Running head: Sense of Self

Sense of Self: Evidence from Neurotypical people and people with Autism Spectrum Conditions

Marchella Smith

Supervised by Professor Heather Ferguson

Thesis submitted for the degree of Doctor of Philosophy

January 2023

Declaration

This work contains original research and writing produced specifically for this thesis. Reasonable care has been taken to ensure that the work does not break any UK law or infringe any Intellectual Property Right. Where information has been derived from third party copyright sources, this has been acknowledged in the thesis in line with American Psychological Association (APA) referencing format.

Funding

This research was supported by a three-year Graduate Teaching Assistant Studentship awarded by the University of Kent, United Kingdom. I thank the University of Kent for their financial support and for the opportunity for me to complete this work.

ACKNOWLDGEMENTS

My sincerest thanks and gratitude go to my supervisor, Heather Ferguson. Heather is inspiring. Her wealth of experience and generosity has provided me with every opportunity to develop my research and professional skills. More importantly, her endless kindness, sincerity, and level headedness, have been invaluable to me. I have had the best possible supervisor, and it is hard to convey how grateful and lucky I feel.

Also, a huge thank you to Amanda Bates for her kindness, support, and enthusiasm – she has been a pleasure to work with over the last two years. Thank you to my second supervisor, Lindsey Cameron, particularly for her support with Chapter 3. Thank you to Alexandra Dinu, and Maria Alonso Fernández for their support with data collection for Chapters 4 and 6. I am sure they both have glowing careers ahead of them, whatever they choose to dedicate their time to. Finally, thank you to all of the participants who took part in this series of experiments. Without their support this research would not have been possible.

ABSTRACT

The self has a significant influence on human cognition across different processing domains (memory, perception, attention, action). For example, research often finds *self-biases* in cognition – the phenomena whereby people show enhanced cognitive processing for information processed in relation to the self, than for information processed in relation to someone else, or in other ways. Given that the self influences cognition, it follows that any individual with an impoverished self-concept should have diminished self-bias capacity/profile. Potentially, this is highly relevant for our understanding of autism spectrum conditions (ASC), which is characterised by atypical development of both the understanding of self (meta-cognition) and of other people (mindreading). One possible reason for this might be that self-other distinction is impaired – e.g., people are processing self-referential information similarly to other-referential information. However, empirical evidence for impaired self-bias in ASC is inconsistent, and importantly, as of yet, no theory or explanation has been put forward that fully accounts for the discrepancy. Therefore, across five experimental chapters (a total of 11 experiments), this thesis investigates typical and atypical self-referential cognition by analysing evidence of self-biases in neurotypical adults and adults with ASC, in the domains of perception, action, and mental states. **Chapter 2 (mental-state level)** found that whereas neurotypical participants (N = 28) showed a moderate-to-large and statistically significant ownership effect in Experiment 1.2, autistic (N = 28) participants showed a small and statistically non-significant effect (i.e. comparable memory for items encoded in relation to the self and in relation to the experimenter). Importantly however, Chi-squared analysis revealed that there was no categorical difference between groups in the number of people who did/did not display an ownership effect, which suggests that although the size of the ownership effect was smaller in ASC (as shown by our continuous variable analysis), autistic participants were no less likely than neurotypical participants to display an ownership effect, thus suggesting that the ownership effect is not definitively impaired in autism. Furthermore, the ownership effect was unrelated to the number of autistic traits in either the case-control sample (Experiment 1.2: N = 56) or in a larger sample of

neurotypical adults (Experiment 1.1: N = 100). **Chapter 3 (mental-state level)** investigated *scene construction* – the ability to mentally generate (i.e., imagine) and maintain a coherent spatial scene/event in mind (critical for episodic simulation), and found that scene construction ability was diminished (i.e., lower in quality) in autistic (N = 55) compared to neurotypical (N=62) adults, and was negatively associated with autistic traits. However, the ability to infer others' mental states (i.e. Theory of Mind; ToM) and to recognise emotions in oneself (i.e. alexithymia) were found to be better predictors of scene construction ability than an autism diagnosis. The frequency of descriptions of sensory experiences (which followed the same pattern in both groups: sight > sound > touch = smell > taste) and self-reference did not differ between groups. **Chapter 4 (perceptual level)** presented two experiments that aimed to test findings of undiminished perceptual self-bias in ASC when using unfamiliar stimuli Both in a case-control sample (pre-registered Experiment 3.2: N = 57), and in a sample of neurotypical adults (Experiment 3.1: N = 59), behavioural results showed that participants experienced both a self-bias and a familiarity effect. Importantly, the number of autistic traits was unrelated to the magnitude of self-bias and familiarity effects on RT or accuracy (in Experiments 3.1 and 3.2) and there was no between-group difference in the speed and accuracy of self-bias and familiarity effects (Experiment 3.2). Analysis of the ERP indices revealed clear evidence of a self-bias in both groups, with self trials eliciting larger N1 and P3 components than familiar ('friend') and unfamiliar ('stranger') other trials. Furthermore, the familiarity effect was largely absent on ERP measures, though the ASC group elicited some modulation of N2 reflecting familiarity conflict. **Chapter 5 (mental-state level)** investigated the malleability of self and other representations. In two neurotypical adult samples, we found evidence of both simulation induced malleability (Experiment 4.1; N = 163) and self-reflection induced malleability (Experiment 4.2; N = 162). Furthermore, the size of simulation induced malleability was positively associated with the number of autistic traits, but only when the simulated other was considered as the "average person" and not a "similar-other". Furthermore, in two pre-registered case-control experiments, there was greater simulation induced malleability in autistic than neurotypical adults (Experiment 4.3; N = 50)

irrespective of similarity of the target, but no self-reflection induced malleability was observed in either diagnosis group (Experiment 4.4; N = 49). **Chapter 6 (action level)** measured participants' ability to distinguish their own motor repertoire from those displayed by other people of differing levels of *similarity* – specifically, the self vs. unfamiliar adult (similar-other) vs. unfamiliar child (dissimilar-other). We found a typical imitative compatibility effect in behavioural performance, which was not associated with the number of autistic traits (Experiments 5.1 and 5.2), and did not differ between neurotypical and autistic participants (Experiment 5.2). Furthermore, associated mu/alpha (8-13Hz) and beta (13-35Hz) desynchronisation – which are thought to reflect an index of mirror neuron activity – also did not reveal consistent differences between groups (Experiment 5.2). However, the key manipulation here, the referent effect, did not show any effect in behavioural performance (Experiment 5.1 and 5.2), and the imitative compatibility effect in mu/alpha and beta did not differ between referents (Experiment 5.2). The theoretical and practical implications of these findings for our understanding of self-awareness/self-referential cognition in ASC and in the neurotypical population, are discussed.

Publications

Smith, M., Cameron, L., & Ferguson, H. J. (in revision). Scene Construction Ability in Neurotypical and Autistic Adults.

Smith, M., & Ferguson, H. J. (in revision). Autistic adults experience greater malleability of the self after simulating other people than neurotypical adults.

Smith, M., Williams, D., Lind, S., & Ferguson, H. J. (in revision). Revisiting the Ownership Effect in Adults with and without Autism.

Smith, M., & Ferguson, H. J. (in revision). Indistinguishable Behavioural and Neural Correlates of Perceptual Self-Other Distinction in Autistic and Neurotypical Adults.

Presentations/Poster Presentations

Smith, M., Cameron, L., & Ferguson, H. J. (2022, January). Scene Construction Ability in people with and without Autism. Poster presented by M.S. at Experimental Psychology Society Conference Virtual.

Smith, M., Dinu, A., Fernández, M. A., & Ferguson, H. J. (2023, January). Neural correlates of self-other distinction in autistic and neurotypical adults. Talk given by H.J.F. at the Experimental Psychology Society Conference.

Smith, M. & Ferguson, H. J. (2023, March). Simulation and Motor Self-Other Distinction. Talk given by M.S. at the International Convention of Psychological Science, as part of a symposium *Exploring Social Mimicry to Understand Its Cognitive Mechanisms.*

Contents

Chapter 1| Introduction 11
Self-Awareness and Self-Other Distinction 12
Self-Bias in Cognition 18
Autism Spectrum Conditions 27
Self-Bias in Autism Spectrum Conditions 29
Self-Bias in Autism Spectrum Conditions: A meta-analysis 37
METHOD 37
RESULTS 43
DISCUSSION 46
Methodological Issues in Self-Bias Research 47
Conclusion and organisation of this thesis 49
Chapter 2| Revisiting the Ownership Effect in Adults with and without Autism 54
ABSTRACT 56
Experiment 1.1 | Individual Differences Approach 61
METHOD 61
RESULTS 64
DISCUSSION 66
Experiment 1.2 | Case-Control Approach 67
METHOD 68
RESULTS 71
DISCUSSION 73
GENERAL DISCUSSION 74
Chapter 3| Scene Construction Ability in Neurotypical and Autistic Adults 80
ABSTRACT 83
METHOD 89
RESULTS 95
DISCUSSION 101
Chapter 4| Indistinguishable Behavioural and Neural Correlates of Perceptual Self-Other Distinction in Autistic and Neurotypical Adults 107
Experiment 3.1| Individual Differences Approach 115
METHOD 115
RESULTS 118
Summary 122
Experiment 3.2| Case-Control Approach 122

METHOD 122
RESULTS 127
DISCUSSION 136
Chapter 5| Autistic adults experience greater malleability of the self after simulating other people than neurotypical adults 143
ABSTRACT 145
Experiment 4.1|Simulation Induced Malleability, Individual Differences Approach 152
METHOD 152
RESULTS 157
DISCUSSION 159
Experiment 4.2| Self-Reflection Induced Malleability, Individual Differences Approach 160
METHOD 160
RESULTS 163
DISCUSSION 165
Experiments 4.3 and 4.4|Case-Control Approach 166
METHOD 166
RESULTS 174
DISCUSSION 180
GENERAL DISCUSSION 181
Chapter 6| Motor Self-Other Distinction in Neurotypical and Autistic Adults 189
ABSTRACT 191
Experiment 5.1| Individual Differences Approach 198
METHOD 198
RESULTS 201
Experiment 5.2|Case-Control Approach 205
METHOD 205
RESULTS 209
GENERAL DISCUSSION 216
Chapter 7| Discussion 222
Summary of Results from Empirical Chapters 226
Integrating the current findings with existing literature 235
Importance of diminished self-other distinction and its significance to mental wellbeing 241
Future Directions 245
REFERENCES 248
APPENDICES 285

Chapter 1| Introduction

Self-Awareness and Self-Other Distinction

The notion of 'levels of consciousness' has been acknowledged for some time (Morin, 2006). Notably, Mead (1935) distinguished between: 1) 'consciousness', referring to one's unreflective experience of their present external environment, whereby, according to James (1890), the self is an existential entity that is the subject of consciousness and the agent of activity (first-order representation; or 'I-Self'); and 2) 'self-awareness' – a higher form of consciousness - referring to the process by which one focuses attention inward toward the self, whereby, according to James (2018), the self becomes the object of consciousness that can be experienced and thought about (second-order self-representation; or 'Me-Self'). Morin (2006; Morin & Everett, 1990) also acknowledged a third level of consciousness: 'meta-self-awareness'. This represents a logical extension of self-awareness whereby one is aware of being self-aware. In other words, whilst self-awareness enables one to determine that "I am happy", more advanced meta-self-awareness allows one to understand that "I am currently analysing my mental state of feeling happy". It is through self-awareness that people gain self-knowledge or self-understanding, and this information in turn can become part of one's *self-concept*: one's enduring, but continually updated mental representation (i.e., set of beliefs) of the self. Therefore, an elaborate and accurate self-concept is thought to rely on having a sufficient level of self-awareness.

Self-awareness is a multidimensional phenomenon involving several domains. Neisser's (1995) *social psychological conception of the self* proposes that people can access five different levels of self-information. These emerge from experience of multiple domains, ranging from perception to cultural experiences, and typically develop sequentially during different stages of ontogeny, with each level being reliant on sufficient development of more fundamental levels. The five levels are:

1. *Ecological self* is the lowest level and refers to the self as it is directly perceived in one's physical environment (i.e., I am the person, in this place, engaged in this action). This includes self-world differentiation and awareness of physical agency. Proprioception is

widely thought to contribute to the specification of the ecological self. Receptors of the somatosensory system convert mechanical energy applied to the body into action potentials which subsequently travel to the central nervous system (CNS). The CNS integrates such proprioceptive input with information from other sensory modalities (e.g., vision to create an overall representation of the position and movement of parts of our body in relation to themselves and the external world. Children at two months old have been observed to pass tests of ecological self (Lee & Aronson, 1974).

2. *Interpersonal self* is the second lowest level and refers to the self as it is *directly perceived* in one's social environment (i.e., I am the person who is engaged in this current social interaction). This includes self-other differentiation and awareness of social agency. The interpersonal self is proposed to be specified through interaction with others, whereby we experience ourselves through our interaction and emotional response to others, and by internalizing others experience of ourselves. Children at one month old have been observed to pass tests of interpersonal self (Tronick et al., 1978). The ecological and interpersonal self-representations encompass self-recognition (Jeannerod, 2003) whereby one recognizes themselves as the owner of a body (ownership) and the agent of activity (agency).
3. *Temporally extended self* refers to the self that incorporates the self as it was in the past (remembered), as it is in the present (currently experienced) and as it will be in the future (as we expect/imagine), as well as one's continuing existence through time. This level of self-information is primarily based on memory. Children typically pass tests of temporally extended self at 4-5 years (Povinelli et al., 1996; Suddendorf & Corballis, 1997).
4. *Private self* refers to the knowledge that we each have conscious experiences, including thoughts, that are exclusively our own, and not directly available to/shared with other agents – e.g., "In principle, I am the only person that can feel this unique and particular …". Children aged four years have been observed to pass tests of private self (Cultice et al., 1983).

5. *Conceptual self* or *'self-concept'* is the last level to fully develop and refers to an objective form of self-knowledge requiring the ability to reflect upon the self (i.e., one's self concept). The conceptual self refers to one's explicit beliefs/theories/assumptions about the self, including for example psychological and physical traits (e.g., kind, selfish, appearance) and social roles (e.g., wife, student, English). Children at two years have been observed to pass *basic* tests of conceptual self (Lewis & Brooks-Gunn, 1979; Lewis & Ramsay, 2004).

Based on the various operationalizations of the self that have been defined in the neuroscientific literature, a practical distinction has been made whereby, at a basic level, self-awareness can focus on physical- or psychological-self-aspects (Gillihan & Farah, 2005); similar to public and private self-aspects outlined by Fernigstein et al. (1975). Physical self-aspects encompass awareness of one's own body (e.g., behaviour, physical appearance, mannerisms etc.), whereas psychological self-aspects are unobservable events and characteristics which include autobiographical memory and conscious experience of first-person perspective (or "egocentric frame of reference", Vogeley & Fink, 2003; e.g., thoughts, emotions, perceptions, physiological sensation, values, traits etc., Fenigstein et al., 1975). Such psychological self-aspects are more conceptual and hence attending to them represents a higher form of self-awareness compared to more concrete physical self-aspects.

As illustrated in Figure 1.1, there are significant overlaps between the social psychological and the neuroscientific conceptions of the self. Like Quesque and Brass (2019), this thesis will primarily draw upon three distinct levels of self that overlap with the existing social psychological and neuroscientific distinction: 1) the *perceptual level* considers the ability to identify one's own body and distinguish it from that of other's; 2) the *action level* considers the ability to experience agency of one's own motor actions and distinguish it from that of others, including the control (or inhibition) of imitation (i.e., distinguishing one's own motor representations from that of others); 3) the *mental state level* considers the ability to identify one's own mental states and distinguish them from those of others (Quesque & Brass, 2019).

Researchers have suggested a hierarchy between the control of shared representations at the perceptual, action and mental state levels (Quesque & Brass, 2019; Bardi & Brass, 2016; Brass et al., 2009). During infancy, a shared representational system (same between agents) develops from repeated exposure to associations between a motor command and its perceptual consequences (Brass & Heyes, 2005). With experience, this develops the ability to dissociate between events that are and are not contingent with one's own actions – i.e., development of sense of self (and others). Last, development of the ability to co-represent and switch between self and other representations constitutes a pre-requisite to the numerous socio-cognitive mechanisms discussed above (Quesque & Brass, 2019).

REDACTED

Figure 1.1. Diagram (from Quesque & Brass, 2019, p. 945) depicts the overlap between Neisser's (1995) five-dimensional social psychological conception of self (light grey) and Gillihan and Farah's (2005) two-dimensional neuroscientific conception of the self (dark grey). Black double-headed arrows illustrate correspondence between associated levels.

The ability to distinguish between what originates from ourselves and what originates from other people is essential for efficient interaction with the social environment. However, it is not yet known whether self-other distinction is a domain-general mechanism which is common to various

social-cognitive functions (e.g., perceptual, action and mental state levels), or whether there are specific self-other distinction mechanisms associated with different social-cognitive functions (Quesque & Brass, 2019; Amodeo et al., 2021). The temporal parietal junction (TPJ; see Figure 1.2) – located at the border of the posterior superior temporal and inferior parietal lobes – is a possible brain region associated with a domain-general self-other distinction mechanism (Farrer et al., 2003). The TPJ has been observed to contribute to attention reorientation (e.g., Corbetta & Shulman, 2002), language processing (Pulvermüller, 2005), episodic memory (Cabeza et al., 2012), and a variety of social cognitive tasks (e.g., Lombardo et al., 2010; Young et al., 2010).

REDACTED

Figure 1.2. Illustration taken from Quesque and Brass (2019, p. 944) "of the anatomical features of the brain that are encompassed under the label "Temporoparietal Junction" (left hemisphere is here depicted). The TPJ is characterized by the red ellipse. Covered brain regions are represented in orange (supramarginal gyrus), cyan (superior temporal gyrus), green (angular gyrus) and pink (superior temporal sulcus). The Wernicke area is depicted as a blue squared grid" (Quesque & Brass, 2019, p. 944).

This thesis will investigate typical and atypical development of self-referential cognition by analysing evidence of self-biases in cognition: whereby people process self-related information preferentially compared to information related to other people. This Introduction chapter will first outline and evaluate evidence from different paradigms tapping into this self-bias to inform our understanding of how a self-bias occurs, and what its implications are for social cognition. Then, this will analyse evidence of self-bias in autism spectrum conditions (ASC), aiming to understand evidence suggesting a paradox whereby individuals with ASC are thought to have a diminished self-bias, yet extreme egocentrism. This will set a comprehensive background to the experimental chapters of this thesis which aim to gain a better understanding of self-awareness/self-referential cognition in ASC, and furthermore, use evidence from ASC as a model for understanding self-awareness/self-referential cognition in the neurotypical population.

Self-Bias in Cognition

The self appears to have a significant influence on human cognition across different processing domains (memory, perception, attention). In the domain of memory, for example, using the self-referential trait memory paradigm (Rogers et al., 1977), research has found that people often have superior recall or recognition memory (select from a list of words) for personality trait adjectives (e.g., kind, clever, tenacious) that they had previously encoded in relation to the self (e.g., "Are you *kind*?") compared to those encoded in relation to someone else (e.g., "Is Brad Pitt *kind*?") or those encoded in other ways (e.g., semantically: "does *kind* mean the same as *friendly*?"; structurally, or phonemically; Klein & Loftus, 1988; Rogers et al., 1977; Symons & Johnson, 1997) (see Figure 1.3). This enhanced memory for self-related information is termed the "self-reference effect" (SRE: Symons & Johnson, 1997) and has been observed in adults (Sui & Humphreys, 2013) and children (Cunningham et al., 2014b). Several theories attempt to explain the SRE, with early researchers suggesting functionally distinct systems for self-related memory and other-related memory. If true, this would suggest the presence of a distinct self-concept that structures self-trait knowledge differently from other-trait knowledge (Rogers et al., 1977).

