). Dissociation between key processes of social cognition in autism: Impaired mentalizing but intact sense of agency. Journal of autism and developmental disorders, 38(4), 593-605. doi:10.1007/s10803-007-0425-x

Dawson, G., & McKissick, F. C. (1984). Self-recognition in autistic children. J Autism Dev Disord, 14(4), 383-394. doi:10.1007/BF02409829

Dichter, G. S., Lam, K. S., Turner-Brown, L. M., Holtzclaw, T. N., & Bodfish, J. W. (2009). Generativity abilities predict communication deficits but not repetitive behaviors in Autism Spectrum Disorders. J Autism Dev Disord, 39(9), 1298-1304. doi:10.1007/s10803-009-0742-3

Dinstein, I., Thomas, C., Humphreys, K., Minshew, N., Behrmann, M., & Heeger, D. J. (2010). Normal Movement Selectivity in Autism. Neuron, 66(3), 461-469. doi:10.1016/j.neuron.2010.03.034

Dipellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events - a neurophysiological study. Experimental Brain Research, 91(1), 176-180. doi:10.1007/BF00230027

Do Lam, A. T., Axmacher, N., Fell, J., Staresina, B. P., Gauggel, S., Wagner, T., . . . Weis, S. (2012). Monitoring the mind: the neurocognitive correlates of metamemory. Plos One, 7(1), e30009. doi:10.1371/journal.pone.0030009

Dunlosky, J., Domoto, P. K., Wang, M. L., Ishikawa, T., Roberson, I., Nelson, T. O., & Ramsay, D. S. (1998). Inhalation of 30% nitrous oxide impairs people's learning without impairing people's judgments of what will be remembered. Experimental and Clinical Psychopharmacology, 6(1), 77-86. doi:10.1037/1064-1297.6.1.77

Dunlosky, J., Kubat-Silman, A. K., & Hertzog, C. (2003). Training monitoring skills improves older adults' self-paced associative learning. Psychology and Aging, 18(2), 340-345. doi:10.1037/0882-7974.18.2.340

Dunlosky, J., & Metcalfe, J. (2008). Metacognition: Sage Publications.

Dunlosky, J., & Nelson, T. O. (1992). Importance of the kind of cue for judgments of learning (JOL) and the delayed-JOL effect. Memory & Cognition, 20(4), 374-380. doi:10.3758/BF03210921

Dunlosky, J., & Nelson, T. O. (1997). Similarity between the cue for judgments of learning (JOL) and the cue for test is not the primary determinant of JOL accuracy. Journal of memory and language, 36(1), 34-49. doi:10.1006/jmla.1996.2476

Elmose, M. & Happé, F. (2014). Being aware of own perfromance: How accurately do children with autism spectrum disorder judge own memory perfromance? *Autism Research, 7*, 712-719. doi: 10.1002/aur.1421

Engelkamp, J. (1998). Memory for actions. Hove: Psychology Press.

Engelkamp, J., & Zimmer, H. D. (1989). Memory for action events - A new field of research. Psychological Research-Psychologische Forschung, 51(4), 153-157. doi:10.1007/BF00309142

Estes, A., Rivera, V., Bryan, M., Cali, P., & Dawson, G. (2011). Discrepancies Between Academic Achievement and Intellectual Ability in Higher-Functioning School-Aged Children with Autism Spectrum Disorder. Journal of autism and developmental disorders, 41(8), 1044-1052. doi:10.1007/s10803-010-1127-3

Farrant, A., Blades, M., & Boucher, J. (1998). Source monitoring by children with autism. Journal of autism and developmental disorders, 28(1), 43-50. doi:10.1023/A:1026010919219

Farrant, A., Boucher, J., & Blades, M. (1999). Metamemory in children with autism. Child development, 70(1), 107-131.

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

Ferrari, M., & Matthews, W. S. (1983). Self-recognition deficits in autism: syndrome-specific or general developmental delay? J Autism Dev Disord, 13(3), 317-324. doi:10.1007/BF01531569

Fisher, N., Happé, F., & Dunn, J. (2005). The relationship between vocabulary, grammar, and false belief task performance in children with autistic spectrum disorders and children with moderate learning difficulties. Journal of Child Psychology and Psychiatry, 46(4), 409-419. doi:10.1111/j.1469-7610.2004.00371.x

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive developmental inquiry. American Psychologist, 34, 906-911. doi:10.1037//0003-066X.34.10.906

Flavell, J. H. (2000). Development of children's knowledge about the mental world. International Journal of Behavioral Development, 24(1), 15-23. doi:10.1080/016502500383421

Fleming, S. M., & Dolan, R. J. (2012). The neural basis of metacognitive ability. Philosophical Transactions of the Royal Society B-Biological Sciences, 367(1594), 1338-1349. doi:10.1098/rstb.2011.0417

Fleming, S. M., Ryu, J., Golfinos, J. G., & Blackmon, K. E. (2014). Domain-specific impairment in metacognitive accuracy following anterior prefrontal lesions. Brain, 137(10), 2811-2822. doi:10.1093/brain/awu221

Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J., & Rees, G. (2010). Relating introspective accuracy to individual differences in brain structure. Science, 329(5998), 1541-1543. doi:10.1126/science.1191883

Freeman, J. E., & Ellis, J. A. (2003). The representation of delayed intentions: A prospective subject-performed task? Journal of Experimental Psychology-Learning Memory and Cognition, 29(5), 976-992. doi:10.1037/0278-7393.29.5.976

Frith, C. D., Blakemore, S.-J., & Wolpert, D. M. (2000). Explaining the symptoms of schizophrenia: Abnormalities in the awareness of action. Brain Research Reviews, 31(2–3), 357-363. doi:10.1016/S0165-0173(99)00052-1

Frith, C. D., & Done, D. J. (1989). Experiences of alien control in schizophrenia reflect a disorder in the central monitoring of action. Psychological Medicine, 19(02), 359-363.

Frith, C. D., & Frith, U. (2006). The neural basis of mentalizing. Neuron, 50(4), 531-534. doi:10.1016/j.neuron.2006.05.001

Frith, U. (1989). Autism: Explaining the Enigma. Oxford: Blackwell.

Frith, U. (2003). Autism: Explaining the Enigma (2nd Edition). Oxford: Blackwell.

