Kallitsounaki, Aimilia (2018) *Lower and High Order Phenomena of Self-Awareness in Autism Spectrum Disorder.* Master of Science by Research (MScRes) thesis, University of Kent,.

Lower and High Order Phenomena of Self-Awareness in Autism Spectrum Disorder

Aimilia Kallitsounaki

Supervisor: Professor David Williams

Word Count: 20,586/ Number of pages:113

A research paper submitted in fulfilment of the degree of M.Sc. by Research in Psychology, in the Department of Psychology, University of Kent, (11/2018).

Abstract

The current project aimed to explore lower and higher order phenomena of self-awareness in Autism Spectrum Disorder (ASD). According to some theoretical approaches (e.g. the alexithymia hypothesis), ASD is characterised by impaired interoceptive accuracy. That is the ability to form first order representations of the physiological conditions of the body and thus to detect body signals. It has also been claimed that both mindreading (ability to attribute mental states to others) and metacognitive monitoring (ability to attribute mental states to one's own self) are diminished in ASD because a deficit in the ability to form metarepresentations lies at the heart of this disorder. Nonetheless, the study of lower and higher order phenomena of self-awareness in ASD is limited and the existing findings conflicting. As such, we conducted five experiments aiming to investigate interoception and metacognitive monitoring in ASD. In Experiment 1.1, we found strong evidence that interoceptive accuracy is undiminished in ASD, exploring both cardiac and respiratory domains and moderate evidence for diminished interoceptive awareness. Despite finding in Experiment 2.1 that autism traits were unrelated to peoples' ability to impute mental states to themselves, in Experiment 2.2, we found that implicit, but not explicit metacognitive monitoring abilities were diminished among individuals with ASD. In Experiment 3.1, we also found that autism traits were unrelated to metacognitive monitoring in the general population, as indexed by accuracy of metacognitive experiences of difficulty. In line with that, we did not find evidence that this ability was diminished among people with a full diagnosis of ASD (Experiment 3.2). In sum, the current project provides strong evidence for intact lower order phenomena of self-awareness in ASD and tentative evidence for impaired higher order phenomena. Replication awaits before strong conclusions can be drawn.

Lower and High Order Phenomena of Self-Awareness in Autism Spectrum Disorder

"Self" describes "the particular being any person is, whatever it is about each of us that distinguishes you or me from others, draws the parts of our existence together, persists through changes, or opens the way to becoming who we might or should be" (Seigel, 2007, p.3). This definition leads us to think that "self" is a non-unitary construct.

Indeed, the multidimensional nature of "self" has been proposed since the period of Renaissance, and to date it remains a generally acknowledged notion (Lewis, 1995; Lind, 2010; Neisser, 1988; Rochat, 2003; Seigel, 2007; Williams, 2010). The philosophical movement towards accepting the existence of many "selves", rather than a single one led to the diffusion of a plethora of taxonomies that made the systematic exploration of "self" a feasible accomplishment (Neisser, 1988). Two salient and particularly important distinctions are those between the "I" and "me" self (James, 1890) and between the "physical" and "psychological" self (Gillihan & Farah, 2005).

"I" self-awareness occurs when an individual experiences themselves as the subject of a particular experience (Butterworth, 1995; James 1890). Whereas, "me" self-awareness, the "mysterious operation", as it has been characterised by James (1890) is ascribed to a person that recognises themselves as the object of their own cognition/perception (Butterworth, 1995; James, 1890; Lind, 2010). "Physical" self-awareness entails a sense of body ownership that extends from its outward phenotype to its internal functions (Gillihan & Farah, 2005). Whereas, "psychological" self-awareness involves the sense of an inner psychological world that comprises traits, emotions, memories, and so on (Gillihan & Farah, 2005). Two different mechanisms underlie these dimensions of self-awareness. That is the capacity to form first order representations and the ability to form metarepresentations (or else second order representations).

On one hand, first order representations are internal mental representations of the world, including the "self", which have been constructed in absolute consistence with the external reality (Dretske, 1995; Leslie, 1987). It is the ability to form fist order representations that enables individuals of get an awareness of their very basic dimensions and functions of self, namely lower order phenomena of self-awareness. Interoception has been defined as one of the lower order phenomena of self-awareness (Craig, 2009; Damasio, 2000; Lind & Bowler, 2008; Seth, 2013). On the other hand, metarepresentations are the representations of the first order representations (Carruthers, 2009; Leslie 1987; Pylyshyn, 1978) and as such enable individuals to get an awareness of more complex and delicate dimensions of self, namely higher order phenomena (Asendorpf, Warkentin, & Baudonnière, 1996; Carruthers, 2008; Lind & Bowler, 2008; Perner, 1991; Suddendorf & Whiten, 2001). Metarepresentational abilities have been considered the epitome of metacognition, and thus this concept has been used as index of higher order phenomena of self-awareness (Carruthers, 2008; 2009).

Interoception

Interoception was originally described as any reaction of sensory nerve receptors triggered by visceral stimuli (Fowler, 2003; Sherrington, 1906), but now this narrow definition has been broadened into any “sense of the physiological conditions of the body” (Craig, 2003, p. 500). A key element of interoception is that these signals can become conscious and thus facilitate the regulation of peoples’ behaviour accordingly (Cameron, 2002). Whilst, the contemporary definition refers to a broad range of signals, research has predominantly focused on signals from the heart (Schandry, 1981).

The heartbeat tracking task is one of the most classic and widely used paradigms employed to measure cardiac interoceptive accuracy (Critchley, Wiens, Rothstein, Öhlman, & Dolan, 2004; Herbert, Pollatos, & Schandry, 2007; Herbert, Ulbrich, & Schandry, 2007; Pollatos, Gramann, & Schandry, 2007; Schandry, 1981; Shah, Hall, Catmur, & Bird, 2016; Wiens, 2005). In this task, participants are asked to estimate the number of their heartbeats, during a given period, without touching any part of their body (hence "from the inside"), while an objective measure of heart rate is taken (e.g., using a finger pulse oximeter or ECG recordings). The degree of correspondence between the objective and the subjective heartbeat estimation provides an index of interoceptive *accuracy*. Despite the objective nature of the heartbeat tracking task, some potential limitations have been reported. It has been argued that this measure is unable to detect individual differences at the lower levels of interoceptive accuracy (Khalsa, Rudrauf, Damasio, Davidson, Lutz, Tranel, & 2008; Murphy et al., 2018) but most importantly, it has been reported that individuals could achieve an accurate heartbeat estimation by perceiving vibrations from the reflection of the heartbeats on parts of their bodies (Khalsa, Rudraufl, Feinstein, & Tranel, 2009, Murphy et al., 2018). Therefore, other organs, including the lungs have started gaining more attention in the field of interoception (Daubenmier, Sze, Kerr, Kemeny, & Mehling, 2013; Faull, Cox, & Pattinson, 2016).

Respiratory interoceptive accuracy has been mainly measured with different types of resistive load tasks (Bogaerts et al., 2008; Daubenmier, Sze, Kerr, Kemeny, & Mehlin, 2013; Garfinkel et al., 2016a). However, these kind of tasks cannot be considered the most appropriate one for testing clinical populations, given their invasive nature. In a recently published study, Murphy, Catmur, and Bird (2018a) employed a blow task and measured individuals' ability to exert control over exhaling, as index of respiratory interoceptive accuracy. Nevertheless, interoception is a multidimensional construct, and as such, it does not include only accuracy. Interoceptive awareness is another facet of interoception, which is distinct from interoceptive accuracy (Ceunen, Van Diest, & Vlaeyen, 2013; Forkmann, et al., 2016; Garfinkel, Seth, Barrett, Suzuki, & Critchley, 2015).

Interoceptive *awareness* concerns the accuracy of one's judgements of the degree of correspondence between estimates of interoceptive accuracy and actual accuracy (Forkmann et al., 2016; Garfinkel et al., 2015, 2016b). In more simple terms, it refers to "metacognitive awareness of objective accuracy" (Garfinkel & Critchley, 2013, p. 233), and as such, it involves forming metarepresentations of one’s own beliefs about their ability to judge their own physical internal states (Garfinkel & Critchley, 2013). The higher the accuracy of judgements of interoceptive accuracy the better the metacognitive awareness. This ability has been assessed with the addition of judgment of confidence tasks in traditional paradigms (e.g. heartbeat tracking task, heartbeat discrimination task) that measure interoceptive accuracy (Ehlers, Breuer, Dohn, & Fiegenbaum, 1995; Garfinkel et al., 2015; Khalsa et al., 2008).

Metacognition

Metacognition describes “any knowledge or cognitive activity that takes as its cognitive object, or regulates, any aspect of any cognitive activity” (Flavell, Miller, & Miller, 1993, p.150). In simple terms, it is the cognition about cognition. That is to say, people’s ability to form thoughts about their own thoughts and cognition (Dunlosky & Metcalfe, 2009; Flavell, 1979). Metacognition has been considered multifaceted. One of its most widely acknowledged dimensions is the awareness of a current mental state or cognitive process. Namely, metacognitive monitoring (Dunlosky & Metcalfe, 2009). Metacognitive monitoring is a powerful cognitive tool that is involved in various aspects of human life.

Individuals’ behaviour in everyday situations is much affected by this ability, as it is the awareness of ones’ own current mental states that allows the regulation of those states (Nelson & Leonesio, 1988). Moreover, it has been found that metacognitive monitoring predicts educational achievement independent of IQ (Hartwig & Dunlosky, 2012; Pishghadam & Khajavy, 2013; Thiede, Anderson, & Therriault, 1999; Veenman & Spaans, 2005), with its predictive effect being especially strong for performance in mathematics (Higgins et al., 2013; Iuculano et al., 2014). In addition, cognitive deficits and difficulties in mathematics can be alleviated through training in metacognitive strategies (Dunlosky, Kubat-Silman, & Hertzog, 2003; Maras, Gamble, & Brosnan, 2017; Murphy, Schmitt, Caruso, & Sanders, 1987; Roebers, Cimeli, Röthlisberger, & Neuenschwander, 2012; Roebers, Krebs, & Roderer, 2014; Teong, 2003).

Metacognitive monitoring entails different metacognitive experiences that reflect the conscious cognitive or affective experience that follows a cognitive process, including judgements of confidence and metacognitive experiences of difficulty (Dunlosky & Metcalfe, 2009; Efklides, 2006; Flavell, 1979; Nelson & Narens, 1994; Schneider, 2008). These experiences enable us to operationalise metacognitive monitoring and thus to measure this ability.

Judgement of confidence (JoC) is one of the classic tasks that have been used to investigate metacognitive monitoring (Dunlosky & Metcalfe, 2009; Koriat, 2007). At first, individuals take a test (e.g. memory or visual discrimination), and then they are asked to indicate their confidence about the correctness of each of the previously given responses. The correspondence between the object-level test performance and the meta-level judgements of confidence provides an index of metacognitive monitoring. The higher the correspondence the better the ability (Dunlosky & Metcalfe, 2009).

Metacognitive experiences of *difficulty* have been explored with a masked priming experiment (Desender, Van Opstal, Hughes, & Van den Bussche, 2016). In this task, an invisible prime arrow and a target one are presented, almost simultaneously to individuals who are instructed to respond to the direction of the target arrow. On congruent trials, the prime arrow has the same direction with the target one, and thus individuals tend to provide fast and correct responses to the target arrow. Whereas, on incongruent trials, the stimuli have different directions, creating a response conflict that make individuals give slower and prone to error responses (Desender, Van Opstal, Hughes, & Van den Bussche, 2014). After each response, individuals are asked to judge the difficulty they experienced when responding to the target arrow, without being aware of which trials were experimentally manipulated to be more difficult (incongruent trials), compared to the other ones (congruent trials) (Desender et al., 2016). The more the subjective experience of difficulty coincides with the actual difficulty of the trial, the better the metacognitive monitoring ability. When Desender et al. (2016) employed this task among neurotypical people, they found that people were indeed able to give accurate metacognitive judgements, labelling incongruent trials as “rather more difficult” and congruent ones as “rather less difficult”.

Regardless of the different ways that “self” has been operationalised, the importance of self-awareness is indisputable; besides “we are what our attention to ourselves makes us be” (Seigel, 2005, p.6). This makes the exploration of self-awareness even more crucial for disorders that are inherently related with this concept, when it comes to their understanding (Kerig, Ludlow, & Wenar, 2012). One of the most prominent ones is Autism Spectrum Disorder (autism = “*autos*” = self).

Autism Spectrum Disorder

Autism Spectrum Disorder (ASD) is a developmental disorder that affects approximately 1% of the childhood population in the UK (Baird et al., 2006; Baron-Cohen et al., 2009). The prognosis of ASD is poor, with the majority of affected individuals dependent on families and social services, socially excluded and unemployed (Howlin, Goode, Hutton, & Rutter, 2004).

Individuals are diagnosed with ASD based on a constellation of significant social-communication difficulties, together with restricted interests and repetitive behaviours (American Psychiatric Association, 2013). This “dyad” of impairments is reflected upon the clinical manifestation of the disorder. That is to say, difficulties in social-emotional reciprocity, absence or idiosyncratic verbal and non-verbal communicative behaviour, deficits in social relationships, insistence on sameness and on stereotyped behaviours, highly restricted interests, and sensory peculiarities (American Psychiatric Association, 2013).

These difficulties have been proposed to be part of a “broad autism phenotype” (Bolton et al., 1994; Goldberg et al., 2005; Le Couteur et al., 1996; Murphy et al., 2000; Pickles et al., 2000; Piven et al., 1997; Szatmari et al., 2000). That is to say, there are personality characteristics in the neurotypical population that are qualitatively similar to the defining features of ASD, reflecting the phenotypic expression of the disorder, henceforth autism traits. Research findings have indicated that autism traits are normally distributed in the general population (Constantino & Todd, 2000; 2003; Ronald, Happé, Price, Baron-Cohen, & Plomin, 2006) but most importantly that “unaffected” relatives of people with ASD have increased autism traits compared to the general population (Frazier et al., 2014; Pickles et al., 2000; Piven et al., 1994; 1997). As such, it has been suggested that the study of the relation between individual differences in autism traits and psychological concepts among people form the general population could be informative about the ASD itself (Williams, Nicholson, & Grainger, 2018a).

According to one well-documented hypothesis, the social-communication difficulties that characterise individuals with ASD could be partially attributed to impaired mindreading abilities (Brunsdon & Happé, 2014; Frith, 1994; Tager-Flusberg, 1999). Mindreading (often called “theory of mind”) is the ability to represent people’s thoughts. In our everyday life, we constantly (and often unconsciously) attribute mental states, such as beliefs, desires, intentions, goals, and feelings to other individuals in an attempt to interpret, predict, and explain their behaviour (Carruthers, 2009; Premack & Woodruff, 1978). Baron-Cohen, Leslie, and Frith (1985) were the first who provided evidence for impaired mindreading abilities in children with ASD. Since then, many researchers have successfully replicated these findings, examining both children and adults with ASD (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001a; Senju, Southgate,White, & Frith, 2009; Yirmiya, Erel, Shaked, & Solomonica-Levi, 1998).

Despite the wealth of studies on social-communication difficulties in ASD, the second half of the “dyad” of impairments remains relatively under-researched. Although, it has been suggested that many of the unexplained clinical manifestations of ASD, such as repetitive behaviours and absence of pretend play could be attributed to diminished self-awareness (Carruthers, 1996; Frith & Happé, 1999; Hobson, 1990), the study of lower and higher order phenomena of self-awareness in ASD has only recently began.

Interoception and ASD.

Recent studies have shown that interoceptive accuracy plays a major role in important psychological functions, including emotion-processing, empathy and mindreading, on the basis that emotions are triggered from the perception and interpretation of physiological changes (Craig, 2003; Damasio, 2000; Garfinkel et al., 2015; Shah, Catmur, & Bird, 2017; Terasawa, Moriguchi, Tochizawa, & Umeda, 2014). As such, individuals with high interoceptive abilities tend to experience their emotions more intensely (Herbert et al., 2007; Wiens, Mezzacappa, & Katkin, 2000) and thus to understand them better (Critchley, 2005; Herbert , Herbert, & Pollatos, 2011). Based on this assumption, diminished interoception can have potentially serious implications in disorders that are characterised by deficits in those abilities, such as ASD (Baron-Cohen & Wheelwright, 2004; Blair, 2008; Philip et al., 2010).

Quattrocki and Friston (2014) argued that the cause of ASD is a profoundly damaged oxytocin system. Oxytocin is a hormone that acts as a neurotransmitter in the brain and plays an important role in sexual reproduction and regulation of related social behaviours. According to the “oxytocin hypothesis”, a dysfunctional oxytocin system mediates the relationship between interoception and social cognition. In the case of ASD, that dysfunctional system transfers diminished interoceptive signals to the affected individual, resulting in a reduced capacity to represent cognitive/emotional states in others (Quattrocki & Friston, 2014). Brewer, Happé, Cook, and Bird (2015), criticising this account claimed that it is not autism *per se* that is characterised by diminished interoception. Instead, the co-occurrence with ASD alexithymia results in interoceptive deficits.

Alexithymia affects 50% of individuals with ASD (Berthoz & Hill, 2005) and describes a difficulty in identifying and describing emotions, as well as in distinguishing feelings from body sensations (Nemiah, Freyberger, & Sifneos, 1976; Sifneos, 1973). According to the “alexithymia hypothesis”, some of the social and emotional difficulties in ASD are explained by the co-occurrence of alexithymia (Bird & Cook, 2013; Bird et al., 2010; Cook, Brewer, Shah, & Bird, 2013; Oakley, Brewer, Bird & Catmur, 2016). In line with this hypothesis, it has been suggested that interoception too has a unique relation to alexithymia and thus individuals with ASD have impaired interoception only when alexithymia co-occurs with ASD (Brewer, Happé, Cook, & Bird, 2015; Hatfield, Brown, Giummarra & Leggenhager, 2017; Herbert et al., 2011; Shah, et al., 2016). The prediction that follows both theoretical accounts is that affected individuals with ASD have impaired interoceptive accuracy, either as a core feature of ASD or as one of the consequences of alexithymia.

The counter argument that interoceptive accuracy is intact in ASD stems from the theoretical claim that there is a distinction between physical and psychological self (Gillihan & Farah, 2005). Based on findings that people with ASD have intact body recognition when they reach an appropriate developmental level (Dawson & McKissick, 1984; Lind & Bowler, 2009; Nielsen, Suddendorf, & Dissanayake, 2006) as well as unimpaired recognition of their own agency (Frith & Hermelin, 1969; Williams & Happé, 2009), it has been argued that physical self-awareness is intact in ASD (Williams, 2010). Given that interoceptive accuracy lies under the sphere of physical self-awareness, it has been suggested that is undiminished in ASD as well (Lind, 2010; Nicholson et al., 2018, Uddin , 2011; Williams, 2010).

Metacognition and ASD.

Whilst a consensus regarding metacognition in ASD could have direct impact on affected individuals' everyday life and education, it remains controversial whether ASD is characterised by impaired or intact metacognition. "One-system" theorists claim that metacognition is impaired in ASD, on the basis that individuals become aware of themselves with the same way they become aware of others (Carruthers, 2009; Gopnik, 1993; Gopnik & Meltzoff, 1994; Hobson, 1990).

Carruthers, who is one of the main proponents of this account, argues in favour of a single metarepresentational faculty (Carruthers, 2009; Leslie, 1987; Moore & Frye, 1991) that enables individuals to be aware of themselves, simply by turning the human capacity of attributing mental states to others to their own selves. This faculty was initially developed to serve the purpose of reading other people's mind, so that human beings could increase their chances of survival (Wegner, 2002; Wilson, 2004). Yet, during the course of evolution, this role became dual with individuals using metarepresentational abilities in order to read their own minds as well (Carruthers, 2009; 2011; 2013). Subsequently, mindreading and metacognition are inherently related and thus, a double dissociation between them could never be the case (Gopnik & Meltzoff, 2006). Given that is well-established that people with ASD have attenuated mindreading abilities (Baron-Cohen et al., 2001a; Senju et al., 2009; Yirmiya et al., 1998), "one-system" theorists predict that individuals with ASD have profound difficulties also in metacognition (Caruthers, 2009; Gopnik, 1993; Williams, 2010).

In contrast, "two-system" theorists postulate that mindreading and metacognition are two independent capacities based on either entirely different or partly different underlying cognitive mechanisms (Couchman, Coutinho, Beran, & Smith, 2009; Goldman, 2006; Nichols & Stich, 2003). As such, they predict that people with ASD have diminished mindreading abilities but intact metacognition (Goldman, 2006; Nichols & Stich, 2003).

According to “two-system” theorists, autobiographical accounts indicate that individuals with ASD are able to reflect upon their mental states and discuss about them, providing evidence for undiminished metacognitive abilities in ASD (Nichols & Stich, 2003; McGeer, 2004; but see Williams, 2010). To date, the theoretical dispute between “one-system” theorists and “two-system” theorists remains unresolved.

The Current Project

In the current project, we conducted five experiments aiming to provide insight on lower and higher order phenomena of self-awareness in ASD (see Appendix A for Experiment 4.1). In Experiment 1.1, we conducted a case-control study in order to explore interoceptive accuracy and interoceptive awareness, within two different interoceptive domains. In Experiment 2.1 and 2.2, we investigated accuracy of implicit and explicit judgments of confidence not only among people with a full diagnosis of ASD but also in relation to autism traits, as measured in the general population. Finally, in Experiment 3.1 and 3.2, we explored accuracy of metacognitive experiences of difficulty following the approach described above. That is a novel methodology that is gaining increasing attention in the study of ASD (Nicholson et al., 2018; Williams, Bergstrom, & Grainger, 2016; Williams et al., 2018a; Williams, Nicholson, Grainger, Lind, & Carruthers, 2018b).

Based on the hypothesis of a “broad autism phenotype”, it has been argued that the relation between cognitive phenomena of interest and autism traits could inform predictions about between-group differences in case-control studies (Williams et al., 2016). However, it has been also claimed that measuring autism traits is not always a valid proxy for deficits in ASD, because qualitative differences in cognitive mechanisms could differentiate people with and without a diagnosis of ASD (Peterson, Wellman, & Liu, 2005; Ruzich et al., 2015). Therefore, conducting both types of experiments appears to provide a better understanding of the phenomena under investigation (Williams et al., 2018b).

1. Exploring Interoception in ASD

As noted in the general introduction, there is a suggestion that ASD is characterised by attenuated interoceptive abilities, either due to a dysfunctional oxytocin system (Quattrocki & Friston, 2014) or as one of the consequences of alexithymia (Brewer et al., 2015) that affects almost half of the individuals with ASD (Berthoz & Hill, 2005). Alternatively, it has been argued that interoceptive accuracy is undiminished among individuals with ASD, as it reflects their intact physical self-awareness (e.g. Nicholson et al., 2018; Williams, 2010). Regardless of the contradictory predictions that stem from these theoretical accounts, all of them raise the importance of examining interoception when it comes to the understanding of ASD. Nonetheless, to date research evidence is surprisingly sparse, with only four studies having examined cardiac interoceptive accuracy employing the heartbeat tracking task among adults with ASD and with only one study having focused on respiratory interoceptive domain. Despite the clear hypotheses that stem from the theoretical accounts described above, we cannot draw firm conclusions about interoceptive accuracy in ASD because the existing research findings are to a great extent contradictory.

Garfinkel et al. (2016b) tested 20 adults with ASD and 20 comparison participants, matched for gender and age, using the heartbeat tracking task. They found that the ASD group showed significantly lower interoceptive accuracy, compared to the control group, and thus they concluded that interoceptive abilities are attenuated in ASD. However, Garfinkel et al. (2016b) did not measure IQ, which has been found to be positively associated with interoceptive accuracy (Mash, Schauder, Cochran, Park, & Cascio, 2017; Murphy et al., 2018c). Therefore, it is unclear whether their results reflect a genuine trough in ASD or have been confounded by the effect of IQ.