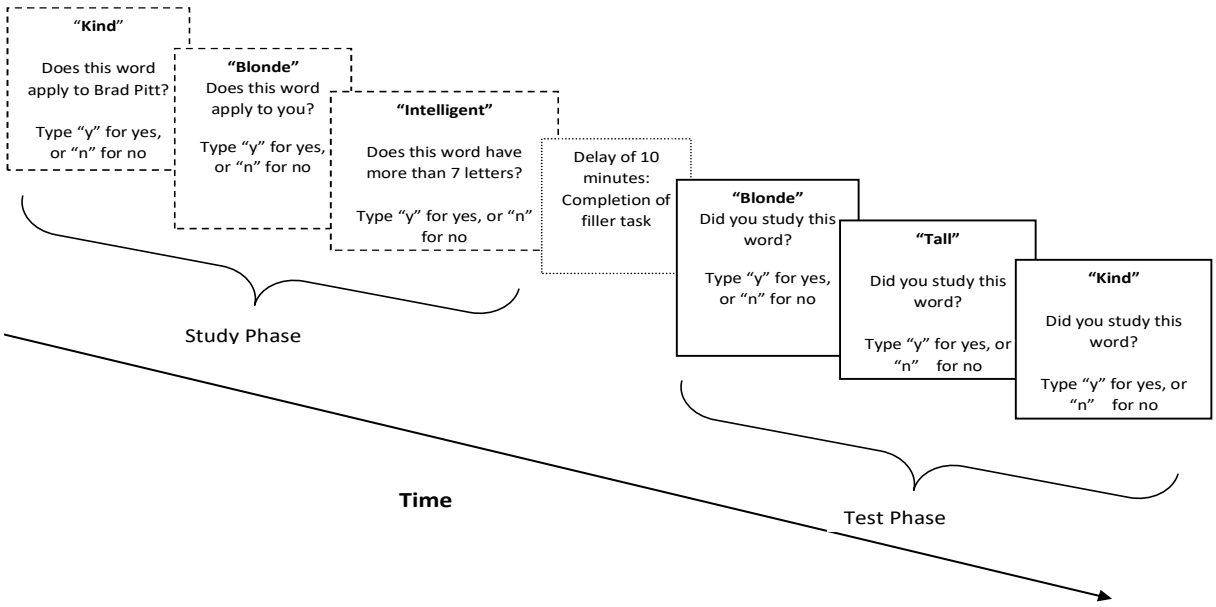


Figure 1.3. Illustration of an example study and test trials from the *general* self-referential trait memory paradigm; note that the recognition test involves previously studied items and lure items.

However, alternative encoding factors have been found to enhance memory of information to the same extent as self-referential processing. Klein and Loftus (1988) found that *semantic elaborative encoding* (thinking of a definition for the word) of semantically *related* words (e.g., a type of music: “Jazz”; “Opera”), and *organisational encoding* (sorting each word into one of three categories) of semantically *unrelated* words (e.g., associated with parties: “weekend”; “band”), led to an enhancement in subsequent memory for items that was not significantly different than if the items were encoded via self-referential encoding (determining whether or not the word induced memory of a personal experience). Therefore, elaborative and organisational encoding strategies appear to enhance memory to the same extent as self-referential encoding depending on the type of information being encoded.

Craik and Tulving's (1975) Depth of Processing (DoP) account explains that greater stimuli *elaboration* – the level of processing at the time of stimuli encoding [ranging from i) *shallow elaboration*, involving physical analysis of stimuli (e.g., detecting the number of letters in a word), to ii) *deep elaboration*, involving semantic analysis of stimuli (whereby the meaning of the word is processed; e.g., detecting whether the word relates to the self)] – and stimuli *organisation* – the extent that stimuli list items can be encoded as associates of one another (e.g., shy and quiet) and/or in relation to a common category label (e.g., Self) – facilitates recall. Therefore, challenging the notion of functionally distinct memory mechanisms for self and other-related processing, the DoP account proposes that a SRE occurs because processing information in relation to the self-concept – which acts as an organizing structure – facilitates more elaborate encoding and subsequent better retrieval of such information than information that has not been encoded in relation to the self-concept (Lind, 2010; Lind et al., 2018).

The notion that the level of semantic encoding rather than distinct memory systems influences a SRE, would predict greater recognition/recall memory for trait adjectives encoded in relation to a familiar-other person (e.g., a friend) who the participant has some knowledge of,

compared to an unfamiliar-other person (e.g., a stranger) who the participant has less knowledge of. Indeed, evidence often finds a linear effect of familiarity (self > familiar-other (e.g., friend) > unfamiliar-other (e.g., stranger) pattern) in SRE (Bower & Gilligan, 1979; Henderson et al., 2009). The more knowledge one has of a person (i.e., the stronger the person-schema), the greater the effect that such person-related processing has on cognition. This supports the DoP account: familiarity effects suggest that self-referential processing is the function of the same processes (i.e., semantic elaboration) involved in general person-related processing, with the self simply being the person we have the strongest schema for (i.e., most stored knowledge).

However, neural correlates of behavioural familiarity effects have undermined the DoP account of the SRE. fMRI has revealed that the left inferior frontal cortex and anterior cingulate – associated with semantic encoding (Devlin et al., 2003) – are significantly more activated during encoding of personality trait adjectives in relation to the self ("Does this adjective describe you?") and a familiar-other ("Does this adjective describe U.S. President George W. Bush?") – i.e., semantic encoding – than during the encoding of personality trait adjectives according to surface presentation ("Is this adjective printed in upper or lower case?") (Kelley et al., 2002). Importantly however, this activation did not differ between self and familiar-other conditions, despite behavioural performance showing a typical SRE in memory (self > familiar-other > surface presentation). If SRE is simply an extension of the DoP account, then we would expect left frontal activity to mediate the difference in memory for adjectives encoded in reference to the self compared to an other.

Instead, the ventral medial prefrontal cortex (vmPFC) – a region associated with emotional, interoceptive and self-processing (Trimble, 1986; Wheeler et al., 1997) - was *selectively* engaged during encoding of adjectives in relation to the self (Kelley et al., 2002), and furthermore, was found to predict which items would be subsequently remembered (Macrae, 2004). Therefore, neural evidence suggests the vmPFC might reflect a mechanism specifically associated with self-related

processing: this would suggest that self is "special" since it is "functionally dissociable from other types of semantic processing in the brain" (Kelley et al., 2002, p. 785).

The vmPFC is part of the default mode network (DMN): a network of brain regions – including the precuneus/posterior cingulate cortex (p/PCC), the medial prefrontal cortex (mPFC), and the medial, lateral, and inferior parietal cortex – which play a major part in the Self-Relevance (i.e., regulatory) System (SRS; Conway & Pleydell-Pearce, 2000). During periods of sustained rest (i.e., the default state), which typically involves daydreaming (e.g., imagining future events and thinking about past events/autobiographical memories), the default system is characterised by sustained activation (and with extensive functional connectivity between regions) in anterior and posterior networks – the same networks that show activity during remembering and imagining (Conway & Loveday, 2015). In short, this default network becomes activated when attention is oriented internally[1]. These memories and imaginings form the basis of preoccupations (Conway et al., 2016), which in turn are thought to drive the emergence of attentional biases (Conway et al., 2016).

"New Wave" Methodologies to assess self-other processing

The notion that attentional biases result from preoccupations related to the default state, might explain self-biases in memory. According to the Self-Memory System (SMS, Conway, 2005; Conway & Pleydell-Pearce, 2000), there is a bi-directional relationship between the working self (a set of active personal goals) and the autobiographical knowledge base (contains knowledge of the self). Retention of information in memory depends on the goal hierarchy of the working self (i.e., as is imagined/remembered). If incoming information is consistent with the goals of the working self –

[1] During goal-directed activity, the default network shows a consistent pattern of deactivation, which negatively correlates with task difficulty: the more demanding the task, the greater the deactivation (Knyazev, 2013). Therefore, explaining Kelley et al.'s (2002) findings, it might be more cognitively demanding – hence requiring greater vmPFC deactivation – when determining the relevance of traits in relation to other people than in relation to the self.

whether that be short term (e.g., complete essay assignment) or long-term (e.g., coherent identity maintenance), it is more likely to be retrieved than information that is not goal relevant. Regarding self-referent processing, monitoring for self-relevant information might be a perpetual goal of the working self (as suggested by the default state which involves self-oriented thought), thus making self-relevant information more easily retrievable. This would argue that a SRE should be observed not only when information is scaffolded by autobiographical knowledge – i.e., as in the self-referential trait memory paradigm – but for any information encoded in association with the self (see Bluck, 2003).

It is not straightforward to test this using the classic self-referential trait memory paradigm, because scaffolding by autobiographical knowledge is likely triggered by participants making self-relevant judgements to determine the personal significance of traits. However, in the Ownership Task (Cunningham et al., 2008; see Figure 1.4), participants are asked to encode objects (pictured on cards) as being either (notionally) owned by the self (e.g., "place all objects pictured on red-bordered cards into your own red basket"), or (notionally) owned by someone else (e.g., "place all objects pictured on blue-bordered cards into the experimenter's blue basket"). Since self-ownership is completely arbitrary and imagined, encoding should not elicit autobiographical self-knowledge to any greater extent than encoding of other-owned items; therefore, the level of scaffolding by autobiographical knowledge should not differ between conditions. Regardless, superior recognition memory – and associated encoding mPFC activity (Turk et al., 2011) – is often observed for self-assigned compared to other-assigned items in children and adults (Cunningham et al., 2008, 2011, 2013; Hartley & Fisher, 2018; Turk et al., 2011b; van den Bos et al., 2010). Therefore, evidence from the ownership task suggests that a SRE is not limited to encoding contexts where self-knowledge is linked to the to-be-remembered information – e.g., as in the trait memory paradigm ("does *shy* apply to you?"): simply associating items with the self appears to enhance memory.

Therefore, it is proposed that the self acts as an "associative glue" (Humphreys & Sui, 2016): when associations are formed in memory (e.g., between self and object), self-reference enhances the binding process – over and above the binding that results from other-reference or semantic encoding – so that different features associated with the self are more strongly linked into a single representation – i.e., a single 'self-Gestalt' (Humphreys & Sui, 2016), and can subsequently be retrieved together. Accordingly, responding to a single self Gestalt is less cognitively demanding than responding to features independently, and this enhances perception and memory.

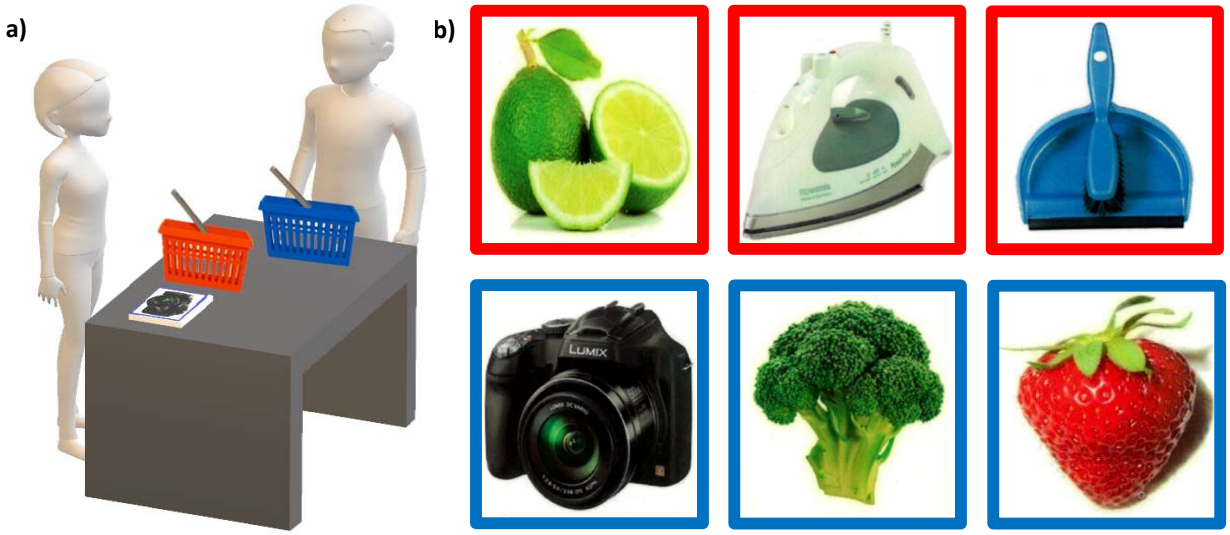


Figure 1.4. Illustration of the Ownership task set up with (a) experimenter/other (male with blue basket), participant/self (female with red basket), and (b) example items owned by the self (red bordered cards) and owned by the other (blue bordered cards).

Furthermore, event related potential (ERP) evidence observed that when the presentation of objects was followed by a cue indicating that the object was owned by the self (a "self" cue) compared to the experimenter (an "other" cue), there was a narrowing of visuospatial attention (lateral occipital P1 ERP component) to the location of the owned object (Turk et al., 2011). This early, automatic response to detection of self-relevance was followed by an increase in automatic attentional processing (P3 ERP component) associated with higher-order, top-down modulation of attention and executive processing (Turk et al., 2011). Based on this, Turk et al. (2011) proposed a

temporal model in which activity in frontal cortex attentional and reward circuits in response to object ownership, was followed by lateral posterior activity associated with attention for action (e.g., planning/imaging how the object can be used). Interestingly, activation of this network was supressed during encoding of other-owned items. Therefore, self-cues appear to trigger perceptual/attentional processing biases that enhances subsequent memory encoding, and serves the ecological function of minimising the risk that information of potential importance is lost (Conway, 2005; Conway & Pleydell-Pearce, 2000; Symons & Johnson, 1997; Cunningham et al., 2013).

Turk et al.'s (2011) ERP findings reflect a key part of the self-relevance system: the Self Attention Network (SAN) (Humphreys & Sui, 2016; Conway & Pleydell-Pearce, 2000; see Figure 1.5). The SAN proposes that perception of self-relevant stimuli (e.g., self-cues, notionally owned items, traits encoded in relation to the self) triggers a response from the vmPFC. The vmPFC then cascades information that stimuli are self-relevant to the left posterior superior temporal sulcus (LpSTS) attentional system, so that it is primed to respond to self-relevant stimuli, which subsequently influences attentional biases.

REDACTED

Figure 1.5. Figure taken from (Humphreys & Sui, 2016, page 12): "Black arrows indicate excitatory connections. Gray arrows indicate inhibitory connections. Dotted arrows (to and from the pSTS) highlight that there is currently little direct evidence for these functional corrections" (Humphreys & Sui, 2016, p. 12).

Such attentional biases – associated with bottom-up self-related activity of the vmPFC and LpSTS – in turn, are thought to be modulated by both inhibitory and excitatory connections from the

dorsolateral prefrontal cortex (DLPFC) top-down attentional control network when making more difficult associations. This top-down control shapes behaviour. For example, greater allocation of attention to self-relevant stimuli, via top-down *excitatory* connections (black arrow) from the attentional control network, can support enhanced memory encoding as is evidenced by the SRE. Similarly, *inhibitory* control is essential, since not all events are highly self-relevant, and those that are might attenuate other processes, such as encoding of memory details (Conway et al., 2015). This is demonstrated by evidence finding that experiences of extreme self-relevance, such as trauma, often result in memories that are low in detail, and with amnesiac gaps (Conway et al., 2004). Therefore, controlling attentional biases caused by high self-relevance might be critical for optimal cognitive functioning (Conway et al., 2015).

Simulation of Others

The self-bias is not only adaptive because it facilitates the creation of a stable sense of self, but in turn, it is also thought to benefit adaptive social functioning. The "simulation" account proposes that inferring the mental states of other people, requires us to imagine our own thoughts, feelings, and behaviours if in the same position (e.g., Goldman, 2006; Mitchell et al., 2006; Shanton & Goldman, 2010). In doing this, we can simulate (i.e., predict) other people's mental states based on direct first-person access to our own (Goldman, 2006). For example, when trying to understand how your friend is feeling about failing an exam, you might recall *how you felt when you failed an exam previously* (e.g., *I felt disappointed*). This final state is then taken "offline" so that it becomes assigned to the subject of simulation, rather than the simulator (Goldman, 2006). Therefore, a better self-representation (and hence self-bias), or metacognition ability (thinking about one's own thoughts), is ontogenetically and phylogenetically prior to, and required for, the ability to attribute mental states to others (Conway et al., 2019; Mitchell, 2009; Nijhof & Bird, 2019) – i.e., Theory of Mind (ToM; aka, mentalising; Baron-Cohen et al., 1985; Premack & Woodruff, 1978).

Arguably, this simulation strategy would only be beneficial if the other person is similar to the self, and hence one's own mental states are considered applicable to the other person. Thus, explaining the linear effect of familiarity typically observed in SRE (self > similar/familiar-other > dissimilar/unfamiliar-other): whilst processes of the SAN might be specialised for self-referential processing (i.e., the vmPFC, Kelley et al., 2002; Turk et al., 2011), if the other-referent is someone who the self is close/similar to (e.g., Mother, best friend) as opposed to unfamiliar/dissimilar to (e.g., U.S. President George W. Bush; Kelley et al., 2002), they have greater relevance to the self and hence elicit enhanced attentional processing in a similar way to object-ownership (e.g., Turk et al., 2011). Indeed, the mPFC distinction between self and other in such cases reduces (Benoit et al., 2010; Mitchell et al., 2006), suggesting that the similar-others might benefit from the SAN automatic enhanced attentional processing, yet to a lesser extent than the self. This shared (neural) representation is thought to allow better understanding of the other person (Mitchell et al., 2006; Pfeifer et al., 2009).

Overall, the evidence suggests that self-bias in cognition is driven by a specialised self mechanism – the SAN (Humphreys & Sui, 2016) – which is primed to attend to self-relevant information. This might serve the ecological function of enhancing subsequent encoding and minimising the risk that information of potential importance to the self is lost (Sui & Humphreys, 2015). Nevertheless, this specialised self-mechanism might also be strategically expanded to simulate other people in circumstances where a self-based inference might be beneficial. Therefore, top-down control of the SAN – both inhibitory and excitatory – appears to be critical for optimal cognitive functioning. Given that the self influences cognition (via the SAN), it follows that any individual with an impoverished self-concept should have atypical cognitive information processing – i.e. diminished self-bias capacity/profile. Potentially, this is highly relevant for our understanding of autism spectrum conditions.

Autism Spectrum Conditions

We acknowledge the debates regarding the appropriate terminology to use when discussing autism. Our goal is to respect the preferences of people on the autism spectrum and recognise that their differences include strengths as well as challenges. For this reason, we have included both person-first (e.g., individual with autism) and identity-first (e.g., autistic individual) language. Furthermore, we acknowledge that autism spectrum *disorder* is the official diagnostic label, and in some cases is the most relevant way to refer to autistic people; however, we use this term selectively, and instead primarily use the term "condition" rather than "disorder". This reflects the diversity in autistic individuals' preferences and abilities, and we hope it minimises stigma (Kenny et al., 2016).

According to the Diagnostic and Statistical Manual of Mental Disorders (DSM-V: American Psychiatric Association, 2013), autism spectrum conditions (ASC) are neurodevelopmental disorders diagnosed based on two core areas of behavioural impairments – there are currently no known genetic or biological markers (Goldani et al., 2014). The first area is difficulties with social communication and interaction across multiple contexts. This might manifest itself as atypical behaviour during social communication (e.g., difficulty maintaining typical back-and-forth conversation, reduced initiation or response to social interactions, and reduced sharing of interests and emotions), deficits using and understanding non-verbal communication (e.g., eye contact, gestures), and difficulty developing, maintaining, and understanding relationships (e.g., adjusting behaviour appropriately to a given context). The second area is a tendency for restrictive and repetitive behaviours and interests. This might manifest itself as stereotyped and repetitive motor actions (e.g., stereotypies; echolalia; idiosyncratic phrases), inflexible adherence to routines (e.g., greeting rituals, the need to take the same route or eat the same food everyday), restricted interests (e.g., strong attachment to certain object or topic), and hyper- or hypo-reactivity to sensory input (e.g., adverse response to sounds, textures, smells, lights etc.). The pattern and severity of ASC symptoms is highly heterogenous, meaning that their severity and extent that they impact daily lives differs from one individual with ASC to another (Chawner et al., 2021; Goldani et al., 2014). Therefore, the diagnostic criterion allows

for more variation – hence the term 'spectrum'. Accompanying the core behavioural impairments, some but not all individuals with ASC also have structural language and/or intellectual impairments.