Frith, U., & Happé, F. (1999). Theory of mind and self-consciousness: What is it like to be autistic? Mind & Language, 14(1), 1-22. doi:10.1111/1468-0017.00100

Frith, U., & Hermelin, B. (1969). The role of visual and motor cues for normal, subnormal and autistic children. Journal of Child Psychology and Psychiatry, 10(3), 153-163. doi:10.1111/j.1469-7610.1969.tb02077

Fuchs, L. S., Fuchs, D., Prentice, K., Burch, M., Hamlett, C. L., Owen, R., & Schroeter, K. (2003). Enhancing third-grade students' mathematical problem solving with self-regulated learning strategies. Journal of Educational Psychology, 95(2), 306-315. doi:10.1037/0022-0663.95.2.306

Gaigg, S. B., Gardiner, J. M., & Bowler, D. A. (2008). Free recall in autism spectrum disorder: The role of relational and item-specific encoding. Neuropsychologia, 46(4), 983-992. doi:10.1016/j.neuropsychologia.2007.11.011

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. Trends Cogn Sci, 2(12), 493-501. doi:10.1016/S1364-6613(98)01262-5

Gallup, G. G. (1970). Chimpanzees: Self-Recognition. Science, 167(3914), 86-87. doi:10.1126/science.167.3914.86

Gardiner, J. M., & Java, R. I. (1990). Recollective experience in word and nonword recognition. Memory & Cognition, 18(1), 23-30. doi:10.3758/bf03202642

Gilbert, S. J., Armbruster, D. J., & Panagiotidi, M. (2012). Similarity between brain activity at encoding and retrieval predicts successful realization of delayed intentions. Journal of cognitive neuroscience, 24(1), 93-105. doi:10.1162/jocn_a_00094

Gillihan, S. J., & Farah, M. J. (2005). Is self special? A critical review of evidence from experimental psychology and cognitive neuroscience. Psychological Bulletin, 131(1), 76-97. doi:10.1037/0033-2909.131.1.76

Goddard, L., Howlin, P., Dritschel, B., & Patel, T. (2007). Autobiographical Memory and Social Problem-solving in Asperger Syndrome. Journal of autism and developmental disorders, 37(2), 291-300. doi:10.1007/s10803-006-0168-0

Goldman, A. (2006). Simulating Minds: The Philosophy, Psychology and Neuroscience of Mindreading. Oxford: Oxford University Press.

Goodman, L. A., & Kruskal, W. H. (1954). Measures of association for cross classifications. Journal of the American Statistical Association, 49(268), 732-764. doi:10.2307/2281536

Goschke, T., & Kuhl, J. (1993). Representation of intentions: Persisting activation in memory. Journal of Experimental Psychology: Learning, Memory, and Cognition, 19(5), 1211. doi:10.1037/0278-7393.19.5.1211

Grainger, C., Williams, D. M., & Lind, S. E. (2014). Metacognition, metamemory, and mindreading in high-functioning adults with autism spectrum disorder. Journal of Abnormal Psychology, 123(3), 650-659. doi:10.1037/a0036531

Hacker, D. J., Bol, L., Horgan, D. D., & Rakow, E. A. (2000). Test prediction and performance in a classroom context. Journal of Educational Psychology, 92(1), 160-170. doi:10.1037//0022-0663.92.1.160

Hala, S., Rasmussen, C., & Henderson, A. E. (2005). Three types of source monitoring by children with and without autism: The role of executive function. Journal of autism and developmental disorders, 35(1), 75-89. doi:10.1007/s10803-004-1036-4

Hamilton, A. F. D., Brindley, R. M., & Frith, U. (2007). Imitation and action understanding in autistic spectrum disorders: How valid is the hypothesis of a deficit in the mirror neuron system? Neuropsychologia, 45(8), 1859-1868. doi:10.1016/j.neuropsychologia.2006.11.022

Happé, F. (1999). Autism: cognitive deficit or cognitive style? Trends Cogn Sci, 3(6), 216-222. doi:10.1016/S1364-6613(99)01318-2

Happé, F., Booth, R., Charlton, R., & Hughes, C. (2006). Executive function deficits in autism spectrum disorders and attention-deficit/hyperactivity disorder: examining profiles across domains and ages. Brain Cogn, 61(1), 25-39. doi:10.1016/j.bandc.2006.03.004

Happé, F., & Ronald, A. (2008). The 'fractionable autism triad': a review of evidence from behavioural, genetic, cognitive and neural research. Neuropsychol Rev, 18(4), 287-304. doi:10.1007/s11065-008-9076-8

Happé, F., Ronald, A., & Plomin, R. (2006). Time to give up on a single explanation for autism. Nat Neurosci, 9(10), 1218-1220. doi:10.1038/nn1770

Happé, F. G. E. (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(2), 129-154. doi:10.1007/bf02172093

Happé, F. G. E. (1995). The role of age and verbal-ability in the theory of mind task-performance of subjects with autism. Child development, 66(3), 843-855. doi:10.2307/1131954

Hare, D. J., Mellor, C., & Azmi, S. (2007). Episodic memory in adults with autistic spectrum disorders: Recall for self- versus other-experienced events. Research in Developmental Disabilities, 28(3), 317-329. doi:10.1016/j.ridd.2006.03.003

Harms, M. B., Martin, A., & Wallace, G. L. (2010). Facial emotion recognition in autism spectrum disorders: a review of behavioral and neuroimaging studies. Neuropsychology review, 20(3), 290-322. doi:10.1007/s11065-010-9138-6

Hart, J. T. (1965). Memory and the feeling-of-knowing experience. Journal of Educational Psychology, 56(4), 208-216. doi:10.1037/h0022263

Hartwig, M. K., Was, C. A., Isaacson, R. M., & Dunlosky, J. (2012). General knowledge monitoring as a predictor of in-class exam performance. British Journal of Educational Psychology, 82(3), 456-468. doi:10.1111/j.2044-8279.2011.02038.x

Hauck, M., Fein, D., Maltby, N., Waterhouse, L., & Feinstein, C. (1998). Memory for faces in children with autism. Child Neuropsychology, 4(3), 187-198. doi:10.1076/chin.4.3.187.3174

Henderson, H. A., Zahka, N. E., Kojkowski, N. M., Inge, A. P., Schwartz, C. B., Hileman, C. M., . . . Mundy, P. C. (2009). Self-referenced memory, social cognition, and symptom

presentation in autism. Journal of Child Psychology and Psychiatry, 50(7), 853-861. doi:10.1111/j.1469-7610.2008.02059.x