When Shah, Hall, Catmur, and Bird (2016) tested 19 adults with ASD and 19 comparisons matched for age, gender, IQ, and alexithymia, they did not find significant between-group differences in interoceptive accuracy. Nonetheless, they concluded that this was the case only because groups were matched for alexithymia, implying that differences would be apparent if participants with ASD had elevated levels of alexithymia.

Nicholson et al. (2018) refuted this hypothesis by testing 46 adults with ASD and 46 comparison participants, closely matched for age, gender, and IQ and finding that a “high alexithymic” ASD group did not differ significantly from a “low alexithymic” group in interoceptive accuracy. Moreover, their complete ASD group showed interoceptive accuracy equivalent with the comparison group, reflecting an intact ability in ASD. Yet, even these results should be interpreted with caution because Nicholson and his colleagues (2018) did not control for the effect of body mass index (BMI), which has been found to be negatively associated with interoceptive accuracy (Murphy, Geary, Millgate, Catmur, & Bird, 2018b; Rouse, Jones, & Jones, 1988). Thus, if their control group had greater BMI, compared to the ASD group, then potential between-group differences in interoceptive accuracy may have been obscured.

Indeed, when Mul, Stagg, Herbelin, and Aspell (2018) controlled for BMI, they found significant between-group differences in interoceptive accuracy that led them to conclude that this ability is attenuated among people with ASD. Nonetheless, they did not measure either verbal or performance IQ. Groups were only equated for their full scale IQ, which does not exclude the possibly that diminished interoceptive accuracy in the ASD group might be due to the confound effect of between-group differences in either verbal or performance IQ (Mash et al., 2017; Murphy et al., 2018c).

Murphy et al. (2018a) examined interoceptive accuracy within the respiratory domain, employing a blow task among people from the general population and measuring their autism traits. In this novel task, participants were asked to blow into a peak flow meter and then to exhale for a second time, aiming to target a particular percent of the intensity of their first exhalation. Murphy et al. (2018a) found that autism traits were not significantly associated with respiratory interoceptive accuracy and based on that they concluded that it is questionable whether individuals with a diagnosis of ASD have diminished respiratory interoceptive abilities. Nonetheless, to date this has never been explored.

Research evidence is also sparse with respect to interoceptive awareness. To our knowledge, only two studies have examined this metacognitive phenomenon among individuals with ASD, and yet the heartbeat tracking task has not been used in either of them. Instead, Garfinkel et al. (2016b) used a heartbeat discrimination task, in which participants had to judge whether a tone was synchronous or asynchronous to their heart rate. Following this task, participants' judgements of confidence were elicited. Their analysis revealed no significant between-group differences in interoceptive awareness. However, it remains unclear whether confidence judgments were predictive of interoceptive accuracy in both groups or in any of them, and whether interoceptive awareness was above or below chance level. Palser, Fotopoulou, Pellicano, and Kilner (2018) employed the same task among children with ASD and neurotypical children, and found that children with ASD were more confident about their interoceptive accuracy, compared to neurotypical children. Nonetheless, this study is uninformative about interoceptive awareness in ASD because judgments of confidence were analysed independently of participants' interoceptive accuracy.

Overall, it may not be possible to draw strong conclusions from studies measuring cardiac interoceptive accuracy and awareness alone among individuals with ASD, given conflicting findings and possible methodological limitations. Nonetheless, there is an almost complete absence of studies in ASD focusing on other interoceptive domains, making the relation between interception and ASD even more nebulous.

Given that it has already been found a non-significant association between autism traits and either cardiac or respiratory interoceptive accuracy in the general population

(Murphy, Catmur, & Bird, 2018a, Nicholson et al., 2018), we conducted only a case-control study among people with a full diagnosis of ASD and neurotypical people. The aim of the current experiment was to examine two different interoceptive domains in ASD, overcoming some of the limitations of previous research, such controlling for the potential confound effects of depressive symptoms and anxiety on interoceptive accuracy (Garfinkel et al., 2016; Pollatos, Traut-Mattausch, & Schandry).

Experiment 1.1: Case-Control

In Experiment 1.1, we explored cardiac and respiratory interoceptive accuracy, using the classic heartbeat tracking task (Schandry, 1981) and a slightly modified version of Murphy et al.'s (2018a) blow task, controlling for the effect of a series of potential confounds. In both tasks, participants' judgments of confidence were elicited in order to measure cardiac and respiratory interoceptive awareness. In addition, participants completed the Autism Spectrum Quotient (AQ; Baron-Cohen, Wheelwright, Skinner, Martin, and Clubley, 2001b), the 20-Item Toronto Alexithymia Scale (TAS-20; Parker, Taylor, & Bagby, 1994), the Reading the Mind in the Eyes task (RMIE; Baron-Cohen et al., 2001a), and the Movie for the Assessment of Social Cognition (MASC; Dziobek et al., 2006), with the last two measuring mindreading abilities.

As it is well-established that individuals with ASD have significant deficits in attributing mental states to others (e.g. Yirmiya, et al., 1998), our first predictions was that participants with ASD would show diminished performance in both mindreading tasks, compared to neurotypical participants. Second, we predicted that we would find no significant between-group differences, in either cardiac or respiratory interoceptive accuracy. This prediction was based on the theoretical claim that interoceptive accuracy entails first order representations of body signals and as such is undiminished in ASD, reflecting intact physical self-awareness (e.g. Damasio, 2000; Williams, 2010). Also based on that, we did not expect significant differences in interoceptive accuracy, even when contrasting ASD participants with low alexithymia with ASD participants with high alexithymia (Bird & Cook, 2013). Despite predicting undiminished accuracy, we expected that participants with ASD would show atypical cardiac and respiratory interoceptive awareness, based on the claim that both mindreading and metacognition are under the umbrella of a single metarepresentational faculty (e.g. Carruthers, 2009). Finally, we aimed to establish the extent to which interoceptive awareness is associated with mindreading abilities. Previous findings have indicated a positive association between metacognition, as indexed by judgements of confidence accuracy and mindreading abilities (Williams et al., 2016), but to date no study has examined the association between interoceptive awareness and mindreading. Based on "one-system" theory (e.g. Carruthers, 2009) and on the suggestion that interoceptive awareness is metarepresentational (e.g. Garfinkel & Critchley, 2013), we expected to find a positive and significant association between interoceptive awareness and mindreading.

Experiment 1.1: Method

Participants

Twenty-two adults with ASD and 20 neurotypical comparison adults participated in the current experiment after they had given written informed consent. All participants with ASD had received formal diagnoses, according to established criteria (DSM-IV-TR, American Psychiatric Association 2000; ICD-10, World Health Organisation 1993). In addition, participants with ASD completed the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000), a widely used semi-structured observational measure that provides an accurate diagnosis of autism and symptom severity.[1] Ethical approval for this study was obtained from the Kent School Psychology Research Ethics Committee (Ethics ID: 201815259101245011).

In terms of participant characteristics, as shown in Table 1 groups were matched for verbal IQ, performance IQ, and full scale IQ, using the Wechsler Abbreviated Scale of

Intelligence (WASI; Wechsler, 1999). They were also equated for chronological age, gender, and BMI but as expected, they differed significantly in AQ, MASC, and TAS-20 scores.

Table 1

Baseline Characteristics and Matching Statistics for Experiment 1.1

	Diagnostic Group		Group Differences			
	ASD (n = 22; 14 male)	Neurotypical (n = 20; 14 male)	t	p	d	BF_{10}
CA: years	36.54 (12.01)	41.95 (13.94)	-1.35	.18	0.42	0.62
VIQ	105.41 (11.18)	104.05 (11.22)	0.39	.70	0.12	0.32
PIQ	106.41 (17.51)	105.60 (15.18)	0.16	.87	0.05	0.31
FSIQ	106.32 (13.27)	105.65 (12.99)	0.16	.87	0.05	0.31
BMI	26.14 (5.73)	26.31 (4.61)	-0.11	.92	0.03	0.31
Heartrate	75.84 (16.20)	69.48 (12.79)	1.40	.17	0.44	0.66
ADOS	9.27 (4.68)	-	-	-	-	-
AQ total	32.86 (8.03)	14.25 (4.56)	9.34	<.001	2.85	>100
RMIE[a]	25.36 (5.57)	27.80 (3.86)	-1.63	.06	0.51	1.60
MASC[a]	28.32 (6.51)	33.75 (5.21)	-2.97	.01	0.92	16.73
TAS-20	61.59 (11.47)	43.50 (9.38)	5.56	<.001	1.73	>100
BDI-II	15.45 (11.64)	8.00 (6.05)	2.64	.01	0.80	3.78
STAI	96.09 (20.63)	71.10 (18.80)	4.09	<.001	1.27	>100
Time estimation	.81 (.17)	.74 (.17)	1.20	.24	0.37	0.54
Memory[b]	.86 (.07)	.87 (.09)	-0.41	.69	0.13	0.33

Note. CA = chronological age; VIQ = verbal IQ; PIQ = performance IQ; FSIQ = full scale IQ; BMI = Body mass index; Heart Rate = Participant mean heartrate during the heartbeat tracking task; ADOS = Autism Diagnostic Observation Schedule; AQ total = Total score on Autism-Spectrum Quotient; RMIE = Reading the Mind in the Eyes; MASC = Movie for the Assessment of Social Cognition; TAS-20 = Twenty-item Toronto Alexithymia Scale; BDI-II = Becks Depression Inventory 2; STAI = State/Trait Anxiety Inventory; Time estimation = Mean accuracy (proportion) in time estimation control task; Memory = Mean accuracy (proportion) in memory control task.

[a] Values for one-tailed tests are reported because of a priori directional predictions.

[b] Memory data is missing for one TD participant.

Materials, Procedures, and Scoring

Experimental interoception tasks.

Heartbeat tracking task.

We employed this widely used task to measure cardiac interoceptive accuracy and cardiac interoceptive awareness (Schandry, 1981). Participants were asked to silently count their heartbeats during four different time intervals (25s, 35s, 45s, & 100s), having their eyes closed and being seated in an upright position. The order of intervals was randomised across participants. Throughout the task, a finger pulse oximeter (Contec Systems CMS-50Dþ; Qinhuangdao, China) was attached to the index finger of their dominant hand, recording their heartrate. An auditory tone signalled the beginning and the end of each trial. Within this time period, participants were instructed to “feel their heart from the inside and count their heartbeats”. They were given strict guidelines not to take their pulse or touch any part of their body that could facilitate the detection of their heartbeats.

At the end of each trial, participants were asked to type in the estimated number of their heartbeats, using a PC keyboard. Subsequently, they gave a confidence judgment (JoC) based on how confident they felt about the accuracy with which they estimated their heartbeats. In order to give a JoC, they used a 5-point scale, ranging from “I don’t know” to “I am sure”.[2]

Cardiac interoceptive accuracy was quantified using the formula: 1 – (actual number of heartbeats – counted number of heartbeats) / ((actual number of heartbeats + counted number of heartbeats)/2), with scores taking values between -1 and 1 (Garfinkel et al., 2016b; Hart, McGowan, Minati, & Critchley, 2013; Nicholson et al., 2018). We opted for this way to quantify interceptive accuracy, as the formula described above takes into account both under and over heartbeat estimations and thus provides scores less prone to positive accuracy bias (Nicholson et al., 2018). Cardiac interoceptive accuracy was extracted for each trial and then a mean score was calculated for each participant, averaging their accuracy scores across the four different time intervals. Large positive scores indicate high cardiac interoceptive accuracy.

Following Garfinkel et al.’s (2015) approach, we quantified cardiac interoceptive awareness calculating a Pearson *r* correlation for each participant, indicating the correspondence between interoceptive accuracy and JoC (see also Forkmann et al., 2016). Large positive correlation coefficients imply high interoceptive awareness. Cardiac interoceptive awareness data is missing for three participants (one ASD/two TD), due to lack of variability in their JoC responses.

Respiratory interoception task.

This task, which is a modified version of a novel paradigm designed by Murphy et al. (2018a), was employed to measure respiratory interoceptive accuracy and respiratory interoceptive awareness. Participants were instructed to exhale with a particular intensity (weak, medium, and firm) into a peak flow meter (Mini-Wright) and after 15s to exhale for a second time (comparator blow), with the intention of matching the intensity of the prior exhalation (actual blow). Disposable mouthpieces were used.

Each participant completed three consecutive trials for each intensity, with the order of intensities being counterbalanced across participants. During the administration of the experimental task, participants neither were informed about their scores nor could see them, as a cardboard was attached to the base of the peak flow meter and thus they were instructed to rely solely on their feelings to recreate the intensity of the actual blow. Between the different intensities, a 40s break was given to participants to allow their breathing to return to normal. Before the beginning of the task, each participant completed at least one practice trial for each intensity.

At the end of each trial, participants marked their level of confidence in a paper and pencil JoC task, using a 5-point scale that ranged from “I don’t know” to “I am sure”. The higher the score the greater the confidence that the intensity of the each comparison blow had been accurately matched with the intensity of the actual blow. Due to an error in the data collection process, confidence judgements were elicited only from 15 participants with ASD and 19 comparisons.

As regards to the scoring, first we calculated an accuracy score for each trial, using the formula: 1 – (score on the actual blow – score on the comparator blow) / ((score on the actual blow + score on the comparator blow)/2). Then, we averaged accuracy scores for each intensity and finally we calculated a grand mean, with large positive scores indicating high respiratory interoceptive accuracy.

Respiratory interoceptive awareness was quantified calculating a Pearson *r* correlation for each participant, indicating the correspondence between respiratory interoceptive accuracy and JoC. Large positive correlation coefficients imply high interoceptive awareness. Respiratory interoceptive awareness data is missing for one TD participant, due to lack of variability in their JoC responses.

“Control” tasks.

Time estimation.

We used this task to control for the potential confound effect of time estimation on cardiac interoceptive accuracy (e.g. Ainley, Brass, & Tsakiris, 2014; Shah et al., 2016). Following the same procedure as in the heartbeat tracking task, participants were asked to count the number of seconds of three different time intervals (19, 37, & 49s), with the order of intervals being randomised across participants.

Accuracy was quantified using the same formula that used to measure cardiac interoceptive accuracy. That is 1 – (actual number of seconds – counted number of seconds) / ((actual number of seconds + counted number of seconds)/2). A mean accuracy score was calculated for each participant. Large positive scores indicate high accuracy in time estimation.

Memory task.

Between-group differences in respiratory interoceptive accuracy could be explained by between-group differences in short-term memory for internal effort. Therefore, we employed a memory control task to compare performance on the respiratory interoception task, in order to ensure that short-term memory would not have any confound effect on our results. In this task, participants were asked to press the “b” key of a keyboard down, producing a tone and then release the key whenever they wanted to. Subsequently, they were asked to repeat the task with the aim of pressing the button down for the same length of time, as they previously did. They were instructed to rely solely on their feelings to recreate the duration of the prior tone and not to count seconds. Participants completed one practice trial and then nine experimental ones.

Accuracy in each trial was calculated using the formula: 1 – (actual number of seconds – comparison number of seconds) / ((actual number of seconds + comparison number of seconds)/2). A mean accuracy score was calculated for each participant, averaging their accuracy scores across the nine trials.

Mindreading tasks.

Reading the Mind in the Eyes (RMIE) test.

The RMIE (Baron-Cohen et al., 2001a) is a reliable and widely used measure of mindreading abilities in general and clinical populations (e.g. Domes, Heinrichs, Michel, Berger, & Herpertz, 2007). In this task, individuals are presented with a series of 36 photographs showing the eye-region of males and females, and they are asked to choose among four different options the emotion/feeling that best describes the mental state of the depicted person. Scores range from 0 to 36, with higher scores indicating better mindreading abilities.

Movie for the Assessment of Social Cognition (MASC).

The MASC (Dziobek et al., 2006) is a reliable and widely used measure of mindreading abilities in general and clinical populations (e.g. Martinez et al., 2017; Shah et al., 2017. In this task, individuals are asked to watch a 15-minute movie about four characters getting together for a dinner party. The video is paused 46 times and questions concerning the characters' feelings, thoughts, and intentions are asked. All of the answers are multiple choice and require one option to be selected from a choice of four. Scores range from 0 to 46, with higher scores indicating better mindreading abilities. The task also involves six control questions, assessing attention and understanding of non-social aspects of the plot.

Self-report measures.

Autism-Spectrum Quotient (AQ).

The AQ (Baron-Cohen et al., 2001b) is a widely used self-report questionnaire that measures reliably ASD traits, in both general and clinical populations (e.g. Reed, Lowe, & Everett, 2011; Williams et al., 2018a). In this task, individuals are asked to indicate the extent to which they agree with each of the 50 statements (e.g., "I find social situations easy") that the questionnaire comprises, using a 4-point Liker scale, ranging from "definitely agree" to "definitely disagree". Scores range from 0 to 50, with higher scores indicating more ASD traits. A score of ≥ 26 is the cut-off point that denotes clinically significant levels of autism traits (Woodbury-Smith, Robinson, Wheelwright, & Baron-Cohen, 2005).

Beck Depression Inventory 2 (BDI-II).

The BDI-II (Beck, Steer, & Brown, 1996) is a reliable and widely used self-report questionnaire that consists of 21 aspects of depression (e.g. O'hara & Swain, 1996). In this questionnaire, individuals are asked to choose one statement for each aspect of depression that best describes their feelings the last two weeks including the day of the testing session (e.g. I am so sad or unhappy that I can't stand it). Each statement is rated on a scale ranging from 0 to 3, with 0 being equivalent to no presence of the aspect and 3 indicating strong presence of the aspect. Scores range from 0 to 63, with higher scores indicating more severe depressive symptoms. A total score was obtained for each participant by summing all their responses.

State-trait Anxiety Inventory (STAI).

The STAI (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983) is a reliable and widely used self-report questionnaire that measures state and trait anxiety (e.g. Bryant, Harvey, Dang, Sackville, & Basten, 1998). The STAI comprises 40 statements; half of them refer to current states, while the rest refer to general trait states. First, individuals are presented with statements such as “ I feel calm”, and they are asked to rate these statements, using a 4-point scale, ranging from “not at all” to “very much so”, based on their current feelings. These measure state anxiety. Then, individuals are presented with the statements that measure trait anxiety, such as “I am a steady person”, and they are asked to rate these statements using a 4-point scale, ranging from “almost never” to “almost always”, based on how they generally feel. Scores range from 40 to 160, with higher scores indicating greater levels of anxiety. We calculated an overall score of anxiety for each participant, by summing their state and trait anxiety subscores.

20-Item Toronto alexithymia scale (TAS-20).

The TAS-20 (Parker et al., 1994) is a reliable and widely used self-report questionnaire that assesses people’s ability to identify and describe feelings and emotions (e.g. Szatmari et al., 2008). It comprises 20 statements (e.g “I am able to describe my feelings easily”) with which individuals have to answer whether they completely disagree or completely agree on a 5-point scale. Scores ranges from 20 to 100. A score of ≥ 61 indicates clinically significant levels of alexithymia.

Body mass index (BMI).

Participants’ BMI was calculated using the following formula, BMI = (weight (kg) / height (cm))/height (cm).[3]

Statistical Analysis

In the current experiment, the alpha level of .05 was the criterion for statistical significance, but when a-priori directional predictions were made, values for one-tailed tests were reported. The analysis approach was to conduct independent samples *t*-tests to examine between group differences and pairwise correlations to examine the strength of association between variables. Given the small number of variables, pairwise correlations were sufficient to let us confirm or reject the hypotheses of the current studies. With respect to indices of effect size, we reported partial eta squared (η_p^2) values when ANOVA tests were conducted (≥ .01 = small effect, ≥ .06 = moderate effect, ≥ .14 = large effect; Cohen, 1969) and Cohen's *d* values (≥ 0.20 = small effect, ≥ 0.50 = moderate effect, ≥ 0.80 = large effect; Cohen, 1969) for *t*-tests. In correlational analyses, coefficients $r \geq 10$ indicated a small effect size, ≥ 30 indicated a moderate effect size, and values ≥ 50 indicated a large effect size (Cohen, 1992).

In addition to *p*-values, we calculated another indicator that allowed us to make inferences based on our data. Bayesian analysis is an alternative to null hypothesis significance testing that steadily gains ground in the field of social sciences and has been used in similar with our projects studies (Nicholson et al., 2018; Williams et al., 2016; 2018a; 2018b). Bayes factors enabled us to determine whether null results were due to sample insensitivity or because the effect under examination was minimal in the population (Wagenmakers, 2007). When a-priori directional predictions were made, BF_{10} values for one-tailed tests were reported. According to Jeffreys's (1961) criteria, Bayes factors > 1 indicate increasing evidence for the alternative hypothesis over the null hypothesis (1 - 3 = anecdotal evidence; 3 - 10 = substantial evidence; 10 - 30 = strong evidence; 30 - 100 = very strong; values > 100 = decisive evidence). Whereas, scores < 1 indicate evidence for the null hypothesis over the alternative hypothesis (1 - 0.33 = anecdotal evidence; 0.33 - 0.10 = substantial evidence; 0.10 - 0.03 = strong evidence; 0.03 - 0.01 = very strong evidence; scores < 0.01 decisive evidence). All the reported Bayesian analyses were performed using the statistical software package JASP 0.8.1.2 (JASP Team, 2016) and thus we used the default prior; that is a zero-centered Cauchy distribution. Please note that the same method of statistical analysis was followed in all the studies included in the current project.

Experiment 1.1: Results

Interoceptive Accuracy

First, a series of correlation analyses revealed small and non-significant associations between interoceptive accuracy for either cardiac or respiratory interoceptive domain and any of the control variables (age, BMI, anxiety symptoms, depressive symptoms, mean heartrate, timing accuracy, and memory accuracy). All *p*s were greater than .05 and all *r*s smaller than .30, *without* applying corrections for multiple correlations. More crucially, Table 1 shows that groups were matched for the most of these key variables. Significant between-group differences were found only in anxiety and depressive symptoms, with participants with ASD scoring significantly higher than control participants in STAI and BDI-II, as expected based on the literature regarding comorbidity in ASD (e.g., Simonoff et al., 2008).

As shown in Figure 1, interoceptive accuracy was undiminished among participants with ASD. An independent sample *t*-test revealed no significant between-group differences in interoceptive accuracy, either in cardiac, $t(40) = 0.26$, $p = .80$, $d = 0.08$, $BF_{10} = 0.31$ or in respiratory domain, $t(40) = -0.35$, $p = .73$, $d = 0.11$, $BF_{10} = 0.32$. However, based on these results we cannot exclude the possibility that there were differences in the case of participants with ASD who had clinically significant levels of alexithymia.

In order to examine the “alexithymia hypothesis”, we followed Nicholson et al.’s approach. (2018). That is to say, we divided the ASD group in two sub-samples. Participants who scored above the cut-off point (≥ 61) on TAS-20 were classified in the “high alexithymic” group ($n = 14$), whereas those who scored lower were classified in the “low alexithymic” group (*n* = 8). These sub-groups were equated for VIQ, PIQ, FSIQ, ADOS total score (all *p*s ≥ .35, all *d*s ≤ 0.42), and sex, x^2 (1) = 0.01, *p* = .93. Although, there was a marginally significant between-group difference in age, *t(19)* = 1.77, *p* = .09, *d* = 0.70, with participants in the “low alexithymic” group being younger (*M* = 31.75, *SD* = 5.92), compared to participants in the “high alexithymic” group (*M* = 39.29, *SD* = 13.86), age did not correlate significantly with interoceptive accuracy, *r* = 12, *p* = .44. Therefore, the between-group difference in age cannot provide an explanation for the results reported below. As shown in Table 2, there was no significant difference between ASD participants with clinically significant levels of alexithymia and without, either in cardiac or in respiratory interoceptive accuracy. One thing to note, groups were matched for FIQ, ADOS total score, and BMI (all *p*s ≥ .10).