Regardless of age or intellectual ability, approximately 1-2% of the population have a diagnosed ASC, and this is more prevalent in males, with the male to female ratio being 4:1 (e.g., Giarelli et al., 2010; Rutherford et al., 2016). A diagnosis requires that the symptoms cause clinically significant impairment in daily functioning (e.g., school, work, relationships), and must not be better explained by intellectual disability or global developmental delay. It is widely accepted that ASC is a neurodevelopmental condition, and hence requires that symptoms have been present from birth – although, symptoms might not manifest until social demands exceed one's limited capacity, and moreover might remain undetected in later life if one is able to use learnt strategies to camouflage symptoms.

Self-Bias in Autism Spectrum Conditions

Using the self-referential trait memory paradigm, research previously found impaired SRE in individuals with ASC (Henderson et al., 2009). Unlike neurotypical children (mean chronological age (CA) = 13.25 years; Henderson et al., 2009) and adolescents (Burrows et al., 2017[2], same laboratory as Henderson et al., 2009), age-, gender- and IQ-matched children and adolescents with ASC did not have enhanced memory for personality trait adjectives (e.g., happy, cranky, clever, funny) previously encoded in relation to the self ("Does this word describe something about you?") than in relation to a familiar/dissimilar-other ("Does this word describe something about Harry Potter?"). This is despite both groups showing a social-semantic depth of processing effect whereby they recognised significantly more person-referenced (i.e., semantically processed) words (i.e., self, and other-referenced) than words encoded according to featural presentation ("Does this word contain seven or more letters?"). Similarly, in adults, research has observed a diminished – although not entirely absent – SRE in ASC (Lombardo et al., 2007). Therefore, there appears to be at least a developmental delay, in using self-information to facilitate cognition. Across these three early studies of self-bias using the self-referential trait memory paradigm (Henderson et al., 2009; Burrows et al., 2017; Lombardo et al., 2007[3]), the mean weighted effect size for the difference in size of the SRE between

[2] Burrows et al.'s (2017) results must be considered with caution. Participant groups were not matched for age, VIQ, PIQ or sex-ratio. Attempting to correct for this potential confound, Burrows et al. (2017) included age, VIQ, and sex as a covariate in in series of ANCOVAs. However, as noted by Lind et al. (2020), this does not correct for the problem of unmatched groups, and moreover, ANCOVA is not appropriate when groups are not matched on the covariates. Therefore, it is difficult to determine whether the unmatchedon groups caused the between group difference in SRE.

[3] In the first published results using the trait memory paradigm in individuals with ASC, Toichi et al. (2002) found that adults with ASC did not have enhanced memory for personality trait adjectives encoded in relation to the self ("Does the word describe you?") compared to personality trait adjectives encoded phonologically (e.g., "Does the word rhyme with – ?") or semantically (e.g., "Is the meaning of the word similar to –?"). Nevertheless, Lombardo et al. (2007) corrected for two key limitations of Toichi et al.'s (2002) study. 1) Lombardo et al. (2007) included comparison conditions for thinking about other people; 2) Whilst Toichi et al. (2002) used a biased measure of memory performance whereby the dependent variable was the percentage of words that were correctly recognised, Lombardo et al. (2007) used a less biased measure of memory sensitivity (d') which corrects for false alarm rate (by subtracting the number of incorrectly recognized words that were in fact not presented previously from the number of correctly recognised words that were presented previously).

ASC and neurotypical participants is Cohen's *d* = 0.65, suggesting a moderate-to-large reduction of this effect in ASC (Lind et al., 2020), and thus impaired self-referential cognition.

This impaired self-referential cognition in the self-referential trait memory paradigm (as well as in the ownership task, see Grisdale et al., 2014) has frequently been interpreted as supporting the notion of at least a diminished – if not "absent" – sense of self in ASC (Scott, 2004): this is a diminished awareness of one's own self-concept, including traits, beliefs, abilities etc. A diminished sense of self would cause difficulty forming associations between external information and the self. Consequently, attention is not biased to automatically attend to self-related information as is suggested by the SAN (Humphreys & Sui, 2016), and subsequently there is no self-bias in cognition (Lind, 2010; Conway, 2005; Conway & Pleydell-Pearce, 2000). Indeed, one of the earliest and strongest predictors of a later ASC diagnosis is a failure to orient (i.e., attentional bias) to one's own name during infancy (Osterling et al., 2002; Elmose, 2016), thus suggesting a lack of binding and/or diminished preferential orienting to the self.

However, counterintuitive to the notion of diminished self-referential cognition, individuals with ASC have also been proposed to have an extreme self-focus at the detriment of processing other-relevant information – i.e., egocentrism. This egocentrism is thought to underlie the core difficulties with social communication and interaction in ASC. For example, a simulation approach is only perfectly reliable if everyone's experiences were identical; but this is not the case. In reality, one must combine their own experience with their perception of the other's experience. Therefore, someone who is entirely self-focused and hence cannot inhibit their own beliefs to accurately understand the different beliefs of other people, would have impaired ToM in cases where a strict simulation approach is not optimal. Therefore, extreme egocentrism means that individuals with ASC do not appear to develop the appropriate context-dependent realization that one is similar to, yet different from others (Brewer, 1991): i.e., "duality of self". Therefore, a paradox arises: how can the

self be both “absent” (i.e., impaired/diminished self-referential cognition) and “too present” (i.e., extreme egocentrism) in ASC?

Lombardo and Baron-Cohen (2010) responded to this paradox by explaining that it might simply arise from the assumption that extreme egocentrism implies enhanced self-referential cognition. They argued that this assumption is questionable given that egocentrism can also arise if the self and other are processed equivalently. Thus, in ASC, egocentrism might simply define instances where individuals do not appropriately distinguish between self and other. Indeed, at a neurological level, Lombardo et al. (2010) observed that when participants were presented with statements which asked them to make reflective judgments regarding the physical characteristics and mental states of themselves (self-representation) or a familiar-other (mentalizing) (e.g., Physical: “how likely are you (self condition)/the Queen (other condition) to have bony elbows”; mental state: “how likely are you (self condition)/the Queen (other condition) to keep a diary”), unlike neurotypical males, males with ASC did not preferentially recruit vmPFC and middle cingulate gyrus (MCC) during self-representation compared to other-representation (Lombardo et al., 2010; see also Kennedy & Courchesne, 2008). The vmPFC responded egocentrically in ASC: equivalently for both self and other; and MCC was activated more when mentalizing about others than about the self. This was characterised by atypical activation specifically for the self condition, but not the other condition (see Di Martino et al., 2009, for a meta-analysis of neuroimaging studies on social processing in ASC, and see Bird & Viding, 2014, for a review).

Therefore, the finding that individuals with ASC do show a social-semantic depth of processing effect in memory whereby they recognise significantly more person-referenced (i.e., semantically processed) words (i.e., self, and other-referenced) than words encoded according to featural presentation (“Does this word contain seven or more letters?”), despite showing equivalent memory for self- and other-referential information (Henderson et al., 2009; Burrows et al., 2017), suggests that self-representation is intact in ASC. However, the mechanisms influencing self-other

switch – which is thought to be instantiated in the temporoparietal junction (TPJ; see Chapter 1a) (Santiesteban et al., 2012; Spengler et al., 2009) – might be impaired. This causes difficulty distinguishing self from other, and consequently, any self-bias is diminished since everything, including the other, is processed equivalently, in relation to the self (i.e., egocentrically).

Self-bias across different cognitive domains and levels of self-awareness in ASC

Research suggests that the magnitude of self-biases across cognitive domains are not correlated, meaning that they rely on independent mechanisms (Amodeo et al., 2021; Nijhof et al., 2020). Therefore, an alternative explanation for an impaired SRE in ASC, is that self-referential cognition is intact, but in some way "blocked" from influencing memory; perhaps due to atypical connectivity between the brain regions underpinning self-representation and those underpinning memory (see Grisdale et al., 2014). If so, then individuals with ASC should have an intact self-bias at the fundamental attentional/perceptual level.

A new line of work on the self-bias in cognition has explored the effect of self-biases at this attentional/perceptual level. Sui et al. (2012) asked participants to learn to associate three different labels ("You", "Friend", and "Stranger") with three neutral geometric shapes (square, circle, triangle) in the learning phase of a so-called "shape-association task" (see Figure 1.6). In the test phase, participants are asked to judge whether the shape-label pair presented briefly (100 milliseconds) are a match (e.g., square is presented with correct label association: "self") or a mismatch (e.g., square is presented with incorrect label association: "stranger"). Despite the elimination of familiarity effects with the use of neutral shapes (as opposed to self-related stimuli), a self-bias has been observed in children and adults (Sui et al., 2012), whereby self-relevant matches are responded to more quickly, more accurately, and furthermore are less affected by perceptual blurring (Sui et al., 2012) than familiar-other or unfamiliar-other matches (*RT*: self < familiar-other < unfamiliar-other; *Accuracy*: self > familiar-other > unfamiliar-other).

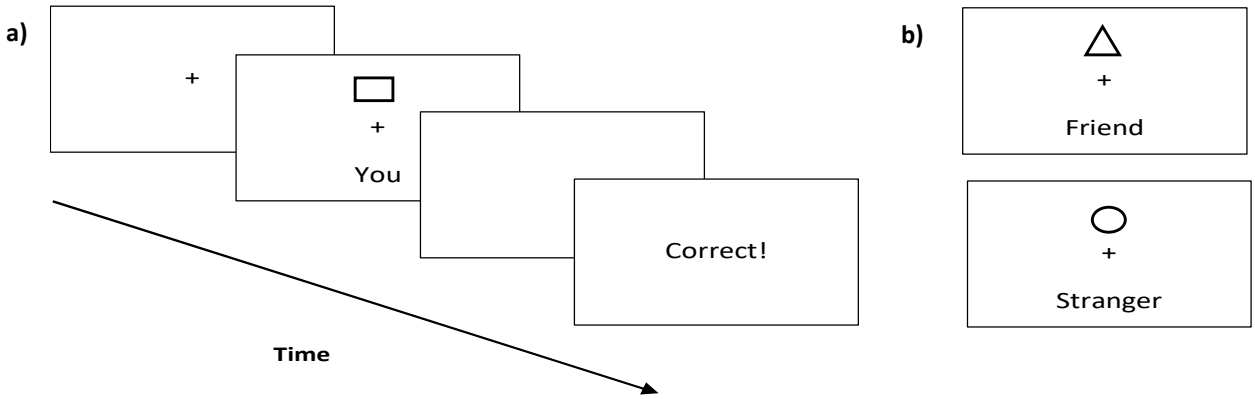


Figure 1.6. Illustration of a correct trial example (a) and other matches (b) in the Shape association task.

Using the shape association task (Sui et al., 2012; see Figure 1.6), research has found that like neurotypical adults (n = 21), adults with ASC (n = 22) show a self-bias in perception, whereby self-relevant (e.g., Self) matches were responded to more quickly and more accurately than familiar-other (e.g., Mother) or unfamiliar-other (e.g., Stranger) matches (*RT*: self < familiar-other < unfamiliar-other; *Accuracy*: self > familiar-other > unfamiliar-other) (Williams et al., 2018; Zhao et al., 2018). Importantly, the speed and accuracy of responses did not differ between diagnostic groups. Furthermore, unlike the self-referential trait memory paradigm (Henderson et al., 2009), the number of ASC traits (measured using Autism Spectrum Quotient; Baron-Cohen et al., 2001) was unrelated to the magnitude of either the RT or accuracy self-bias in this case-control sample or in an additional sample of neurotypical participants (n = 124) (Williams et al., 2018). Therefore, self-bias at the perceptual level appears to be intact in ASC. This challenges the notion of an impaired self-other switch in ASC: if this were the case, then self-bias would be universally impaired regardless of what cognitive domain the task relies on.

However, Williams et al. (2018) also found that successful performance on the shape-association task in adults with ASC contrasted with their significantly poorer performance on two tests of ToM compared to neurotypical participants. This challenges the one-mechanism theory (e.g., Carruthers, 2009), by suggesting that self-awareness is not dependent on the same functional

system/process employed when becoming aware of someone else's mental states (i.e., mindreading) – at least not in individuals with ASC. If this were the case, then intact self-awareness (i.e., characterised by intact SRE) ability would predict intact ToM in ASC. Rather this suggests that self-awareness and ToM might rely on separate mechanisms (Nichols & Stich, 2003).

Williams et al. (2018) proposed that the discrepant findings regarding self-awareness ability might reflect different levels of self-awareness required by different tasks. Whereas a self-bias in the self-referential trait-memory and ownership paradigms requires a second-order representation of the self (Me-self; James, 1890) at encoding – whereby the self becomes the object of ones thought (e.g., "am I kind?") – to consider whether/how personality-trait adjectives/objects apply to oneself; a self-bias in the perceptual shape task might require only a more superficial first-order representation of the self (I-self; James, 1890) – i.e., simply tagging something as self-relevant (e.g., "self = square") – to make a speeded perceptual judgement about whether shape-label pairs match the learned contingency. Therefore, the discrepant findings across tasks might reflect an intact I-self but impaired Me-self in ASC (Williams et al., 2018). This is consistent with metacognitive impairments observed in ASC (e.g., Carpenter et al., 2019) since metacognition requires a second-order representation of one's own mental states.

If so, then in an incidental encoding version of the self-referential trait memory paradigm, whereby encoding requires participants to simply determine if trait adjectives were presented above or below their own (self-condition) or a celebrity's (other condition) name (Cunningham et al., 2014; Turk et al., 2008), a self-bias should be undiminished in ASC since encoding of self-relevant information requires only a more subjective first-order orientation to information related to the self (spatial judgement, i.e., iconic memory), and not explicit reflection on the self (i.e., deep semantic encoding; Williams et al., 2017). However, Lind et al. (2020) found that adults and children with ASC showed a significant SRE in *both* the incidental and the classic evaluative processing conditions, similar to neurotypical comparison participants. Importantly, there was no significant association

between the size of the SRE and autistic traits in either condition, and no significant difference between size of SRE between adults and children with ASC and NT comparison participants, thus suggesting that self-representation is unrelated to social functioning difficulties. Moreover, participant group demographics were matched to those of Henderson et al. (2009) and Lombardo et al. (2007) to allow for direct comparison of results. Therefore, Lind et al.'s (2020) results appear to challenge the validity of the previous experiments (Henderson et al., 2009; Burrows et al., 2017; Lombardo et al., 2007).

However, on closer inspection, there are some differences between the experimental procedures in Lind et al. (2020) and the earlier three experiments (Henderson et al., 2009; Burrows et al., 2017; Lombardo et al., 2007). First, Henderson et al. (2009), Burrows et al. (2017) and Lombardo et al. (2007) required participants to encode a larger set of words (42/42/120 words respectively) than Lind et al. (2020; 48 test words in the explicit condition), with more distractor words at retrieval (84/84/120 words respectively vs. 24 in Lind et al., 2020); both of which reduce the difficulty of the task. Second, whereas the three earlier studies tested each level of referent-processing (self, familiar-other and unfamiliar-other) in *separate blocks* of the experiment, Lind et al. (2020) included each level of referent-processing *within each block*. Third, the identity of the "other" differed between experiments. Whereas the three earlier experiments (Henderson et al., 2009; Burrows et al., 2017; Lombardo et al., 2007) used Harry Potter (a fictional character) as the "other", Lind et al. (2020) used a member of the Royal family (The Queen or Prince Charles; gender-matched with participant) or celebrities (e.g., Taylor Swift or Justin Bieber; gender-matched with participant) for adult and children/adolescent samples, respectively. Acquiring knowledge about a celebrity arguably requires stronger social communicative understanding to decipher the often biased information portrayed in the media[4]. Alternatively, acquiring knowledge about fictional characters

[4] Indeed Lombardo et al. (2007) found that adults with ASC showed poorer memory relative to NT adults for the self (ASC: *M* = 1.38; *SD* = 0.51; NT: *M* = 1.68; *SD* = 0.51.; difference, *p* = 0.027) and similar-other (best friend) (ASC: *M* = 1.14; *SD* = 0.43; NT: *M* = 1.44; *SD* = 0.40.; difference, *p* = 0.007) conditions – of which both

(e.g., Harry Potter) is dependent on an un-biased "all-knowing" narrator. Therefore, in line with familiarity effects observed in neurotypical individuals (self>familiar-other>unfamiliar-other) self-other distinction and subsequent self-bias might be easier when one has little knowledge of the other (e.g., celebrity, e.g., Lind et al., 2020) than when they have strong knowledge of the other (e.g., Harry Potter, Henderson et al., 2009; Burrows et al., 2017; Lombardo et al., 2007).

Intact SRE in ASC only under certain conditions (e.g., Lind et al., 2020) might reflect Bowler et al.'s (2014; 2004) Task Support Hypothesis. This proposes that performance difficulties in ASC can be minimised by procedures that i) support memory retrieval at *test* by providing cues to the studied information – e.g., category labels (Bowler et al., 2010), word fragments (Boucher & Warrington, 1976), rhymes (Tager-Flusberg, 1991) and recognition tasks (Beversdorf et al., 2000); and/or ii) support memory encoding at study, by presenting categories that have minimal overlap (e.g., self, and dissimilar-, rather than similar-other), and fewer study/distractor items. Therefore, reduced task difficulty might simply better support individuals with ASC to compensate for their impaired self-referential cognition.

Overall, it appears that there is a growing amount of evidence finding that self-bias might be intact in ASC (e.g., Nijhof et al., 2022; Nijhof & Bird, 2019; Williams et al., 2018; Lind et al., 2020), contrary to earlier conclusions (e.g., Lombardo et al., 2010; Henderson et al., 2009; Burrows et al., 2017; Grisdale et al., 2014). This has led researchers to consider the potential that this area of research has been affected by the "file drawer effect" (e.g., Lind et al. 2020), whereby some results/research on self-bias in ASC – particularly ones that do not seem to fit with previous findings – remains unpublished, thus potentially causing a publication bias. Therefore, we conducted a formal meta-analysis on all relevant self-bias experiments comparing an ASC and neurotypical sample, to compare the size of the self-bias between ASC and neurotypical groups. One possibility is

require social communicative experience/understanding – but not the dissimilar-other (Harry Potter) (ASC: *M* = 0.83; *SD* = 0.47; NT: *M* = 1.02*; SD* = 0.49; difference, *p* = 0.450) condition.

that self-bias in tasks requiring a more basic first-order representation of self is intact in ASC, however difficulties become apparent during tasks requiring second-order representation of self if sufficient task support is not provided. Therefore, we also conducted a meta-analysis on first-order and second-order tasks separately.

Self-Bias in Autism Spectrum Conditions: A meta-analysis

METHOD

Literature Search

Searches were conducted on PubMed and Google Scholar for the terms "autistic"/ "autism" and the terms "self-awareness" / "self-bias"/ "self-reference"/ "SRE"/ "self-referential"/ "self-referenced", finding papers published up till November 2020. The 54 empirical papers were grouped according to cognitive domain (e.g., perception/attention, memory) and level of self-awareness (e.g., First level/I self, vs. Second level/Me self). The six papers (which includes 10 experiments/conditions) identified for review are presented in Table 1 according to group.

Inclusion Criteria: We included studies published in peer-reviewed journals and written in English that compared a group of participants with an official diagnosis of ASC with a group of neurotypical participants. Conference presentations, Master's and Doctoral theses were not included (see Figure 1.7 for selection process). Articles that were not directly relevant (e.g., reviews, tasks that do not measure self-bias) and those in which behavioural data were not (fully) reported, were excluded.

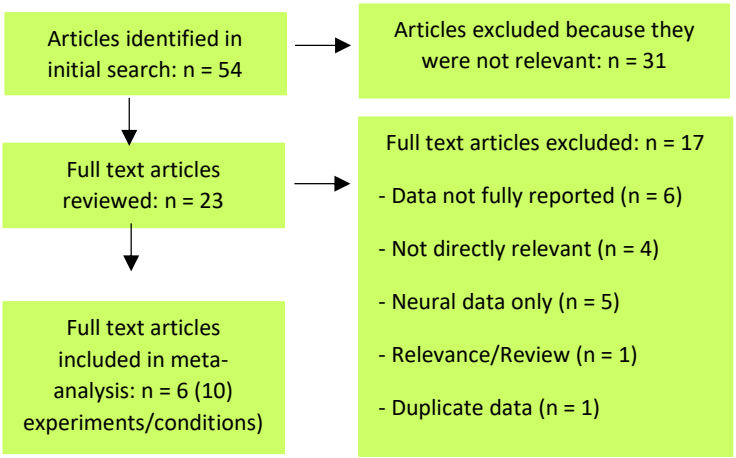


Figure 1.7. Article selection process. 23 full text articles were reviewed, six (including 10 experiments/conditions) of which were included in the final meta-analysis.