Hicks, J. L., & Marsh, R. L. (2002). On predicting the future states of awareness for recognition of unrecallable items. (vol 30, pg 60, 2002). Memory & Cognition, 30(3), 488-488. doi:10.3758/bf03194949

Hill, E. L. (2004). Evaluating the theory of executive dysfunction in autism. Developmental Review, 24(2), 189-233. doi:10.1016/j.dr.2004.01.001

Hill, E. L., Berthoz, S., & Frith, U. (2004). Brief report: Cognitive processing of own emotions in individuals with autistic spectrum disorder and in their relatives. Journal of autism and developmental disorders, 34(2), 229-235. doi:10.1023/B:Jadd.0000022613.41399.14

Hill, E. L., & Russell, J. (2002). Action memory and self-monitoring in children with autism: Self versus other. Infant and Child Development, 11(2), 159-170. doi:10.1002/Icd.303

Hobson, R. P. (1990). On the origins of self and the case of autism. Development and Psychopathology, 2(02), 163-181. doi:10.1017/S0954579400000687

Hobson, R. P., Lee, A., & Hobson, J. A. (2010). Personal Pronouns and Communicative Engagement in Autism. Journal of autism and developmental disorders, 40(6), 653-664. doi:10.1007/s10803-009-0910-5

Hobson, R. P., Ouston, J., & Lee, A. (1988). Emotion recognition in autism: Coordinating faces and voices. Psychological medicine, 18(04), 911-923. doi:10.1017/S0033291700009843

Hogrefe, G. J., Wimmer, H., & Perner, J. (1986). Ignorance versus false belief - a developmental lag in attribution of epistemic states. Child development, 57(3), 567-582. doi:10.2307/1130337

Hurlburt, R. T., Happé, F., & Frith, U. (1994). Sampling the form of inner experience in 3 adults with Asperger syndrome. Psychological medicine, 24(2), 385-395.

Iacoboni, M. (2005). Understanding others: Imitation, language, empathy. In S. Hurley & N. Chater (Eds.), Perspectives on imitation: From cognitive neuroscience to social science (Vol. 1, pp. 77-99). Cambridge: MIT Press.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. Science, 286(5449), 2526-2528. doi:10.1126/science.286.5449.2526

Jahn, P., & Engelkamp, J. (2003). Design-effects in prospective and retrospective memory for actions. Experimental Psychology, 50(1), 4-15. doi:10.1021//1618-3169.50.1.4

James, W. (1890). The principles of psychology. New York: Holt.

Janowsky, J. S., Shimamura, A. P., & Squire, L. R. (1989). Memory and metamemory - comparisons between patients with frontal-lobe lesions and amnesic patients. Psychobiology, 17(1), 3-11. doi:10.3758/BF03337811

Johnson, S., Filliter, J., & Murphy, R. (2009). Discrepancies Between Self- and Parent-Perceptions of Autistic Traits and Empathy in High Functioning Children and Adolescents on the Autism Spectrum. Journal of autism and developmental disorders, 39(12), 1706-1714. doi:10.1007/s10803-009-0809-1

Jolliffe, T., & Baron-Cohen, S. (1997). Are people with autism and Asperger syndrome faster than normal on the Embedded Figures Test? J Child Psychol Psychiatry, 38(5), 527-534. doi:10.1111/j.1469-7610.1997.tb01539.x

Jones, C. R. G., Happé, F., Golden, H., Marsden, A. J. S., Tregay, J., Simonoff, E., . . . Charman, T. (2009). Reading and Arithmetic in Adolescents With Autism Spectrum Disorders: Peaks and Dips in Attainment. Neuropsychology, 23(6), 718-728. doi:10.1037/a0016360

Jones, C. R. G., Swettenham, J., Charman, T., Marsden, A. J. S., Tregay, J., Baird, G., . . . Happé, F. (2011). No Evidence for a Fundamental Visual Motion Processing Deficit In Adolescents with Autism Spectrum Disorders. Autism Research, 4(5), 347-357. doi:10.1002/aur.209

Jordan, R. R. (1989). An experimental comparison of the understanding and use of speaker-addressee personal pronouns in autistic children. International Journal of Language & Communication Disorders, 24(2), 169-179. doi:10.3109/13682828909011954

Kadesjo, B., Gillberg, C., & Hagberg, B. (1999). Brief report: autism and Asperger syndrome in seven-year-old children: a total population study. J Autism Dev Disord, 29(4), 327-331. doi:10.1023/A:1022115520317

Kagan, J. (1982). The emergence of self. Journal of Child Psychology and Psychiatry and Allied Disciplines, 23(4), 363-381. doi:10.1111/j.1469-7610.1982.tb00084.x

Kanner, L. (1943). Autistic disturbances of affective contact. Nervous Child(2), 217-250.

Kao, Y. C., Davis, E. S., & Gabrieli, J. D. (2005). Neural correlates of actual and predicted memory formation. Nat Neurosci, 8(12), 1776-1783. doi:10.1038/nn1595

Kazak, S., Collis, G. M., & Lewis, V. (1997). Can young people with autism refer to knowledge states? Evidence from their understanding of "know" and "guess". Journal of Child Psychology and Psychiatry and Allied Disciplines, 38(8), 1001-1009. doi:10.1111/j.1469-7610.1997.tb01617.x

Kelemen, W. L., Frost, P. J., & Weaver, C. A. (2000). Individual differences in metacognition: Evidence against a general metacognitive ability. Memory & Cognition, 28(1), 92-107. doi:10.3758/BF03211579

Kennedy, D. P., & Courchesne, E. (2008). Functional abnormalities of the default network during self- and other-reflection in autism. Social Cognitive and Affective Neuroscience, 3(2), 177-190. doi:10.1093/scan/nsn011

Klein, A. M., Zwickel, J., Prinz, W., & Frith, U. (2009). Animated triangles: An eye tracking investigation. Quarterly Journal of Experimental Psychology, 62(6), 1189-1197. doi:10.1080/17470210802384214

Klein, S. B., Chan, R. L., & Loftus, J. (1999). Independence of episodic and semantic self-knowledge: The case from autism. Social Cognition, 17(4), 413-436. doi:10.1521/soco.1999.17.4.413

Koriat, A. (1993). How do we know what we know - The accessibility model of feeling-of-knowing. Psychological Review, 100(4), 609-639. doi:10.1037/0033-295X.100.4.609