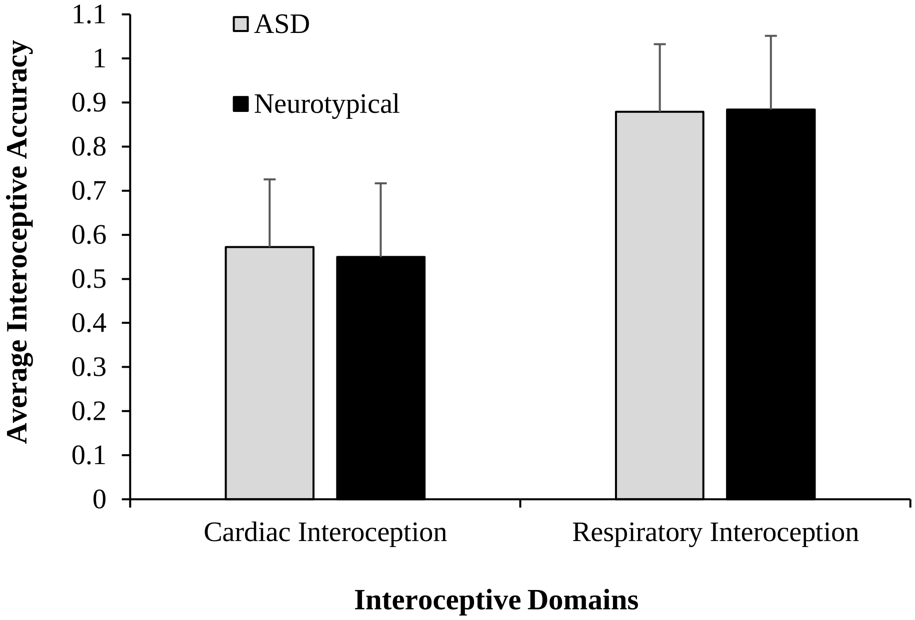

Figure 1. Average accuracy in cardiac and respiratory interoceptive domain among ASD and neurotypical participants. Error bars represent standard error of measurement.

Table 2

Means (SDs) and Inferential Statistics for Differences among "High Alexithymic" and "Low Alexithymic" Group in Cardiac and Respiratory Interoceptive Accuracy

Interoceptive Accuracy	ASD group		Group Differences			
	High Alexithymia	Low Alexithymia	*t*	*p*	*d*	BF_{10}
Cardiac Domain	.57 (.29)	.58 (.16)	-0.19	.85	0.08	0.40
Respiratory Domain	.88 (.04)	.88 (.06)	-0.12	.91	0.06	0.40

Interoceptive Awareness

In both interoceptive domains, the mean correlation coefficient (Pearson *r*) between interoceptive accuracy and JoC was higher among neurotypical participants than ASD participants, yet that difference did not reach the level of statistical significance (see Table 3). Nonetheless, a series of one-sample *t*-tests showed that the mean coefficient for cardiac interoceptive awareness was significantly different from zero in the comparison group, *t*(17) = 1.96, *p* = .03, *d* = 0.46, BF_{10} = 2.21 (one - tailed) whereas it was at chance level in the ASD group *t*(20) = 0.28, *p* = .78, *d* = 0.06, BF_{10} = 0.24. That was also the case within the respiratory interoceptive domain. In the comparison group, awareness was significantly above chance, *t*(17) = 2.38, *p* = .02, *d* = 0.56, BF_{10} = 4.30 (one - tailed), but it remained at chance level in the ASD group *t*(14) = 1.52, *p* = .15, *d* = 0.39, BF_{10} = 0.68.

Table 3

Means (SDs) and Inferential Statistics for Group Differences in Cardiac and Respiratory Interoceptive Awareness

Interoceptive Awareness	Diagnostic Group		Group Differences			
	ASD	Neurotypical	*t*	*p*	*d*	BF_{10}
Cardiac Domain[a]	.03 (.57)	.28 (.60)	-1.30	.10	0.42	1.06
Respiratory Domain[b]	.16 (.41)	.26 (.46)	-0.62	.27	0.22	0.55

Note. Values for one-tailed tests are reported because of a priori directional predictions.
[a] ASD: $n = 21$; Neurotypical: $n = 18$.
[b] ASD: $n = 15$; Neurotypical: $n = 18$.

Association Analyses

A series of correlation analyses was conducted exploring the relations between mindreading abilities and interoceptive awareness. As shown in Figure 2, the mean correlation coefficient, used as index of respiratory interoceptive awareness was positively and significantly associated with performance on MASC among participants with ASD, $r = .56$, $p = .02$, $BF_{10} = 5.32$ (one-tailed), yet it did not reach the level of statistical significance among neurotypical participants, $r = .13$, $p = .30$, $BF_{10} = 0.45$ (one-tailed). However, a Fisher's Z test indicated that there was not a significant difference in the association observed among ASD and comparison participants, $Z = 1.30$, $p = .19$. In cardiac interoceptive domain, the mean correlation coefficient between accuracy and JoC was not significantly associated with performance on MASC either among ASD or comparison participants (all rs $\leq .19$, all ps $\geq .22$, , and all BF_{10}s $\leq .42$, one tailed). In addition, we did not find significant associations between either cardiac or respiratory interoceptive awareness and mindreading abilities, when measured using the RMIE task among ASD or comparison participants (, all rs $\leq .15$, all ps $\geq .27$and all BF_{10}s ≤ 0.47, one tailed).

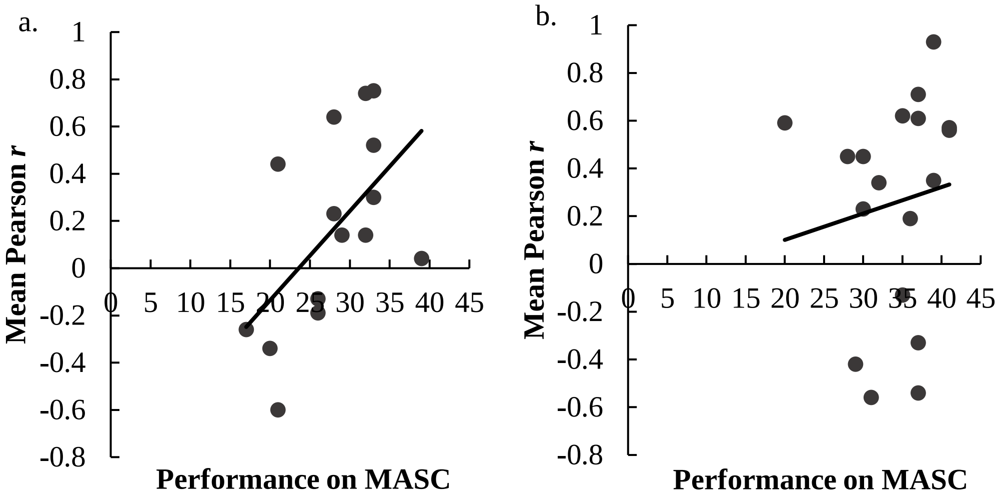

Figure 2. Scatterplot of respiratory interoceptive awareness vs mindreading abilities among (a) ASD and (b) neurotypical participants.

Experiment 1.1: Discussion

To date, no study had examined the extent to which individuals diagnosed with ASD can sense accurately the physiological conditions of their body and provide accurate beliefs about their ability to judge their own internal states, by exploring two different interoceptive domains within a single experiment. As such, our primary aim was to examine interoceptive accuracy and interoceptive awareness of individuals with ASD, in both cardiac and respiratory domains.

In terms of the first central experimental finding, the current experiment did not find any evidence to support the hypothesis that interoceptive accuracy is diminished in ASD. The difference in cardiac interoceptive accuracy between participants with ASD and neurotypical participants was non-significant and very small (*d* = 0.08). In keeping with this finding, the between-group difference in respiratory interoceptive accuracy was also non-significant and very small (*d* = 0.11). In addition, a Bayesian analysis of the between-group difference in cardiac and respiratory interoceptive accuracy indicated that the data provided substantial evidence in favour of the null hypothesis for both interoceptive domains. Together, these findings suggest that interoceptive accuracy is intact in ASD.

Moreover, these results are consistent with prior studies that indicated undiminished cardiac interoceptive accuracy among individuals with ASD (Nicholson et al., 2018; Shah et al., 2016) and no significant relation between autism traits and either cardiac or respiratory interoceptive accuracy (Murphy et al., 2018a; Nicholson et al., 2018).

One point to note here is that a series of extraneous factors has been suggested to have a confound effect on performance on the heartbeat tracking task. For example, high body mass index (BMI) and high levels of anxiety and depression have been found to adversely affect cardiac interoceptive accuracy (De Pascalis, Alberti, & Pandolfo, 1984; Garfinkel et al., 2016b; Herbert et al., 2014; Murphy et al., 2018c; Rouse et al., 1988). Crucially in our study, participants with ASD were closely matched with neurotypical participants in terms of their body mass index, age, and intelligence, and thus we overcame a common limitation in previous research (Garfinkel et al., 2016b; Nicholson et al., 2018). In addition, we measured participants' level of depression and anxiety. As expected, participants with ASD group showed significantly higher levels of depression and anxiety, compared to neurotypical participants (McManus, Bebbington, Jenkins, & Brugha, 2016; Simonoff et al., 2008). Nonetheless, their relation to interoceptive accuracy was non-significant. Additionally, it has been found that the ability of accurate time estimation can increase performance on the heartbeat tracking task (Ainley et al., 2014; Shah et al., 2016). Using the same line of reasoning, it could be argued that performance on the respiratory interoception task can be influenced by short-term memory for internal effort. Nevertheless, in the current study groups had equivalent performance in both control tasks (time estimation/ memory). Overall, we can argue that none of these extraneous factors had a confound effect on the current findings, providing reassurance that we did not erroneously conclude that interoceptive accuracy is undiminished in ASD.

From a theoretical perspective, the current experiment challenges the idea that mindreading deficits in ASD stem from impaired interoceptive accuracy, due to a damaged oxytocin system (Quatrocki & Friston, 2014). As expected, we found a large and significant ASD-specific impairment in mindreading, as measured using the MASC (alongside a marginally significant and moderate between-group difference in performance on the RMIE). These results are in keeping with previous research findings (e.g. Baron-Cohen et al., 2001a; Dziobek et al., 2006; Yirmiya et al., 1998) and indicate that participants with ASD who took part in the current experiment had significant mindreading deficits. Based on that, and if Quatrocki and Friston's (2014) hypothesis was correct we should have found impaired interoceptive accuracy among participants with ASD. Nevertheless, this was not the case in the current experiment.

Moreover, our findings provide a significant challenge to the "alexithymia hypothesis" (e.g. Bird & Cook, 2013; Bird et al., 2010; Brewer et al., 2015; Cook et al., 2013). According to this hypothesis, individuals with ASD have interoceptive deficits only when alexithymia co-occurs with ASD. This theoretical account has been partly built upon the speculation that a significant difference in interoceptive accuracy between individuals with ASD and neurotypical individuals would be apparent if individuals with ASD had elevated levels of alexithymia (Shah et al., 2016). In the current experiment, we examined the extent to which this speculation can be confirmed by research evidence, adopting Nicholson et al.'s (2018) approach. That is to say, we directly contrasted an ASD "high alexythimic" group with an ASD "low alexythimic" group in terms of their cardiac and respiratory interoceptive accuracy. The between-group differences in both interoceptive domains were non-significant and very small (Cardiac: d = 0.08; Respiratory d = 0.06), with the Bayes factors indicating that the data provided anecdotal evidence for the null hypotheses.

Overall, we can argue that the findings of the current experiment are more consistent with the theoretical claim that individuals with ASD have intact physical self-awareness because their ability to make first order representations is undiminished (Lind, 2010; Nicholson et al., 2018; Uddin, 2011; Williams, 2010). However, this intact ability speaks only for the accurate detection of physiological body sensations, and by no means implies their correct interpretation (Nicholson et al., 2018).

In terms of the second central experimental finding, the current experiment found moderate evidence to support the hypothesis that interoceptive awareness is diminished in ASD. Although neurotypical participants had higher interoceptive awareness, compared to participants with ASD, the between-group difference was marginally significant and small ($d = 0.42$) within the cardiac interoceptive domain and non-significant and small ($d = 0.22$) within the respiratory domain. According to Bayesian analyses, the data provided just anecdotal evidence in favour of the null hypothesis for the respiratory domain. Nonetheless, we crucially found that beliefs of participants with ASD about their ability to judge their own internal states were random. In both interoceptive domains, awareness was at chance level, among participants with ASD but significantly above chance among neurotypical participants. Given that the study of interoceptive awareness in ASD is almost non-existent, the current findings provide the first evidence that the ability to represent first order representations of body sensations upon cognition appears to be diminished in ASD. This is important because this is the mechanism that has been proposed to be involved in the correct interpretation of physical internal states (Nicholson et al., 2018).

In terms of associations, this was the first study that examined the relation between interoceptive awareness and mindreading abilities. As predicted, we found a significant and large ($r = .56$) association between performance on MASC and respiratory interoceptive awareness in the ASD group. Despite not replicating this finding among neurotypical participants, we found that the size of the shared variance between performance on MASC and respiratory interoceptive awareness was the same in both groups, indicating that metarepresentational abilities provide a link between mindreading and interoceptive awareness (e.g. Carruthers, 2009). Non-significant associations were found when mindreading abilities were measured with the RMIE task, as well as within the cardiac interoceptive domain.

In sum, the current experiment provides further evidence of undiminished interoceptive accuracy in ASD, but also of diminished interoceptive awareness in adults with ASD. As such, we could argue that lower order phenomena of self-awareness are intact in ASD, but higher order phenomena appear to be impaired. Nonetheless, in order to increase our confidence in the last assumption, in the next chapter we also explored metacognition in ASD, conducting two experiments and using one of the most widely used indices of metacognitive monitoring abilities. That is judgements of confidence accuracy.

2. Exploring Implicit and Explicit Judgements of Confidence in ASD

Much of the evidence regarding metacognitive monitoring in ASD has emerged from studies that explored judgments of confidence accuracy (JoC) among children with ASD. Nonetheless, strong conclusions cannot be drawn from these studies, given that their results have been to a great extent contradictory (Grainger, Williams, & Lind, 2016; McMahon, Henderson, Newell, Jaime, & Mundy, 2016; Wilkinson, Best, Minshew, & Strausset, 2010; Williams et al., 2016; Wojcik, Allen, Brown, & Souchay, 2011). Having said that, when metacognitive monitoring abilities were examined in the real-world setting of mathematics learning, students with ASD showed clear deficits in their ability to judge their performance and discrepancy between pre-test and post-test intentions about their performance, indicating diminished metacognitive monitoring abilities (Brosnan et al., 2016). Nonetheless, to date, only three studies have investigated JoC accuracy in the population of adults with ASD and one study has explored the relation between autism traits and accuracy of JoC in the general population.

Williams, Bergstrom, and Grainger (2016) explored the associations among autism traits, mindreading abilities and explicit JoC accuracy among individuals from the general population. Despite finding a significant association between mindreading and explicit JoC accuracy, indicating that both rely on metarepresentational abilities, the association between autism traits and JoC accuracy was non-significant. As such, Williams et al. (2016) concluded that people's ability to impute mental states to themselves is unrelated their autism traits. Nonetheless, this did not exclude the possibility that individuals with ASD would show diminished JoC accuracy.

Wilkinson et al. (2010) employed an explicit JoC task, following a facial recognition task among 16 adults with ASD and 15 neurotypical adults matched for age and IQ. Participants' judgments of confidence were elicited using a 3-point scale, ranging from "certain" to "guessing". Wilkinson et al. (2010) found that whilst JoC accuracy in the

ASD group was not significantly different between “certain” and “somewhat certain” judgements, there was a significant difference in the comparison group. Furthermore, in the ASD group JoC accuracy was 72% for “certain” judgements, whereas the same percentage was 85% for the comparisons. That difference may not reach the level of statistical significance, but it was moderate in size (d = 0.53). This evidence led Wilkinson et al. (2010) to conclude that subtle metacognitive monitoring atypicalities are present in ASD. Nonetheless, we cannot draw firm conclusions from this study because it is well-established that individuals with ASD have diminished face processing abilities (Riby, Doherty-Sneddon, & Bruce, 2008; Riby & Hancock, 2008), and thus diminished object-level performance in the facial recognition task could have a confound effect in JoC accuracy among the ASD group (Dunlosky & Metcalfe, 2009).

Sawyer, Williamson, & Young (2014) employed an explicit JoC task and two different tasks (emotion recognition & general knowledge task) to measure object-level performance, among 30 adults with ASD and 52 neurotypical participants, matched for gender, verbal IQ, and full scale IQ. Sawyer et al. (2014) did not find significant between-group differences in JoC accuracy, when ASD participants showed diminished performance in the emotion recognition task. Whereas, people with ASD showed diminished JoC accuracy compared to the comparison group, with the between-group difference being marginally significant (*p* = .06) when groups were matched for object-level performance in the general knowledge task.

Cooper, Plaisted-Grand, Baron-Cohen, and Simons (2016) employed an explicit JoC to measure metamemory, among 24 adults with ASD and 24 neurotypical participants, matched for age, verbal and non-verbal ability, and phonological and semantic fluency. They found that metamemory was diminished in the ASD group. Noteworthy, groups were also equated for their object-level performance.

Whilst absolute conclusions cannot be drawn from the studies described above, it could be argued that neither of them found large between-group differences that could be considered clinically significant. One possible explanation might be that explicit JoC tasks are not sensitive to detect metacognitive atypicalities, among high-functioning adults with ASD. This is plausible, given it has already been found that there are people with ASD who perform equivalently with neurotypical people in explicit mindreading tests. Nonetheless, their significantly diminished performance in implicit mindreading tests indicates that this achievement is more of the product of compensatory learning, rather than the effect of an intact inherent mindreading ability (Callenmark, Kjellin, Rönnqvist, & Bölte, 2014; Senju, 2012; Senju et al., 2009).

As such, an implicit (non-verbal) measure might be more informative about metacognitive monitoring abilities in ASD. Implicit JoC tasks have already been used in comparative psychology in order to explore metacognitive monitoring among non-human primates (Beran, Smith, Coutinho, Couchman, & Boomer, 2009; Couchman, Coutinho, Beran, & Smith, 2010; Kornell, Son, & Terrace, 2007; Smith, Shields, & Washburn, 2003, Smith & Washburn, 2005; Washburn, Gulledge, Beran, & Smith, 2010) as well as in neurotypical children (Paulus, Proust, & Sodian, 2013). Nonetheless, to date accuracy of implicit JoC has never been examined in ASD.

Experiment 2.1: Individual Differences

Experiment 2.1 addresses this issue investigating the extent to which metacognitive monitoring abilities, as indexed by both explicit and implicit JoC accuracy scores, are associated with mindreading abilities and autism traits in the general population. In the current experiment, participants performed either the implicit or the explicit version of the gambling paradigm described below (Son & Kornell, 2005). All of them completed the AQ (Baron-Cohen et al., 2001b; see p. 24 of the current project for a detailed description), the RMIE task (Baron-Cohen et al., 2001a; see p. 23 of the current project for a detailed description), and the theory of mind clips (ToM) of the animations task (Abell, Happé & Frith, 2000), with the last two assessing mindreading abilities. In addition, we measured individual differences in risk aversion, with a view to controlling this in correlation analyses because based on Dienes and Seth's findings (2009), it would be reasonable to suggest that any significant correlation might have been confounded by the degree of risk aversion manifested by an individual.

The first aim of the current experiment was to examine the extent to which metacognitive monitoring is associated with mindreading abilities. Based on the theoretical claim that mindreading and metacognition are under the umbrella of the same metarepresentational faculty (Caruthers, 2009; Gopnik, 1993; Gopnik & Meltzoff, 1994; Hobson, 1990), and as both versions of the gambling paradigm tap metacognition (Son and Kornell, 2005; Kornell, Son, & Terrace, 2007), we expected to find a positive and significant association between mindreading and both explicit and implicit JoC accuracy.

The second aim of the current experiment was to examine the extent to which metacognitive monitoring is associated with the number of ASD traits reported by individuals from the general population. Based on the claims of "one-system" theorists that metarepresentational abilities underlie both mindreading and metacognitive monitoring (e.g. Caruthers, 2009), and as it has been found that mindreading abilities are negatively associated with autism traits (e.g. Baron-Cohen, et al., 2001b; Williams et al., 2016), we predicted that autism traits would be negatively and significantly associated with both implicit and explicit JoC accuracy.

Experiment 2.1: Method

Participants

Fifty-six undergraduate students (48 female) participated in this experiment, after they had given written, informed consent. Their average age was 19.46 (*SD* = 2.83; range = 18 to 39) years. All participants had normal or corrected to normal vision and none of them was colour blind. In addition, all were native English speakers, or spoke English to native-level proficiency, and none had a history of ASD, language impairment, or dyslexia, according to self-report. Participants were recruited via the Research Participation Scheme (RPS) of the UKC and they were offered course credits in partial fulfilment of their degree, for taking part in the experiment. Additionally, they received a performance-related monetary prize. Ethical approval for this study was obtained from the Kent School Psychology Research Ethics Committee (Ethics ID: 201715129893444795).

Materials, Procedures, and Scoring

As in Experiment 1.1, participants completed the RMIE task (Baron-Cohen et al., 2001a) and the AQ (Baron-Cohen et al., 2001b). We used a between-subjects design for the administration of the gambling task, in order to deter any possible confound effects from participants' exposure to both versions (Charness, Gneezy, Michael, & Kuhnc, 2012). Thus, 28 participants completed the explicit version, and the rest of them completed the implicit one (n = 28). There were no significant between-group differences either in age, $t(54) = 0.94$, $p = .35$, $d = 0.25$, $BF_{10} = 0.39$ or in sex ratio, $x^2 (1) = 2.33$, $p = .13$, $BF_{10} = 1.29$.

Gambling task.

The gambling task is a modified version of the gambling paradigm designed by Son and Kornell (2005). First, participants were asked to make a psychophysical discrimination within a period of 3s, touching either the most densely pixelated of two boxes in one condition or the deepest blue of two boxes in the other one. The order of conditions was counterbalanced across participants. Following each discrimination, participants responded to different stimuli within 3s either as part of the implicit version of the gambling task (Son & Kornell, 2005) or as part of the explicit one. All stimuli were presented on a 22-in computer screen (see Figure 3 for a graphical representation of the task).

Participants were instructed that their goal would be to accumulate as many points as possible, over a 10-minute test period. At the end of the experiment, these points were converted into money. Both versions comprised 60 trials and had the same payoff structure. The "payment rules" were explained to participants at the beginning of the testing session, prior to a set of 10 warmup trials. At the end of the task, all participants completed an oral memory test on the payment rules.

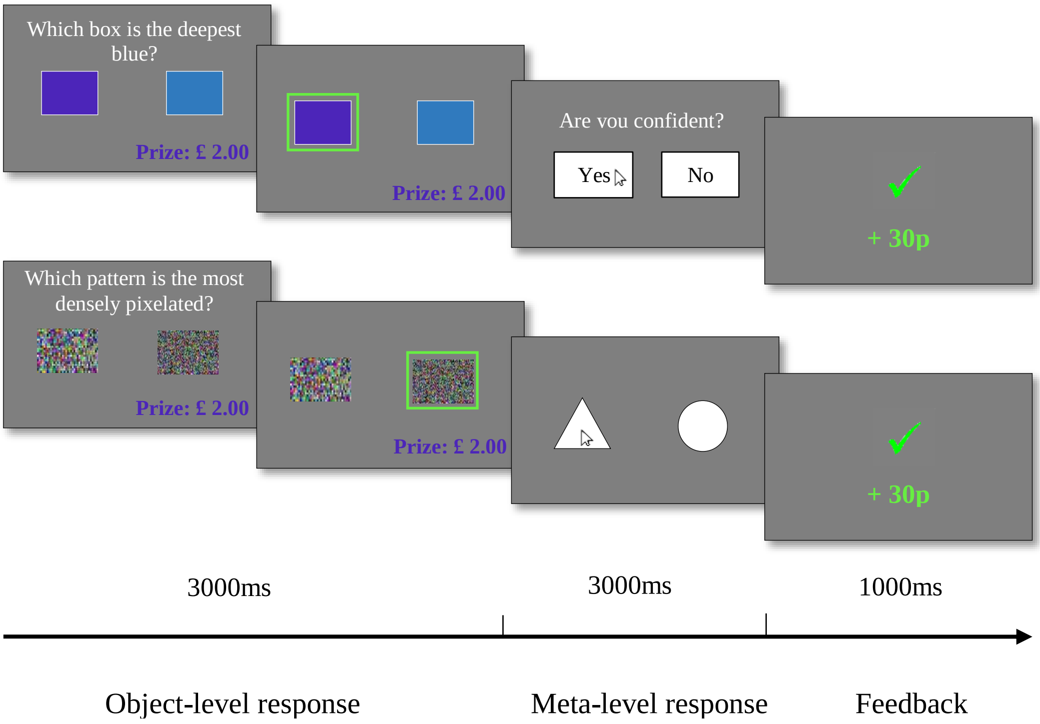

Figure 3. Graphical representation of the procedure and stimuli used in the explicit-colours (top) and implicit-pixels (bottom) version of the gambling task.