Coding and Analysis

Studies were coded for the following variables:

1) Number of male and female participants in the ASC and neurotypical groups.
2) Diagnosis and diagnostic criteria used for diagnosis of participants with ASC.
3) Mean age and standard deviation of participants in the ASC and neurotypical groups.
4) Participant's level of function, including mental age (verbal and non-verbal), intelligence quotient [Full-scale IQ (FSIQ), Verbal IQ (VIQ), and Performance IQ (PIQ)] and whether groups were matched according to these variables.
5) Task: i.e., trait-memory paradigm, shape-label matching task; and the cognitive domain that this task requires: i.e., memory, perception/attention, agency.
6) Highest level of self-awareness that the task is thought to require.

Some papers did not report full data (e.g., the data was plotted on graphs but not reported as numbers), so the authors were contacted to obtain the means and standard deviations of each group/condition. We are very grateful to all those authors who kindly responded to our request. We calculated Hedges' *g* – a measure of effect size which corrects for small sample sizes – using the mean and standard deviation data (Hedges, 1981). Papers that did not report the means and standard deviations for behavioural performance could not be included. If papers reported behavioural performance data for more than one condition – i.e., recognition of own face and recognition of own names – these data were recorded in a separate row in Table 1.

Table 1. Participant group demographics, task type and results for each study identified.

Paper	**Task**	**Cognitive domain**	***n*** **Male, Female**		**Mean Age (SD)**		**Mean FSIQ (SD)**		**Mean VIQ (SD)**		**RESULTS: Mean (SD) Self-Bias**		**Hedges' *g* (SE) for group difference in self-bias**
			ASC	**NT**	**ASC**	**NT**	**ASC**	**NT**	**ASC**	**NT**	**ASC**	**NT**	
Lombardo et al. (2007)	Trait-memory paradigm	**Memory (second-order)**	23, 7	23, 7	29.93 (7.40)	29.93 (7.83)	117.23 (13.11)	117.10 (8.65)	Not reported		0.45 (0.38)	0.66 (0.46)	-0.49 (0.26)
Burrows et al. (2017)	Trait-memory paradigm	**Memory (second-order)**	68, 11	53, 20	12.55 (2.57)	13.23 (2.18)	Not reported	101.51 (14.91)	109.73 (12.28)		0.13 (0.71)	0.62 (0.68)	-0.70 (0.17)
Lind et al. (2020)	Trait-memory paradigm explicit	**Memory (second-order)**	22, 8	24, 6	34.89 (11.32)	39.31 (13.97)	Not reported		103.33 (12.80)	105.70 (10.08)	0.42 (0.65)	0.29 (0.58)	0.21 (0.26)
Lind et al. (2020)	Trait-memory paradigm implicit	**Memory (first-order)**	22, 8	24, 6	34.89 (11.32)	39.31 (13.97)	Not reported		103.33 (12.80)	105.70 (10.08)	0.17 (0.45)	0.2 (0.42)	-0.07 (0.26)

Lind et al. (2020)	Trait-memory paradigm explicit	**Memory (second-order)**	19, 7	18, 8	12.64 (1.52)	13.28 (1.62)	Not reported		105.21 (10.18)	110.23 (11.83)	0.41 (0.54)	0.29 (0.67)	0.19 (0.28)
Lind et al. (2018)	Trait-memory paradigm implicit	**Memory (first-order)**	19, 7	18, 8	12.64 (1.52)	13.28 (1.62)	Not reported		105.21 (10.18)	110.23 (11.83)	0.15 (0.49)	0.14 (0.64)	0.02 (0.28)
Grisdale et al. (2014)	Ownership paradigm	**Memory (second-order)**	13, 3	12, 4	31.33 (9.91)	35.31 (14.70)	112.31 (14.49)	116.63 (12.44)	110.19 (13.99)	113.07 (13.78)	-0.01 (0.17)	0.09 (0.12)	-0.67 (0.36)
Gillespie-Smith et al. (2018)	Ownership paradigm; Chronological age-matched	**Memory (second-order)**	20, 2	16, 6	12.30 (3.1)	12.0 (3.6)	Not reported		Not reported		0.17 (0.20)	0.06 (0.13)	0.66 (0.31)
Gillespie-Smith et al. (2018)	Ownership paradigm	**Memory (second-order)**	20, 2	13, 9	12.30 (3.1)	10.5 (3.6)	Not reported		Not reported		0.17 (0.20)	0.05 (0.07)	0.80 (0.31)

	; Verbal age-matched												
Williams et al. (2018)	Shape-Label Association	**Perception (first-order)**	18, 4	16, 5	35.84 (11.55)	36.32 (12.01)	101.41 (17.21)	107.10 (9.55)	101.95 (15.31)	107.00 (9.97)	0.20 (0.16)	0.21 (0.16)	-0.06 (0.31)

*Indicates studies/tasks that were included in the overall meta-analysis.

RESULTS

Self-bias difference between diagnostic groups: Comparing the self-biases between the ASC groups and the neurotypical groups, a random effect analysis (Restricted Maximum Likelihood) revealed that the overall effect size for the between-group difference was -0.03 (*SE* = 0.16; with 95% confidence limits from -0.34 to 0.29), *z* = -0.17, *p* = 0.87, suggesting no significant between-group difference in self-bias. However, the studies were very heterogeneous, *Q*(9) = 34.29, *p* = <.001, suggesting that effect sizes across studies significantly differed from each other (Hedges et al., 1999). See figure 1.8.

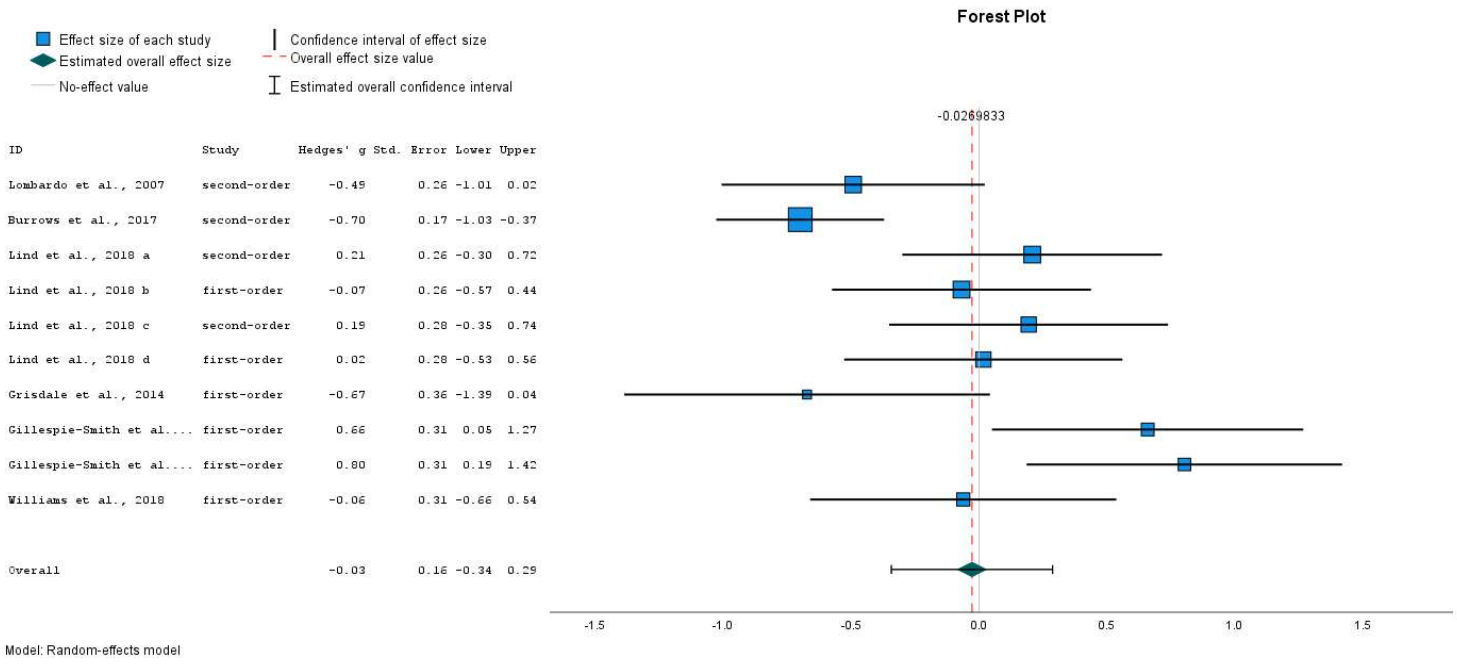


Figure 1.8. Forest plot for depicting between-group difference in self-bias. Effect sizes for each experiment are marked (blue square) along the x-axis (whiskers depict 95% confidence intervals). The size of each blue mark is proportional to the weight that each study has in determining the RE model combined effect size estimate (green diamond shape, whiskers depict 95% confidence interval for the estimate). Statistics are displayed numerically on the left side of the plot (effect size, SE, and confidence intervals). Vertical grey line depicts null-effect size value, and vertical red dashed line depicts observed effect size value.

Impact of Age and IQ: When including mean age of the ASC participants as a covariate in the meta-analysis, a random effect analysis (Restricted Maximum Likelihood) revealed that the overall effect size for the between-group difference was -0.02 (*SE* = 0.02; with 95% confidence limits from -0.05 to 0.02), *z* = -0.92, *p* = 0.36. When including mean VIQ of the ASC participants as a covariate in the meta-analysis, a random effect analysis (Restricted Maximum Likelihood) revealed that the overall effect size for the between-group difference was -0.01 (*SE* = 0.07; with 95% confidence limits from -0.14 to 0.12), *z* = -0.19, *p* = .85. Therefore, neither the mean age nor mean VIQ of the ASC participants predicted the magnitude of the between group difference in self-bias (see figure 1.9).

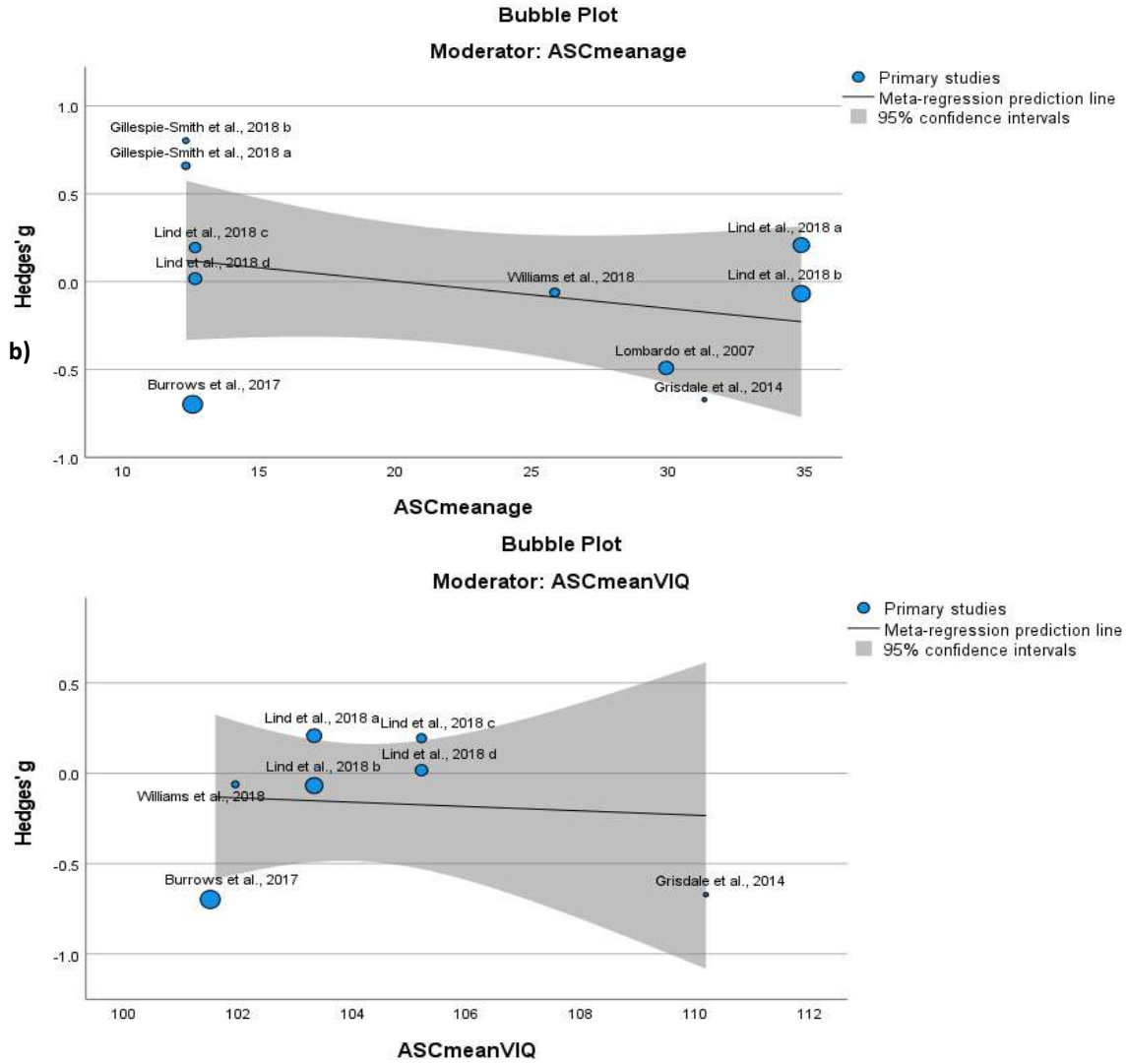


Figure 1.9. Bubble plots depicting non-significant negative relationships between effect size and ASC participants mean age (a) and mean VIQ (b). The size of each "bubble" is proportional to the weight that each study has in predicting the overall effect size.

Level of consciousness: Self-bias comparison between the ASC groups and the neurotypical groups was also conducted in tasks requiring a first-order level of consciousness and second-order level of consciousness separately. A random effect analysis (Restricted Maximum Likelihood) revealed that the overall effect size for the between-group difference in *first-order tasks* was 0.12 (*SE* = 0.208; with 95% confidence limits from -0.29 to 0.53), *z*=.59, *p* = .55. A random effect analysis (Restricted Maximum Likelihood) revealed that the overall effect size for the between-group difference in second-order tasks was -0.23 (*SE* = 0.24; with 95% confidence limits from -0.69 to 0.24), *z* = -0.94, *p* = .35. Both the first-order (*Q*(5) = 13.65, *p* =0.02) and second order (*Q*(3) = 12.87, *p* .005) studies were very heterogeneous, suggesting that effect sizes across studies significantly differed from each other (Hedges et al., 1999).

Impact of Publication Bias: To test the possibility of a publication bias, from the nine published studies we plotted the study precision (1 / standard error) against observed effect sizes (see *Figure* 10). This shows a slight asymmetry**:** the majority of primary (observed) studies (*blue-filled* points in *Figure* 1.10) are towards the right of the funnel plot, which suggests the possibility of a publication bias. Therefore, we used Duval and Tweedie's (2000) trim and fill method which allows an assessment of whether there might be an issue of unpublished data in this area of research. This attempts to add potential missing studies, by adding points that make the funnel plot symmetric (e.g., *green-filled* points in *Figure* 1.10). This imputed one additional missing study (see *Figure* 1.10). We then repeated the standard random effects analysis with this additional study, which estimated an overall effect size for the between-group difference was -0.11 (SE = 0.17; with 95% confidence limits from -0.44 to 0.22), z = -.65, p = 0.52. Therefore, even after the trim and fill, the effect size for the between-group difference in self-bias still shows no significant between group difference.

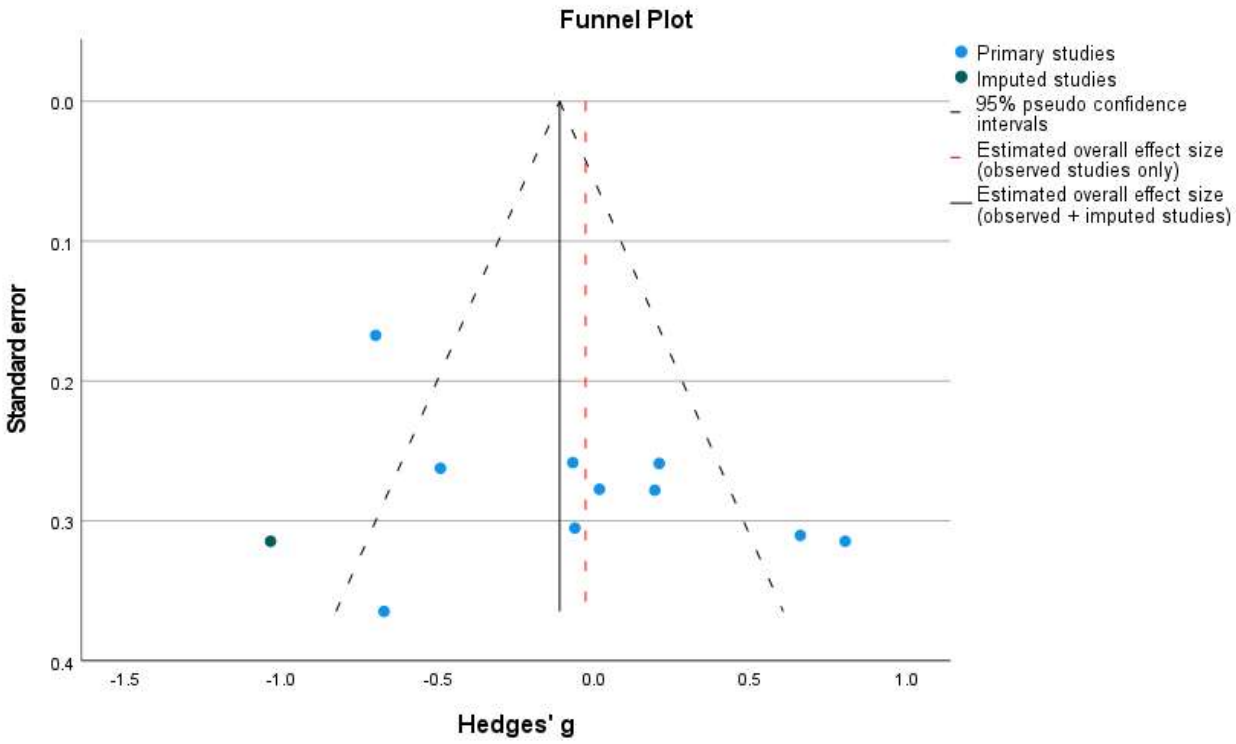


*Figure 1.10. F*unnel plot which shows that most published studies (blue-filled points in) are towards the right of the funnel plot, suggesting the possibility of a publication bias. Trim and fill analysis imputed one new data point (green-filled point), which make the funnel plot more symmetrical. Vertical red dashed line depicts estimated overall effect size of observed studies, and black line depicts estimated overall effect size of observed and imputed studies together.

DISCUSSION

The current meta-analysis aimed to compare the size of the self-bias between neurotypical and autistic groups across studies. Over the six papers (including 10 experiments/conditions) included in the analysis, which included 225 autistic and 240 neurotypical participants in total, we found no evidence of diminished self-bias in ASC overall, or in either the first-order or second-order levels of consciousness tasks separately, and when controlling for age and VIQ of the autistic group. Even when we applied a trim-and-fill analysis to consider the possibility of a publication bias, the addition of one new study depicting diminished self-bias in autistic compared to neurotypical groups, did not

change the results: the difference in self-bias between neurotypical and autistic groups was not significant. These findings increase our confidence in the suggestion that this area of research might have been misled by a few early studies finding significantly diminished self-bias in ASC. Later studies have not replicated this impairment. In the remainder of this chapter, we first outline what an intact self-bias might suggest about sense of self and self-other distinction in ASC. Then we consider methodological issues which cause us to question the validity of intact self-bias in ASC, and consider the possibility that underlying impairments are compensated for by alternative learnt strategies.