Koriat, A., Ben-Zur, H., & Nussbaum, A. (1990). Encoding information for future action: memory for to-be-performed tasks versus memory for to-be-recalled tasks. Mem Cognit, 18(6), 568-578. doi:10.3758/BF03197099

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. Psychological Review, 103(3), 490-517. doi:10.1037//0033-295X.103.3.490

Krebs, S. S., & Roebers, C. M. (2010). Children's strategic regulation, metacognitive monitoring, and control processes during test taking. British Journal of Educational Psychology, 80(3), 325-340. doi:10.1348/000709910X485719

Kristen, S., Rossmann, F., & Sodian, B. (2014). Theory of own mind and autobiographical memory in adults with ASD. Research in Autism Spectrum Disorders, 8(7), 827-837. doi:10.1016/j.rasd.2014.03.009

Kucera, H., & Francis, W. N. (1967). Computational analysis of present day American English. Providence, RI: Brown University Press.

Kuhn, M. H., & McPartland, T. S. (1954). An empirical investigation of self-attitudes. American sociological review, 68-76. doi:10.2307/2088175

Lee, A., & Hobson, R. P. (1998). On Developing Self-concepts: A Controlled Study of Children and Adolescents with Autism. Journal of Child Psychology and Psychiatry, 39(8), 1131-1144. doi:10.1111/1469-7610.00417

Lee, A., Hobson, R. P., & Chiat, S. (1994). I, you, me, and autism - an experimental-study. Journal of autism and developmental disorders, 24(2), 155-176. doi:10.1007/bf02172094

Leonesio, R. J., & Nelson, T. O. (1990). Do different metamemory judgments tap the same underlying aspects of memory. Journal of Experimental Psychology-Learning Memory and Cognition, 16(3), 464-470. doi:10.1037//0278-7393.16.3.464

Lerner, M. D., Calhoun, C. D., Mikami, A. Y., & De Los Reyes, A. (2012). Understanding Parent-Child Social Informant Discrepancy in Youth with High Functioning Autism Spectrum Disorders. Journal of autism and developmental disorders, 42(12), 2680-2692. doi:10.1007/s10803-012-1525-9

Leslie, A. M. (1994). ToMM, ToBy and agency: Core architecture and domain specificity. In L. Hirschfeld & S. Gelman (Eds.), Mapping the Mind: Domain Specificity in Cognition and Culture. Cambridge: Cambridge University Press.

Leslie, A. M., & Thaiss, L. (1992). Domain specificity in conceptual development: Neuropsychological evidence from autism. Cognition, 43(3), 225-251. doi:10.1016/0010-0277(92)90013-8

Lewis, M. (2003). The Development of Self-Consciousness. In J. Roessler & N. Eilan (Eds.), Agency and self-awareness. Oxford: Oxford University Press.

Lewis, M., & Ramsay, D. (2004). Development of Self-Recognition, Personal Pronoun Use, and Pretend Play During the 2nd Year. Child development, 75(6), 1821-1831. doi:10.1111/j.1467-8624.2004.00819.x

Lind, S. E. (2010). Memory and the self in autism: A review and theoretical framework. Autism, 14(5), 430-456. doi:10.1177/1362361309358700

Lind, S. E., & Bowler, D. M. (2009a). Delayed self-recognition in children with autism spectrum disorder. Journal of autism and developmental disorders, 39(4), 643-650. doi:10.1007/s10803-008-0670-7

Lind, S. E., & Bowler, D. M. (2009b). Language and Theory of Mind in Autism Spectrum Disorder: The Relationship Between Complement Syntax and False Belief Task Performance. Journal of autism and developmental disorders, 39(6), 929-937. doi:10.1007/s10803-009-0702-y

Lind, S. E., & Bowler, D. M. (2009c). Recognition memory, self-other source memory, and theory-of-mind in children with autism spectrum disorder. Journal of autism and developmental disorders, 39(9), 1231-1239. doi:10.1007/s10803-009-0735-2

Lind, S. E., & Bowler, D. M. (2010). Episodic memory and episodic future thinking in adults with autism. J Abnorm Psychol, 119(4), 896-905. doi:10.1037/a0020631

Lind, S. E., Bowler, D. M., & Raber, J. (2014). Spatial navigation, episodic memory, episodic future thinking, and theory of mind in children with autism spectrum disorder: Evidence for impairments in mental simulation? Frontiers in Psychology, 5. doi:10.3389/fpsyg.2014.01411

Lind, S. E., Williams, D. M., Bowler, D. M., & Peel, A. (2014). Episodic Memory and Episodic Future Thinking Impairments in High-Functioning Autism Spectrum Disorder: An Underlying Difficulty With Scene Construction or Self-Projection. Neuropsychology, 28(1), 55-67. doi:10.1037/neu0000005

Lind, S. E., Williams, D. M., Raber, J., Peel, A., & Bowler, D. M. (2013). Spatial navigation impairments among intellectually high-functioning adults with autism spectrum

disorder: exploring relations with theory of mind, episodic memory, and episodic future thinking. J Abnorm Psychol, 122(4), 1189-1199. doi:10.1037/a0034819

Linden, W., Wen, F., & Paulhus, D. L. (1995). Measuring alexithymia: reliability, validity, and prevalence. In J. Butcher & C. Spielberg (Eds.), Advances in personality assesment (Vol. 10, pp. 51-95). New York: Psychology Press.

Lombardo, M. V., Barnes, J. L., Wheelwright, S. J., & Baron-Cohen, S. (2007). Self-Referential Cognition and Empathy in Autism. Plos One, 2(9), e883. doi:10.1371/journal.pone.0000883

Lopez, B. R., Lincoln, A. J., Ozonoff, S., & Lai, Z. (2005). Examining the relationship between executive functions and restricted, repetitive symptoms of Autistic Disorder. J Autism Dev Disord, 35(4), 445-460. doi:10.1007/s10803-005-5035-x

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., . . . Rutter, M. (2000). The Autism Diagnostic Observation Schedule-Generic: A standard measure of social and communication deficits associated with the spectrum of autism. Journal of autism and developmental disorders, 30(3), 205-223. doi:10.1023/A:1005592401947

Losh, M., & Capps, L. (2003). Narrative ability in high-functioning children with autism or Asperger's syndrome. J Autism Dev Disord, 33(3), 239-251. doi:10.1023/A:1024446215446

Loveland, K. A., & Landry, S. H. (1986). Joint attention and language in autism and developmental language delay. Journal of Autism and Developmental Disorders, 16(3), 335-349.