Implicit version.

In the implicit version of the gambling task, the follow-up screen after each psychophysical discrimination choice displayed two shapes. That is a triangle and a circle. Clicking the triangle that represented the "high risk" option, participants would receive 30p if their previous discrimination were correct, but would lose 30p if it were incorrect. Clicking the circle that represented the “low risk” option, participants would gain just 10p for a previous correct discrimination, but would lose 10p for an incorrect one. When instructing participants, great care was taken not to use metacognitive language.

Explicit version.

In the explicit version of the gambling task, the follow-up screen after each psychophysical discrimination choice presented the following question: “Are you confident?”. Participants stated their judgment of confidence (JoC) in each of their previously given responses, by choosing one of the following binary items: “Yes” - “No”. Clicking the “Yes” option would give 30p if their previous discrimination were correct but would lose 30p if it were incorrect. Clicking the “No” option would give 10p for a previous correct discrimination but would lose 10p for an incorrect one.

Scoring.

In both versions, we calculated JoC accuracy scores using gamma correlations (Goodman & Kruskal, 1954). This measure has been recommended by Nelson (1984), and Nelson, Narens, and Dunlosky (2004) and has been extensively used in the study of metacognitive monitoring (e.g. Grainger et al., 2016; Sawyer, Williamsonn, & Young, 2014; Williams et al., 2016). Gamma scores were used to measure the degree of association between object-level task performance and either explicit confident judgements or implicit ones. Gamma correlations take values between -1 and +1. Large positive scores indicate a good correspondence between the two variables, while a gamma of zero indicates no association. We calculated one gamma score for each participant using the formula (ad - bc)/(ad + bc), with “ab” representing concordant pairs (correct discriminations-high confidence/risk and incorrect discriminations-low confidence/risk) and “bc” discordant pairs (correct discriminations-low confidence/risk and incorrect discriminations-high confidence/risk). Gamma scores cannot be calculated in cases when at least one of the pairs equals to zero. Therefore, following Snodgrass and Corwin's (1988) recommendation, we corrected the raw data by adding 0.5 to each frequency. Furthermore, object-level performance was calculated as the proportion of correct psychophysical discriminations participants made in the first phase of each trial.

Animations task.

The animations task is based on Heider and Simmel's (1944) animated triangles and is widely used in research to trigger the mindreading process and elicit individuals' statements about mental states, actions, and interactions (e.g. Bird, Castelli, Malik, Frith, & Husain, 2004). The task originally comprises three conditions. That is the goal-directed, the random, and the ToM one. All feature two triangles moving around and interacting in different ways (Abell et al., 2000, Castelli, Frith, Happé, & Frith, 2002; Castelli, Happé, Frith, & Frith, 2000). However, only the four ToM clips were used for the purposes of the current experiment. These clips feature a triangle responding to the mental states of another triangle. Each clip lasted between 34s and 45s and was presented twice on a computer screen. The first time participants were instructed to watch the clip silently, but the second time they were asked to describe what was happening in each clip while they were watching it. Their descriptions were audio-recorded, transcribed, and uploaded to a secure server of the University of Kent. We scored participants' responses on each clip using a scale from 0 to 2, following Abell Happe, & Frith's (2000) criteria. A ToM score was calculated for each participant by summing scores on each clip.

Test of risk aversion.

A modified version of the Arrow-Pratt measure was used to assess individual differences in aversion to risk (Dienes & Seth, 2009). Participants were asked to state the reservation price for a lottery ticket in two hypothetical scenarios (Dienes & Seth, 2009). The first trial included the following scenario: "Imagine there was a lottery for a £10 prize, which will be given to one of the ten ticket holders. How much would you pay for a ticket?”. The second trial included the same scenario with the only difference that the prize was £100. An individual score of risk aversion was calculated for each participant by summing their two numerical answers. The higher the amount of money a participant was willing to pay the less their aversion to risk. Thus, high scores indicate low risk aversion. Following Dienes and Seth’s (2009) scoring, a total score > 11 indicates no risk aversion.[4]

Experiment 2.1: Results

Between-Group Differences

Table 4 shows means (*SD*) for performance on the explicit and implicit version of the gambling task. In both versions, the mean gamma score was significantly above chance, indicating accurate explicit and implicit judgements of confidence, $t(27) = 16.90$, $p < .001$, $d = 3.19$, $BF_{10} = 8.09$ and $t(27) = 8.10$, $p < .001$, $d = 1.53$, $BF_{10} = 1.25$ respectively.

Moreover, Table 4 indicates that groups can be considered equivalent when exploring correlations. Neither object-level performance, $t(54) = 0.38$, $p = .70$, $d = 0.10$, $BF_{10} = 0.29$ nor response bias, $t(54) = 1.54$, $p = .13$, $d = 0.42$, $BF_{10} = 0.72$ differ significantly between-groups. Groups were also equated for the proportion of “payment rules” that were recalled, $t(54) = -0.72$, $p = .48$, $d = 0.19$, $BF_{10} = 0.33$, and as shown in Table 5, for their autism traits, mindreading abilities, and aversion to risk.

Table 4

Descriptive Statistics for Performance on the Explicit and Implicit Version of the Gambling Task

Variable	Explicit	Implicit
	Mean (SD)	
Proportion of correct discriminations	.74 (.08)	.73 (.11)
Proportion of high confidence/risk responses	.75 (.15)	.68 (.18)
Gamma score	.67 (.21)	.53 (.35)
Proportion of “payment rules” recalled	.88 (.26)	.92 (.20)

Table 5

Descriptive Statistics for Major Study Variables and Results from t-Tests

Variable	Explicit	Implicit	Results			
	Mean (SD)		*t*	*p*	*d*	BF_{10}
AQ	16.64 (7.01)	14.86 (6.83)	0.97	.34	0.26	0.40
RMIE	26.18 (4.80)	26.14 (4.87)	0.03	.98	0.01	0.27
Animations	5.82 (1.57)	5.25 (2.10)	1.17	.25	0.31	0.54
Risk Aversion	10.70 (10.47)	9.93 (6.58)	0.33	.74	0.09	0.28

Note. AQ: Autism-Spectrum Quotient; RMIE: Reading the Mind in the Eyes; Animations: ToM animation clips.

Association Analyses

A series of correlation analyses was conducted separately for the explicit and implicit version of the gambling task (see Table 6 and Table 7 respectively) exploring the relations among JoC accuracy, mindreading abilities and autism traits. Our sample was representative of the general population with respect to autism traits and mindreading abilities (see Appendix B). Given that the shared variance between risk aversion and explicit as well as implicit JoC accuracy was negligible, there was no need to control for the effect of risk aversion on any of the correlation analyses reported below.

Contrary to our predictions, the association between either the implicit or the explicit JoC accuracy and performance in mindreading tasks (RMIE and ToM clips) was small and non-significant. However, a Fisher's *Z* test revealed that neither the explicit JoC gamma × RMIE correlation coefficient nor the implicit JoC gamma × RMIE one differ significantly from the correlation coefficient ($r = .25$) reported by Williams et al. (2016) for a significant association between JoC gamma and performance on the RMIE task, $Z = -0.18$, $p = .85$ and $Z = -0.28$, $p = .78$, respectively. Next, we explored the association between autism traits and JoC accuracy. Even though there was a significant correlation between AQ and RMIE (see Table 6), neither the explicit nor the implicit JoC accuracy correlated significantly with AQ.

Table 6

Bivariate correlations in the Explicit Version of the Gambling Task

	1	2	3	4	5	6	7
1. Gamma: Explicit	–	-.21[b]	.35[c]	.15[a]	.04[a]	.21[a]	-.08[a]
2. Proportion of high confidence responses		–	-.44*[d]	-.17[a]	.18[b]	-.04[a]	.46*[d]
3. Object-level performance			–	-.13[a]	.01[a]	.44*[d]	-.32[b]
4. AQ total				–	.12[a]	-.44*[d]	.03[a]
5. Risk Aversion total					–	-.23[b]	-.22[b]
6. RMIE						–	-.04[a]
7. Animations							–

Note. $N = 28$.
$*p < .05$.
[a]BF_{10} < 0.33 (supports the null hypothesis); [b]BF_{10} = 0.34-0.99 (anecdotal evidence for the null hypothesis); [c]BF_{10} = 1-2.99 (anecdotal evidence for the alternative hypothesis); [d] BF_{10} = 3-99 (supports the alternative hypothesis).

Table 7

Bivariate correlations in the Implicit Version of the Gambling Task

	1	2	3	4	5	6	7
1. Gamma: Implicit	–	.31[b]	.31[b]	-.01[a]	.01[a]	.19[a]	-.05[a]
2. Proportion of high risk responses		–	.11[a]	-.02[a]	.25[b]	-.15[a]	.02[a]
3. Object-level performance			–	.22[b]	.01[a]	.16[a]	.19[b]
4. AQ total				–	.17[b]	-.19[b]	-.08[a]
5. Risk Aversion total					–	.07[a]	-.25[b]
6. RMIE						–	-.02[a]
7. Animations							–

Note. $N = 28$.
[a]BF_{10} < 0.33 (supports the null hypothesis); [b]BF_{10} = 0.34-0.99 (anecdotal evidence for the null hypothesis).

Experiment 2.1: Discussion

In contrast with our predictions, we did not find a significant association between mindreading and metacognitive monitoring abilities, as indexed by performance in both versions of the gambling task. Nonetheless, the size of the association between performance on RMIE task and performance on the explicit version of the gambling task was almost identical to that reported by Williams et al. (2016) who found a significant association between explicit JoC accuracy and RMIE. That was also the case for the size of the association between performance on RMIE task and performance on the implicit version of the gambling task. Given that Williams et al.'s (2016) sample size (N=83) was almost three times larger than the sample size of the current experiment, low power could provide an explanation for the non-significant results.

Despite predicting a significant association between autism traits and metacognitive monitoring abilities neither the relation between explicit gamma and AQ was significant nor the relation between implicit gamma and AQ. According to Bayesian correlation analysis, our data provided anecdotal evidence in favour of the null hypothesis for both associations. When we examined the relation between mindreading abilities and autism traits, we found the typical result. That is to say, the association between AQ and RMIE was moderate and significant among participants who completed the explicit version of the gambling task. This finding comes to add to prior research evidence that poor mindreading abilities are associated with higher autism traits in the general population (e.g. Baron-Cohen, 2001b, Williams et al., 2016). One thing to note here is that performance on the ToM clips of the animation task did not correlate either with RMIE or with AQ. A closer examination of participant performance reveals that their mean score *(M* = 5.54, *SD* = 1.84) was significantly lower compared to that (*M* = 7) reported by Abell et al. (2000) for neurotypical adults, *t* (55) = -5.96, *p* <.001. Given that 12.5% of our participants were not native English speakers, linguistic difficulties may have confounded their performance leading to non-significant results in terms of the correlations. Nonetheless, based on the link between AQ and RMIE it can be argued that individuals with high levels of autism traits have difficulties in imputing mental states to others but their ability to attribute mental states to themselves is unrelated to their autism traits (Williams et al., 2016).

However, the extent to which this is also the case among people diagnosed with ASD remains unclear until is directly explored. Previous research has indicated that a non-significant association between autism traits and metacognitive monitoring abilities does not exclude the possibility that there is a significant difference between people with ASD and neurotypical people (Williams et al., 2016). This issue was addressed in Experiment 2.2.

Experiment 2.2: Case-Control

In Experiment 2.2, we examined metacognitive monitoring abilities among individuals with a full diagnosis of ASD and neurotypical people, employing both versions of the gambling task described above (Son & Kornell, 2005). In addition, participants competed the RMIE (Baron-Cohen et al., 2001a) and the theory of mind clips (ToM) of the animations task (Abell et al., 2000). As it is well-established that individuals with ASD have attenuated mindreading abilities (e.g. Baron-Cohen et al., 2001; Senju et al., 2009; Yirmiya et al., 1998), we first predicted that participants with ASD would show diminished performance in both mindreading tasks, compared to comparison participants. Second, in keeping with predictions that stem from "one-system" theory (e.g. Carruthers, 2009), we expected significantly lower explicit and implicit gamma scores among participants with ASD, compared to neurotypical participants, indicating poor metacognitive monitoring abilities in ASD. Please note that the data reported in this experiment was collected as part of an Economic and Social Research Council grant awarded to my MSc supervisor, David Williams. Thus, the data were not collected by me, but I have analysed them independently.

Experiment 2.2: Method

Participants

Twenty-three adults with ASD and 23 neurotypical adults participated in the current experiment, after they had given written informed consent. All participants had normal or corrected to normal vision and none of them was colour blind. Participant characteristics and between-group differences are presented in Table 8. Groups were closely matched for verbal IQ, performance IQ, and full scale IQ, using the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999). They were also matched for chronological age and gender.

All ASD participants had received formal diagnoses, according to established criteria (DSM-IV-TR, American Psychiatric Association 2000; ICD-10, World Health Organisation 1993). Meanwhile, their ASD severity was assessed using the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000). Four participants scored below the ASD cut-off point (≥ 7), based on Lord et al.'s criteria (1989), but remained in the sample because results did not change substantially when excluded.[5] In addition, five participants scored below the Autism Quotient cut-off score (≥ 26), but all of them scored above the ASD cut-off on ADOS and therefore were not excluded from the analyses. Ethical approval for this study was obtained from the Kent School Psychology Research Ethics Committee.

Materials, Procedures, and Scoring

As in Experiment 2.1, all participants completed the RMIE (Baron-Cohen et al., 2001a), the ToM clips of the animations task (Abell et al., 2000), the AQ (Baron-Cohen et al., 2001b) and both versions of the gambling task. Materials, procedure and scoring remained identical across experiments, with the only difference that in the current experiment a within-subjects design was used. The implicit version of task was always presented and competed before the explicit one.[6]

Table 8

Baseline Characteristics and Matching Statistics for Experiment 2.2

	Diagnostic group		Group Differences			
	ASD (*n*=23; 17 male)	Neurotypical (*n*= 23; 17 male)	*t*	*p*	*d*	BF_{10}
CA: years	37.54 (12.62)	37.94 (12.55)	-0.40	.92	0.03	0.29
VIQ	104.65 (12.62)	107.22 (9.43)	-0.78	.44	0.23	0.37
PIQ	102.35 (19.68)	105.13 (11.71)	-0.58	.56	0.17	0.34
FSIQ	103.52 (15.65)	106.61 (9.89)	-0.80	.43	0.24	0.38
ADOS[a]	9.73 (4.58)	-	-	-	-	-
AQ total	32.00 (8.14)	15.70 (5.62)	7.90	<.001	2.33	2.57
Animations[b]	3.91 (2.00)	5.13 (1.87)	-2.14	.02	0.63	3.44
RMIE[b]	24.26 (5.76)	27.70 (3.72)	-2.40	.01	0.71	5.52

Note. CA = chronological age at testing WASI; VIQ = verbal IQ; PIQ = performance IQ; FSIQ = full scale IQ; ADOS: Autism Diagnostic Observation Schedule; AQ total = Total score on Autism Spectrum Quotient; RMIE: Reading the Mind in the Eyes; Animations: ToM animations clips.
[a] ADOS data is missing for one ASD participant.
[b] Values for one-tailed tests are reported because of a priori directional predictions.

Experiment 2.2: Results

As shown in Table 8, an independent sample *t*-test revealed significant between-group differences in both RMIE task and ToM clips, with participants with ASD showing diminished mindreading abilities.

Table 9 shows descriptive statistics and results from a series of *t*-tests for variables associated with performance on the implicit version of the gambling task. In line with our predictions, an independent sample *t*-test revealed that participants with ASD showed diminished implicit gamma scores, compared to the control group.

Noteworthy, as shown in Table 9 there was no significant between-group differences in the proportion of visual stimuli correctly discriminated, indicating that groups were matched for object-level task performance. Groups were also equated for response bias as well as for the proportion of "payment rules" that were correctly recalled after the end of the experiment. Thus, the significant between-group difference in implicit gamma scores is unlikely to be explained by the effect of these variables.

Table 9

Means (SDs) and Inferential Statistics for Group Differences in Performance in the Implicit Version of the Gambling Task

Variable	Group		Group Differences			
	ASD	Neurotypical	*t*	*p*	*d*	BF_{10}
Proportion of correct discriminations	.68 (.10)	.71 (.10)	-1.1	.28	0.30	0.48
Proportion of high risk responses	.74 (.27)	.77 (.20)	-0.38	.71	0.13	0.31
Gamma score[a]	.36 (.40)	.58 (.29)	-2.09	.02	0.63	3.19
Proportion of “payments rules” recalled [b]	.93 (.23)	.97 (.11)	-0.78	.44	0.22	0.38

[a] Values for one-tailed tests are reported because of a priori directional predictions.

[b] Data is missing for two ASD participants.

Table 10

Means (SDs) and Inferential Statistics for Group Differences in Performance on the Explicit Version of the Gambling Task

Variable	Group		Group Differences			
	ASD	Neurotypical	*t*	*p*	*d*	BF_{10}
Proportion of correct discriminations	.70 (.09)	.72 (.09)	-0.50	.62	0.22	0.32
Proportion of high confidence responses	.77 (.20)	.76 (.18)	0.19	.85	0.10	0.30
Gamma score[a]	.60 (.32)	.63 (.27)	-0.24	.41	0.10	0.35
Proportion of “payment rules” recalled[b]	.92 (.16)	.98 (.07)	-1.58	.13	0.48	0.86

[a] Values for one-tailed tests are reported because of a priori directional predictions.

[b] Data is missing for two ASD participants.

Table 10 shows descriptive statistics and results from a series of *t*-tests for variables associated with performance on the explicit version of the gambling task. Contrary to our predictions, an independent sample *t*-test revealed no significant between-group differences in explicit gamma scores. Noteworthy, groups were matched for object-level task performance, response bias and for the proportion of "payment rules" that were recalled after the end of the experiment (see Table 10). Thus, our unexpected finding is unlikely to be explained by the effect of these variables.

Experiment 2.2: Discussion

In keeping with previous research findings, mindreading abilities were found to be diminished among this sample of adults with ASD (e.g. Yirmiya et al., 1998), indicating that our sample was representation of the ASD population. The between-group difference in performance on RMIE task was significant and moderate (*d* = 0.71), as it was also the between-group difference (*d* = 0.63) in scores on ToM clips.

In terms of the difference between participants with ASD and neurotypical participants in metacognitive monitoring abilities as expected, the former showed diminished implicit gamma scores. The between-group difference in the size of these scores was significant and moderate (*d* = 0.63). Surprisingly though, when we assessed JoC accuracy using the explicit version of the gambling task, participants with ASD performed equivalently with neurotypical participants. The between-group difference was non-significant and very small (*d* = 0.10), but with the Bayesian analysis indicating that the data provided only anecdotal evidence in favour of the null hypothesis.

2. General Discussion

In the current chapter, we investigated explicit as well as implicit metacognitive monitoring abilities in ASD. In Experiment 2.1, despite not finding a significant association between mindreading abilities and metacognitive monitoring, as measured either with the explicit or with the implicit version of the gambling task, the size of the shared variance between performance on RMIE and either the explicit or implicit JoC accuracy did not differ significantly from that reported by Williams et al. (2016). This is important because Williams et al. (2016) had a much larger sample than that tested here and thus more statistical power to detect a significant effect (*N* = 83). It is also worth mentioning that to date no study had explored the association between implicit JoC accuracy and mindreading abilities. Although further research needs to be done, our results provide tentative evidence that implicit ways of measuring metacognitive monitoring are metarepresenational and thus can be reliably used, as indices of higher order phenomena of self-awareness (Fleming & Dolan, 2010; Koch & Preuschoff, 2007; Persaud & McLeod, 2008; Persaud, McLeod, & Cowey, 2007).

In contrast with our predictions, we also did not find a significant association between either the explicit or the implicit JoC accuracy and the number of autism traits, as measured in the general population. These findings were keeping with previous evidence that the ability of neurotypical individuals to imputing mental states to themselves is unrelated to their autism traits (Williams et al., 2016). Nonetheless, this did not seem to be the case among individuals with ASD.

In Experiment 2.2, we found diminished gamma scores among individuals diagnosed with ASD, when we employed the implicit version of the gambling task. This finding was in line with our predictions and consistent with “one-system” theorists’ claims that severe deficits in metacognition lie in the core of ASD because mindreading and metacognition share the same underling metarepresentational mechanism (Carruthers, 2009; Gopnik, 1993; Gopnik & Meltzoff, 1994; Hobson, 1990). Despite that, unexpected results arose when we examined metacognitive accuracy employing the explicit version of the gambling task. That is to say, participants with ASD showed undiminished abilities.

One thing to note here is that in Experiment 2.2 ASD participants were closely matched with neurotypical participants in terms of verbal IQ, performance IQ, and full scale IQ. Moreover, we overcame a common limitation of previous studies of metacognition in ASD (Sawyer et al., 2014; Wilkinson et al., 2010), by equating groups for their object-level performance. As such, the unexpected results described above cannot be explained by between-group differences in these extraneous factors. Compensation though might be a more plausible explanation (Livingston & Happé, 2017).

Based on the "compensation hypothesis", we could speculate that participants with ASD might have used a compensatory strategy in order to achieve explicit JoC accuracy equivalent with neurotypical participants. To date, clear evidence for compensation in ASD has been found only in the study of mindreading abilities (Schneider, Slaughter, Bayliss, & Dux, 2013; Senju et al., 2009). With respect to metacognition, even though compensation has been proposed in prior studies as an explanation for non-significant differences between ASD and neurotypical people (Grainger et al., 2016), there were no findings to support this hypothesis.

Crucially, this is the first study that contrasted explicit with implicit metacognitive monitoring abilities in ASD and found a difference between those two, which could be attributed to compensation. This is important because the study of compensation in ASD is very limited and thus, the current findings indicate that research in the area of metacognition might be fruitful for the understanding of compensation in ASD. Nevertheless, as it has already been noted metacognitive monitoring abilities do not entail only judgements of confidence; in the next chapter we explored metacognitive experiences of difficulty in ASD, as an alternative index of metacognitive monitoring.

3. Exploring Metacognitive Experiences of Difficulty in ASD

In the current chapter, two experiments were conducted aiming to explore the extent to which ASD is characterised by atypical metacognitive experiences of difficulty. That is a relatively new area of metacognition that has never been explored before in ASD.

Experiment 3.1: Individual Differences

In Experiment 3.1, we investigated the extent to which accuracy of metacognitive experiences of difficulty relates to mindreading abilities or to autism traits. Individuals from the general population completed a masked priming paradigm (Desender et al., 2016), the RMIE (Baron-Cohen et al., 2001a) task, and the AQ (Baron-Cohen et al., 2001b).

The first aim of the current experiment was to replicate Desender et al.'s (2016) findings (see p. 7 of the current project). Second, we explored the association between accuracy of metacognitive experiences of difficulty and mindreading abilities. "One-system" theorists claim that both mindreading and metacognition are under the umbrella of the same metarepresentational faculty (e.g. Carruthers, 2009). If this is the case, then accuracy of metacognitive experiences of difficulty should be positively and significantly associated with performance in mindreading tasks. Third, we aimed to examine the extent to which there is an association between accuracy of metacognitive experiences of difficulty and autism traits. As it is well-established that mindreading abilities are negatively associated with autism traits in the general population (e.g. Baron-Cohen, 2001b; Williams et al., 2016), and based on the theoretical claim that mindreading and metacognition are dependent with each other (e.g. Carruthers, 2009), autism traits should be negatively and significantly associated with accuracy of metacognitive experiences of difficulty.