Evidence of no difference in self-bias between autistic and neurotypical groups might challenge the notion of a diminished – if not "absent" – sense of self in ASC (Scott, 2004). Monitoring for self-relevant information is thought to be a perpetual goal of the working self (as suggested by the default state which involves self-oriented thought), and so a SRE should be observed not only when information is scaffolded by autobiographical knowledge – i.e., as in the self-referential trait memory paradigm – but for any information encoded in association with the self (see Bluck, 2003). This is driven by the self-attention network (SAN; Humphreys & Sui, 2016; Conway & Pleydell-Pearce, 2000; see Figure 1.5), and would explain self-bias across a range of tasks. For this to occur, one must be able to monitor for self-relevant information, and hence sufficient awareness of one's own self-concept is required to perceive the association between external information and the self. Therefore, this might suggest that self-referential cognition, and hence self-awareness, is intact in ASC. Furthermore, intact self-bias challenges the notion of impaired self-other distinction. If self-other distinction is impaired, then self-bias would be diminished since everything is processed equivalently.

Methodological Issues in Self-Bias Research

It is possible that lack of between group difference in self-bias reflects a type 2 error. There are at least two reasons why this might occur. First, power analysis suggests that to detect a large (effect size = 0.8) between group difference with a power of 0.95, 35 participants are required in each

group (Uljarevic & Hamilton, 2013). In the current meta-analysis, the average number of participants in each study was 30 in the ASC group and 30 in the neurotypical group. Therefore, studies might be underpowered to detect a between-group difference. Only Burrows et al. (2017) included a sample of more than 35 participants per group (79 ASC and 73 neurotypical participants), and this was one of the only two studies included in this meta-analysis to find significantly smaller self-bias in autistic than neurotypical groups (Hedges' g = -0.70; SE = 0.17). Second, experiments with autistic participants typically include participants with an average or above average IQ, and sufficient language abilities to complete the tasks. Therefore, their performance is considered reflective of, but perhaps not representative of, all people with an autism diagnosis; approximately 50% of whom have below average IQ and language impairments (e.g., Bertrand et al., 2001; Charman et al., 2011; Fombonne, 2009). Therefore, it is possible that the few studies finding impaired self-bias in ASC might simply reflect a proportion of the ASC population that have more severe impairments/specific impairment of abilities required for self-bias, and who are more reflective of the autism phenotype (Mottron, 2021). The studies finding intact self-bias reflect high functioning abilities of participants typically included in research experiments, who are able to compensate for their underlying impairments associated with ASC.

Evidence on self-processing in ASC currently relies on only a few (or often a single) studies (Nijhof & Bird, 2019). In the first instance, replication might be particularly useful for studies that found an effect that differed significantly from other studies. In the current meta-analysis, Grisdale et al.'s (2014) findings depicting a moderate (Hedges' *g* = -0.67) reduction in self-bias in autistic adults is particularly striking given that more recently, Gillespie-Smith et al. (2018) observed intact self-bias in children with ASC using a similar task(Gillespie-Smith et al., 2018). The issue of small sample size might point to the increased use of an individual differences approach. This involves an entirely neurotypical participant sample which makes it easier and often more cost-effective to recruit a large sample size (e.g., Psychology undergraduate students). Since clinical autism is considered the extreme of a continuous distribution of traits in the general population, it is thought

to be possible to learn about autistic people by analyzing the correlation between social-cognitive task performance and the number of ASC traits possessed by people in the general population. As such, a negative correlation between autistic traits and self-bias (self condition score minus other score condition score) would indicate that people with more autistic traits – and hence those with an ASC diagnosis – have a lower self-bias than people with fewer autistic traits. This might help to overcome the issue of small sample sizes; however subsequent replication in a case-control design would support the validity of any conclusions.

Typically, in meta-analyses, if an identified paper reported more than one analysis for the same group of participants – i.e., all participants with ASC completed the recognition of own face and recognition of own name conditions – then it is typically best to not include data from the same group of ASC participants twice. Therefore, typically only one of the results – the one that used the more common task/stimuli to minimise heterogeneity of the results – is included in the meta-analysis (e.g., Uljarevic & Hamilton, 2013). However, given the small number of experiments on self-bias in autism, the current meta-analysis included maximal data – and hence some of the groups of participants make up more than one of the 10 data points. Nevertheless, even when we take the former approach, the results are the same: no between group difference in self-bias in ASC.

Conclusion and organisation of this thesis

Overall, the meta-analysis suggests that self-bias is intact in ASC. In both the overall meta-analysis and in the first-order and second order tasks separately, the effect size for the between-group difference in self-bias was not significant. This diminishes the validity of earlier conclusions of impairment in ASC (e.g., Lombardo et al., 2010; Henderson et al., 2009; Burrows et al., 2017; Grisdale et al., 2014). However, given the small number of experiments, we might be underpowered to detect an effect, thus highlighting the importance of further investigation in this area, including replication attempts. One possibility is that the discrepant findings across studies is caused by high functioning participant having learnt to use compensation strategies which successfully mask

impairments in some tasks but not others (Nijhof & Bird, 2019; Hull et al., 2020; Livingston & Happé, 2017). Therefore, further investigation of which aspects of self-related processing are and are not impaired (in the absence of compensation) is required to better understand the associated social communication and interaction abilities in ASC and in the neurotypical population more generally.

Therefore, across five experimental chapters (a total of 11 experiments), this thesis investigates typical and atypical development of self-referential cognition by analysing evidence of self-biases in neurotypical adults and adults with autism spectrum conditions (ASC). This took a multi-modal approach, whereby we tested behavioural indices of self-referential processing in the cognitive domains of perception, action, mental states, and furthermore, for two experiments, we used ERP/EEG to measure the online mechanisms underlying self-bias. We included individual difference approach experiments whereby we investigated the association between task performance and autistic traits in neurotypical adults; we also included case-control experiments to test whether task performance differs between neurotypical and autistic adults. This aimed to gain a better understanding self-awareness/self-referential cognition in ASC, and furthermore, use evidence from ASC as a model for understanding self-awareness/self-referential cognition in the neurotypical population. We will then re-run the meta-analysis, including the new data from the case-control experiments in the current thesis.

Chapter 2 (mental-state level) presents two experiments investigating whether the self-bias in memory using the ownership task (Cunningham et al., 2008) is associated with autistic traits (Experiment 1.1), and is intact in autistic relative to neurotypical participants (Experiment 1.2). This aims to determine the reliability of previous findings (e.g., Grisdale et al., 2014; Gillespie-Smith et al., 2018; Wuyun et al., 2020): we improved on previous studies that have examined this question by using well-powered sample sizes, pre-registered plans and more sensitive (Bayesian) statistical methods.

Chapter 3 (mental-state level) investigates whether *scene construction* – the ability to mentally generate (i.e., imagine) and maintain a coherent spatial scene/event in mind (critical for episodic simulation) – is intact in autistic relative to neurotypical participants (Experiment 2). Episodic memory and semantic knowledge, are thought to provide the building blocks and conceptual framework respectively for episodic simulation (Conway et al., 2019) – the ability to mentally simulate or imagine events (Schacter et al., 2012). Therefore, in ASC, atypical influence of the self on memory (e.g., investigated in Chapter 2) might influence differences in episodic simulation (Conway et al., 2019).

Chapter 4 (perceptual level) presents two experiments aiming to extend findings (Williams et al., 2018; Zhao et al., 2018) of undiminished perceptual self-bias in ASC in a perceptual-association task (Sui et al., 2012) using unfamiliar faces that have been matched with labels of differing levels of *familiarity* (e.g., self vs. friend (familiar-other) vs. stranger (unfamiliar-other)). Event related potential (ERP) technique is used which allowed us to test for real-time evidence of compensatory mechanisms that might enable autistic adults to mask difficulties with self-other distinction at a behavioural level.

Chapter 5 (mental-state level) presents two experiments investigating whether malleability of self-representations is associated with autistic traits (Experiment 4.1), and is intact in autistic relative to neurotypical adults (Experiment 4.3), using a task in which simulating the knowledge state of other people has been found to change self-knowledge, causing one to incorporate information about the simulated person into the self (termed, simulation induced malleability). Furthermore, an additional two experiments test whether self-reflection might also change other knowledge, causing one to incorporate information about the self into the other (termed, self-reflection induced malleability). We test whether this is associated with autistic traits (Experiment 4.2), and whether this differs between neurotypical and autistic adults (Experiment 4.4). As such, malleability of self and/or other representations have the potential to explain differences in the extent of self-other distinction.

Chapter 6 (action level) presents two experiments using a stimulus response compatibility (SRC) paradigm (Brass et al., 2000) to investigate whether the ability to distinguish our own motor repertoire from those displayed by other people of differing levels of similarity – specifically, the self vs. unfamiliar adult (similar-other) vs. unfamiliar child (dissimilar-other) – is associated with autistic traits (Experiment 5.1), and is intact in autistic relative to neurotypical adults (Experiment 5.2). Power spectrum analysis of the electroencephalogram (EEG) is used to measure the online mechanisms underlying action mirroring, which should reveal whether autistic people adopt compensation strategies which obscure evidence of impaired self-other distinction in this task.

Open Science All experimental chapters in this thesis take an open science approach; the method and analysis plans for the case-control experiments in Chapters 2, 4, 5 and 6, were pre-registered on the open science framework (OSF), and furthermore, data files for all experiments are deposited on OSF. This aims to make our research more transparent and accessible to all levels, and encourages others to attempt to replicate our studies to determine the reliability of our results.

Power analyses using G*Power 3.1 (Vinet & Zhedanov, 2010) was conducted for each experiment to determine the required sample size to detect a between group difference in task performance, using Cohen's (1992) criterion. **Statistical significance** for all analyses was determined based on an alpha level of 0.05. Where t-tests were used, effect sizes are reported as Cohen's *d* values (≥0.20 = small effect; ≥0.5 = moderate effect; ≥0.80 = large effect; Cohen, 1992). When ANOVAs were used, effect sizes are reported as η_p^2 (≥0.01 = small effect; ≥0.06 = moderate effect; ≥0.14 = large effect; Cohen, 1969). Where correlations were used, ≥0.30 = small effect, ≥0.50 = moderate effect, and ≥0.70 = large effect (Cohen, 1969). **Bayesian analysis** is used to supplement null hypothesis significance testing. Bayesian analysis enables a more graded interpretation of data compared to simply using p values or effect sizes, by estimating the relative strength of the alternative over the null hypothesis or vice versa. (e.g., Rouder et al., 2009). We include Bayesian analysis to interpret our results (conducted using JASP 0.8.1; JASP team, 2020): Bayes factor (BF^{10}) < 1 provides evidence for the null

hypothesis (<0.33 provides firm evidence), whereas Bayes factors > 3, >10, > 30, and >100 provides firm, strong, very strong, and decisive evidence for the alternative hypothesis respectively (Rouder et al., 2009).

Chapter 2| Revisiting the Ownership Effect in Adults with and without Autism

Revisiting the Ownership Effect in Adults with and without Autism

Marchella Smith[1], David Williams[1], Sophie Lind[2], & Heather J Ferguson[1]

[1]School of Psychology, Keynes College, University of Kent, Canterbury CT2 7NP, UK

[2]Department of Psychology, City, University of London, Northampton Square, London, EC1V 0HB, UK

Corresponding author

Marchella Smith (mmls2@kent.ac.uk)

Author contributions

MS designed the study, data collection, data analysis and interpretation, and drafted the manuscript; DW conceived of the study, designed the study, and drafted the manuscript; SL conceived of the study, and drafting the manuscript; HJF data analysis and interpretation, and drafting and revising the manuscript. All authors gave final approval for publication.

Acknowledgements

The datasets supporting this article are available on the Open Science Framework (https://osf.io/y2mf9/?view_only=dbaf2b8d791b4386befa6beca75a844c)

ABSTRACT

Self-owned items are better remembered than other-owned items; this ownership effect reflects privileged processing of self-related information. The size of this ownership effect has been shown to decrease in neurotypical adults as the number of autistic traits increases, and is reduced in autistic adults. However, emerging evidence has questioned the reliability of these findings. This paper aimed to replicate previous work using well-powered, pre-registered designs, and Bayesian analyses. Experiment 1.1 (N=100) found a significant ownership effect in neurotypical adults; however, the size of this was unrelated to individual differences in autistic traits. Experiment 1.2 (N=56) found an ownership effect in neurotypical but not autistic adults. We discuss how these findings can be explained by impairments in psychological self-awareness in autism.

Keywords: autism; ownership effect; self-reference effect; self-awareness; memory

The Self is widely thought to affect memory processing in a variety of ways (Conway & Pleydell-Pearce, 2000; Craik & Tulving, 1975; Humphreys & Sui, 2016). This influence has been demonstrated empirically by research into self-biases in performance on memory tests. It is widely observed among neurotypical people that memory for self-referential information is superior to memory for information relating to other people. This self-bias, or "self-reference effect", is thought to be driven by preferential processing of self-relevant information relative to other-relevant information (Tversky & Kahneman, 1974). One of the most widely studied effects is the so-called "ownership effect", which reflects superior recall or recognition memory for objects previously encoded as being (notionally) owned by the self than objects previously encoded as being (notionally) owned by someone else (Cunningham et al., 2008; Grisdale et al., 2014). This self-bias is thought to reflect distinct cognitive processing of items considered extensions of the self (Beggan, 1992). Given that memory is influenced by the self, it follows that any individual with an impoverished self-concept should have diminished and/or atypical memory capacity/profile. Potentially, this is highly relevant for our understanding autism spectrum conditions (ASC)[5].

ASC is a neurodevelopmental condition diagnosed on the basis of social-communication impairments, and a restricted and repetitive repertoire of behaviour and interests (Diagnostic and Statistical Manual of Mental Disorders; DSM-5: American Psychiatric Association, 2013). At the psychological level, ASC is thought to be characterised by difficulties in aspects of both self-awareness and memory, and it has been suggested that these two atypicalities are linked (Lind, 2010). If self-awareness is diminished among autistic people, then the self should not structure memory in the same way that it does for most neurotypical people, and so autistic people should show a reduced (or absent) ownership effect on traditional paradigms.

[5] We acknowledge the debates about the terminology used to describe autism, and therefore use a mixture of commonly endorsed terms to reflect the fact that "there is no single way of describing autism that is universally accepted and preferred" (Kenny et al., 2016).

Indeed, in a seminal study, Grisdale et al. (2014) found that autistic adults showed a significantly diminished ownership effect relative to that shown by age- and IQ-matched neurotypical adults. Whereas neurotypical participants showed a highly significant 9% advantage in memory for items they owned over items owned by an experimenter, autistic participants only showed a slight (non-significant) advantage in memory for items owned by the experimenter over items owned by themselves. The between-group difference in the size of the ownership effect was large and statistically significant, in keeping with the theory that self-referential processing has a reduced/absent effect on memory processing among autistic adults. Furthermore, Grisdale et al. (Experiment 1) adopted an individual differences approach and found that the size of the ownership effect was associated moderately and significantly with the number of autistic traits reported among a group of 40 neurotypical individuals (i.e., the more autistic traits, smaller the ownership effect). Since ASC is seen as a continuous spectrum rather than categorical, autistic traits are thought to vary continuously in the population. Therefore, it is possible to learn about ASC by studying individual differences in autistic traits in a neurotypical population (although further analysis in a case-control sample is also necessary to check the validity).

These results all appeared in keeping with the theory that self-awareness is diminished in ASC and that this contributes to memory difficulties experienced by many autistic people (Lind & Bowler, 2008; Lind, 2010). However, more recent findings are somewhat more difficult to interpret. Firstly, in neurotypical adults studies have found no significant association between autisitc traits (10-item AQ, Allison et al., 2012; or 50-item AQ, Baron-Cohen et al., 2001) and self-bias magnitude in memory (Amodeo et al., 2021; Lind et al., 2019), perception, or attention (Amodeo et al., 2021; Nijhof et al., 2020; Williams, Nicholson, & Grainger, 2018), which suggest that self-bias is not affected by autistic tendencies in neurotypical people. Furthermore, Gillespie-Smith et al. (2018) used the ownership task in a case-control design and reported a significantly *enhanced* ownership effect in autistic children (mean chronological age = 12.3 years) compared to that observed in either neurotypical children matched for chronological age (mean = 12.0 years) or neurotypical children

matched for verbal ability (mean = 13.9 years). Clearly, this is the exact opposite of the pattern found by Grisdale et al. (2014, Experiment 2). However, Gillespie-Smith et al. subsequently performed a median split on the number of parent-reported autistic traits among participants in the ASC group. They then compared the size of the ownership effect in the group with "severe" ASC (i.e. above-the-median scores) and those with "mild" ASC (i.e. below-the-median scores), and found that only the sub-group with mild ASC showed a statistically significant ownership effect. In one respect, the finding that the severe ASC sub-sample did not show a significant ownership effect appears to replicate Grisdale et al.'s results. However, it is not clear whether the results should actually be interpreted as showing that even the severe ASC group showed a diminished ownership effect. The severe ASC sub-sample had only 11 participants and so had low power to detect a true ownership effect. In fact, the effect size for the contrast between self- and other-owned items in the severe ASC sample was large (d = 0.78) and borderline significant ($p < .08$). In other words, even the severe ASC sub-sample preferentially recognised self- over other-owned items to a statistically large extent. Indeed, the size of the ownership effect in the severe ASC sub-sample was actually *larger* than in either of the two neurotypical sub-samples (who showed only small ownership effects with associated effect sizes of d = 0.45 and d = 0.47, respectively) and not significantly smaller than that observed in the mild ASC sub-sample. Arguably, therefore, it is not appropriate to draw the conclusion that autistic participants in Gillespie-Smith et al.'s study showed anything other than a normal – or even enhanced – ownership effect, and this certainly contradicts the pattern observed by Grisdale et al. (2014).

An even more recent case-control study of the ownership effect in ASC was conducted by Wuyun et al. (2020). In their Experiment 1, autistic children were compared to young neurotypical children and age- and IQ-matched children with intellectual disability on a traditional ownership task. Children watched as an experimenter sorted pictures of items into a basket designated as the child's own, or a separate basket designated as the experimenter's own. Subsequent memory for items was assessed and the results indicated that both control groups showed a significant

ownership effect, but the autistic group did not. This appears to replicate the findings of Grisdale et al. (2014), but arguably not those of Gillespie-Smith et al. (2018). In a second experiment, however, three new groups were tested on a version of the ownership paradigm that involved the child, rather than the experimenter, sorting items into their own and the other's baskets. Results showed that all three groups, including the ASC group, showed a significant ownership effect, contrary to findings from their Experiment 1 and the findings from Grisdale et al. (2014). The authors suggested that they expected these results on the basis that "the hands-on experience of identifying ownership would strengthen the connection between self and the self-related items". It is unclear why merely self-sorting would increase the salience of self-owned items, because participants sorted *both* self-owned and other-owned items and so motor traces associated with action should have scaffolded memory for both types of item (Oberman & Ramachandran, 2007; Rizzolatti & Sinigaglia, 2016; Williams, 2008). More importantly, the self-sorting procedure used by Wuyun et al. in their Experiment 2 was similar to the procedure used by Grisdale et al. who observed an absence of the ownership effect in ASC, contrary to Wuyun et al.'s Experiment 2 results.

The point in the analysis of previous research on the ownership effect in ASC is not to criticise the studies or suggest methodological flaws, but rather to highlight that findings are inconsistent across studies. Despite our own previous arguments that self-biases of various kinds are diminished in ASC, the inconsistency of results across studies has led us to be more cautious in drawing conclusions about the nature of self-awareness (and its relation to memory) in ASC. Therefore, the current study aimed to replicate the study by Grisdale et al. (2014), which produced the clearest evidence of a diminished ownership effect in ASC, but using well-powered sample sizes, pre-registered plans and more sensitive (Bayesian) statistical methods. We felt it important to try and replicate Grisdale et al. (2014) given the risk of drawing false conclusions based on limited studies and the potential that this area of research might have been misled by publication biases (Lind et al., 2019).