Maras, K., Memon, A., Lambrechts, A., & Bowler, D. (2012). Recall of a live and personally experienced eyewitness event by adults with autism spectrum disorder. Journal of autism and developmental disorders, 43(8), 1-13. doi:10.1007/s10803-012-1729-z

Marsh, R. L., Hicks, J. L., & Bink, M. L. (1998). Activation of completed, uncompleted, and partially completed intentions. Journal of Experimental Psychology-Learning Memory and Cognition, 24(2), 350-361. doi:10.1037//0278-7393.24.2.350

Maylor, E. A., Chater, N., & Brown, G. D. A. (2001). Scale invariance in the retrieval of retrospective and prospective memories. Psychonomic Bulletin & Review, 8(1), 162-167. doi:10.3758/bf03196153

Mazzoni, G., & Nelson, T. O. (1995). Judgments of learning are affected by the kind of encoding in ways that cannot be attributed to the level of recall. Journal of Experimental Psychology-Learning Memory and Cognition, 21(5), 1263-1274. doi:10.1037//0278-7393.21.5.1263

McDaniel, M. A., & Einstein, G. O. (2007). Prospective memory: An overview and synthesis of an emerging field: Sage Publications.

Meins, E., & Fernyhough, C. (2010). Mind-mindedness coding manual, Version 2.0.

Mervis, C. B., & Klein-Tasman, B. P. (2004). Methodological issues in group-matching designs: alpha levels for control variable comparisons and measurement characteristics of control and target variables. Journal of autism and developmental disorders, 34(1), 7-17. doi:10.1023/B:JADD.0000018069.69562.b8

Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The Cue-Familiarity Heuristic In Metacognition. Journal of Experimental Psychology-Learning Memory and Cognition, 19(4), 851-861. doi:10.1037//0278-7393.19.4.851

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. Journal of Abnormal Psychology, 110(1), 40-48. doi:10.1037//0021-843X.110.1.40

Millward, C., Powell, S., Messer, D., & Jordan, R. (2000). Recall for self and other in autism: Children's memory for events experienced by themselves and their peers. Journal of autism and developmental disorders, 30(1), 15-28. doi:10.1023/A:1005455926727

Modinos, G., Ormel, J., & Aleman, A. (2009). Activation of anterior insula during self-reflection. Plos One, 4(2), 26. doi:10.1371/journal.pone.0004618

Neisser, U. (1988). Five kinds of self-knowledge. Philosophical Psychology, 1(1), 35-59. doi:10.1080/09515088808572924

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. Psychological Bulletin, 95(1), 109-133. doi:10.1037//0033-2909.95.1.109

Nelson, T. O., & Dunlosky, J. (1991). When Peoples Judgments Of Learning (JOLs) Are Extremely Accurate At Predicting Subsequent Recall - The Delayed-Jol Effect. Psychological Science, 2(4), 267-270. doi:10.1111/j.1467-9280.1991.tb00147.x

Nelson, T. O., & Narens, L. (1990). Metamemory A Theoretical Framework And New Findings. The Psychology of Learning and Motivation, 26, 125-174. doi:10.1016/S0079-7421(08)60053-5

Nelson, T. O., Narens, L., & Dunlosky, J. (2004). A revised methodology for research on metamemory: Pre-judgment recall and monitoring (PRAM). Psychological Methods, 9(1), 53-69. doi:10.1037/1082-989X.9.1.53

Neuman, C. J., & Hill, S. D. (1978). Self-recognition and stimulus preference in autistic children. Dev Psychobiol, 11(6), 571-578. doi:10.1002/dev.420110606

Nichols, S., & Stich, S. (2003). Mindreading: An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds. Oxford: Oxford University Press.

Nielsen, M., Suddendorf, T., & Slaughter, V. (2006). Mirror self-recognition beyond the face. Child Dev, 77(1), 176-185. doi:10.1111/j.1467-8624.2006.00863.x

Nishitani, N., Avikainen, S., & Hari, R. (2004). Abnormal imitation-related cortical activation sequences in Asperger's syndrome. Annals of neurology, 55(4), 558-562. doi:10.1002/ana.20031

Oberman, L. M., Hubbard, E. M., McCleery, J. P., Altschuler, E. L., Ramachandran, V. S., & Pineda, J. A. (2005). EEG evidence for mirror neuron dysfunction in autism spectrum disorders. Cognitive Brain Research, 24(2), 190-198. doi:10.1016/j.cogbrainres.2005.01.014

Owens, J. S., Goldfine, M. E., Evangelista, N. M., Hoza, B., & Kaiser, N. M. (2007). A critical review of self-perceptions and the positive illusory bias in children with ADHD. Clinical Child and Family Psychology Review, 10(4), 335-351. doi:10.1007/s10567-007-0027-3

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(7), 1081-1105. doi:10.1111/j.1469-7610.1991.tb00351.x

Paynter, C. A., Reder, L. M., & Kieffaber, P. D. (2009). Knowing we know before we know: ERP correlates of initial feeling-of-knowing. Neuropsychologia, 47(3), 796-803. doi:10.1016/j.neuropsychologia.2008.12.009

Pellicano, E. (2007). Links between theory of mind and executive function in young children with autism: clues to developmental primacy. Dev Psychol, 43(4), 974-990. doi:10.1037/0012-1649.43.4.974

Perner, J., Aichhorn, M., Kronbichler, M., Staffen, W., & Ladurner, G. (2006). Thinking of mental and other representations: The roles of left and right temporo-parietal junction. Social Neuroscience, 1(3-4), 245-258. doi:10.1080/17470910600989896

Perner, J., Frith, U., Leslie, A. M., & Leekam, S. R. (1989). Exploration of the autistic child's theory of mind: knowledge, belief, and communication. Child development, 60(3), 688-700. doi:10.2307/1130734

Phillips, W., Baron-Cohen, S., & Rutter, M. (1998). Understanding intention in normal development and in autism. British Journal of Developmental Psychology, 16(3), 337-348. doi:10.1111/j.2044-835X.1998.tb00756.x

Povinelli, D. J. (2001). The self: elevated in consciousness and extended in time. In C. Moore & K. Lemmon (Eds.), The self in time: developmental perspectives. Mahwah, NJ: Lawrence Erlbaum Associates.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? Behavioral and Brain Sciences, 1(04), 515-526. doi:10.1017/S0140525X00076512