Experiment 3.1: Method

Participants

Seventy-nine undergraduate students (65 female) participated in this study after they had given written, informed consent. Their average age was 19.56 (*SD* = 2.23; range = 18 to 36) years. All participants had normal or corrected-to-normal vision, were native English speakers or spoke English to native-level proficiency and had no history of ASD, language impairment, or dyslexia, according to self-report. Participants were recruited via the Research Participation Scheme (RPS) of the University of Kent, UK, and they received course credits in partial fulfilment of their degree for their participation. Ethical approval for this study was obtained from the Kent School Psychology Research Ethics Committee (Ethics ID: 201815215553564983).

Materials, Procedures, and Scoring

As in Experiment 2.1, participants completed the RMIE task (Baron-Cohen et al., 2001a) and the AQ (Baron-Cohen et al., 2001b). In addition, they performed a masked priming task (Desender et al., 2016).

Masked priming task.

We employed this task to measure accuracy of metacognitive experiences difficulty. In each trial, following a fixation cross displayed for 1000ms on a 22" computer screen, a prime arrow (1.5° wide and 0.7° high) was presented for 34ms pointing either to the left or to the right side of the screen. Then, following a blank screen displayed for 34ms, a target arrow (3.3°wide and1.4° high) was presented for 116ms pointing again either to the left or to the right. A blank screen followed, and participants had to respond as fast and accurately as possible to the direction of the target arrow. If the arrow pointed to the left, they should press the “d” key of a keyboard, using the middle finger of their left hand, but if the arrow pointed to the right, they should press the “k” key, using the middle finger of their right hand.

Subsequently, when they responded within 3000ms, a blank screen was presented for 516ms, and the following metacognitive question was displayed: “How much difficulty did you experience when responding to the arrow?”. Participants were instructed to press the key “o” with the ring finger of their right hand if they experienced “rather more difficulty” or the key “m” with the index finger of their right hand if they experienced “rather less difficulty”. No time limit was imposed on participants to give their metacognitive responses, and they had been instructed to use all the sources of information available in their mind before they gave a judgement of difficulty.

The experiment took place in a dimly lit room. Participants first completed 20 practice trials in which the metacognitive questions were omitted. Next, they performed 20 additional ones with the metacognitive questions being included. Each participant completed two blocks of 80 trials, with the trial order being randomised within blocks. On half the trials, the direction of the prime arrow was in congruency with the direction of the target arrow, rendering these trials easier for participants. On the other half, the direction of the two arrows was incongruent, creating a response conflict that made these trials more difficult. In both occasions, the prime arrow was invisible to participants, as it fitted exactly with the target arrow. Thus, participants were not consciously aware of which trials had an increased difficulty, compared to the other tones (see Figure 4 for a graphical representation of the task).

Following the approach adopted by Desender et al. (2016), we excluded all the incorrect responses to the target arrow, along with the first trial of each block, and each trial that followed an incorrect response. On average, 4% of the trials were eliminated per participant. All the analyses reported below refer to that trimmed dataset.

In terms of scoring, Desender et al. (2016) calculated conflict-*d* scores, as index of accuracy of metacognitive experience of difficulty. Likewise, we used the same approach, employing the formula: conflict-*d* = Z(Hit Rate) – Z(False Alarm Rate). Hit rate (H) was calculated dividing the number of incongruent trials that participants experienced as “rather more difficult” (hits) by the number of hits plus the number of incongruent trials experienced as “rather less difficult” (misses). False alarm rate (FA) was calculated dividing the number of congruent trials that participants experienced as “rather more difficult” (false alarms) by the number of false alarms plus the number of congruent trials that participants experienced as “rather less difficult” (correct rejections). The higher the conflict-*d* the greater the accuracy.

Conflict-*d* scores cannot be calculated either when the hit rate is equal to one, or when the false alarm rate is equal to zero. As such, raw data were corrected using the formula 1 – 1/(2N) for hit rates, with *N* representing the maximum possible number of hits and the formula 1/(2N) for false alarm rates, with *N* representing the maximum possible number of false alarms (Macmillan & Kaplan, 1985). As in Desender et al. (2016), we also calculated median RTs to the target arrow.

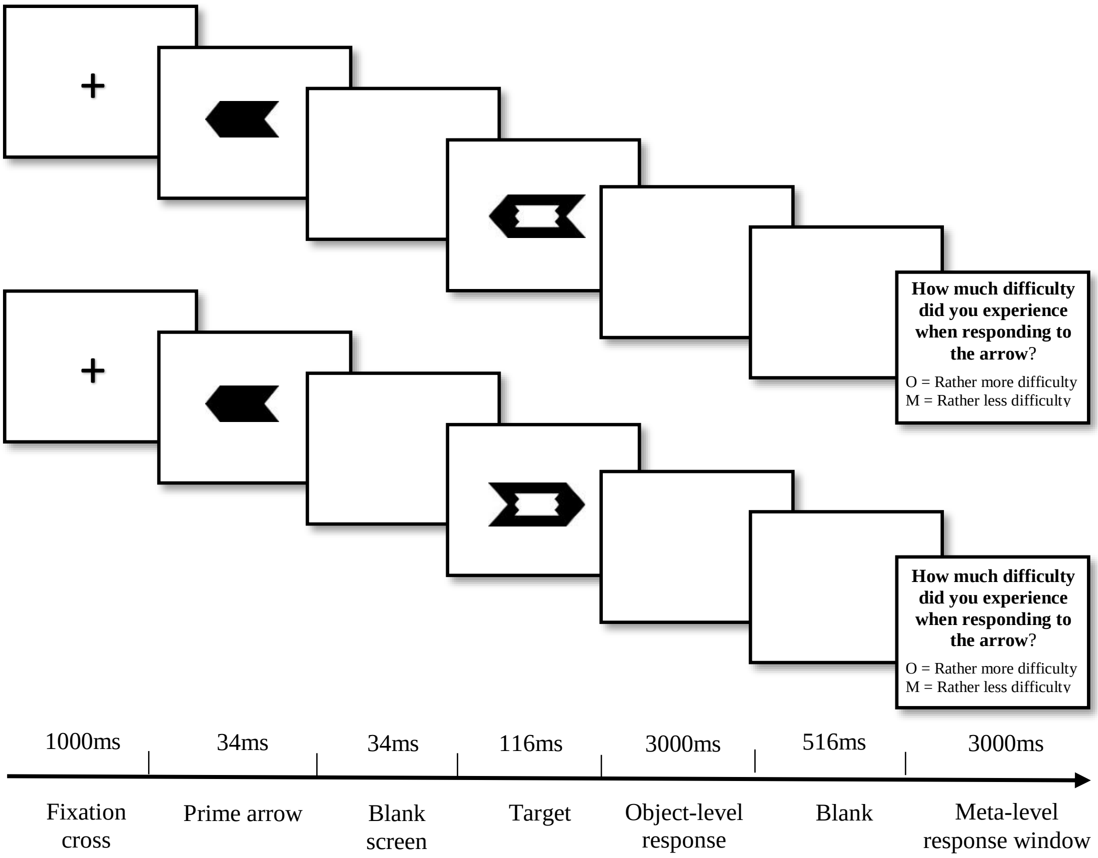

Figure 4. Graphical representation of the procedure and stimuli used in congruent (top) and incongruent (bottom) trials of the masked priming task.

Experiment 3.1: Results

Metacognitive Experiences of Difficulty

A one-sample *t*-test showed that the average conflict-*d* (*M* = .93, *SD* = .83) was significantly above chance, indicating that participants showed the expected metacognitive effect, $t(78) = 9.91$, $p < .001$, $d = 1.11$, $BF_{10} > 100$ (one-tailed). That is to say, they accurately classified incongruent trials as “rather more difficult” and congruent trials as “rather less difficult”. In addition, the average conflict-*d* was below ceiling, $t(78) = -1058.2$, $p < .001$, $BF_{10} > 100$.

Next, we examined the effect of trial difficulty on median RTs to the target arrow as well as the effect of these RTs on metacognitive responses. A 2 (congruency: incongruent/ congruent) x 2 (metacognitive judgement: “rather more difficult” / “rather less difficult”) repeated measures ANOVA was conducted on median RTs to the target arrow (see Figure 5). The analysis revealed that the main effect of congruency was non-significant, $F(1,78) = 2.34$, $p = .13$, $\eta_p^2 = .03$, whereas the main effect of metacognitive judgement was significant reflecting slower RTs to the target arrow in trials that participants labelled as “rather more difficult”, than in trials labelled as “rather less difficult”, $F(1,78) = 53.52$, $p < .001$, $\eta_p^2 = .41$. Additionally, the interaction between congruency and metacognitive judgement was marginally significant, $F(1,78) = 3.70$, $p = .06$, $\eta_p^2 = .05$. This interaction was further investigated, using paired sample *t*-tests.

Results from paired sample *t*-tests established the main effect of metacognitive judgement. That is “rather less difficult” judgements were made after quick responses, and “rather more difficult” judgements were made after slow responses in both congruent and incongruent trials, $t(78) = -5.34$, $p < .001$, $BF_{10} = 3.56$ and $t(78) = -7.56$, $p < .001$, $BF_{10} = 6.18$ respectively. There was also a significant difference in median RTs between congruent and incongruent trials for “rather less difficult” judgements. In incongruent trials, participants gave significantly slower responses, than in congruent trials, $t(78) =$ -

6.57, $p < .001$, $BF_{10} > 100$. However, as shown in Figure 5 the difference between congruent and incongruent trials was only 40ms. Congruency did not have any effect on median RTs for “rather more difficult” judgements, $t(78) = -0.21$, $p = .84$, $BF_{10} = 0.13$.

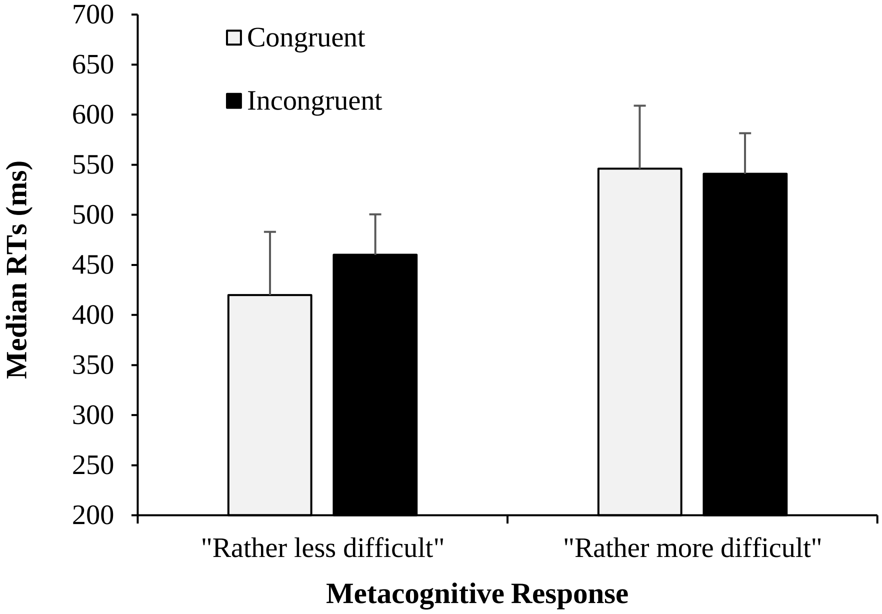

Figure 5. Median RTs to the target arrow in the masked priming task as a function of congruency and metacognitive response. Error bars represent standard error of measurement.

Association Analyses

A series of correlation analyses was conducted to examine the associations among accuracy of metacognitive experiences of difficulty (indexed by conflict-*d*), mindreading abilities, and autism traits. Our sample was representative of the general population with respect to autism traits and representative of student population in terms of mindreading abilities (see Appendix c).

In line with our predictions, the association between conflict-*d* and RMIE was positive and marginally significant, $r = .17$, $p = .07$, $BF_{10} = 0.74$ (one-tailed). In addition, A Fisher’s *Z* test revealed that the effect size of this association did not differ significantly either from the correlation coefficient JoC gamma × RMIE ($r = .25$) reported by Williams et al. (2016), $Z = -0.55$, $p = .58$ or from the coefficient JoC gamma × RMIE reported in Experiment 2.1, $Z = -0.20$, $p = .84$. Nonetheless, in contrast with our predictions, the correlational analysis revealed a no significant association between conflict-*d* and AQ, $r = -.03$, $p = .41$, $BF_{10} = 0.17$ (one-tailed), indicating that accuracy of metacognitive experiences of difficulty was unrelated to autism traits, as measured in the general population.

Experiment 3.1: Discussion

In Experiment 3.1, we successfully replicated Desender et al.'s (2016) findings. That is to say, we found that participants from the general population classified congruent trials as "rather less difficult" and incongruent trials as "rather more difficult" at above chance level, indicating that their experiences of difficulty coincided to a great extent with the actual difficulty of each trial of the task. Nonetheless, it could be argued that this effect can be better explained by introspection on reaction times and visibility of the prime arrow, rather than by intact metacognitive monitoring abilities (Marti et al., 2010).

In terms of the first alternative explanation, reaction time analyses showed that congruency had a significant effect on reaction times only in the case of trials labelled as "rather less difficult", with participants responding 40ms faster in congruent trials, than in incongruent ones. As such, it is doubtful whether such a short difference in reaction times could even be noticeable in order to affect metacognitive responses. Therefore, we can conclude that in the current experiment accuracy of metacognitive experiences of difficulty was not achieved by introspection on reaction times to the target arrow.

The second alternative explanation seems more plausible, compared to the first one, given we did not examine the extent to which the prime arrow was visible to participants. This was something we could have controlled for employing a simple detection task, as Desender et al. (2016) did. Nevertheless, previous research findings have shown that either with zero visibility or with a mean visibility of $d = 0.55$, accuracy of metacognitive experiences of difficulty was above chance level (Desender et al., 2014; 2016), indicating that the awareness of the prime arrow is distinct from the awareness of the conflict experience (Desender et al., 2014). As such, we can be confident enough that our results were not confounded by this factor and thus to use accuracy of metacognitive experiences of difficulty as index of metacognitive monitoring abilities.

In terms of associations, as predicted accuracy of metacognitive experiences of difficulty was positively associated with mindreading abilities. Even though, the association was marginally significant and its effect size small (r = .17), their shared variance was equal to that reported in Experiment 2.2 for the association between explicit JoC accuracy and RMIE as well as to that reported by Williams et al. (2016). Crucially, to date this is the first study of the relation between metacognitive experiences of difficulty and mindreading abilities and arguably the current findings appear to provide some evidence in favour of the theoretical claim that metarepresentational abilities underlie not only mindreading but also metacognition (e.g. Carruthers, 2009).

Furthermore, in contrast with our predictions the association between accuracy of metacognitive experiences of difficulty and autism traits was non-significant and trivial (r = -.03). The Bayesian correlation analysis suggested that the data provided substantial evidence in favour of the null hypothesis. Therefore, our findings indicate that the extent to which an individual from the general population reports high or low levels of autism traits is unrelated to their ability to impute mental states to themselves. Despite making this conclusion, it is still questionable whether this applies to people with a full diagnosis of ASD.

Williams et al. (2016) found that the lack of a significant association between autism traits and an index of metacognition in the general population, by no means implies that the ASD folk will not be impaired. Qualitative differences in cognition could differentiate people with and without a diagnosis of ASD (Peterson et al., 2005; Ruzich et al., 2015). As such, whilst we did not find a significant association between accuracy of metacognitive experiences of difficulty and individual differences in the number of autism traits, among people from the general population, in Experiment 3.2 we further explored this facet of metacognitive monitoring abilities among individuals with a full diagnosis of ASD.

Experiment 3.2: Case-Control

In Experiment 3.2 we conducted a case-control study employing among individuals with ASD as well as neurotypical people the masked priming task used in Experiment 3.1, the RMIE (Baron-Cohen et al., 2001a) task and the MASC (Dziobek et al., 2006). As it is well-established that individuals with ASD have attenuated mindreading abilities (e.g. Baron-Cohen, 2001; Senju et al., 2009; Yirmiya et al., 1998), we first predicted that participants with ASD would show diminished performance in both mindreading tasks, compared to comparison participants. Second, in keeping with the theoretical claims of "one-system" theorists that mindreading and metacognition are inherently related with each other (e.g. Carruthers, 2009), we predicted that we would find a significant between-group difference in performance on the masked priming task, with participants with ASD showing diminished accuracy of metacognitive experiences of difficulty.

Experiment 2: Method

Participants

Twenty-two adults with ASD and 20 neurotypical comparison adults participated in the current experiment, after they had given written informed consent.[7] All participants with ASD had received formal diagnoses, according to established criteria (DSM-IV-TR, American Psychiatric Association 2000; ICD-10, World Health Organisation 1993), while their ASD severity was assessed using the ADOS .[8] All neurotypical participants scored below the AQ cut-off score. Two ASD participants were excluded because their error rate in the masked priming task, which was the main experimental task of the current experiment, was higher than 25% (following criteria set out by Desender et al., 2014). In the final sample, 20 participants with ASD and 20 neurotypical participants were included. As shown in Table 11 groups were matched for verbal IQ, performance IQ and full scale IQ, using the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999). They were also matched for chronological age and sex ratio. Ethical approval for this study was obtained from the Kent School Psychology Research Ethics Committee (Ethics ID: 201815259101245011).

Table 11

Baseline Characteristics and Matching Statistics for Experiment 3.2

	Diagnostic group		Group Differences			
	ASD (*n* = 20; 12 male)	Neurotypical (*n* = 20; 14 male)	*t*	*p*	*d*	BF_{10}
CA: years	36.80 (12.60)	41.95 (13.94)	-1.23	.23	0.39	0.56
VIQ	106.80 (10.61)	104.05 (11.22)	0.80	.43	0.25	0.40
PIQ	108.45 (15.62)	105.60 (15.18)	0.59	.56	0.19	0.35
FSIQ	108.20 (11.47)	105.65 (12.99)	0.66	.51	0.21	0.37
ADOS	9.45 (4.86)	-	-	-	-	-
AQ total	33.20 (8.32)	14.25 (4.56)	8.93	<.001	2.82	>100
RMIE[a]	25.50 (5.80)	27.80 (3.86)	-1.48	.07	0.47	1.31
MASC[a]	28.25 (6.83)	33.75 (5.21)	-2.86	.00	0.91	13.32

Note. CA = chronological age; VIQ = verbal IQ; PIQ = performance IQ; FSIQ = full scale IQ; ADOS = Autism Diagnostic Observation Schedule; AQ total = Total score on Autism-Spectrum Quotient; RMIE = Reading the Mind in the Eyes; MASC = Movie for the Assessment of Social Cognition.

[a] Values for one-tailed tests are reported because of a priori directional predictions.

Materials, Procedures, and Scoring

As in Experiment 3.1, all participants completed the RMIE (Baron-Cohen et al., 2001a), and the masked priming task (Desender et al., 2016). Materials, procedures, and scoring remained identical across experiments with the only difference that in the current experiment, participants completed only the second practice session of the masked priming task in which the metacognitive questions were included. In addition, participants completed the MASC (Dziobek et al., 2006; see pp. 23-24 X of the current project for a detailed description of the task).

Experiment 3.2: Results

Between-Group Differences

With respect to performance on the masked priming task, as shown in Table 12 an independent samples *t*-test indicated that groups were equated for object-level task performance. That is to say, the proportion of correct responses to the target arrow among ASD participants was equivalent to neurotypical participants. Next, we examined the extent to which there was a significant between-group difference in accuracy of metacognitive experiences of difficulty (see Table 12 for descriptive and inferential statistics).

Table 12

Means (SDs) and Inferential Statistics for Group Differences in Performance on the Masked Priming Task

Variable	Group		Group Differences			
	ASD	Neurotypical	*t*	*p*	*d*	BF_{10}
Object-level performance	0.97 (0.01)	0.98 (0.03)	-1.19	.24	0.45	0.54
Proportion of "rather more difficult" responses	0.23 (0.18)	0.43 (0.13)	-3.88	<.001	1.27	66.13
Conflict-*d*[a]	0.69 (1.20)	0.57 (0.82)	0.36	.36	0.12	0.25

[a] Values for one-tailed tests are reported because of a priori directional predictions.

An independent samples *t*-test showed that in contrast with our predictions, there was not a significant between-group difference in conflict-*d*, indicating that participants with ASD had undiminished accuracy of metacognitive experiences of difficulty.

Although, it became clear during testing that there was a substantial sex difference among autistic participants in performance on the experimental task a post-hoc analysis of sex differences has not been presented in the current thesis due to the small n per each group.

Association Analyses

A series of correlation analyses was conducted exploring the relations between mindreading abilities and accuracy of metacognitive experiences of difficulty. Among the ASD participants neither the RMIE nor MASC correlated significantly with conflict-*d*, $r = .06$, $p = .40$, $BF_{10} = 0.34$ (one-tailed) and $r = -.03$, $p = .45$, $BF_{10} = 0.25$(one-tailed) respectively. When these associations were examined among neurotypical participants, conflict-*d* correlated significantly with MASC, $r = .54$, $p = .01$, $BF_{10} = 9.60$ (one-tailed) but not with RMIE, $r = -.03$, $p = .45$, $BF_{10} = 0.25$ (one-tailed).

Experiment 3.2: Discussion

As expected, we found a large and significant ASD-specific impairment in mindreading, as measured using the MASC (alongside a marginally significant and small between-group difference in performance on the RMIE). These results are in keeping with previous research findings (e.g. Baron-Cohen et al., 2001a; Dziobek et al., 2006; Yirmiya et al., 1998) and indicate that participants with ASD who took part in the current experiment were representative of the ASD population with respect to mindreading abilities.

In contrast with our predictions, accuracy of metacognitive experiences of difficulty was found to be undiminished among this sample of adults with ASD. The between-group difference in the size of accuracy of metacognitive experiences of difficulty was non-significant and very small ($d = 0.12$), with the Bayes factor suggesting that the data provided substantial evidence in favour of the null hypothesis. Based on these findings, it would be reasonable to suggest that the current experiment provides evidence for intact metacognitive monitoring abilities in ASD.

In terms of associations, as predicted accuracy of metacognitive experiences of difficulty was positively and significantly associated with mindreading abilities, as measured employing the MASC among neurotypical participants. Noteworthy, the size of the association was large ($r = .54$), suggesting that metarepresantional abilities were the source of their shared variance (29%). Nonetheless, all the other correlation analyses we conducted to examine the association between mindreading and accuracy of metacognitive experiences of difficulty produced non-significant results.

3. General Discussion

In the current chapter, were conducted two experiments aiming to explore metacognitive experiences of difficulty in ASD. In Experiment 3.1, we replicated previous research findings showing that neurotypical people can feel when they experience either increased or decreased difficulty, and that this experience coincides with the actual difficulty of the task they perform (Chambon & Haggard, 2012; Desender et al., 2014, Desender et al., 2016; Wenke, Fleming, & Haggard, 2010). Arguably, this ability taps metacognitive monitoring, as individuals need to form metarepresentations of the first order representations of the task difficulty in order to state accurately their experience of difficulty (Desender et al., 2016). Results from association analyses increased our confidence in this suggestion.

In Experiment 3.1, we found a marginally significant association between performance on the RMIE task and accuracy of metacognitive experiences of difficulty. Likewise, in Experiment 3.2 we found a significant correlation between performance on MASC and accuracy of metacognitive experiences of difficulty. These findings indicate that peoples' ability to impute mental states to themselves is metarepresentational and thus increase our confidence to theoretical claims that mindreading and metacognition are under the umbrella of the same metarepresentational faculty (Carruthers, 2009; Leslie, 1987; Moore & Frye, 1991).