Experiment 1.1 | Individual Differences Approach

METHOD

Experiment 1.1 was fully pre-registered on the Open Science Framework (https://osf.io/y2mf9/?view_only=dbaf2b8d791b4386befa6beca75a844c)

Participants

A volunteer sample of 100 (16 male; 84 female) undergraduate students, aged 18-58 years (*M* = 20.11; *SD* = 4.80), from the University of Kent took part. An *a priori* power analysis using G*Power (Vinet & Zhedanov, 2010) revealed that 69 participants would be enough to detect a correlation between AQ and level of ownership effect with effect size 0.33 (Grisdale et al., 2014, Exp. 1) and power of .80. All participants had normal or corrected-to-normal hearing and vision and had no history of psychiatric or neurological conditions (self-report). All participants provided written informed consent and following completion of the tasks were fully debriefed and received partial course credits for their participation in the current (and an additional unrelated) experiment. Ethical approval was provided by the School of Psychology's Research Ethics Committee at the University of Kent.

Ownership Task

Materials and Stimuli

The stimuli used were identical to those used by Grisdale et al. (2014). This comprised of 222 laminated picture cards (855mm x 685mm) depicting colour photographs of items commonly sold in a supermarket (e.g., pepper, kettle, washing-up liquid; see Item 1 in Appendices). The 222 items were divided into three groups of 74 items. The mean word length and number of syllables of items did not significantly differ between the three groups (see Grisdale et al., 2014). For counterbalancing

purposes, the condition allocated to each group of items was rotated across participants so that each group of items was used for self-owned, other-owned and "lure" item conditions (lure items were only shown in the recognition phase). In the study phase, participants were either presented with red-bordered cards depicting self-owned items and blue-bordered cards depicting other-owned items or vice-versa. Therefore, there were twelve versions of the task, of which participants were randomly allocated to one version. In the recognition test phase, all three sets of items (self-owned; other-owned; lure items) were presented on a card with no coloured border surrounding. Identical red and blue shopping baskets were used to place self-owned and other-owned items during the study phase.

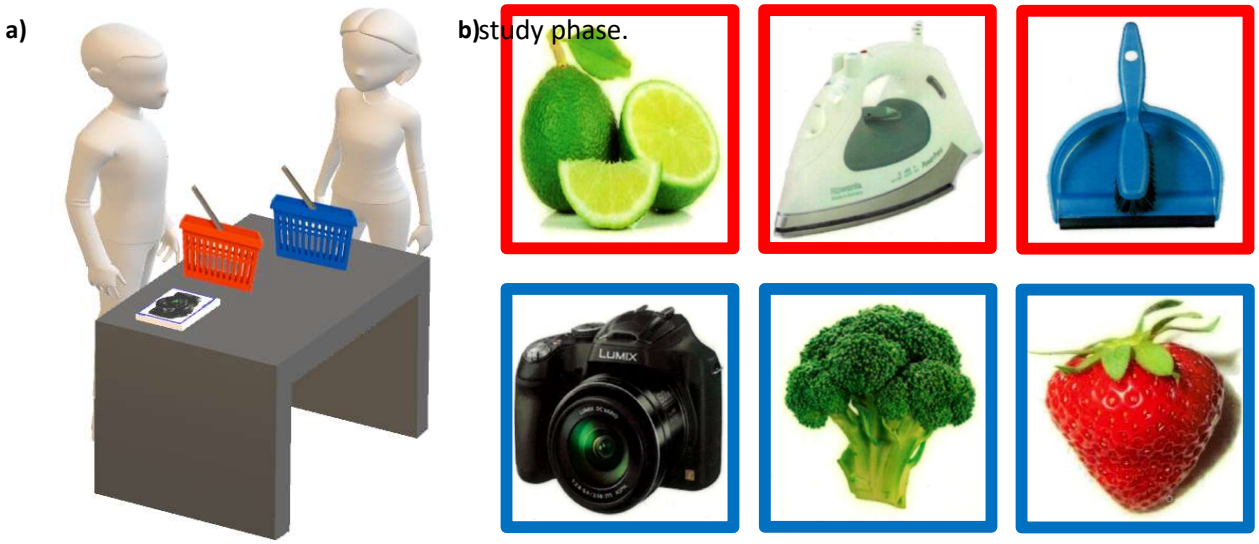


Figure 2.1. Illustration of the Ownership task set up (created in Microsoft Paint 3D) with (a) experimenter/other (female with blue basket) and participant/self (male with red basket), and (b) example items owned by the self (red bordered cards) and owned by the other (blue bordered cards). The colour association was counterbalanced across participants (i.e. half the participants were allocated the red items, and half the participants were allocated the blue items).

Procedure

The procedure replicated that of Grisdale et al. (2014, Exp. 1). In the *study phase*, participants were instructed to sort a pile of 148 (74 red-bordered and 74 blue-bordered) laminated picture cards into their own or the experimenter's shopping basket according to who owned the item (see Figure 2.1).

Participants were asked to attend to the items on each card, and sort them into the associated basket upside-down, one at time. They were told that, for the duration of the experiment, they "owned" the red basket and all of the items presented on a red-bordered card, whereas the experimenter "owned" the blue basket and all the items presented on a blue-bordered card (colour association was counterbalanced across participants). The cards were presented in a random order for each participant with the exception that no more than four cards with the same coloured border could appear in a row. In the subsequent unexpected *recognition test phase*, participants were presented with 222 picture cards, 148 of which depicted the items presented in the study phase and 74 of which were "lure" items that had not previously been studied. None of the picture cards in the recognition test phase had a coloured border and participants were asked to answer whether they did or did not see each item at study (yes-no recognition test).

Autism-spectrum Quotient

Participants completed the Autism-spectrum Quotient (AQ; Baron-Cohen et al., 2001). This is a standardized 50-item self-report questionnaire used to measure people's number of autistic traits. Participants were asked the extent to which they agree (definitely agree, slightly agree, slightly disagree, definitely disagree) with each statement about the self (e.g., "I find social situations easy"). The dependent variable is the total AQ score (ranging from 0-50), with higher scores indicating a greater number of autistic traits (Woodbury-Smith et al., 2005). Good test-retest reliability scores have been observed (range between r = .70 and r = .95) (Broadbent et al., 2013). In the current sample, the mean AQ score was 16.98 (SD = 6.51; ranging from 2-43).

Scoring and Statistical Analysis Procedure

In line with Grisdale et al., each participant's hit rate (proportion of items correctly recognised as having been presented at study), false alarm rate (proportion of lure items that were incorrectly judged as having been presented at study), and corrected hit rate (hit rate *minus* false alarm rate) were calculated for each condition (self-owned/other-owned). Higher corrected hit rates reflected

participant's better recognition memory. The size of the ownership effect was calculated as the corrected hit rate for self-owned items minus the corrected hit rate for other-owned items. The larger the resulting value, the greater the self-bias/ownership effect.

As well as conducting standard inferential statistical analyses, we included Bayesian analysis to interpret our results (conducted using JASP 0.14.1; JASP Team, 2020) because Bayesian analysis enables a more graded interpretation of data compared to simply using *p* values or effect sizes, by estimating the relative strength of the alternative hypothesis over the null hypothesis or vice versa (e.g., Dienes, 2014; Rouder et al., 2009). Bayes factor (BF^{10}) < 1 provides evidence for the null hypothesis (<0.33 provides firm evidence), whereas Bayes factors > 3, >10, > 30, and >100 provides firm, strong, very strong, and decisive evidence for the alternative hypothesis respectively.

RESULTS

Table 1. Descriptive Statistics for proportion of hit rates, false alarm rates and corrected hit rates in each condition in Experiment 1.1.

	Condition	Mean	SD
Hit Rate	Self-owned	.658	.184
	Other-owned	.637	.180
False Alarm Rate	Lure items	.112	.105
Corrected Hit Rate	Self-owned	.547	.214
	Other-owned	.525	.213

To investigate the effect of ownership on corrected hit rate, a repeated measures ANOVA was conducted with Referent (self-owned/other-owned) as the within-subjects independent variable and

corrected hit rate as the dependent variable[6] (see Table 1 for descriptive statistics). All assumptions for the current analysis were met, and alpha was set to .05. The ANOVA revealed a significant main effect of Referent, $F(1, 99) = 6.01$, $p = .02$, $\eta_p^2 = 0.06$, $BF^{10} = 2.31$, reflecting significantly greater corrected hit rate for self-owned compared to other-owned items. The data was also analysed categorically. If participants corrected hit rate for self-owned items was higher than that of other-owned items, then they were considered to have shown an ownership effect. By this definition, 52/100 (52%) participants showed an ownership effect.

A Pearson's correlation analysis examined the degree to which the size of the ownership effect was associated with individual differences in autistic traits (using AQ scores; see Figure 2.2). This revealed no significant correlation between the size of the ownership effect and total AQ score, $r = -.02$, $p = .81$, $BF^{10} = 0.13$. We also directly compared the correlation between size of the ownership effect and the AQ score from the current sample with that of Grisdale et al. (2014). A one-tailed Fisher Z test revealed that the correlation was significantly weaker in our sample than Grisdale et al.'s (2014; r (40) = -.33) sample, $Z = 1.67$, $p = .047$.

[6] When using d-prime (*d'*) as the dependent variable instead of corrected-hit rate, a repeated measures ANOVA with Referent (self-owned/other-owned) as the within subjects independent variable revealed a significant main effect of Referent, $F(1, 99) = 8.07$, $p = .01$, $\eta_p^2 = 0.08$, $BF_{10} = 5.76$ reflecting significantly greater *d'* for self-owned ($M = 1.92$, $SD = .99$) compared to other-owned ($M = 1.84$, $SD = .99$) items.

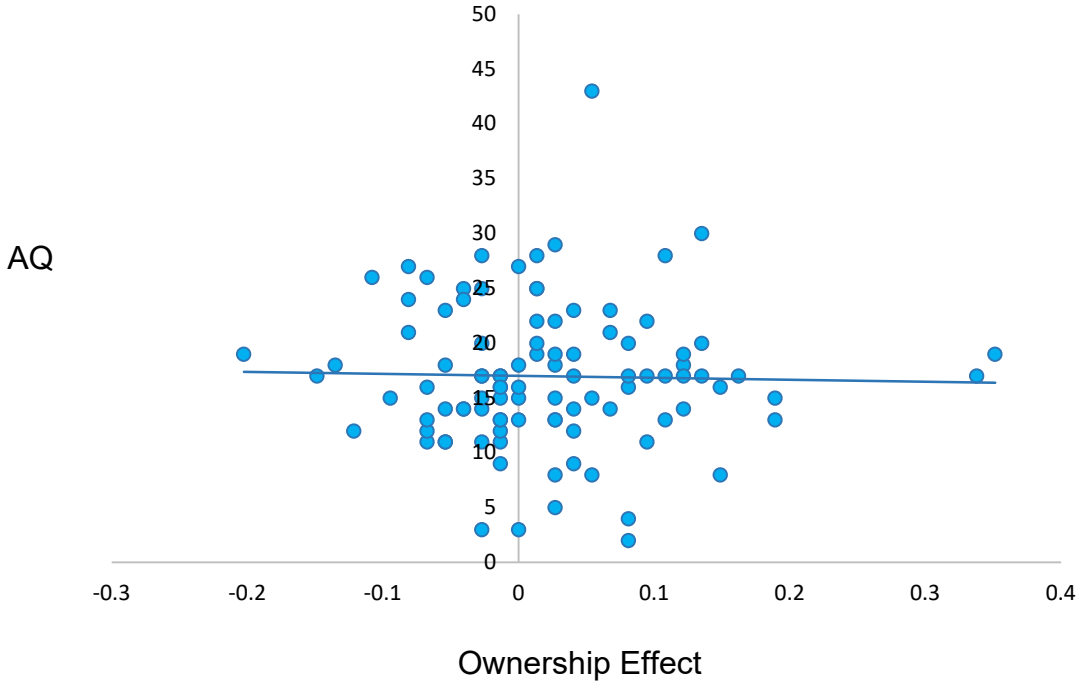


Figure 2.2. Scatter plot depicting non significant correlation between size of ownership effect (corrected hit rate for self-owned items *minus* other-owned items) and AQ. Line depicts line of best fit.

DISCUSSION

The aim of Experiment 1.1 was first to test whether the ownership effect and the association between the size of the ownership effect and self-reported autistic traits (AQ) observed by Grisdale et al. (2014, Exp. 1) could be replicated in a well-powered sample with pre-registered design and analysis. Secondly, like Grisdale et al. (2014), our Experiment 1.1 took an individual differences approach to test whether the size of the ownership effect was associated with the number of autistic traits. The use of a larger sample of participants allows us to determine the reliability of this task in influencing an ownership effect in neurotypical individuals, before using it to draw conclusions about an intact or reduced ownership effect in a smaller sample of autistic participants.

Experiment 1.1 found a significant ownership effect in neurotypical adults whereby significantly more self-owned items were successfully recognised than other-owned items. This replicates the patterns reported in multiple studies of the ownership effect among neurotypical

individuals and supports the notion that items thought to be (in this case notionally) owned by the self, become a psychological extension of the self, which encourages enhanced cognitive processing (Beggan, 1992).

Crucially, we did *not* observe a significant association between the size of the ownership effect and AQ. Bayesian analysis indicated that the null hypothesis was strongly supported by the data. Furthermore, the size of this correlation was significantly smaller than that of Grisdale et al. (2014). Hence this does not replicate Grisdale et al. (2014) and does not support our hypothesis. However, these results are in line with other emerging evidence that a) the ownership effect may not necessarily be diminished among autistic people (Gillespie-Smith et al., 2018; Wuyun et al., 2020), and b) the size of other forms of memory self-bias are not significantly associated with the number of autistic traits manifest by neurotypical individuals, contrary to previous suggestions (Amodeo et al., 2021; Dinulescu et al., 2021; Lind et al., 2019; Williams et al., 2018). Therefore, results from Experiment 1.1 appear to call into question the reliability of Grisdale et al.'s (2014, Exp. 1) findings and more generally, the claim that the ownership effect is significantly affected by autistic traits.

Experiment 1.2 | Case-Control Approach

Whilst autistic traits vary within the neurotypical population, and hence an individual differences approach can inform us about the effects of autistic traits on the ownership effect, it does not rule out the possibility that individuals with clinical levels of autistic traits (i.e. an ASC diagnosis) produce outcomes that are qualitatively different – i.e. facilitated by alternative cognitive processes – compared to neurotypical individuals with high autistic traits. To address this possibility, in Experiment 1.2 we used the same ownership task, but employed a case-control approach as Grisdale

et al. (2014) did, to compare the size of the ownership effect between neurotypical and autistic adults.

METHOD

Experiment 1.2 was fully pre-registered on the Open Science Framework (https://osf.io/y2mf9/?view_only=dbaf2b8d791b4386befa6beca75a844c).

Participants

Participants were recruited from the Autism Research Kent participant database and included 28 neurotypical adults with no history of psychiatric or neurological conditions (self-report) and 28 adults with an ASC clinical diagnosis. Based on Grisdale et al. (2014), an *a priori* power analysis using G*Power (Faul et al., 2007) revealed that 54 participants (27 per group) would be enough to detect a significant group × referent interaction with effect size 0.39 and power of 0.80.

All participants had normal or corrected to normal hearing and vision. The two participant groups were matched on age, gender, and Intelligence Quotient (IQ, Wechsler Adult Intelligence Scale, WAIS-III or WAIS-IV; Wechsler, 1997; Wechsler, 2008), but were not matched on Autism spectrum Quotient (AQ; Baron-Cohen et al., 2001) (see Table 2 for participant demographics).

Participants in the ASC group received a clinical diagnosis of either autistic disorder (N = 6), autism spectrum disorder (N = 4) or Asperger's syndrome (N = 12) or not known (N = 6), of which official diagnostic information was checked. To assess the current autistic characteristics, 25/28 autistic participants were also assessed on module 4 of the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000) by a trained, research-reliable researcher, and videos were double coded to ensure reliability of scoring (inter-rater reliability was found to be excellent with intraclass correlation of .89). In the autistic group, ADOS scores ranged from 1-21, with fifteen individuals

scoring higher than 7 (i.e. the clinical cut off score). AQ scores ranged from 19-45 in the autistic group and from 3-28 in the neurotypical group.

Participants provided written informed consent and were fully debriefed following completion of the experiment. They received £10 per hour of their participation, plus additional travel expenses. Ethical approval was provided by University of Kent Research Ethics Committee.

Table 2. Comparison of participant characteristics between groups in Experiment 1.2.

	Autistic (n = 28; 19 male)	Neurotypical (n = 28; 18 male)			
	Mean (SD)	Mean (SD)	*t*	*p*	*d*
Age (years)	38.77 (12.69)	39.47 (14.18)	0.19	.847	0.05
IQ: Full Scale	106.39 (16.27)	110.59 (10.06)	1.15	.257	0.31
IQ: Verbal	104.61 (12.67)	109.26 (10.97)	1.45	.152	0.39
IQ: Performance	108.36 (21.14)	111.22 (11.16)	0.63	.535	0.17
AQ	30.79 (7.55)	14.39 (6.36)	8.60	<.001	2.34
ADOS: Total	8.71 (5.23)	-	-	-	-

IQ is based on 27 neurotypical and 28 autistic participants.

Design

This was a 2 (Group: autistic/neurotypical) × 2 (Referent: self-owned/other-owned) mixed design with repeated measures on the second factor.

Stimuli and Procedure

The stimuli and procedure were identical to those described for Experiment 1.1, which were identical to those used by Grisdale et al. (2014). As in Experiment 1.1, all participants completed the AQ (Baron-Cohen et al., 2001).

RESULTS

As in Experiment 1.1, each participant's hit rate (proportion of items that were correctly recognised as having been presented at study), false alarm rate (proportion of lure items that were incorrectly judged as having been presented at study) and corrected hit rate (hit rate *minus* false alarm rate) was calculated for self-owned and other-owned items separately (see Table 3 for descriptive statistics).

Table 3. Descriptive statistics for proportion of hit rates, false alarm rates and corrected hit rates in the self-owned and other-owned conditions in autistic and neurotypical participants (Experiment 1.2).

		Autistic		neurotypical	
	Condition	*M*	*SD*	*M*	*SD*
Hit Rate	Self-owned	.492	.241	.680	.180
	Other-owned	.476	.257	.614	.178
False Alarm Rate	Lure items	.111	.107	.129	.104
Corrected Hit Rate	Self-owned	.380	.269	.551	.182
	Other-owned	.365	.276	.485	.191

To investigate the effect of Group and Referent on corrected hit rate a 2 (Group: autistic/neurotypical) × 2 (Referent: self-owned/other-owned) mixed ANOVA was conducted, with repeated measures on the second factor[7]. This revealed a significant main effect of Referent, $F(1, 54) = 11.84$, $p = .001$, $\eta_p^2 = .18$, $BF^{10} = 19.23$, with a significantly greater corrected hit rate for self-owned ($M = .47$, $SD = .24$) compared to other-owned ($M = .43$, $SD = .24$) items. There was also a significant main effect of Group, $F(1, 54) = 5.62$, $p = .02$, $\eta_p^2 = .09$, $BF^{10} = 2.52$, showing that corrected hit rate was greater for the neurotypical ($M = .52$, $SD = .18$) than the autistic ($M = .37$, $SD = .27$) group.

Crucially, there was a significant Group × Referent interaction, $F(1, 54) = 4.54$, $p = .04$, $\eta_p^2 = .08$, $BF^{10} = 39.77$. Bonferroni corrected paired-samples t-tests revealed that neurotypical participants had a significantly greater corrected hit rate for self-owned compared to other-owned items, $t(27) = 3.84$, $p = <.001$, $d = .73$. In contrast, autistic participants[8] showed no significant difference in corrected hit rate according to referent, $t(27) = .95$, $p = .35$, $d = .18$. Furthermore, the corrected hit rate for self-owned items was significantly greater in the neurotypical than the autistic group, $t(54) = 2.78$, $p = .008$, $d = .74$; however, the corrected hit rate for other-owned items did not differ between groups, $t(54) = 1.90$, $p = .06$, $d = .51$.

The data was also analysed categorically, as in Experiment 1.1. This revealed that 20/28 (71.43%) neurotypical and 18/28 (65.29%) autistic participants showed an ownership effect. The between group difference was not statistically significant, $X^2 = 0.33$, $p = .57$.