Ramachandran, V. S., & Oberman, L. M. (2006). Broken mirrors: a theory of autism. Scientific American, 295(5), 62-69. doi:10.1038/scientificamerican1106-62

Reder, L. M. (1987). Strategy selection in question answering. Cognitive Psychology, 19(1), 90-138. doi:10.1016/0010-0285(87)90005-3

Reder, L. M., Nhouyvanisvong, A., Schunn, C. D., Ayers, M. S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of remember-know judgments in a continuous recognition paradigm. Journal of Experimental Psychology-Learning Memory and Cognition, 26(2), 294-320. doi:10.1037/0278-7393.26.2.294

Rhee, E., Uleman, J. S., Roman, R. J., & Lee, H. K. (1995). Spontaneous self-descriptions and ethnic identities in individualistic and collectivistic cultures. Journal of Personality and Social Psychology, 69(1), 142-152. doi:10.1037/0022-3514.69.1.142

Rizzolatti, G., & Fabbri-Destro, M. (2010). Mirror neurons: from discovery to autism. Experimental Brain Research, 200(3-4), 223-237. doi:10.1007/s00221-009-2002-3

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. Cognitive Brain Research, 3(2), 131-141. doi:10.1016/0926-6410(95)00038-0

Rochat, P. (2003). Five levels of self-awareness as they unfold early in life. Consciousness and Cognition, 12(4), 717-731. doi:10.1016/s1053-8100(03)00081-3

Roebers, C. M., Schmid, C., & Roderer, T. (2009). Metacognitive monitoring and control processes involved in primary school children's test performance. British Journal of Educational Psychology, 79(Pt 4), 749-767. doi:10.1348/978185409X429842

Rouse, H., Donnelly, N., Hadwin, J. A., & Brown, T. (2004). Do children with autism perceive second-order relational features? The case of the Thatcher illusion. Journal of Child Psychology and Psychiatry, 45(7), 1246-1257. doi:10.1111/j.1469-7610.2004.00317.x

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. Trends in cognitive sciences, 11(6), 251-257. doi:10.1016/j.tics.2007.04.004

Russell, J. (1996). Agency: Its role in mental development. Hove, UK: Taylor & Francis.

Russell, J., & Hill, E. L. (2001). Action-monitoring and intention reporting in children with autism. Journal of Child Psychology and Psychiatry and Allied Disciplines, 42(3), 317-328. doi:10.1111/1469-7610.00725

Russell, J., Hill, E. L., & Franco, F. (2001). The role of belief veracity in understanding intentions-in-action: Preschool children's performance on the transparent intentions task. Cognitive Development, 16(3), 775-792. doi:10.1016/S0885-2014(01)00057-0

Russell, J., & Jarrold, C. (1998). Error-correction problems in autism: Evidence for a monitoring impairment? Journal of autism and developmental disorders, 28(3), 177-188. doi:10.1023/A:1026009203333

Russell, J., & Jarrold, C. (1999). Memory for actions in children with autism: Self versus other. Cognitive Neuropsychiatry, 4(4), 303-331. doi:10.1080/135468099395855

Sacher, M., Taconnat, L., Souchay, C., & Isingrini, M. (2009). Divided attention at encoding: effect on feeling-of-knowing. Conscious Cogn, 18(3), 754-761. doi:10.1016/j.concog.2009.04.001

Salter, G., Seigal, A., Claxton, M., Lawrence, K., & Skuse, D. (2008). Can autistic children read the mind of an animated triangle? Autism, 12(4), 349-371. doi:10.1177/1362361308091654

Samson, D., Apperly, A., Chiavarino, C., & Humphreys, G. (2004). Left temporoparietal junction is necessary for representing someone else's belief. Nat Neurosci, 7(5), 499-500. doi:10.1038/Nn1223

Santiesteban, I., Banissy, Michael J., Catmur, C., & Bird, G. (2012). Enhancing Social Ability by Stimulating Right Temporoparietal Junction. Current Biology, 22(23), 2274-2277. doi:10.1016/j.cub.2012.10.018

Sawyer, A. C., Williamson, P., & Young, R. (2014). Metacognitive processes in emotion recognition: are they different in adults with Asperger's disorder? J Autism Dev Disord, 44(6), 1373-1382. doi:10.1007/s10803-013-1999-0

Saxe, R., & Powell, L. J. (2006). It's the thought that counts: Specific brain regions for one component of theory of mind. Psychological Science, 17(8), 692-699. doi:10.1111/j.1467-9280.2006.01768.x

Schneider, W., Visé, M., Lockl, K., & Nelson, T. O. (2000). Developmental trends in children's memory monitoring: Evidence from a judgment-of-learning task. Cognitive Development, 15(2), 115-134. doi:10.1016/S0885-2014(00)00024-1

Schnyer, D. M., Nicholls, L., & Verfaellie, M. (2005). The role of VMPC in metamemorial judgments of content retrievability. Journal of cognitive neuroscience, 17(5), 832-846. doi:10.1162/0898929053747694

Schnyer, D. M., Verfaellie, M., Alexander, M. P., LaFleche, G., Nicholls, L., & Kaszniak, A. W. (2004). A role for right medial prefrontal cortex in accurate feeling-of-knowing judgments: evidence from patients with lesions to frontal cortex. Neuropsychologia, 42(7), 957-966. doi:10.1016/j.neuropsychologia.2003.11.020

Schwartz, B. L. (1994). Sources of information in metamemory: Judgments of learning and feeling of knowing. Psychonomic Bulletin & Review, 1(3), 357-375. doi:10.3758/BF03213977

Sebanz, N., Knoblich, G., Stumpf, L., & Prinz, W. (2005). Far from action-blind: Representation of others' actions in individuals with autism. Cognitive Neuropsychology, 22(3-4), 433-454. doi:10.1080/02643290442000121

Selman, R. (1980). The growth of interpersonal understanding. New York: Academic Press.