Nonetheless, it could be argued that these were just spurious correlations, given that all the other correlation analyses we conducted in order to examine this relation produced non-significant results. Before strong conclusion can be drawn, we should take into account that we did find a significant association, when we measured mindreading abilities using the MASC. Compared with other mindreading tasks, this is considered a quite sensitive measure that taps many aspects of mindreading and not only emotion recognition, such as the RMIE (Turner & Felisberti, 2017). Moreover, we should take into account that lack of statistical power in our sample could provide a plausible explanation for the non-significant results (see pp. 75-76 of the current project for statistical power considerations).

In Experiment 3.1, we also found that the association between autism traits and accuracy of metacognitive experiences of difficulty was non-significant. Nonetheless, this is in keeping with results from Experiment 2.1 and with previous research findings (Williams et al., 2016), suggesting that the ability to impute mental states to one's own self is unrelated to autism traits. In Experiment 3.2, we examined whether this applies to individuals diagnosed with ASD.

Results from Experiment 3.2 were in absolute contrast with our predictions. Despite showing diminished mindreading abilities, individuals with ASD nonetheless showed undiminished accuracy of metacognitive experiences of difficulty. This provides a significant challenge to "one-system theory" and is more consistent with theoretical claims that mindreading and metacognition rely on different mechanisms (Goldman, 2006; Nichols & Stich, 2003). Nevertheless, we should note that replication awaits before strong conclusions can be drawn.

General Conclusion

The current project provides further evidence of intact interoceptive accuracy in ASD and challenges theoretical claims that relate the ASD phenotype with deficits in interoception (e.g. Brewer et al., 2015; Quattrocki & Friston, 2014). Our findings indicate that people with ASD are able to form first order representations of their physiological conditions of their body and as such, they can detect precisely their body signals. Even though this ability appears to be intact, it does not necessarily mean that is a prerequisite for the correct interpretation of those signals. An individual should be able to reflect the first order representations of the physiological conditions of their body upon their cognition in order to interpret them correctly. In other words, their ability to form metarepresentations should be intact and as such, we can argue that metacognition is crucial for this process (Nicholson et al., 2018).

In the current project, we explored four different indices of metacognition. That is interoceptive awareness, accuracy of implicit and explicit judgments of confidence, and accuracy of metacognitive experiences of difficulty. Not all metacognitive results were in the predicted direction indicating atypical metacognitive monitoring abilities among individuals with ASD. One thing to note here is that our sample was always small and this could plausibly explain why some of our findings provided evidence that was inconsistent with our predictions.

Statistical Power Considerations

Given that the study of metacognition in ASD is very limited and the existing findings conflicting, estimates of effect sizes were difficult to make precisely, and if made they would be to a great extent arbitrary. Based on the argument of the “minimally interesting value” (Dienes, 2014), we could claim that ideally in the study of ASD studies should be powered to find an effect of 0.50 (or, and .25– equivalent to *d* of 0.50– and above). It is questionable whether between-group differences, which are only small in magnitude, could imply clinical significance (Grainger et al., 2014; Nicholson et al., 2018; Williams et al., 2016).

As such, assuming an effect size of 0.50 for the between-group difference in cardiac interoceptive awareness and using one-tailed tests, a power analysis using G*Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009) indicated that a total sample of 102 participants was required to achieve Cohen’s recommended power of 0.80 (Cohen, 1992). When $N = 39$ statistical power was 0.31, indicating that our sample was underpowered to detect a significant between-group difference in cardiac interoceptive awareness. Using the same criterion, we can argue that none of our studies of metacognition in ASD was sufficiently powered. That is a common limitation in the study of ASD, given the difficulty in recruiting and testing participants from the ASD population. Nonetheless, we cannot underestimate the implications that arise from the findings of the current project.

Implications

In terms of the theoretical implications of the current project, our findings about undiminished interoceptive accuracy in ASD increase our confidence in theoretical claims that physical and psychological self-awareness are distinct and independent from each other (Gillihan & Farah, 2005). Based on this account, it has been argued that although individuals with ASD have impaired psychological self-awareness, their physical self-awareness is intact because their ability to make first order representations of their body outward phenotype and its internal functions is undiminished (Lind, 2010; Nicholson et al., 2018; Uddin, 2011; Williams, 2010). Our results fit also within the major theoretical debate between “one-system” and “two-system” theorists.

In addition to strong evidence that mindreading abilities are diminished among individuals with ASD, we found either moderate or tentative evidence to support the hypothesis that metacognitive monitoring abilities are diminished in ASD as well. Arguably, the current project provides more evidence in favour of “one-system” over “two-system” theorists, indicating that both mindreading and metacognition are under the control of a single metarepresentational faculty. Along with theoretical implications, major clinical and educational implications arise too.

First, results about diminished interoceptive awareness in ASD should be taken into account in medical environments. A deficit in the ability of individuals with ASD to think about the physiological conditions of their body or to feel themselves feeling them, could potentially makes them to ignore or give a wrong interpretation to a body signal that is out of the ordinary. If this is the case, the diminished interoceptive awareness in ASD could provide an explanation for their “apparent indifference to pain/heat/cold’ (American Psychiatric Association, 2013, p. 50). Arguably, this difficulty in ASD could lead to late diagnosis and treatment of serious medical conditions (Frith & Happé, 1999). This is important because it has been found that children with ASD present high comorbidity with conditions related to pain (Bottos & Chambers, 2006) and that mortality risk is elevated in ASD (Bilder et al., 2013). As such, training programs in physiology and health designed for individuals with ASD, along with the psychoeducation of medical staff and carers could prevent serious delays in identifying medical conditions among children and adults with ASD.

Second, results about diminished implicit metacognitive monitoring abilities in ASD should be taken into account in educational environments. This is crucial given that metacognitive monitoring predicts educational achievement independent of IQ and has a unique relation with mathematics learning (Ashburner, Ziviani, & Rodger, 2010; Pishghadam & Khajavy, 2013; Thiede, 1999; Veenman et al., 2005; Veenman & Spaana, 2005). As such, deficits in this domain can provide an explanation for the academic underachievement that characterise children with ASD and is unrelated to their intellectual abilities (Estes, Rivera, Bryan, Cali, & Dawson, 2011). Research has already established a direct link between research evidence about deficits in metacognitive monitoring abilities in educational settings and positive outcomes after training.

Bronsan et al. (2016) found that metacognitive deficits were prominent in students with ASD when assessed their performance in mathematics. Specifically, they overestimated their performance and they showed lack of cohesion in their intentions before answering each mathematic question and after knowing whether they made it right or wrong. Based on these findings, Maras, Gamble and Brosnan (2017) designed and employed a computer-based programme in metacognitive support among children with ASD and neurotypical children. Results indicated that children with ASD who received metacognitive support showed increased performance in mathematics, compared to children with ASD group who completed the programme without receiving support.

As such, the design of interventions to support metacognitive monitoring abilities and train individuals to metacognitive strategies is of major importance for the alleviation of difficulties in areas that students with ASD struggle, such as mathematics (Mayes & Calhoun, 2006) and reading comprehension (Minshew et al. 1994; Nation, Clarke, Wright, & Williams, 2006; O’Connor & Klein 2004), given that achievement in school is an early indicator of a successful career in the future (Kell, Lubinski, & Benbow, 2013).

Directions for Future Research

In sum, the findings of the current project provide strong evidence that lower order phenomena of self-awareness are intact in ASD and some evidence that higher order phenomena of self-awareness are impaired. Nonetheless, due to the novelty of the tasks we used and the small sample sizes of the current project, replication of our findings awaits before firm conclusion can be drawn.

Future research might usefully replicate Experiment 2.2 employing both the implicit and the explicit JoC tasks among children with ASD and neurotypical children. To our knowledge, evidence about compensation in mindreading have been found only among adults with ASD (Senju et al., 2009). Arguably, children with ASD are less prone to make use of compensatory strategies and as such, results will be possible more informative about metacognitive monitoring abilities in ASD. In addition, if children with ASD showed diminished performance in both JoC tasks, then our confidence in the compensation hypothesis would be increased. Overall, despite being in its infancy, we can argue that the study of self-awareness appears to be promising when it comes to the understanding of many of the unexplained aspects of Autism Spectrum Disorder.

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,* 1-16. doi:10.1016/S0885-2014(00)00014-9

Ainley, V., Brass, M., & Tsakiris, M. (2014). Heartfelt imitation: high interoceptive awareness is linked to greater automatic imitation. *Neuropsychologia, 60,* 21-28. doi:10.1016/j.neuropsychologia.2014.05.010

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.).Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Asendorpf, J. B., Warkentin, V., & Baudonnière, P. M. (1996). Self-awareness and other-awareness. II: Mirror self-recognition, social contingency awareness, and synchronic imitation. *Developmental Psychology, 32,* 313-321. doi:10.1037/0012-1649.32.2.313

Ashburner, J., Ziviani, J., & Rodger, S. (2010). Surviving in the mainstream: Capacity of children with autism spectrum disorders to perform academically and regulate their emotions and behavior at school. *Research in Autism Spectrum Disorders, 4,* 18-27. doi:10.1016/j.rasd.2009.07.002

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 East Thames: the Special Needs and Autism Project (SNAP). *Lancet, 368,* 210-215. doi:10.1016/S0140-6736(06)69041-7

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

Baron-Cohen, S., Scott, F. J., Allison, C., Williams, J., Bolton, P., Matthews, F. E., & Brayne, C. (2009). Prevalence of autism-spectrum conditions: UK school-based population study. *The British Journal of Psychiatry, 194,* 500-509. doi:10.11 92/bjp.bp.108.059345

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders, 34,* 163-175. doi:10.1023/B:JADD .0000022607.19833.00

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001ba). The “reading the mind in the eyes” test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, 42,* 241-251. doi: 10.1111/1469-7610.00715

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001b). 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,* 5-17. doi:10.1023 /A:1005653 4 11471

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). Beck depression inventory-II. *San Antonio, 78*(2), 490-498.

Beran, M. J., Smith, J. D., Coutinho, M. V., Couchman, J. J., & Boomer, J. (2009). The psychological organization of “uncertainty” responses and “middle” responses: A dissociation in capuchin monkeys (Cebus apella). *Journal of Experimental Psychology: Animal Behavior Processes, 35,* 371-381. doi:10.1037/a0014626

Berthoz, S., & Hill, E. L. (2005). The validity of using self-reports to assess emotion regulation abilities in adults with autism spectrum disorder. *European Psychiatry, 20,* 291-298. doi:10.1016/j.eurpsy.2004.06.013

Bilder, D., Botts, E. L., Smith, K. R., Pimentel, R., Farley, M., Viskochil, J., ... Coon, H. (2013). Excess mortality and causes of death in autism spectrum disorders: a follow up of the 1980s Utah/UCLA autism epidemiologic study. *Journal of Autism and Developmental Disorders, 43,* 1196-1204. doi:10.1007/s10803-012-1664-z

Bird, C. M., Castelli, F., Malik, O., Frith, U., & Husain, M. (2004). The impact of extensive medial frontal lobe damage on 'Theory of Mind'and cognition. *Brain, 127,* 914-928. doi:10.1093/brain/awh108

Bird, G., & Cook, R. (2013). Mixed emotions: the contribution of alexithymia to the emotional symptoms of autism. *Translational Psychiatry, 3,* 1-8. doi:10.1038/tp.2013.61

Bird, G., Silani, G., Brindley, R., White, S., Frith, U., & Singer, T. (2010). Empathic brain responses in insula are modulated by levels of alexithymia but not autism. *Brain, 133,* 1515-1525. doi.org/10.1093/brain/awq060

Blair, R. J. R. (2008). Fine cuts of empathy and the amygdala: dissociable deficits in psychopathy and autism. *The Quarterly Journal of Experimental Psychology, 61,* 157-170. doi:10.1080/17470210701508855

Bogaerts, K., Millen, A., Li, W., De Peuter, S., Van Diest, I., Vlemincx, E., Fannes, S, Van den Bergh, O. (2008). High symptom reporters are less interoceptively accurate in a symptom-related context. *Journal of Psychosomatic Research, 65,* 417-424. doi:10.1016/j.jpsychores.2008.03.019

Bolton, P., Macdonald, H., Pickles, A., Rios, P., Goode, S., Crowson, M., … Rutter, M. (1994). A case-control family history study of Autism. *Journal of Child Psychology and Psychiatry, 35,* 877-900. doi:10.1111/j.1469-7610.1994.tb02300.x

Bottos, S., & Chambers, C. T. (2006). The epidemiology of pain in developmental disabilities. In T. F. Oberlander, & F. J. Symons (Eds.), *Pain in children and adults with developmental disabilities* (pp. 67-87). Baltimore, MD: Brookes.

Brewer, R., Happé, F., Cook, R., & Bird, G. (2015). Commentary on "Autism, oxytocin and interoception": Alexithymia, not autism spectrum disorders, is the consequence of interoceptive failure. *Neuroscience & Biobehavioral Reviews, 56,* 348-353. doi:10.1016/j.neubiorev.2015.07.006

Brosnan, M., Johnson, H., Grawemeyer, B., Chapman, E., Antoniadou, K., & Hollinworth, M. (2016). Deficits in metacognitive monitoring in mathematics assessments in learners with autism spectrum disorder. *Autism, 20,* 463-472. doi:10.1177/13623 61315589477

Brunsdon, V. E., & Happé, F. (2014). Exploring the ‘fractionation’ of autism at the cognitive level. *Autism, 18,* 17-30. doi:10.1177/1362361313499456

Bryant, R. A., Harvey, A. G., Dang, S. T., Sackville, T., & Basten, C. (1998). Treatment of acute stress disorder: A comparison of cognitive-behavioral therapy and supportive counseling. *Journal of Consulting and Clinical Psychology, 66,* 862-866. doi:10.103 7/0022-006X.66.5.862

Butterworth, G. (1995). The self as an object of consciousness in infancy. In P. Rochat (Ed.), *The self in infancy: Theory and research* (pp. 35-51). Amsterdam, Netherlands: North Holland/Elsevier Science. doi:10.1016/S0166-4115(05)80005-6

Callenmark, B., Kjellin, L., Rönnqvist, L., & Bölte, S. (2014). Explicit versus implicit social cognition testing in autism spectrum disorder. *Autism, 18,* 684-693. doi:10.11 77/136 2361313492393

Cameron, O. G. (2002). *Visceral sensory neuroscience: interoception*. New York, NY: Oxford University Press.

Carruthers, P. (1996). Autism as mind-blindness: an elaboration and partial defence. In P. Carruthers & P. Smith (Eds.), *Theories of Theories of Mind* (257-276). Cambridge: Cambridge University Press.

Carruthers, P. (2008). Meta-cognition in Animals: A Skeptical Look. *Mind & Language, 23,* 58-89. doi:10.1111/j.1468-0017.2007.00329.x

Carruthers, P. (2009). How we know our own minds: The relationship between mindreading and metacognition. *Behavioral and Brain Sciences 32,* 1-18. doi:10.1017/S 0140525X 09000545

Carruthers, P. (2011). *The opacity of mind: An integrative theory of self-knowledge.* Oxford: Oxford University Press

Carruthers, P. (2013). Mindreading in infancy. *Mind & Language, 28*, 141-172. doi:10.1 111/mila.12014

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,*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*, 314-325. doi:10.1006/nimg.2000.0612

Ceunen, E., Van Diest, I., & Vlaeyen, J. W. S. (2013). Accuracy and awareness of perception: Related, yet distinct (commentary on Herbert et al., 2012). *Biological Psychology, 92*, 426-427. doi:10.1016/j.biopsycho.2012.09.012

Chambon, V., & Haggard, P. (2012). Sense of control depends on fluency of action selection, not motor performance. *Cognition, 125*, 441-451. doi:10.1016/j.cognition. 2012.07.011

Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization, 81,* 1-8. doi:10.1016/j.jebo.2011.08.009

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science, 1*(3), 98-101. Retrieved from http://www.jstor.org/stable/20182143

Cohen, J. (1969). *Statistical power analysis for the behavioural sciences.* San Diego, CA: Academic Press.

Constantino, J. N., & Todd, R. D. (2003). Autistic traits in the general population: A twin study. *Archives of General Psychiatry, 60,* 524-530. doi:10.1001/archps yc.60.5.524

Cook, R., Brewer, R., Shah, P., & Bird, G. (2013). Alexithymia, not autism, predicts poor recognition of emotional facial expressions. *Psychological Science, 24,* 723-732. doi:org/10.1177/0956797612463582

Cooper, R. A., Plaisted-Grant, K. C., Baron-Cohen, S., & Simons, J. S. (2016). Reality monitoring and metamemory in adults with autism spectrum conditions. *Journal of Autism and Developmental Disorders*, *46,* 2186-2198. doi:10.1007/s10803-016-2749-x

Couchman, J. J., Coutinho, M. V., Beran, M. J., & Smith, J. D. (2009). Metacognition is prior. *Behavioral and Brain Sciences, 32*, 142-142. doi:10.1017/S0140525X0900 0594

Couchman, J. J., Coutinho, M. V., Beran, M. J., & Smith, J. D. (2010). Beyond stimulus cues and reinforcement signals: a new approach to animal metacognition. *Journal of Comparative psychology, 124,* 356-368. doi:10.1037/a0020129

Constantino, J. N., & Todd, R. D. (2000). Genetic structure of reciprocal social behavior. *American Journal of Psychiatry, 157,* 2043-2045. doi: 10.1176/appi.ajp.157.12.2043

Craig, A. D. (2003). Interoception: the sense of the physiological condition of the body. *Current Opinion in Neurobiology, 13,* 500-505. doi:10.1016/S0959-4388 (03)00090-4

Craig, A. D., & Craig, A. B. (2009). How do you feel-now? The anterior insula and human awareness. *Nature Reviews Neuroscience, 10,* 59-70. doi:10.1038/nrn2555

Critchley, H. D. (2005). Neural mechanisms of autonomic, affective, and cognitive integration. *Journal of Comparative Neurology, 493,* 154-166. doi:10.1002/cne.20 749

Critchley, H. D., Wiens, S., Rotshtein, P., Öhman, A., & Dolan, R. J. (2004). Neural systems supporting interoceptive awareness. *Nature Neuroscience, 7,* 189-195. doi:10.1038/ nn1176

Damasio, A. (2000). *The feeling of what happens: Body and emotion in the making of consciousness.* San Diego, CA: Harcourt/Harvest Books.

Daubenmier. J., Sze, J., Kerr, C.E., Kemeny, M.E., Mehling, W. (2013). Follow your breath: respiratory interoceptive accuracy in experienced meditators. *Psychophysiology, 50,* 777-789. doi:10.1111/psyp.12057

Dawson, G. & McKissick, F.C. (1984) 'Self-Recognition in Autistic Children'. *Journal of Autism and Developmental Disorders, 14,* 383-394. doi:10.1007/BF02409829

De Pascalis, V., Alberti, M. L., & Pandolfo, R. (1984). Anxiety, perception, and control of heart rate. *Perceptual and Motor Skills, 59,* 203-211. doi:10.2466/pms.1984.59.1.203

Desender, K., Van Opstal, F., Hughes, G. & Van den Bussche, E. (2016). "The temporal dynamics of metacognition: Dissociating task-related activity from later metacognitive processes." *Neuropsychologia, 8,* 54-64. doi:10.1016/j.neuropsychologia.2016.01.003

Desender, K., Van Opstal, F. & Van den Bussche, E. (2014). Feeling the Conflict: The Crucial Role of Conflict Experience in Adaptation. *Psychological Science, 25,* 1-9. doi:10.1177/0956797613511468

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology, 5,* 1-17. doi:10.3389/fpsyg.2014.00781

Dienes, Z., & Seth, A. (2010). Gambling on the unconscious: A comparison of wagering and confidence ratings as measures of awareness in an artificial grammar task. *Consciousness and Cognition, 19,* 674-681. doi:10.1016/j.concog.2009.09.009

Domes, G., Heinrichs, M., Michel, A., Berger, C., & Herpertz, S. C. (2007). Oxytocin improves "mind-reading" in humans. *Biological Psychiatry, 61,* 731-733. doi:10.1

016/j.biopsych.2006.07.015

Dretske, F. (1995). Meaningful perception. In S. M. Kosslyn & D. N. Osherson (Eds.), *An invitation to cognitive science-2nd edition: Vol. 2: Visual cognition* (331-352). Cambridge: MIT Press.

Dunn, B. D., Stefanovitch, I., Evans, D., Oliver, C., Hawkins, A., & Dalgleish, T. (2010).Can you feel the beat? Interoceptive awareness is an interactive function ofanxiety- and depression-specific symptom dimensions. *Behaviour Research and Therapy, 48*(11), 1133–1138. http://dx.doi.org/10.1016/j.brat.2010.07.006

Dunlosky, J., Kubat-Silman, A. K., & Hertzog, C. (2003). Training monitoring skills improves older adults' self-paced associative learning. *Psychology and Aging, 18,* 340-345. doi:10.1037/0882-7974.18.2.340

Dunlosky, J., & Metcalfe, J. (2009). *Metacognition.* Thousand Oaks, CA, US: Sage Publications, Inc.

Dziobek, I., Fleck, S., Kalbe, E., Rogers, K., Hassenstab, J., Brand, M., ... & Convit, A. (2006). Introducing MASC: a movie for the assessment of social cognition. *Journal of Autism and Developmental Disorders*, *36*, 623-636. doi:10.1007/s10803-006-0107-0

Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process?. *Educational Research Review, 1,* 3-14. doi :10.1016/j.edurev.2005.11.001

Ehlers, A., Breuer, P., Dohn, D., & Fiegenbaum, W. (1995). Heartbeat perception and panic disorder: Possible explanations for discrepant findings. *Behaviour Research and Therapy, 33,* 69-76. doi:10.1016/0005-7967(94)E0002-Z

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,* 1044-1052. doi:10.1007/s10803-010-1127-3

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41,* 1149-1160. doi:10.3758/BRM.41.4.1149

Faull, O. K., Cox, P. J., & Pattinson, K. T. S. (2016). Psychophysical Differences in Ventilatory Awareness and Breathlessness between Athletes and Sedentary Individuals. *Frontiers in Physiology, 7,* 231. doi:10.3389/fphys.2016.00231

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., Miller, P.H., & Miller, S.A. (1993). *Cognitive development* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Fleming, S. M., & Dolan, R. J. (2010). Effects of loss aversion on post-decision wagering: Implications for measures of awareness. *Consciousness and Cognition, 19,* 352-363. doi:10.1016/j.concog.2009.11.002

Forkmann T, Scherer A, Meessen J, Michal M, Schächinger H, Vögele C, Schulz A. (2016). Making sense of what you sense: Disentangling interoceptive awareness, sensibility and accuracy. *International Journal of Psychophysiology, 109,* 71-80. doi:10.1016/j. ijpsycho.2016.09.019

Fowler, C. J. (2003). Visceral sensory neuroscience: interoception. *Brain, 126,* 1505-1506. doi:10.1093/brain/awg120

Frazier, T. W., Ratliff, K. R., Gruber, C., Zhang, Y., Law, P. A., & Constantino, J. N. (2014). Confirmatory factor analytic structure and measurement invariance of quantitative autistic traits measured by the social responsiveness scale 2. *Autism, 18,* 31-44. doi:10.1177/1362361313500382

Frith, U. (1994). Autism and theory of mind in everyday life. *Social Development, 3,* 108-124. doi:10.1111/j.1467-9507.1994.tb00031.x

Frith, U., & Happé, F. (1999). Theory of mind and self-consciousness: What is it like to be autistic?. *Mind & Language, 14,* 82-89. 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,* 153-163. doi:10.1111/j.1469-7610.1969.tb02077.x

Garfinkel, S. N., & Critchley, H. D. (2013). Interoception, emotion and brain: new insights link internal physiology to social behaviour. *Social Cognitive and Affective Neuroscience, 8,* 231-234. doi: 10.1093/Scan/Nss140

Garfinkel, S. N., Manassei, M. F., Hamilton-Fletcher, G., den Bosch, Y. I., Critchley, H. D., & Engels, M. (2016a). Interoceptive dimensions across cardiac and respiratory axes. *Philosophical Transactions of the Royal Society B, 371,* 1-10. doi: 10.1098/rstb.2 016.0014

Garfinkel, S. N., Seth, A. K., Barrett, A. B., Suzuki, K., & Critchley, H. D. (2015). Knowing your own heart: Distinguishing interoceptive accuracy from interoceptive awareness. *Biological Psychology, 104,* 65-74. doi: 10.1016/j.biopsycho.2014.11.004 S0301- 0511(14)00229-4

Garfinkel, S. N., Tiley, C., O'Keeffe, S., Harrison, N. A., Seth, A. K., & Critchley, H. D. (2016b). Discrepancies between dimensions of interoception in autism: implications for emotion and anxiety. *Biological Psychology, 114,* 117-126. doi:10.1016/j.biopsycho .2015.12.003

Gillihan, S. J., & Farah, M. J. (2005). Is self special? A critical review of evidence from experimental psychology and cognitive neuroscience. *Psychological Bulletin, 131,* 76 -97. doi:10.1037/0033-2909.131.1.76

Goldberg, W. A., Jarvis, K. L., Osann, K., Laulhere, T. M., Straub, C., Thomas, E., ... & Spence, M. A. (2005). Brief report: Early social communication behaviors in the younger siblings of children with autism. *Journal of Autism and Developmental Disorders, 35,* 657-664. doi:10.1007/s10803-005-0009-6

Goldman, A. I. (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,* 732-764. doi:10.1080/01621459.1954.10501231

Gopnik, A. (1993). How We Know Our Minds – The Illusion of First-Person Knowledge of Intentionality. *Behavioral and Brain Sciences, 16,* 1-14. doi:10.1017/S014052 5X00028636

Gopnik, A., & Meltzoff, A. N. (2006). Minds, bodies, and persons: Young children's understanding of the self and others as reflected in imitation and theory of mind. In S. T. Parker, R. W. Mitchell, & M. L. Boccia (Eds.), *Self-awareness in animals and humans: Developmental perspectives* (pp. 166-186). Cambridge: Cambridge University Press.