Association Analysis

An exploratory Pearson's correlation analysis replicated the analysis in Experiment 1.1 and Grisdale et al. to examine the degree to which the size of the ownership effect was associated with individual differences in autistic traits (using AQ scores; see Figure 2.3). This revealed no significant correlation between the size of

[7] When using d-prime (*d'*) as the dependent variable instead of corrected-hit rate, a mixed ANOVA with Referent (self-owned/other-owned) as the within subjects independent variable, and Group (autistic/neurotypical) as the between subjects variable revealed a significant main effect of Referent, $F(1, 51) = 13.88$, $p = <.001$, $\eta_p^2 = 0.21$, $BF_{10} = 0.64$ reflecting significantly greater *d'* for self-owned ($M = 1.53$, $SD = .83$) compared to other-owned ($M = 1.37$, $SD = .80$) items. There was no significant main effect of group, $F(1, 51) = 3.63$, $p = .06$, $\eta_p^2 = 0.07$, $BF_{10} = 0.02$, and no referent x group interaction, $F(1, 51) = 2.68$, $p = .11$, $\eta_p^2 = 0.05$, $BF_{10} = 0.63$.

[8] When splitting the autistic group according to those who did and did not score higher than 7 (i.e., the cut-off) on the ADOS, this revealed that neither the above cut-off, $t(14) = 1.83$, $p = .09$, $d = .47$ or below cut-off group, $t(8) = 1.07$, $p = .32$, $d = .36$, displayed a significant ownership effect.

the ownership effect and total AQ score in the neurotypical group, $r(26) = .12$, $p = .56$, $BF^{10} = 0.29$, or the autistic group, $r(28) = .03$, $p = .90$, $BF^{10} = 0.24$[9]. We compared the correlation between size of the ownership effect and the AQ score from the current samples with that of Grisdale et al. (2014, Exp2). As in Experiment 1.1, a one-tailed Fisher Z test revealed that the correlation was significantly weaker in our neurotypical sample than Grisdale et al.'s (2014; $r(16) = -.11$) neurotypical sample, $Z = 0.67$, $p = .25$. The correlation was also significantly weaker in our autistic sample than Grisdale et al.'s (2014; $r(16) = -.10$) autistic sample, $Z = 0.38$, $p = .35$.

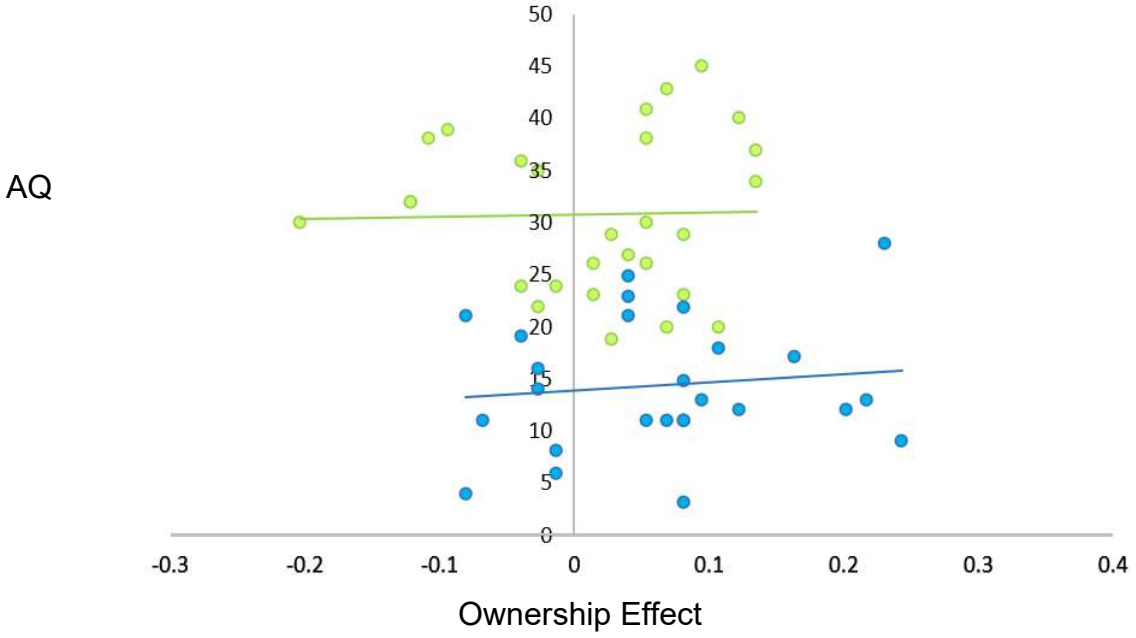


Figure 2.3. Scatter plot depicting non-significant correlations between size of the ownership effect (corrected hit rate for Self *minus* Other) and AQ in neurotypical (blue) and autistic (green) groups. Lines depict lines of best fit.

DISCUSSION

The significant interaction between group and referent revealed that unlike the neurotypical group, the autistic group did not demonstrate a statistically significant ownership effect; rather the autistic group

[9] There was also no significant correlation between the size of the ownership effect and total AQ score when both groups were combined into a larger sample, $r(54) = -.16$, $p = .24$, $BF^{10} = 0.34$.

showed comparable memory for items encoded in relation to the self and in relation to the experimenter. Importantly, whilst there was no group difference in memory for other-owned items, the neurotypical participants were more accurate at recalling self-owned items than autistic participants. However, this was not associated with autistic traits (AQ). Since the categorical analysis revealed that there was no between group difference in the number of people who did/did not display memory biases in the direction of an ownership effect, this suggests that neurotypical participants were no more likely than autistic participants to display an ownership effect, and therefore the group difference reflects a reduced size of ownership effect in ASC rather than an absent one.

GENERAL DISCUSSION

The current study aimed to investigate whether self-referential cognition is impaired in autistic adults relative to neurotypical adults. We tested the self-reference effect in memory – more specifically the ownership effect (Cunningham et al., 2008) – whereby neurotypical people typically display a memory advantage for items encoded in relation to the self (i.e., self-referentially) relative to those encoded in relation to other people (Amodeo et al., 2021; Nijhof et al., 2020; Rogers et al., 1977; Symons & Johnson, 1997). Experiment 1.1 took an individual differences approach by testing whether the ownership effect is associated with the number of autistic traits (measured by AQ; Baron-Cohen et al., 2001) possessed by neurotypical adults; Experiment 1.2 tested whether the ownership effect differs between neurotypical and autistic adults. We improved on previous studies that have examined this question (e.g., Grisdale et al., 2014; Gillespie-Smith et al., 2018; Wuyun et al., 2020) by using well-powered sample sizes, pre-registered plans and more sensitive (Bayesian) statistical methods.

Arguably, the key finding from the current experiments was that autistic participants in Experiment 1.2 showed a significantly reduced ownership effect compared to age- and IQ-matched neurotypical comparison participants. A significant group x referent interaction revealed that whereas neurotypical participants showed a moderate-to-large and statistically significant ownership effect in Experiment 1.2, autistic participants showed a small and statistically non-significant effect (i.e. comparable memory for

items encoded in relation to the self and in relation to the experimenter). Moreover, whilst there was no group difference in memory for other-owned items, the neurotypical participants had higher memory for self-owned items than autistic participants. It is reassuring that these findings replicate closely those of Grisdale et al. (2012, Experiment 2) and add weight to the suggestion that this effect is reliably reduced among people with a clinical diagnosis of ASC. This is especially so, because the current study was a gold standard replication attempt, which was fully pre-registered, and included the same materials used by Grisdale et al. but with a sample that was almost 65% larger than the original investigation.

The ownership effect observed in the neurotypical group is in line with previous research that has used the ownership paradigm in neurotypical adults (Cunningham et al., 2008; van den Bos et al., 2010) and children (Cunningham et al., 2013, 2014; Saylor et al., 2011). Among neurotypical people, memory for self-referent information is superior to memory for information relating to other people. This self-bias, or "self-reference effect", is thought to be driven by preferential processing of self-relevant information relative to other-relevant information (Tversky & Kahneman, 1974). According to the extended-self hypothesis, one's self-concept integrates property items (e.g., car, house) that represent their personal identity into ones extended self-concept (Belk, 1988; Carrier & Dittmar, 1994). Subsequently through 'contamination' (Argo et al., 2008), an abstract trace of the self might transfer to the object. As such, the psychological connection to owned items causes them to benefit from the self-bias, making them more memorable than non-owned items (Beggan, 1992).

At the psychological level, ASC is thought to be characterised by difficulties in aspects of both self-awareness and memory, and it has been suggested that these two limitations are linked (Lind, 2010). If psychological self-awareness – and associated extended self-concept – is diminished among autistic people, then the self should not structure memory in the same way that it does in most neurotypical people. Consequently, we hypothesised that autistic people would show a reduced (or absent) ownership effect on traditional paradigms. The diminished ownership effect observed in the current study supports the proposal that psychological self-awareness is diminished in ASC, and is consistent with research that has shown atypical self-processing in ASC across other domains (e.g. Nijhof et al., 2018; Nijhof & Bird, 2019).

However, some caution is warranted before concluding that the reduced ownership effect observed among autistic participants in Experiment 1.2 reflects a reliable limitation in psychological self-awareness. While it is reassuring that this key finding replicated the original finding of Grisdale et al. (2012, Exp. 2), other results in the current study did not replicate the original results.

First, results from Experiment 1.1 challenged our pre-registered hypothesis. Although there was a clear ownership effect on memory, the size of this effect did not correlate with the number of autistic traits (measured using the AQ; this lack of correlation was replicated in our Experiment 1.2). Bayesian analysis suggested that the data supported the null hypothesis, that no reliable association between the variables existed in the current sample. This suggests that the ownership effect is not affected by autistic tendencies in a neurotypical population (echoing findings by Amodeo et al., 2021; Nijhof et al., 2020). This pattern contrasts with the original finding reported by Grisdale et al. (Exp. 1), who reported a moderate and statistically significant association between the size of the ownership effect and AQ total score. The sample of neurotypical participants in our Experiment 1.1 was over twice the size of Grisdale et al.'s Experiment 1.1 sample, thus low statistical power is unlikely to explain the non-significant relation in the current study. Instead, it may be that the findings of an association between the ownership effect and number of ASC traits among non-autistic participants reported by Grisdale et al. is not reliable.

Second, in Experiment 1.2 Chi-squared analysis revealed that there was no categorical difference between groups in the number of people who did/did not display an ownership effect, which suggests that although the size of the ownership effect was smaller in ASC (as shown by our continuous variable analysis), autistic participants were no less likely than neurotypical participants to display an ownership effect. This pattern also contradicts Grisdale et al. (2014; Exp. 2), who found that the ownership effect was more likely in neurotypical (100% *vs.* 71% in the current study) than autistic (44% *vs.* 65% in the current study) participants. It is notable, however, that the sample size in Grisdale et al.'s (2014) Experiment 1.2 (16 participants per diagnostic group) was smaller than the current Experiment 1.2 (28 per diagnostic group), and this difference might account for the discrepancy.

If our current finding of a *non*-significant association between the ownership effect and ASC traits is reliable, then this causes us to question how this can be reconciled with the findings of a significantly reduced ownership effect among autistic participants in both Grisdale et al. (Exp 2) and the current Exp. 2. One possibility is that qualitative differences exist between autistic traits in the general population and autistic traits in those diagnosed with ASC. This is not the prevailing view in the field currently, which is why Grisdale et al. and we opted to explore an individual differences approach in a first experiment. However, there is an emerging view that there may be differences in the underlying basis of autistic traits in diagnosed and undiagnosed individuals (Mottron, 2021), and there are findings that suggest distinct phenotypes can be observed (e.g., Chawner et al., 2021). It is therefore theoretically possible that no reliable association between autistic traits and the ownership effect exists in the general population, but that individuals with a clinical diagnosis of ASC nonetheless show a reliably diminished effect.

One possible consequence of impaired psychological self-awareness in ASC may be that it causes people to value self-owned items differently, and this contributes to the reduced ownership effect seen among autistic adults. Hartley and Fisher (2018) found that when children were randomly assigned one of three toys to own, neurotypical children showed a preference to retain their assigned toy (i.e. they rarely traded it for a different item, thus showing an ownership effect) but autistic children frequently traded for a different toy that they preferred. This suggests that unlike neurotypical children, autistic children do not over-value self-owned items and do not experience a mnemonic advantage for these self-owned items. However, when children *self-selected* a toy to own, neurotypical and autistic children rarely traded it, showing that both groups over-valued their new toy relative to a toy owned by someone else, and experienced an ownership effect (Hartley & Fisher, 2018; Experiment 2). Importantly, while neurotypical children continued to show a preference for their original self-selected toy compared to a replica, autistic children showed no preference. Therefore, when both items have identical material qualities, mere ownership of one item does not influence the preferences of autistic children as it does in neurotypical children (see also Hartley et al., 2020). Together, these findings suggest that autistic people elicit a more economically rational strategy that bases evaluation more on material qualities rather than ownership: they appear more concerned about *what* the object is, rather than *who* owns it (Hartley & Fisher, 2018).

Another potential explanation for the reduced ownership effect in ASC is that autistic people might be less aware of their own and other people's ownership rights (Hartley et al., 2021). Diagnosis-defining deficits in social cognition – including impaired theory of mind (ToM) and reduced social motivation (American Psychiatric Association, 2013; Chevallier et al., 2012) – might reduce the frequency and quality of social interactions through which people learn ownership norms (Hartley et al., 2021). Specifically, impairments in ToM are likely to increase the difficulty of mentally representing and tracking relationships between people and their property (Hartley & Fisher, 2018). These difficulties are further exacerbated by pragmatic language atypicalities in ASC, including the use of pronoun reversals (e.g., saying "yours" instead of "mine", and vice versa; Kanner, 1943; Tager-Flusberg & Joseph, 2005), since language scaffolds children's identification of ownership and enables them to communicate their relationship to property.

The current finding that self-referential cognition is impaired in autistic adults is a direct contrast to the patterns reported among autistic children, where either no group difference or an enhanced effect in the autistic group was found (Gillespie-Smith et al., 2018; Wuyun et al., 2020, Experiment 2). It is hard to reconcile these differences in terms of developmental differences (i.e. that psychological self-awareness is intact in childhood but becomes impaired in adulthood), so it is likely that the discrepant findings across studies reflect differences in task format and sample characteristics. For example, in Gillespie-Smith et al.'s (2018) modified ownership task, dyads included pairs of participants who took it in turns to sort cards into the appropriate basket, rather than a participant and experimenter (as used in our studies, Grisdale et al., 2014, and Wuyun et al., 2020). This means that different social contexts were established during the tasks (i.e. peer-to-peer *versus* more formal adult-to-child), and this may have influenced the success with which cognitive processes were deployed or managed (e.g., Barzy et al., 2020; Crompton et al., 2020; Hess, 2014). Furthermore, Gillespie-Smith et al. (2018) only found intact ownership in children with *low* social difficulties, and not those with *high* social difficulties, so it might be that this latter group is more comparable to the autistic samples used here and in Grisdale et al. (2014). Importantly, since there were only 11 participants in each of the low and high social difficulties subsamples, Gillespie-Smith et al.'s (2018) sub-group analysis is likely to be poorly powered to reliably detect effects. The results of Wuyun et al.'s (2020) Experiment 2 are harder to reconcile with the current study given that task instructions and set-up

appeared to be identical. The key difference between studies was that Wuyun and colleagues reduced the number of items to encode (6 items allocated each to the self and other, *versus* 74 here) to lower task demands in their child sample. Therefore, Wuyun et al.'s (2020) study might not have elicited the level of encoding that would reveal an impairment in ASC; self-bias might be intact in ASC when the cognitive demands of the task are lower (Bowler et al., 2015).

Conclusion

We have presented two pre-registered experiments that examined whether self-referential cognition is impaired in autistic adults. In Experiment 1.1, the ownership effect was replicated in a large neurotypical sample, but the size of this ownership effect was not associated with autistic traits. This suggests that individual differences in autistic traits do not impact the ownership effect, but does not rule out the possibility that individuals with clinical levels of autistic traits (i.e. an ASC diagnosis) produce outcomes that are qualitatively different – i.e. facilitated by alternative cognitive processes – compared to neurotypical individuals with high autistic traits. Experiment 1.2 used a case-control design and replicated Grisdale et al.'s (2014; Exp. 2) key findings, showing that the ownership effect was significantly smaller (but just as likely to occur) in autistic adults. This suggests that psychological self-awareness is impaired in ASC, which diminishes integration of items into one's extended-self relative to neurotypical people. Autistic people may therefore elicit a more economically rational strategy (Hartley & Fisher, 2018) that emphasises material qualities rather than ownership (i.e. *what* the object is rather than *who* owns it). The current study challenges previous studies that have found an intact or even enhanced ownership effect in autistic children (Gillespie-Smith et al., 2018; Wuyun et al., 2020). This discrepancy is hard to reconcile in terms of developing cognitive processes, and is likely to reflect methodological and demographic (e.g., social ability) differences between studies.

Chapter 3| Scene Construction Ability in Neurotypical and Autistic Adults

Scene Construction Ability in Neurotypical and Autistic Adults

Marchella Smith, Lindsey Cameron, & Heather J Ferguson

School of Psychology, Keynes College, University of Kent, Canterbury CT2 7NP, UK

Corresponding author

Marchella Smith (mmls2@kent.ac.uk)

Author contributions

MS designed the study, data coding, analysis and interpretation, and led on drafting and revising the manuscript; LC assisted with data interpretation, and revising the manuscript; HJF conceived of the study, assisted with study design, data collection, data analysis and interpretation, and drafting and revising the manuscript. All authors gave final approval for publication.

Acknowledgements
Data in this paper was collected as part of a grant to HJF, funded by the Leverhulme Trust (Ref: RPG-2014-298). The datasets supporting this article are available on the Open Science Framework (https://osf.io/qdyvk/?view_only=9498ceaa69464656b230f51b0915d092).

LAY ABSTRACT

People with autism spectrum conditions (ASC) have difficulties imagining events, which might result from difficulty mentally generating and maintaining a coherent spatial scene. The current study compared this scene construction ability between autistic (N=55) and neurotypical (N=63) adults. Results showed that scene construction was diminished in autistic compared to neurotypical participants. Furthermore, participants with fewer autistic traits had better scene construction ability. However, analysis of the cognitive predictors suggested that the ability to understand our own and other people's mental states predicted scene construction ability more than an ASC diagnosis or autistic traits. Furthermore, ASC diagnosis did not influence the frequency of mentions of the self or of sensory experiences.

ABSTRACT

People with autism spectrum conditions (ASC) have difficulties mentally simulating events, perhaps due to a difficulty mentally generating and maintaining a coherent spatial scene– i.e., "scene construction". The current study compared scene construction ability between autistic adults (N=55) and matched neurotypical adults (N=63), using a task in which participants were asked to vividly imagine and describe fictitious scenes. Results showed that scene construction was diminished in autistic compared to neurotypical participants, and was negatively associated with autistic traits. However, when scene construction was entered into a regression model alongside cognitive predictors, individual differences in Theory of Mind and alexithymia predicted performance, and the influence of diagnosis and autistic traits was no longer significant. ASC diagnosis did not influence the frequency of self-reference or sensory experiences, which followed the same pattern in both groups: sight > sound > touch = smell > taste.

Keywords: autism; scene construction; self-awareness; self-projection; social cognition

Episodic simulation is the ability to mentally simulate or imagine events (Schacter et al., 2012). It is a key component of episodic memory (Maguire, 2001; Rugg et al., 2002; Svoboda et al., 2006), episodic future thinking (Addis et al., 2007; Atance & O'Neill, 2001; Szpunar et al., 2007), spatial navigation (Burgess et al., 2002), theory of mind (Frith & Frith, 2003) and 'mind-wandering' (Mason et al., 2007). The capacity for "self-projection" – the ability to shift one's current state of self into the past (episodic remembering) or future (episodic future thinking) to mentally simulate an alternative perspective (i.e., mental time travel of the self, or autonoetic consciousness; Tulving, 2005) – is thought to be a crucial common cognitive process underlying all of these skills (Buckner & Carroll, 2007). Episodic simulation and self-projection capacity appear to emerge in parallel during early development (e.g., Quon & Atance, 2010; Suddendorf, 2010), and are disrupted in some developmental disorders. In this paper, we examine episodic simulation in adults with autism spectrum conditions (ASC)[10].