Senju, A. (2012). Spontaneous theory of mind and its absence in autism spectrum disorders. Neuroscientist, 18(2), 108-113. doi:10.1177/1073858410397208

Senju, A., Southgate, V., White, S., & Frith, U. (2009). Mindblind Eyes: An Absence of Spontaneous Theory of Mind in Asperger Syndrome. Science, 325(5942), 883-885. doi:10.1126/science.1176170

Shah, A., & Frith, U. (1993). Why do autistic individuals show superior performance on the block design task? J Child Psychol Psychiatry, 34(8), 1351-1364. doi:10.1111/j.1469-7610.1993.tb02095.x

Shimamura, A. P., & Squire, L. R. (1986). Memory and metamemory - a study of the feeling-of-knowing phenomenon in amnesic patients. Journal of Experimental Psychology-Learning Memory and Cognition, 12(3), 452-460. doi:10.1037//0278-7393.12.3.452

Silani, G., Bird, G., Brindley, R., Singer, T., Frith, C., & Frith, U. (2008). Levels of emotional awareness and autism: An fMRI study. Social Neuroscience, 3(2), 97-112. doi:10.1080/17470910701577020

Sitko, B. M. (1998). Knowing how to write: Metacognition and writing instruction. In D. J. Hacker, J. Dunlosky, & A. Graesser (Eds.), Metacognitionin educational theory and practice (pp. 93-115). London: Lawrence Erlbaum Associates.

Skavhaug, I.-M., Wilding, E. L., & Donaldson, D. I. (2010). Judgments of learning do not reduce to memory encoding operations: Event-related potential evidence for distinct metacognitive processes. Brain research, 1318, 87-95. doi:10.1016/j.brainres.2009.11.047

Skavhaug, I. M., Wilding, E. L., & Donaldson, D. I. (2013). Immediate Judgments of Learning Predict Subsequent Recollection: Evidence From Event-Related Potentials. Journal of Experimental Psychology-Learning Memory and Cognition, 39(1), 159-166. doi:10.1037/a0028885

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory - applications to dementia and amnesia. Journal of Experimental Psychology-General, 117(1), 34-50. doi:10.1037//0096-3445.117.1.34

Sommer, W., Heinz, A., Leuthold, H., Matt, J., & Schweinberger, S. R. (1995). Metamemory, distinctiveness, and event-related potentials in recognition memory for faces. Memory & Cognition, 23(1), 1-11. doi:10.3758/BF03210552

Song, C., Kanai, R., Fleming, S. M., Weil, R. S., Schwarzkopf, D. S., & Rees, G. (2011). Relating inter-individual differences in metacognitive performance on different perceptual tasks. Consciousness and Cognition, 20(4), 1787-1792. doi:10.1016/j.concog.2010.12.011

Souchay, C., & Isingrini, M. (2012). Are feeling-of-knowing and judgment-of-learning different? Evidence from older adults. Acta Psychologica, 139(3), 458-464.

Souchay, C., Isingrini, M., Clarys, D., Taconnat, L., & Eustache, F. (2004). Executive functioning and judgment-of-learning versus feeling-of-knowing in older adults. Experimental Aging Research, 30(1), 47-62. doi:10.1080/03610730490251478

Souchay, C., Moulin, C. J., Clarys, D., Taconnat, L., & Isingrini, M. (2007). Diminished episodic memory awareness in older adults: evidence from feeling-of-knowing and recollection. Conscious Cogn, 16(4), 769-784. doi:10.1016/j.concog.2006.11.002

Souchay, C., Wojcik, D. Z., Williams, H. L., Crathern, S., & Clarke, P. (2013). Recollection in adolescents with Autism Spectrum Disorder. Cortex, 49(6), 1598-1609. doi:10.1016/j.cortex.2012.07.011

Southgate, V., & Hamilton, A. F. d. C. (2008). Unbroken mirrors: Challenging a theory of autism. Trends in cognitive sciences, 12(6), 225-229. doi:10.1016/j.tics.2008.03.005

Spiker, D., & Ricks, M. (1984). Visual self-recognition in autistic children: developmental relationships. Child Dev, 55(1), 214-225. doi:10.2307/1129846

Stokes, M., Thompson, R., Cusack, R., & Duncan, J. (2009). Top-down activation of shape-specific population codes in visual cortex during mental imagery. The Journal of Neuroscience, 29(5), 1565-1572. doi:10.1523/JNEUROSCI.4657-08.2009

Stuss, D. T., & Knight, R. T. (Eds.). (2002). Principles of frontal lobe function. New York: Oxford Univeristy Press.

Suddendorf, T. (1999). Children's Understanding of the Relation between Delayed Video Representation and Current Reality: A Test for Self-Awareness? Journal of Experimental Child Psychology, 72(3), 157-176. doi:10.1006/jecp.1998.2485

Suddendorf, T. (2010). Episodic memory versus episodic foresight: Similarities and differences. Wiley Interdisciplinary Reviews: Cognitive Science, 1(1), 99-107. doi:10.1002/wcs.23

Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? Behav Brain Sci, 30(3), 299-313; discussion 313-251. doi:10.1017/s0140525x07001975

Summers, J. A., & Craik, F. I. M. (1994). The effects of subject-performed tasks on the memory performance of verbal autistic-children. Journal of autism and developmental disorders, 24(6), 773-783. doi:10.1007/BF02172285

Surian, L., & Leslie, A. M. (1999). Competence and performance in false belief understanding: A comparison of autistic and normal 3-year-old children. British Journal of Developmental Psychology, 17, 141-155. doi:10.1348/026151099165203

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta-analysis. Psychological Bulletin, 121(3), 371-394. doi:10.1037//0033-2909.121.3.371

Tanweer, T., Rathbone, C. J., & Souchay, C. (2010). Autobiographical memory, autonoetic consciousness, and identity in Asperger syndrome. Neuropsychologia, 48(4), 900-908. doi:10.1016/j.neuropsychologia.2009.11.007

Terrett, G., Rendell, P. G., Raponi-Saunders, S., Henry, J. D., Bailey, P. E., & Altgassen, M. (2013). Episodic future thinking in children with autism spectrum disorder. J Autism Dev Disord, 43(11), 2558-2568. doi:10.1007/s10803-013-1806-y

Théoret, H., Halligan, E., Kobayashi, M., Fregni, F., Tager-Flusberg, H., & Pascual-Leone, A. (2005). Impaired motor facilitation during action observation in individuals with autism spectrum disorder. Current Biology, 15(3), R84-R85. doi:10.1016/j.cub.2005.01.022

Thiede, K. W., Anderson, M. C. M., & Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. Journal of Educational Psychology, 95(1), 66-73. doi:10.1037/0022-0663.95.1.66