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,* 650-659. doi: 10.1037/a0036531

Grainger, C., Williams, D. M., & Lind, S. E. (2016). Metacognitive monitoring and control processes in children with autism spectrum disorder: Diminished judgement of confidence accuracy. *Consciousness and Cognition, 42,* 65-74. doi:10.1016/j.conc og.2016.03.003

Hart, N., McGowan, J., Minati, L., & Critchley, H. D. (2013). Emotional regulation and bodily sensation: interoceptive awareness is intact in borderline personality disorder. *Journal of Personality Disorders, 27,* 506-518. doi:10.15 21/pedi_2012_26 _049

Hartwig, M. K., & Dunlosky, J. (2012). Study strategies of college students: Are self-testing and scheduling related to achievement?. *Psychonomic Bulletin & Review, 19,* 126-134. doi: 10.3758/s13423-011-0181-y.

Hatfield, T. R., Brown, R. F., Giummarra, M. J., & Lenggenhager, B. (2017). Autism spectrum disorder and interoception: Abnormalities in global integration?. *Autism,* E-pub ahead of print, November 15. doi:10.1177/1362361317738392

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology, 57,* 243-259. doi:10.2307/1416950

Herbert, B. M., Herbert, C., & Pollatos, O. (2011). On the relationship between interoceptive awareness and alexithymia: is interoceptive awareness related to emotional awareness?. *Journal of Personality, 79,* 1149-1175. doi: 10.1111/ j.1467-6494.2011.00717.x

Herbert, B. M., Pollatos, O., & Schandry, R. (2007). Interoceptive sensitivity and emotion processing: an EEG study. *International Journal of Psychophysiology, 65,* 214-227. doi:10.1016/j.ijpsycho.2007.04.007

Herbert, B. M., Ulbrich, P., & Schandry, R. (2007). Interoceptive sensitivity and physical effort: Implications for the self-control of physical load in everyday life. *Psychophysiology, 44,* 194-202. doi:10.1111/j.1469-8986.2007.00493.x

Higgins, S., Katsipataki, M., Kokotsaki, D., Coe, R., Major, L. E., & Coleman, R. (2013). *The Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit: Technical Appendices*. London, UK:Education Endowment Foundation

Hobson, R. P. (1990). On the origins of self and the case of autism. *Development and Psychopathology, 2,* 163-181. doi:10.1017/S0954579400000687

Howlin, P., Goode, S., Hutton, J., & Rutter, M. (2004). Adult outcome for children with autism. *Journal of Child Psychology and Psychiatry, 45,* 212-229. doi:10.1111/j.1469-7610.2004.00215.x

Iuculano, T., Rosenberg-Lee, M., Supekar, K., Lynch, C. J., Khouzam, A., Phillips, J., ... & Menon, V. (2014). Brain organization underlying superior mathematical abilities in children with autism. *Biological Psychiatry, 75,* 223-230. doi:10.1016/j.biopsych.2013.06.018

James, W. (1890). *The principles of psychology*. New York, NY: Holt.

JASP Team. (2016). *JASP* (Version 0.8) [Computer software]. Amsterdam: University of Amsterdam.

Jeffreys, H. (1961). *Theory of probability*. Oxford, UK: Oxford University Press.

Kell, H. J., Lubinski, D., & Benbow, C. P. (2013). Who rises to the top? Early indicators. *Psychological Science, 24,* 648-659. doi:10.1177/0956797612457784

Kerig, P., Ludlow, A., & Wenar, C. (2012). *Developmental psychopathology: From infancy through adolescence.* Berkshire, UK: McGraw-Hill.

Khalsa, S.S., Rudrauf, D., Damasio, A.R., Davidson, R.J., Lutz, A., Tranel, D. (2008). Interoceptive awareness in experienced meditators. *Psychophysiology, 45,* 671-677. doi:10.1111/j.1469-8986.2008.00666.x

Khalsa, S. S., Rudrauf, D., Feinstein, J. S., & Tranel, D. (2009). The pathways of interoceptive awareness. *Nature Neuroscience*, *12*(12), 1494. doi: 10.1038/nn.2411

Koch, C., & Preuschoff, K. (2007). Betting the house on consciousness. *Nature neuroscience, 10,* 140-141. doi:10.1038/nn0207-140

Koriat, A. (2007). Metacognition and consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *The Cambridge handbook of consciousness* (pp. 289-325). New York, NY: Cambridge University Press.

Kornell, N., Son, L. K., & Terrace, H. S. (2007). Transfer of metacognitive skills and hint seeking in monkeys. *Psychological Science*, *18*, 64-71. doi:10.1111/j.1467-9280.2007.01850.x

Kreiser, N. L., & White, S. W. (2014). ASD in females: are we overstating the gender difference in diagnosis?. *Clinical Child and Family Psychology Review*, *17*, 67-84. doi:10.1007/s10567-013-0148-9

Lai, M. C., Lombardo, M. V., Ruigrok, A. N., Chakrabarti, B., Auyeung, B., Szatmari, P., ... & MRC AIMS Consortium. (2017). Quantifying and exploring camouflaging in men and women with autism. *Autism*, *21*, 690-702. doi:10.1177/1362361316671012

Le Couteur, A., Bailey, A., Goode, S., Pickles, A., Gottesman, I., Robertson, S., & Rutter, M. (1996). A broader phenotype of autism: the clinical spectrum in twins. *Journal of Child Psychology and Psychiatry*, *37*, 785-801. doi:10.1111/j.1469-7610.1996.tb01475.x

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., ... & Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: the Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology*: *Applied*, *8,* 75-84. doi:10.1037/1076-898X.8.2.75

Leslie, A. M. (1987). Pretense and representation: The origins of “theory of mind”. *Psychological Review*, *94,* 412-426. doi:10.1037/0033-295X.94.4.412

Lewis, M. (1995). Aspects of self: From systems to ideas. In P. Rochat (Ed.), *The self in infancy: Theory and research* (pp. 95-115). Amsterdam, Netherlands: North Holland/Elsevier Science. doi:10.1016/S0166-4115(05)80008-1

Lind, S. E. (2010). Memory and the self in autism: A review and theoretical framework. *Autism*, *14,* 430-456. doi:10.1177/1362361309358700

Lind, S.E. & Bowler, D.M. (2008). Episodic memory and autonoetic consciousness in autistic spectrum disorders: The roles of self-awareness, representational abilities and temporal cognition. In J. Boucher & D. Bowler (Eds.), *Memory in autism: Theory and evidence* (pp. 166-187). Cambridge: Cambridge University Press. doi:10.1017/ CBO978051149 0101

Lind, S.E. & Bowler, D.M. (2009). Delayed Self-Recognition in Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders, 39,* 643–650. doi:10.1007/s10803-008-0670-7

Livingston, L. A., & Happé, F. (2017). Conceptualising compensation in neurodevelopmental disorders: reflections from autism spectrum disorder. *Neuroscience & Biobehavioral Reviews, 80,* 729-742. doi:10.1016/jneubiorev.2 017.06.005

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*, 205-223. doi:10.1023/A:10055924019 47

Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., & Schopler, E. (1989). The autism diagnostic observation schedule: A standardized observation of communicative and social behavior. *Journal of Autism and Developmental Disorders, 19,* 185-212. doi:10.1007/BF02211841

McMahon, C. M., Henderson, H. A., Newell, L., Jaime, M., & Mundy, P. (2016). Metacognitive awareness of facial affect in higher-functioning children and adolescents with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 46*, 882-898. doi:10.1007/s10803-015-2630-3

Macmillan, N. A., & Kaplan, H. L. (1985). Detection theory analysis of group data: estimating sensitivity from average hit and false-alarm rates. *Psychological Bulletin, 98,* 185-199. doi:10.1037/0033-2909.98.1.185

Mandy, W., & Lai, M. C. (2017). Towards sex-and gender-informed autism research. *Autism, 21,* 643-645. doi:10.1177/1362361317706904

Maras, K., Gamble, T., & Brosnan, M. (2017). Supporting metacognitive monitoring in mathematics learning for young people with autism spectrum disorder: A classroom-based study. *Autism,* 1-11. doi:10.1177/1362361317722028

Marti, S., Sackur, J., Sigman, M., and Dehaene, S. (2010). Mapping introspection's blind spot: reconstruction of dual-task phenomenology using quantified introspection. *Cognition, 115,* 303-313. doi: 10.1016/j.cognition.2010.01.003

Martinez, G., Alexandre, C., Mam-Lam-Fook, C., Bendjemaa, N., Gaillard, R., Garel, P., ... & Krebs, M. O. (2017). Phenotypic continuum between autism and schizophrenia: evidence from the movie for the assessment of social cognition (MASC). *Schizophrenia Research, 185,* 161-166. doi:10.1016/j.schres.2017.01.012

Mayes, S. D., & Calhoun, S. L. (2006). Frequency of reading, math, and writing disabilities in children with clinical disorders. *Learning and individual Differences, 16,* 145-157. doi:10.1016/j.lindif.2005.07.004

McCartney, K., Burchinal, M. R., & Bub, K. L. (2006). Introduction to the monograph. *Monographs of the Society for Research in Child Development, 71,* 1-8. doi:10.1111/j.1540-5834.2006.07103001.x

McGeer, V. (2004). Autistic self-awareness. *Philosophy, Psychiatry, & Psychology, 11,* 235-251. doi:10.1353/ppp.2004.0066

McManus, S., Bebbington, P., Jenkins, R., & Brugha, T. (2016). *Mental Health and Wellbeing in England: Adult Psychiatric Morbidity Survey 2014: a Survey Carried Out for NHS Digital by NatCen Social Research and the Department of Health Sciences, University of Leicester*. NHS Digital.

Minshew, N. J., Goldstein, G., Taylor, H. G., & Siegel, D. J. (1994). Academic achievement in high functioning autistic individuals. *Journal of Clinical and Experimental Neuropsychology, 16,* 261-270. doi:10.1080/01688639408402637

Moore, C., and Frye, D. (1991). The Acquisition and Utility of Theories of Mind. In D. Frye and C. Moore (Eds.), *Children's Theories of Mind: Mental States and Social Understanding* (pp. 1-14). Mahwah, NJ: Erlbaum.

Mul, C. L., Stagg, S. D., Herbelin, B., & Aspell, J. E. (2018). The Feeling of Me Feeling for You: Interoception, Alexithymia and Empathy in Autism. *Journal of Autism and Developmental Disorders, 48,* 1-15. doi: 10.1007/s10803-018-3564-3.

Murphy, M., Bolton, P. F., Pickles, A., Fombonne, E., Piven, J., & Rutter, M. (2000). Personality traits of the relatives of autistic probands. *Psychological Medicine, 30,* 1411-1424. doi:10.1017/S0033291799002949

Murphy J, Catmur C, Bird G. (2018a). Alexithymia is associated with a multidomain, multidimensional failure of interoception: Evidence from novel tests. *Journal of Experimental Psychology: General, 147,* 398-408. doi:10.1037/xge0000366

Murphy, J., Geary, H., Millgate, E., Catmur, C., & Bird, G. (2018b). Direct and indirect effects of age on interoceptive accuracy and awareness across the adult lifespan. *Psychonomic Bulletin & Review, 25,* 1193-1202. doi:10.3758/s13423-017-1339-z

Murphy, J., Millgate, E., Geary, H., Ichijo, E., Coll, M. P., Brewer, R., ... & Bird, G. (2018c). Knowledge of resting heart rate mediates the relationship between intelligence and the heartbeat counting task. *Biological Psychology, 133*, 1-3. doi:10.1016/j.biopsycho.2018.01.012

Murphy, M. D., Schmitt, F. A., Caruso, M. J., & Sanders, R. E. (1987). Metamemory in older adults: the role of monitoring in serial recall. *Psychology and Aging, 2,* 331-339. doi:10.1037/0882-7974.2.4.331

Mussey, J. L., Ginn, N. C., & Klinger, L. G. (2017). Are males and females with autism spectrum disorder more similar than we thought?. *Autism, 21*, 733-737. doi:10.1177 /1362361316682621

Nation, K., Clarke, P., Wright, B., & Williams, C. (2006). Patterns of reading ability in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 36,* 911-919. doi:10.1007/s10803-006-0130-1

Neisser, U. (1988). Five kinds of self-knowledge. *Philosophical Psychology, 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*, 109-133. doi:10.1037/0033-2909.95.1.109

Nelson, T. O., & Leonesio, R. J. (1988). Allocation of self-paced study time and the “labor-in-vain effect”. *Journal of Experimental Psychology: Learning, Memory, and Cognition,14*(4), 676-686.

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1-25). Cambridge, MA: MIT Press.

Nelson, T. O., Narens, L., & Dunlosky, J. (2004). A revised methodology for research on metamemory: Pre-judgment Recall and Monitoring (PRAM). *Psychological methods, 9,* 53-69. doi:10.1037/1082-989X.9.1.53

Nemiah, J. C., Freyberger, H., & Sifneos, P. E. (1976). Alexithymia: A view of the psychosomatic process. In O. W. Hill (Ed.), *Modern trends in psychosomatic medicine* (Vol. 3; pp. 430-439). London, UK: Butterworths.

Nicholson, T. M., Williams, D. M., Grainger, C., Christensen, J. F., Calvo-Merino, B., & Gaigg, S. B. (2018). Interoceptive impairments do not lie at the heart of autism or alexithymia. *Journal of Abnormal Psychology, 127,* 612-622. doi:10.1037/abn0000370

Nichols, S., & Stich, S. P. (2003). *Oxford cognitive science series. Mindreading: An integrated account of pretence, self-awareness, and understanding other minds.* New York, NY, US: Clarendon Press/Oxford University Press. doi:/10.1093/0198236107.00 1.0001

Nielsen, M., Suddendorf, T. & Dissanayake, C. (2006). Imitation and Self-Recognition in Autism. In S. Rogers & J.W.Williams (Eds.), *Imitation and the Development of the Social Mind: Autism and Typical Development* (pp. 138-156). New York, NY: Guilford Press.

Oakley, B.F.M., Brewer, R., Bird, G. & Catmur, C. (2016). Theory of mind is not theory of emotion: A cautionary note on the Reading the Mind in the Eyes Test. *Journal of Abnormal Psychology, 125,* 818-823. doi:10.1037/abn0000182

O'connor, I. M., & Klein, P. D. (2004). Exploration of strategies for facilitating the reading comprehension of high-functioning students with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 34*, 115-127. doi:10.1023/B:JADD.000002260 3.44077.6b

O'hara, M. W., & Swain, A. M. (1996). Rates and risk of postpartum depression-a meta-analysis. *International Review of Psychiatry, 8,* 37-54. doi:10.3109/0954026960 9037816

Palser, E. R., Fotopoulou, A., Pellicano, E., & Kilner, J. M. (2018). The link between interoceptive processing and anxiety in children diagnosed with autism spectrum disorder: Extending adult findings into a developmental sample. *Biological Psychology, 136,* 13-21. doi:10.1016/j.biopsycho.2018.05.003

Parker, J. D., Taylor, G. J., & Bagby, R. M. (2003). The 20-Item Toronto Alexithymia Scale: III. Reliability and factorial validity in a community population. *Journal of Psychosomatic Research, 55,* 269-275. doi:10.1016/S0022-3999 (02)0057 8-0

Paulus, M., Proust, J., & Sodian, B. (2013). Examining implicit metacognition in 3.5-year-old children: An eye-tracking and pupillometric study. *Frontiers in Psychology, 4,* 1-7 doi:10.3389/fpsyg.2013.00145

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

Persaud, N., & McLeod, P. (2008). Wagering demonstrates subconscious processing in a binary exclusion task. *Consciousness and Cognition, 17,* 565-575. doi:10.1016 /j.conco g.2007.05.003

Peterson, C. C., Wellman, H. M., & Liu, D. (2005). Steps in theory-of-mind development for children with deafness or autism. *Child Development, 76,* 502-517. doi:10.1111/j.1467-8624.2005.00859.x

Philip, R. C. M., Whalley, H. C., Stanfield, A. C., Sprengelmeyer, R., Santos, I. M., Young, A. W., ... & Hall, J. (2010). Deficits in facial, body movement and vocal emotional processing in autism spectrum disorders. *Psychological Medicine, 40,* 1919-1929. doi:10.1017/S0033291709992364

Pishghadam, R. & Khajavy, G. H. (2013). Intelligence and metacognition as predictors of foreign language achievement: A structural equation modeling approach. *Learning and Individual Differences, 24,* 176-181. doi:10.1016/j.lindif.2012.12.004

Pickles, A., Starr, E., Kazak, S., Bolton, P., Papanikolaou, K., Bailey, A., ... & Rutter, M. (2000). Variable expression of the autism broader phenotype: findings from extended pedigrees. *The Journal of Child Psychology and Psychiatry and Allied Disciplines, 41,* 491-502. doi:10.1111/1469-7610.00634

Piven, J., Palmer, P., Landa, R., Santangelo, S., Jacobi, D., & Childress, D. (1997). Personality and language characteristics in parents from multiple-incidence autism families. *American Journal of Medical Genetics, 74,* 398-411. doi:10.1002/(SICI)1096-8628(19970725)74:4<398::AID-AJMG11>3.0.CO;2-D

Piven, J., Wzorek, M., Landa, R., Lainhart, J., Bolton, P., Chase, G. A., & Folstein, S. (1994). Personality characteristics of the parents of autistic individuals. *Psychological Medicine, 24,* 783-795. doi:10.1017/S0033291700027938

Pollatos, O., Gramann, K., & Schandry, R. (2007). Neural systems connecting interoceptive awareness and feelings. *Human Brain Mapping, 28,* 9-18. doi:10.1002/hbm.20258

Pollatos, O., Traut-Mattausch, E., & Schandry, R. (2009). Differential effects of anxiety and depression on interoceptive accuracy. *Depression and anxiety, 26*(2), 167-173.

Porges, S. (1993). *Body perception questionnaire*. Laboratory of Developmental Assessment, University of Maryland.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind?. Behavioral and Brain Sciences, 1, 515-526. doi:10.1017/ S0140 525X00076512

Pylyshyn, Z. W. (1978). When is attribution of beliefs justified? *The Behavioral and Brain Sciences, 1,* 592-593. doi:10.1017/S0140525X00076895

Quattrocki, E., & Friston, K. (2014). Autism, oxytocin and interoception. *Neuroscience & Biobehavioral Reviews, 47,* 410-430. doi:10.1016/j.neubiorev.2014.09.012

Reed, P., Lowe, C., & Everett, R. (2011). Perceptual learning and perceptual search are altered in male university students with higher Autism Quotient scores. *Personality and Individual Differences, 51,* 732-736. doi:10.1016/j.paid.2011.06.016

Riby, D. M., Doherty-Sneddon, G., & Bruce, V. (2008). Exploring face perception in disorders of development: Evidence from Williams syndrome and autism. *Journal of Neuropsychology*, *2*, 47-64. doi:10.1348/174866407X255690

Riby, D. M., & Hancock, P. J. (2008). Viewing it differently: Social scene perception in Williams syndrome and autism. *Neuropsychologia*, *46*, 2855-2860. doi:10.1016/j.neuropsychologia.2008.05.003

Rochat, P. (2003). Five levels of self-awareness as they unfold early in life. *Consciousness and Cognition, 12,* 717-731. doi:10.1016/S1053-8100(03)00081-3

Roebers, C. M., Cimeli, P., Röthlisberger, M., & Neuenschwander, R. (2012). Executive functioning, metacognition, and self-perceived competence in elementary school children: An explorative study on their interrelations and their role for school achievement. *Metacognition and Learning, 7,* 151-173. doi:10.1007/s11409-012-9089-9

Roebers, C. M., Krebs, S. S., & Roderer, T. (2014). Metacognitive monitoring and control in elementary school children: Their interrelations and their role for test performance. *Learning and Individual Differences*, *29*, 141-149. doi:10.1016/j.lindif.2012.12.003

Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology, 35*, 677-688. doi:10.1037/0022-3514.35.9.677

Ronald, A., Happé, F., Price, T. S., Baron-Cohen, S., & Plomin, R. (2006). Phenotypic and genetic overlap between autistic traits at the extremes of the general population. *Journal of the American Academy of Child & Adolescent Psychiatry, 45*, 1206-1214. doi:10.1097/01.chi.0000230165.54117.41

Rouse, C. H., Jones, G. E., & Jones, K. R. (1988). The effect of body composition and gender on cardiac awareness. *Psychophysiology, 25,* 400-407. doi:10.1111/j.1469-8986.1988.tb01876.x

Ruzich, E., Allison, C., Smith, P., Watson, P., Auyeung, B., Ring, H., & Baron-Cohen, S. (2015). Measuring autistic traits in the general population: a systematic review of the autism-spectrum Quotient (AQ) in a nonclinical population sample of 6,900 typical adult males and females. *Molecular Autism, 6,* 1-12. doi: 10.1186/2040-2392-6-2

Rydén, G., Rydén, E., & Hetta, J. (2008). Borderline personality disorder and autism spectrum disorder in females: A cross-sectional study. *Clinical Neuropsychiatry, 5,* 22-30. doi:10.1177/1754073912469685

Sawyer, A. C., Williamson, P., & Young, R. (2014). Metacognitive processes in emotion recognition: Are they different in adults with Asperger's disorder?. *Journal of Autism and Developmental Disorders, 44,* 1373-1382. doi:10.1007/s10803-013-1999-0

Schandry, R. (1981). Heart beat perception and emotional experience. *Psychophysiology, 18,* 483-488. doi: 10.1111/j.1469-8986.1981.tb02486.x

Schneider, W. (2008). The Development of Metacognitive Knowledge in Children and Adolescents: Major Trends and Implications for Education. *Mind, Brain, and Education, 2,* 114-121. doi:10.1111/j.1751-228X.2008.00041.x

Schneider, D., Slaughter, V. P., Bayliss, A. P., & Dux, P. E. (2013). A temporally sustained implicit theory of mind deficit in autism spectrum disorders. *Cognition, 129,* 410-417. doi:10.1016/j.cognition.2013.08.004

Seigel, J. (2005). *The Idea of the Self: Thought and Experience in Western Europe since the Seventeenth Century*. Cambridge, UK: Cambridge University Press. doi:10.1017/C BO9780780511818141

Senju, A. (2012). Spontaneous theory of mind and its absence in autism spectrum disorders. *The Neuroscientist, 18,* 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*, 883-885. doi:10.1126/science.1176170

Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. *Trends in Cognitive Sciences, 17,* 565-573. doi:10.1016/j.tics.2013.09.007

Shah, P., Hall, R., Catmur, C., & Bird, G. (2016). Alexithymia, not autism, is associated with impaired interoception. *Cortex, 81,* 215-220. doi:10.1016/j.cortex.2016.0 3.021

Shah, P., Catmur, C., & Bird, G. (2017). From heart to mind: linking interoception, emotion, and theory of mind. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior, 93,* 220-223. doi:10.1016/j.cortex.2017.02.010

Sherrington, C.S. (1906). *The Integrative Action of the Nervous System.* New York, NY: Scribners.