Toichi, M., Kamio, Y., Okada, T., Sakihama, M., Youngstrom, E. A., Findling, R. L., & Yamamoto, K. (2002). A lack of self-consciousness in autism. American Journal of Psychiatry, 159(8), 1422-1424. doi:10.1176/appi.ajp.159.8.1422

Wechsler, D. (1999). Wechsler abbreviated scale of intelligence. New York, NY: The Psychological Corporation: Harcourt Brace & Company.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. Child development, 72(3), 655-684. doi:10.1111/1467-8624.00304

White, S., Hill, E., Happé, F., & Frith, U. (2009). Revisiting the Strange Stories: Revealing Mentalizing Impairments in Autism. Child development, 80(4), 1097-1117. doi:10.1111/j.1467-8624.2009.01319.x

White, S. J., Coniston, D., Rogers, R., & Frith, U. (2011). Developing the Frith-Happé Animations: A Quick and Objective Test of Theory of Mind for Adults with Autism. Autism Research, 4(2), 149-154. doi:10.1002/aur.174

Wilkinson, D. A., Best, C. A., Minshew, N. J., & Strauss, M. S. (2010). Memory Awareness for Faces in Individuals with Autism. Journal of autism and developmental disorders, 40(11), 1371-1377. doi:10.1007/s10803-010-0995-x

Williams , D. L., Goldstein, G., & Minshew, N. J. (2005). Impaired memory for faces and social scenes in autism: clinical implications of memory dysfunction. Archives of Clinical Neuropsychology, 20(1), 1-15. doi:10.1016/j.acn.2002.08.001

Williams, D. M. (2010). Theory of own mind in autism. Autism, 14(5), 474-494. doi:10.1177/1362361310366314

Williams, D. M., Boucher, J., Lind, S., & Jarrold, C. (2013). Time-Based and Event-Based Prospective Memory in Autism Spectrum Disorder: The Roles of Executive Function

and Theory of Mind, and Time-Estimation. Journal of autism and developmental
disorders, 43(7), 1555-1567. doi:10.1007/s10803-012-1703-9

Williams, D. M., & Bowler, D. M. (2014). Autism spectrum disorder: Fractionable or
coherent? Autism, 18(1), 2-5. doi:10.1177/1362361313513523

Williams, D. M., & Happé, F. (2009a). Pre-conceptual aspects of self-awareness in autism
spectrum disorder: The case of action-monitoring. Journal of autism and
developmental disorders, 39(2), 251-259. doi:10.1007/s10803-008-0619-x

Williams, D. M., & Happé, F. (2009b). What did I say? versus What did I think? Attributing
false beliefs to self amongst children with and without autism. Journal of autism and
developmental disorders, 39(6), 865-873. doi:10.1007/s10803-009-0695-6

Williams, D. M., & Happé, F. (2010). Representing intentions in self and other: studies of
autism and typical development. Developmental Science, 13(2), 307-319.
doi:10.1111/j.1467-7687.2009.00885.x

Williams, D. M., Jarrold, C., Grainger, C., & Lind, S. E. (2014). Diminished Time-Based, but
Undiminished Event-Based, Prospective Memory Among Intellectually High-
Functioning Adults With Autism Spectrum Disorder: Relation to Working Memory
Ability. Neuropsychology, 28(1), 30-42. doi:10.1037/neu0000008

Williams, J., Waiter, G. D., Gilchrist, A., Perrett, D. I., Murray, A. D., & Whiten, A. (2006).
Neural mechanisms of imitation and 'mirror neuron' functioning in autistic spectrum
disorder. Neuropsychologia, 44(4), 610-621.
doi:10.1016/j.neuropsychologia.2005.06.010

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

Witkin, H. A., Oltman, P. K., Raskin, E., & Karp, S. (1971). A manual for the embedded figures test. Palo Alto, CA: Consulting Psychologists Press.

Wojcik, D. Z., Allen, R. J., Brown, C., & Souchay, C. (2011). Memory for actions in autism spectrum disorder. Memory, 19(6), 549-558. doi:10.1080/09658211.2011.590506

Wojcik, D. Z., Moulin, C. J., & Souchay, C. (2013). Metamemory in children with autism: exploring "feeling-of-knowing" in episodic and semantic memory. Neuropsychology, 27(1), 19-27. doi:10.1037/a0030526

Wojcik, D. Z., Waterman, A. H., Lestie, C., Moulin, C. J., & Souchay, C. (2014). Metacognitive judgments-of-learning in adolescents with autism spectrum disorder. Autism, 18(4), 393-408. doi:10.1177/1362361313479453

Woodbury-Smith, M., Robinson, J., & Baron-Cohen, S. (2005). Screening adults for Asperger Syndrome using the AQ : diagnostic validity in clinical practice. Journal of autism and developmental disorders, 35(3), 331-335. doi:10.1007/s10803-005-3300-7

World Heath Organisation. (1993). International classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines (10th edn.) Geneva, Switzerland: World Heath Organisation.

Yirmiya, N., Erel, O., Shaked, M., & Solomonica-Levi, D. (1998). Meta-analyses comparing theory of mind abilities of individuals with autism, individuals with mental retardation, and normally developing individuals. Psychological Bulletin, 124(3), 283-307. doi:10.1037//0033-2909.124.3.283

Yokoyama, O., Miura, N., Watanabe, J., Takemoto, A., Uchida, S., Sugiura, M., . . . Nakamura, K. (2010). Right frontopolar cortex activity correlates with reliability of retrospective rating of confidence in short-term recognition memory performance. Neuroscience research, 68(3), 199-206. doi:10.1016/j.neures.2010.07.2041

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. Journal of memory and language, 46(3), 441-517. doi:10.1006/jmla.2002.2864

Zahavi, D. (2010). Complexities of self. Autism, 14(5), 547-551. doi:10.1177/1362361310370040

Zalla, T., Daprati, E., Sav, A.-M., Chaste, P., Nico, D., & Leboyer, M. (2010). Memory for self-performed actions in individuals with Asperger syndrome. Plos One, 5(10), 1-6. doi:10.1371/journal.pone.0013370

Zwickel, J., White, S. J., Coniston, D., Senju, A., & Frith, U. (2011). Exploring the building blocks of social cognition: spontaneous agency perception and visual perspective taking in autism. Social Cognitive and Affective Neuroscience, 6(5), 564-571. doi:10.1093/scan/nsq088