Sifneos, P.E. (1973). The prevalence of 'alexithymic' characteristics in psychosomatic patients. *Psychotherapy and Psychosomatics, 22,* 255-262. doi: 10.1159/00 0 286529

Simonoff, E., Pickles, A., Charman, T., Chandler, S., Loucas, T., & Baird, G. (2008). Psychiatric disorders in children with autism spectrum disorders: prevalence, comorbidity, and associated factors in a population-derived sample. *Journal of the American Academy of Child & Adolescent Psychiatry, 47,* 921-929. doi:10.1097/C HI.0b013e318179964f

Smith, J. D., Shields, W. E., & Washburn, D. A. (2003). The comparative psychology of uncertainty monitoring and metacognition. *Behavioral and Brain Sciences, 26,* 317-339. doi:10.1017/S0140525X03000086

Smith, J. D., & Washburn, D. A. (2005). Uncertainty monitoring and metacognition by animals. *Current Directions in Psychological Science, 14,* 19-24. doi:10.1111 /j.0963-7214.2005.00327.x

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117,* 34-50. doi:10.1037/0096-3445.117.1.34

Son, L. K., & Kornell, N. (2005). Metaconfidence judgments in rhesus macaques: Explicit versus implicit mechanisms. In Terrace, H. S. & Metcalfe, J. (Eds.), *The missing link in cognition: Origins of self-reflective consciousness* (pp. 296-320). New York, NY: Oxford University Press.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press

Suddendorf, T. & Whiten, A. (2001). Mental evolution and development: Evidence for secondary representation in children, great apes and other animals. *Psychological Bulletin, 127,* 629-650. doi:10.1037/0033-2909.127.5.629

Sui, J., He, X., & Humphreys, G. W. (2012). Perceptual effects of social salience: evidence from self-prioritization effects on perceptual matching. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 1105-1117. doi:10.1037/a0029792

Sun, Y., Fuentes, L. J., Humphreys, G. W., & Sui, J. (2016). Try to see it my way: Embodied perspective enhances self and friend-biases in perceptual matching. *Cognition, 153,* 108-117. doi:10.1016/j.cognition.2016.04.015

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: a meta-analysis. Psychological Bulletin, 121, 371-394. doi:10.1037/0033-2909.1 21.3.371

Szatmari, P., Georgiades, S., Duku, E., Zwaigenbaum, L., Goldberg, J., & Bennett, T. (2008). Alexithymia in parents of children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 38,* 1859-1865. doi:10.1007/s10803-008-0576-4

Szatmari, P., MacLean, J. E., Jones, M. B., Bryson, S. E., Zwaigenbaum, L., Bartolucci, G., ... & Tuff, L. (2000). The familial aggregation of the lesser variant in biological and nonbiological relatives of PDD probands: a family history study. *The Journal of Child Psychology and Psychiatry and Allied Disciplines, 41,* 579-586. doi:10.1111/1469-7610.00644

Tager-Flusberg, H. (1999). A psychological approach to understanding the social and language impairments in autism. *International Review of Psychiatry, 11,* 325-334. doi:10.1080/09540269974203

Teong, S. K. (2003). The effect of metacognitive training on mathematical word-problem solving. *Journal of Computer Assisted Learning, 19,* 46-55. doi:10.1046/j.0266-4909.2003.00005.x

Terasawa, Y., Moriguchi, Y., Tochizawa, S., & Umeda, S. (2014). Interoceptive sensitivity predicts sensitivity to the emotions of others. *Cognition and Emotion, 28,* 1435-1448. doi:10.1080/02699931.2014.888988

Thiede, K.W., Anderson, M.C., & Therriault, D.J. (2003). Accuracy of Metacognitive Monitoring Affects Learning of Texts. *Journal of Educational Psychology, 95,* 66-73 doi:10.1037/0022-0663.95.1.66

Turner, R., & Felisberti, F. M. (2017). Measuring mindreading: A review of behavioral approaches to testing cognitive and affective mental state attribution in neurologically typical adults. *Frontiers in Psychology, 8,* 1-7. doi:10.3389/fpsyg.2017.00047

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* (4157), 1124-1131.

Uddin, L. Q. (2011). The self in autism: an emerging view from neuroimaging. *Neurocase, 17,* 201-208. doi:10.1080/13554794.2010.509320

Veenman, M. V. J. & Spaans, M. A. (2005). Relation between intellectual and metacognitive skills: Age and task differences. *Learning and Individual Differences, 15,* 159-176 doi:10.1016/j.lindif.2004.12.001

Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., ... & Meerhoff, F. (2016). Bayesian inference for psychology. part II: Example applications with JASP. *Psychonomic Bulletin & Review, 25,* 1-19. doi:10.3758/s13423-017-1323-7

Washburn, D., Gulledge, J., Beran, M., and Smith, J.D. (2010). With his memory magnetically erased, a monkey knows he is uncertain. *Biology Letters, 6,* 160-162. doi:10.1098/rsbl.2009.0737

Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence*. New York, NY: The Psychological Corporation: Harcourt Brace & Company.

Wegner, D. M. (2002). *The illusion of conscious will.* Cambridge, MA: MIT Press.

Wells, A., & Cartwright-Hatton, S. (2004). A short form of the metacognitions questionnaire: properties of the MCQ-30. *Behaviour Research and Therapy, 42,* 385-396. doi:10.1016/S0005-7967(03)00147-5

Wenke, D., Fleming, S. M., & Haggard, P. (2010). Subliminal priming of actions influences sense of control over effects of action. *Cognition, 115,* 26-38. doi:10.1016/j.cognition. 2009.10.016

Wiens, S. (2005). Interoception in emotional experience. *Current Opinion in Neurology, 18,* 442-447. doi:10.1097/01.wco.0000168079.92106.99

Wiens, S., Mezzacappa, E. S., & Katkin, E. S. (2000). Heartbeat detection and the experience of emotions. *Cognition & Emotion, 14,* 417-427. doi:10.1080/026 999300378905

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,* 1371-1377. doi:10.1007/s10803-010-0995-x

Williams, D. (2010). Theory of own mind in autism: Evidence of a specific deficit in self-awareness?. *Autism, 14,* 474-494. doi:10.1177/1362361310366314

Williams, D. M., Bergström, Z., & Grainger, C. (2016). Metacognitive monitoring and the hypercorrection effect in autism and the general population: Relation to autism (-like) traits and mindreading. *Autism, 22,* 259-270. doi:/10.1177/1362361316680178

Williams, D.M. & Happé, F. (2009). Pre-Conceptual Aspects of Self-Awareness in Autism Spectrum Disorder: The Case of Action-Monitoring. *Journal of Autism and Developmental Disorders, 39,* 251-259. doi:10.1007/s10803-008-0619-x

Williams, D. M., Nicholson, T., & Grainger, C. (2018a). The Self-Reference Effect on Perception: Undiminished in Adults with Autism and No Relation to Autism Traits. *Autism Research, 11,* 331-341. doi:10.1002/aur.1891

Williams, D. M., Nicholson, T., Grainger, C., Lind, S. E. and Carruthers, P. (2018b). Can you spot a liar? Deception, mindreading, and the case of autism spectrum disorder. *Autism Research.* doi:10.1002/aur.1962

Wilson, T. D. (2004). *Strangers to ourselves*. USA: Harvard University Press.

Wojcik, D. Z., Allen, R. J., Brown, C., & Souchay, C. (2011). Memory for actions in autism spectrum disorder. *Memory, 19,* 549-558. doi:10.1080/09658211.2011.590506

Woodbury-Smith, M. R., Robinson, J., Wheelwright, S., & Baron-Cohen, S. (2005). Screening adults for Asperger syndrome using the AQ: A preliminary study of its diagnostic validity in clinical practice. *Journal of Autism and Developmental Disorders, 35,* 331-335. doi:10.1007/s10803-005-3300-7

World Health Organization. (1993). *The ICD-10 classification of mental and behavioural disorders: diagnostic criteria for research (Vol. 2)*. World Health Organization.

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*, 283-307. doi:10.103 7/0033-2909.124.3.283

Footnotes

[1] Six participants scored below seven on ADOS, which is the ASD cut-off point based on Lord et al.'s criteria (1989). However, three of them were females who scored well above the cut-off on the Autism Quotient (AQ; ≥ 26) and based on recent findings about diagnostic bias against females, they were not excluded from the analyses reported below (Mandy & Lai, 2017; Mussey, Ginn, & Klinger, 2017). The rest (three males) also remained in the sample because results did not change substantially, when excluded.[1] Four participants with ASD scored below the AQ cut-off score, but all of them scored above the ASD cut-off on ADOS and therefore were not excluded from the analyses. All TD participants scored below the AQ cut-off score. Results from the main analyses of Experiment 1 did not change substantially when we excluded male participants with ASD who scored below the cut-off score on ADOS ($n = 3$). An independent sample t-test revealed no significant between-group differences in interoceptive accuracy, either in cardiac, $t(37) = -0.05$, $p = .96$, $d = 0.02$, $BF_{10} = 0.31$, or in respiratory domain, $t(37) = -0.06$, $p = .95$, $d = 0.02$, $BF_{10} = 0.31$. In addition, an independent sample t-test showed no significant between-group differences in interoceptive awareness, either in cardiac, $t(34) = -1.27$, $p = .21$, $d = 0.42$, $BF_{10} = 0.60$, or in respiratory domain, $t(28) = -0.49$, $p = .63$, $d = 0.18$, $BF_{10} = 0.38$. A series of one-sample t-tests showed that the mean coefficient for both cardiac and respiratory interoceptive awareness was significant at chance level in the ASD group, $t(17) = 0.20$, $p = .85$, $d = 0.05$, $BF_{10} = 0.25$ and $t(11) = 1.34$, $p = .21$, $d = 0.39$, $BF_{10} = 0.59$ respectively.

[2] Initially, it was designed all participants to start the task having their heartrate elevated. Before the administration of the task, we would measure participants' resting heartrate, and then we would ask them to climb some stairs in order to elevate their heartrate to a maximum of 90 beats per minute. Once their heartrate was elevated, they would start the task. This procedure was designed to be followed for all participants who

did not have heart conditions. However, after we employed it upon 8 participants with ASD and 9 comparisons, it was evident this this would not serve the purposes of the initial design, as participant heartrate dropped to the baseline faster than we expected. Thus, following Professor Williams's indications we omitted this procedure for the rest of our sample.

[3] Participants competed a self-reference task, the awareness subscale of the Body Perception Questionnaire (Porges, 1993), the ToM clips of the animations task (Abell et al., 2000) as part of a wider project entitled "Metacognition and mindreading: One system or two?", with principal investigator Professor David Williams. In addition, participants completed two metacognitive tasks, and the Balloon Analogue Risk Task (BART; Lejuez et al., 2002) as part of a PhD project. These measurements were not central for the purposes of the current experiment and thus they were not included in the analyses.

[4] In addition, participants competed the TAS-20 (Bagby et al., 1994), as part of the wider project "Metacognition and mindreading: One system or two?", but as it was not central for the purposes of the current project we did not included it in the analyses.

[5] Results from the main analyses of Experiment 2.2 did not change substantially when we excluded participants with ASD who scored below the cut-off score on ADOS ($n = 4$). An independent sample *t*-test revealed that participants with ASD showed significantly lower gamma scores in the implicit version of the gambling task compared to neurotypical participants, $t(40) = -1.94$, $p = .03$, $d = 0.59$, $BF_{10} = 2.53$ (one-tailed). Nonetheless, there were no significant between-group differences in gamma scores in the explicit version of the gambling task, $t(40) = -0.31$, $p = .38$, $d = 0.10$, $BF_{10} = 0.38$ (one-tailed).

[6] In addition, participants completed the Meta-cognitions Questionnaire (MCQ; Wells & Cartwright-Hatton, 2004), the Empathy Quotient (EQ; Baron-Cohen & Wheelwright, 2004) and the 20 Item Toronto Alexithymia Scale (TAS-20; Bagby et al.,

1994), as part of the wider project “Metacognition and mindreading: One system or two?”. These measurements were not central for the purposes of the current experiment and thus they were not included in the analyses.

[7] Both Experiment 1.1 and Experiment 3.2 include the same sample. Participants completed all the tasks described in the method section of each experiment in one testing session. However, data was analysed separately for each of these experiments.

[8] Results from the main analysis of Experiment 3.2 did not change substantially when we excluded all male participants with ASD who scored below the cut-off score on ADOS ($n = 2$). An independent sample t-test revealed no significant between-group differences in accuracy of metacognitive experiences of difficulty, $t(36) = 0.53$, $p = .30$, $d = 0.17$, $BF_{10} = 0.23$ (one-tailed).

Appendix A

Experiment 4.1: Individual Differences

The self-reference effect (Rogers, Kuiper, & Kirker, 1977) was also explored to enhance our understanding of lower order phenomena of self-awareness in ASD. That is the phenomenon whereby information related to ones' own self is encoded more deeply to information related to other people. This explains why individuals show superior memory and accurate perceptual judgments for information related to themselves over information related to others (Sui, He, & Humphreys, 2012; Symons & Johnson, 1997; Tversky & Kahneman, 1974). Given that in a recently published study, it was found that individuals with ASD show the typical self-reference effect in perception, providing evidence that lower order phenomena of self-awareness are intact in ASD (Williams et al., 2018b), we aimed to further examine this phenomenon. Sun, Fuentes, Humphreys and Sui (2016) found that the self-referent effect in perception was enhanced when stimuli were presented in an embodied with individuals perspective. Thus, we initially aimed to replicate these findings and examine its relation to autism traits and mindreading abilities in the general population and then to explore this effect among individuals with ASD. Nonetheless, our findings indicated that the embodied perspective did not enhance the self-reference effect. Therefore, we did not precede with employing this task among individuals with ASD. The method we followed to replicate Sun et al.'s (2016) study as well as our basic findings are both described briefly below.

Experiment 4.1: Method

Participants

Fifty-two undergraduate students (40 female) participated in this study, after they had given written, informed consent. Their average age was 19.63 (*SD* = 2.33; range = 18 to 30) years. Participants were recruited via the Research Participation Scheme (RPS) of the University of Kent, UK and they were offered course credits in partial fulfilment of their degree, as compensation for their participation in the experiment. Ethical approval for this study was obtained from the Kent School Psychology Research Ethics Committee (Ethics ID: 201815185209334939).

Materials, Procedures, and Scoring

As in all the experiments of the current project, participants completed the RMIE (Baron-Cohen et al., 2001a), and the AQ (Baron-Cohen et al., 2001b). In addition, they performed the colours task, as described by Sun et al., (2016).

Colours task.

Participants first were instructed to associate three different colours (red, blue, and green) with three different person labels. These labels related to either themselves (“You”), a named best friend (“Friend”), or an unknown person (“Stranger”). Each colour was associated with a different person label. The start of each experimental trial was signalled with a fixation cross appearing on 22-in computer screen for 500ms. Subsequently, two avatars were presented for 300ms, before a person label (“You”, “Friend”, or “Stranger”) appeared in between the two avatars for 300ms. One of the avatars was being displayed in first-person perspective, having its back to the participant, while the other one was being presented in third-person perspective, facing the participant. The target avatar was wearing a t-shirt coloured either blue, green, or red, while the other one was wearing a shallow grey t-shirt. Following this, a blank screen was presented for a variable time between 800 and 1200ms. Within this time frame, participants had to respond to whether the contingency between the target avatar’s t-shirt colour and the person label matched that learnt in the training phase or mismatched, by pressing either the “c” key of a keyboard for a match or the “m” key for a mismatch. After each trial, participants received feedback (correct, incorrect, or slow) based upon their response (see Figure A1 for a graphical representation of the task). Following a minimum of 12 practice trials, each participant performed three blocks of 120 experimental trials with the trial order being randomised within blocks. The task had six versions, representing each possible t-shirt colour and label combination, so each participant completed only one version of the task (counterbalanced across participants).

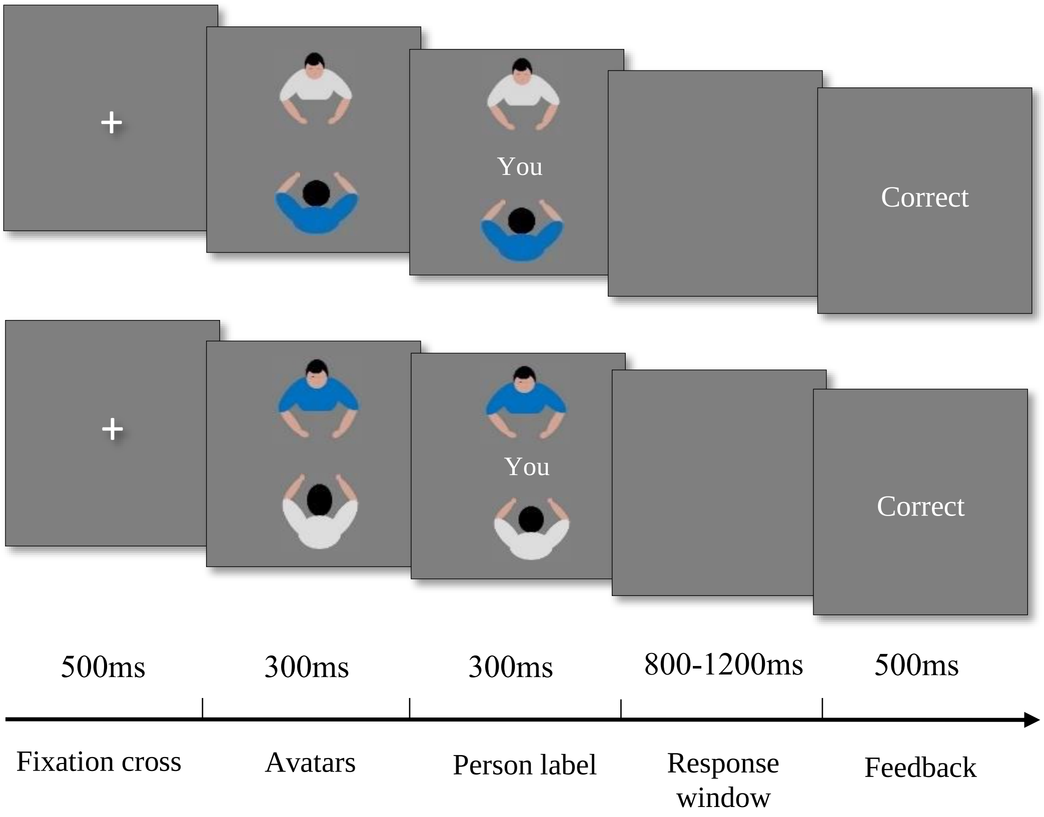

Figure A1. Graphical representation of the procedure and stimuli used in the colours task for the "You" person-label presented in first (top) and third (bottom) perspective.

Following Sun at al. (2016) we calculated d-prime scores (d′), to indicate the accuracy with which participants distinguished matches from mismatches for the three person labels per each perspective; thus six d′ scores were calculated for each participant, using the formula d′ = Z(H) – Z(FA). The higher the d′ score the greater the accuracy. We also calculated median reaction times (RTs) to the three person labels for the two different perspectives. Matches and mismatches were analysed separately. Thus, six median RTs were calculated for each participant.

Experiment 4.1: Results

A 2 (perspective: first/ third) × 3 (person label: you/ friend/ stranger) ANOVA was conducted on d' scores. The analysis revealed only a significant main effect of person, $F(2,102) = 22.13$, $p < .001$, $\eta_p^2 = .30$, with participants responding more accurately to stimuli related to themselves compared to stimuli related to other people. Crucially, neither the main effect of Perspective nor the interaction Person Label x Perspective was significant (all ps $\geq .90$, all $\eta_{pS}^2 \leq .00$), indicating no significant differences in accuracy in terms of the perspective. Next, a 2 (perspective: first/ third) × 3 (person label: you/ friend/ stranger) repeated measures ANOVA was conducted on median RTs for matches. The analysis revealed a significant main effect of person, indicating that participants responded faster to stimuli related to themselves compared to stimuli related to others, $F(2,102) = 34.40$, $p < .001$, $\eta_p^2 = .44$. In addition, contrary to Sun et al.'s (2016) findings, the analysis revealed a marginally significant main effect of perspective, indicating relatively faster responses to stimuli presented in third perspective, indicating relatively faster responses to stimuli presented in third person perspective responses to stimuli presented in third person perspective rather than in first person perspective $F(1,51) = 3.46$, $p < .07$, $\eta_p^2 = .06$. Nonetheless, the interaction between Person Label x Perspective was non-significant, $F(2,102) = 0.68$, $p = .51$, $\eta_p^2 = .01$. Finally, a 2 (perspective: first/ third) × 3 (person label: you/ friend/ stranger) ANOVA was also conducted on median RTs for mismatches. Even though there was a significant Person Label x Perspective interaction, $F(2,102) = 13.24$, $p < .001$, $\eta_p^2 = .21$, results from paired sample *t*-tests were in absolute contrast with Sun et al.'s (2016) findings. That is to say, the analysis revealed that participants responded significantly faster when the label "Friend" was presented in third person perspective than in first person perspective, $t(51) = 8.75$, $p < .001$. All the other *t*-tests produced non-significant results, all *p*s $\geq .62$.

Appendix B

A one sample *t*-test indicated that the average RMIE score of our sample (*M* = 26.16; *SD* =4.79) did not differ significantly from the average score of the general population (*M* = 26.20; *SD* =3.60), reported by Baron-Cohen et al. (2001a), *t*(55) = -0.06, *p* = .95, BF_{10} = 0.15. Likewise, the difference between the average AQ score in the current experiment (*M* = 15.75; *SD* = 6.92) and the average AQ score of the general population (*M* = 16.40; *SD* = 6.30), as reported by Baron-Cohen et al. (2001b), was not significant *t*(55) = 0.70 *p* = .49, BF_{10} = 0.19.

Appendix C

A one sample *t*-test showed that the average AQ score of current sample (*M* = 15.71, *SD* = 6.88) did not differ significantly from the average AQ score of the general population (*M* = 16.40, *SD* = 6.30), as reported by Baron-Cohen et al. (2001b), *t*(78) = -0.89, *p* = .38, BF_{10} = 0.18. An one sample *t*-test also indicated that although the average RMIE score across participants (*M* = 27.37, *SD* =3.55) was significantly higher compared to the average score of the general population (*M* = 26.20, *SD* =3.60), *t*(78) = 2.92, *p* = .005, BF_{10} = 6.28, it was not significantly higher when compared with student population (*M* = 28.00, *SD* =3.5), *t*(78) = -1.59, *p* = .12, BF_{10} = 0.41, as reported by Baron-Cohen et al. (2001a).

Acknowledgments

The author would like to thank all the participants who took part in the studies of the current project. Without the support of these people, this project would not have been possible. Thanks to Professor David Williams for inspiring discussions and useful comments on the earlier drafts of the manuscript. The author would also like to thank Dr Toby Nicholson for assistance with data analysis and Katie Carpenter for assistance with participant recruitment and data collection. This project was funded by an Economic and Social Research Council Research Grant awarded to Professor David Williams.