Research suggests that episodic simulation is impaired in ASC. For example, Lind and Bowler (2010) observed that when asked to recall particular time periods from one's own past (ranging from today to 10 years ago), and when asked to imagine events from corresponding time points in one's own future (ranging from today to 10 years later), autistic adults produced verbal descriptions that were significantly lower in quality (i.e. less specificity) than age- and IQ-matched neurotypical adults (medium effect size: r = .43) (also see Lind et al., 2014a). This suggests self-projection capacity is diminished in autism.

Furthermore, although past and future oriented thinking descriptions were positively correlated in neurotypical participants (r = .72), they were not correlated in autistic participants (r = -.25). Lind and Bowler (2010) suggested that this might reflect the fact that, compared to neurotypical individuals, autistic individuals draw less on episodic memory – phenomenological details such as feelings, emotions, and sensory details stored in memory from past scenarios (e.g., last weekend at the beach, I got sun burnt) – to simulate or pre-experience a possible future scenario (e.g., next weekend at the beach, it will be sunny; Atance & O'Neill, 2005; Suddendorf & Corballis, 1997; Wheeler et al., 1997). Rather, they rely more on

[10] We use a mixture of commonly endorsed terms to reflect the fact that "there is no single way of describing autism that is universally accepted and preferred" (Kenny et al., 2016).

semantic memory – i.e. general knowledge (e.g., memory for facts) – to imagine future events, and perhaps also to recall past events (e.g., Ciaramelli et al., 2018; Goddard et al., 2014; Terrett et al., 2013; McDonnell et al., 2017; Crane et al., 2012**;** Bon et al., 2013).

However, to imagine a *contextually rich* future scenario, we need to extract episodic and contextual details from past episodes, and flexibly recombine them into new configurations. Therefore, episodic future thinking is dependent on the integrity of the episodic memory system (Irish et al., 2012). Autistic people have impaired episodic memory (Bowler et al., 2011; Gaigg et al., 2015; Bowler et al., 2014; Lind & Bowler, 2008), and this might lead to impairments in aspects of self-awareness – specifically an awareness of one's personal continuity through time – i.e. temporally extended self-awareness (Lind & Bowler, 2010). If so, this might impact the use of self-experience to construct and mentally project oneself through time to identify with an anticipated future or past state of self (Lind, Bowler, et al., 2014; Lind & Bowler, 2010). Therefore, reliance on semantic knowledge – which does not require self-projection – might be used to compensate for self-projection difficulties (Ciaramelli et al., 2018; Terret et al., 2013), but could compromise the ability to simulate *contextually rich* future events (e.g., Lind & Bowler, 2010).

Scene Construction

Scene construction is the ability to mentally generate (i.e., imagine) and maintain a coherent spatial scene/event in mind; it is therefore critical for episodic simulation. It is typically operationalized by asking participants to imagine atemporal and impersonal fictitious scenes (e.g., "Imagine you're lying on a beach in a tropical bay"), and asking them to "describe what you can see, hear, smell and feel in as much detail as possible" (e.g. Hassabis, Kumaran, & Maguire, 2007). Using this approach, Lind et al. (2014b) found that autistic adults produced significantly lower quality descriptions for episodic memories and episodic future episodes which were both self-relevant (i.e., measuring self-projection, e.g., "Imagine a possible future meeting that you might have with a friend") and self-irrelevant (i.e., measuring scene construction, e.g., "Imagine you're lying on a sandy beach in a tropical bay"). Lind et al. (2014b) observed this impaired scene construction in autistic relative to neurotypical adults (N=56; large effect size, Cohen's *d* = 0.85), and this pattern has been replicated by Black et al. (2018; N=50)**.** Importantly, this impairment in autism was

comparable to that observed for self-projection and was independent of general narrative ability (Lind et al., 2014b). Since scene construction does not require self-projection, this shows that diminished episodic simulation ability in autism cannot be attributed solely to difficulties in self-projection, and that impairments in scene construction might be the critical factor.

These findings of impaired scene construction map onto Bowler et al.'s (2011) relational binding account, which proposes that relational binding capacity is reduced in autism, characterised by difficulties linking elements of experience to one another and/or to their spatial and/or temporal context. This proposal echoes the notion of "weak central coherence" in autism (Frith, 1989; Frith & Happé, 1994): a cognitive/perceptual processing style that is characterised by difficulty processing environmental stimuli as a coherent whole (global processing), instead focusing on each individual element (feature-based processing). The importance of relational binding for episodic simulation is supported by Gaigg et al. (2015) who observed that, in both neurotypical and autistic adults, contextually rich (episodic) recollection (i.e., *remembering*) of word triads increased with the number of category relations shared between words; this was not observed for familiarity-based (semantic) retrieval (i.e., *knowing*). This suggests that category relations support contextually rich recollection in both neurotypical and autistic participants. So, under real-world conditions, when individuals are required to form and store associations between items themselves, the less contextually rich recollection in autism – i.e., relying less on episodic *remembering* and more on semantic *knowing* (e.g., Lind & Bowler, 2010) – might reflect their reduced capacity to form associations between items at encoding. Consequently, the ability to flexibly retrieve multisensory information as an integrated whole and mentally simulate a coherent episode is diminished. In this paper, we use a scene construction task to compare episodic simulation capacity in a large sample of autistic and neurotypical adults (N=118), and examine whether differences exist in sensory experiences and self-referential cognition in these mentally simulations.

Sensory Experiences

According to Viberg (1983), in neurotypical people the frequency of sense modalities in spontaneous conversation follows a hierarchy of sight > sound > touch > taste > smell. However, sensory perception is

estimated to be atypical in 69-100% of autistic people (e.g., Baranek et al., 2006; Hilton et al., 2010; Leekam et al., 2007), characterised by hyper- and/or hypo- attention, perception, memory, etc. This atypicality is explained by the intense world theory (Markram & Markram, 2010) in terms of hyperactivation of brain circuits, which causes inefficient multisensory integration (Waterhouse et al., 1996; Brock et al., 2002; Foss-Feig et al., 2010). The 'predictive coding theory' of autism (van Boxtel & Lu, 2013; Van de Cruys et al., 2014) proposes that impaired meta-learning abilities in autism cause difficulty distinguishing between important and less important prediction errors. Consequently, autistic individuals struggle to contextualize incoming information and make predictions based on previous experience. Therefore, an extended temporal binding window in autistic *versus* neurotypical individuals might create a "fuzzier" sensory environment (Wallace & Stevenson, 2014) in which unrelated stimuli become integrated (McGurk & MacDonald, 1976; Shams et al., 2000; Stevenson et al., 2012; Thye et al., 2018).

If level of involvement of a sensory experience (such as sound, touch, smell, taste, sight) predicts the likelihood of it being encoded (e.g., hyper-perception of sound increases the probability of it being encoded), then a "fuzzier" sensory environment in autism might predict that descriptions of episodic simulation have atypical reliance on sensory modalities (Wallace & Stevenson, 2014). Indeed, Anger et al. (2019) observed that when asked to describe episodic memories and episodic future thinking in as much multimodal detail as possible, autistic adolescents (M_{age} = 13.5years) produced significantly lower quality descriptions – including fewer and lower intensity sensory details (colour, smell, sound, and tactile feeling) – than matched neurotypical adolescents during *free recall*. Nevertheless, autistic adolescents performed comparably to neurotypical adolescents when provided with visual cues that prompted for details such as: what, how when, where, who, emotions, Likert scales, perceptions [e.g., colours]; perspective [allocentric vs. egocentric]. This suggests that visual scaffolding can support autistic individuals to produce coherent and detailed narratives.

The benefits of task support suggests that the underlying cognitive mechanisms that allow processing and storing of relevant sensory information (bottom-up process) are preserved in autism, but the ability to use prior experience and context to *efficiently* understand and integrate/modulate such sensory/contextual information (top-down process) for social/narrative purposes is impaired. Therefore, the Theory of Mind

(ToM) impairments that are typically observed in autism (Baron-Cohen, 1989; Baron-Cohen et al., 1985; Castelli, 2002; Deschrijver et al., 2016; Tager-Flusberg, 2007) might impact the narration of episodic events due to difficulties knowing what social/sensory information to provide (e.g., Hamilton, 2013**;** Southgate & Hamilton, 2008; Wang & Hamilton, 2012). Autistic individuals, who have reduced social motivation (Chevallier et al., 2012) are more likely to provide task relevant concrete/factual information but diminished sensory/contextual details, since the latter is intended for social rather than functional purposes (e.g., Southgate & Hamilton, 2009; Wang & Hamilton, 2012). However, when task support provides cues to the desired information (e.g., sensory detail), autistic individuals might provide similar information to neurotypical individuals, suggesting typical memory encoding. In our task, we explicitly prompted participants to include sensory details in their scene construction descriptions, compared the relative mention of sensory experiences in the five sub-domains, and tested the degree to which any differences in episodic scene construction and sensory experience could be attributed to individual differences in cognition (including ToM).

Self-Referential Cognition

Among neurotypical individuals, a *self-bias* or "self-reference effect" is commonly observed, which is characterised by preferential processing of self-relevant information relative to other-relevant information (Tversky & Kahneman, 1974)**.** In ASC, self-bias has been found to be reduced or absent in some contexts (Burrows et al., 2017; Grisdale et al., 2014; Henderson et al., 2009; Lombardo et al., 2007; Smith et al., submitted), suggesting that self-referential cognition is impaired relative to neurotypical people. However, self-bias has been found to be intact in autistic adults and children in the domains of perception/attention (Williams et al., 2018; c.f. Navon & Makovski, n.d.) and memory (Lind et al., 2019), and seems to be unrelated to the number of autistic traits (Lind et al., 2019; Smith et al., submitted; Williams et al., 2018). Furthermore, research suggests that autistic people are *not* impaired at imagining the self in the episodic future, but are in the episodic past (Lind & Bowler, 2010). This contradictory evidence suggests that self-biases might operate differently across different cognitive domains, thus further research is needed to fully understand this area.

Self-referential cognition is likely to have important implications for scene construction. The neural-level model of spatial memory and imagery (Bicanski & Burgess, 2018; Burgess et al., 2001; Byrne et al., 2007) proposes that *egocentric* (i.e., body-centred; viewpoint-dependent) representations of the local sensory environment – which uses the self as a reference and corresponds to a specific point of view (i.e., head/gaze direction) – are transformed into *allocentric* (i.e., world-centred; viewpoint-independent) representations. Allocentric representations are dependent on processing the relations between landmarks independent of a single view-point. The reverse process allows reconstruction of egocentric representations from stored allocentric representations, which supports contextually rich recollection.

If we extract episodic and contextual details from past episodes to construct a *contextually rich* (i.e., episodic) fictitious scene, then in line with the neural-level model of spatial memory and imagery (Bicanski & Burgess, 2018; Burgess et al., 2001; Byrne et al., 2007), one must go beyond the *allocentric* representations stored in memory and use these to construct an *egocentric* representation of a fictitious scene which uses the self as a centre of reference. In this paper, we examine whether self-referential cognition is impaired in autism by testing participants' use of the egocentric perspective in scene construction – indexed by first-person pronoun use (e.g., I, me, my, mine, we, our).

Present Study and Hypotheses

The current study aimed to investigate scene construction ability, including analysis of sensory experiences and self-reference, in a larger sample of autistic adults and age- and IQ-matched neurotypical adults than previous studies have used. We also tested some of the cognitive predictors of scene construction ability, including ToM, autistic traits, and emotions. Since scene construction is required for episodic simulation – e.g., theory of mind (Frith & Frith, 2003) and 'mind-wandering' (Mason et al., 2007) – we predicted an association between these variables. Therefore, we included a measure of ToM, and also alexithymia, which is characterised by difficulty identifying and describing ones own emotions. Our hypotheses are detailed in the Supplementary Materials.

METHOD

Participants

The data were collected as part of a battery of tasks given to participants in the departments Autism database and included 63 (41 male; 22 female) neurotypical adults with no history of psychiatric or neurological conditions (self-report) and 55 (37 male; 18 female) adults with a clinical diagnosis of ASC, of which official diagnostic information was checked. All participants self-reported as native English speakers. Participant groups were matched on age, gender, and Intelligence Quotient (IQ, Wechsler Adult Intelligence Scale, WAIS-III or WAIS-IV; Wechsler, 1999a; Wechsler, 1999b), but were not matched on Autism spectrum Quotient (AQ; Baron-Cohen et al., 2001) (see Table 1). Current autistic characteristics were assessed for all autistic participants using module 4 of the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000; see Supplementaty Materials).

Table 1. Participant characteristics and statistical comparisons between diagnostic groups (autistic/neurotypical). Note: some tasks have missing data, so statistics are reported on maximal data for each task.

	Neurotypical	ASC	Neurotypical	ASC				
	N	N	Mean (SD)	Mean (SD)	t	df	*p*	*d*
Age (years)	63	55	31.52 (10.27)	30.50 (11.52)	0.51	116	.61	0.09
IQ: Full Scale	63	54	105.29 (10.41)	103.20 (15.01)	0.88	115	.38	0.16
IQ: Verbal	63	54	103.02 (10.75)	102.81 (12.20)	0.10	115	.93	0.02
IQ: Performance	63	54	106.56 (12.83)	103.37 (18.61)	1.09	115	.28	0.20
AQ	62	55	17.08 (6.2)	31.00 (8.80)	10.00	115	<.001***	1.85
ADOS	-	51	-	7.27 (4.67)	-	-	-	-
Animations	22	32	5.55 (1.90)	4.56 (2.17)	1.72	52	.046*	0.48
TAS-20	31	40	45.55 (9.41)	59.08 (12.53)	5.01	69	<.001***	1.19

*p <.05; ***p < .001.

Materials and Procedure

Scene Construction Task

The scene construction task, described by Hassabis, Kumaran, and Maguire (2007), measured the ability to mentally construct a visual scene. The experimenter asked the participant to close their eyes and vividly imagine three ordinary fictitious scenes (e.g., "Imagine you're lying on beach in a tropical bay"). They were then given two to three minutes to "Describe what you can see, hear, smell and feel in as much detail as possible". Importantly, participants were asked not to recount an actual memory of an experienced event, or something they planned to do, but rather, to create something new. As outlined in Hassabis, Kumaran, and Maguire (2007), a probing protocol was followed whereby *general* prompts were given if a description could not be provided or lacked detail (e.g., "Tell me more about this scene"). Participants were probed to move on if they became fixated on a particular aspect of a scene and were probed to elaborate further if they provided only poor detail. They were encouraged to continue with their descriptions until their account concluded or they were unable to elaborate further.

Participant Ratings: After each description, participants were asked to complete a questionnaire rating their imagined scenes for perceived salience/vividness (salience of the imager: 1, not vivid, to 5, very vivid), and for sense of presence/difficulty (how hard was the trial: 1, easy, to 5, hard).

Spatial Coherence Index: Participants were given a post-description questionnaire consisting of 12 statements, each providing a possible qualitative description of the newly constructed imagined scene/episode. Participants were asked to tick any of the statements that they thought accurately applied to their imagined scene/episode. Eight of the statements were 'integrated' and suggested that aspects of the constructed scene were contiguous (e.g., 'I could see the whole scene in my mind's eye'), and four of the statements were 'fragmented' and suggested that aspects of the constructed scene were not contiguous (e.g., 'It was a collection of separate images'). For each integrated statement that was selected, one point was awarded, and each fragmented statement

that was selected, one point was subtracted. This gave a score between -4 and +8, which was then normalised around zero to give a final spatial coherence index between -6 (completely fragmented) and +6 (completely integrated).

Content: The recordings of each scenario description were transcribed and blind coded by a trained experimenter. Within each scenario description, statements were coded as belonging to one of four main categories: spatial reference, entity presence, sensory description, or thought/emotion/action. Repeated statements, and other tangential information that could not be coded as belonging to one of the four categories were omitted. Hassabis et al. (2007) identified that seven details per category should be considered optimal performance. Therefore, each category has a maximum possible score of seven, which yielded a maximum possible content score of 28 for each scenario description.

Quality Judgement: Each scenario description was rated by the researcher for "how well they felt the description evoked a detailed 'picture' of the experience in their own mind's eye" on a scale of 0 (no picture at all) to 10 (vivid, extremely rich picture).

Experiential Index Score: As in Hassabis et al. (2007), all the description and questionnaire scores detailed above were combined to provide an imagination index (an "experiential index score"), ranging from 0 (not experienced at all) to 60 (extremely richly experienced), with higher scores indicating greater imagination ability. Like Hassabis et al. (2007), participant ratings of vividness and difficulty were both rescaled from 1-5 to 0-4, and quality ratings were rescaled from 0-10 to 0-18 to calculate the experiential index. 4% of the transcriptions were also coded by the second author to check the inter-rater reliability of the coding. Good inter-rater reliability was found, with intraclass correlation of .89 ($F(14, 14) = 11.42$, $p < .001$).

In addition to the main imagination criteria based on existing coding of the data, we defined two additional coding criteria to address our specific questions.

Self-Related Centre of Reference: Within each scenario description, every reference to a first-person pronoun (e.g., "I, me, my, mine, myself, we, us, our") was scored one point, and a sum total of self-references was calculated for each scenario, then averaged across the three scenarios for each participant.

Sensory Elements: Within each scenario description, every reference to a sensory element (sight/sound/taste/touch/smell, e.g., touch: "It felt warm") was scored one point. This was calculated separately for each of the five key sensory modalities (sight/touch/sound/taste/smell). To avoid ceiling effects on *sight* references, simply mentioning an entity (e.g., "I could see the sea") was not awarded any points. Rather, only statements describing the properties of an entity were scored one point (e.g., "the sea was blue and green"). Transcriptions were also coded by the second author to check the inter-rater reliability of the coding. Good inter-rater reliability was found, with intraclass correlation of .97 ($F(19, 19) = 35.49$, $p < .001$).

Measures of Theory of Mind, Autistic traits, and Emotions

Three established measures were used to assess Theory of Mind (Animations Task; Abell et al., 2000), Autistic traits (Autism-spectrum Quotient (AQ); Baron-Cohen et al., 2001), and Emotions (Toronto Alexithymia Scale (TAS-20); Bagby et al., 1994); see Supplementary Materials.

Bayesian Analysis

Bayesian analysis was also used to interpret the results (conducted using JASP 0.14.1; JASP Team, 2020), which enables a more graded interpretation than only using *p* values or effect sizes. This estimates the relative strength of the alternative hypothesis over the null hypothesis, or vice versa (e.g., Dienes, 2014; Rouder et al., 2009). For the Bayesian analyses, we adopted the default Cauchy priors as recommended by Wagenmakers *et al.* (2018). Bayes factor (BF^{10}) < 1 is evidence supporting the null hypothesis (<0.33 is firm evidence), and Bayes factors > 3, >10, > 30, and >100 are firm, strong, very strong, and decisive evidence supporting the alternative hypothesis respectively.

RESULTS

First, we ran analyses to check that the length of scene construction descriptions were comparable across the two diagnosis groups. Effects were non-significant ($t(116) = 1.59$, $p = .12$), meaning that any subsequent group differences on content was not likely to be due to differences in length of descriptions. Second, we identified and excluded statistically significant outliers – i.e. those that were more than three standard deviations away from the overall sample mean – on the experiential index, sensory experience, and/or self-reference measures. This procedure identified two participants from the ASC group with a frequency of self-reference that was substantially greater than 3 standard deviations from the sample mean (75.67 and 68.67 *versus* M=8.78 for the whole group) and these participants were excluded from further analysis on this measure.

1. a) How does autism diagnosis influence scene construction ability?

To investigate the effect of diagnosis on scene construction ability, a one-way between subjects (Diagnosis: ASC/NT) ANOVA was conducted on the experiential index score (see Figure 3.1). This revealed a significant main effect of diagnosis, $F(1, 116) = 4.21$, $p = .04$, $\eta_p^2 = 0.04$, $BF^{10} = 1.29$, whereby participants in the neurotypical group had a significantly higher experiential index score ($M = 39.00$) than participants in the ASC group ($M = 35.59$).

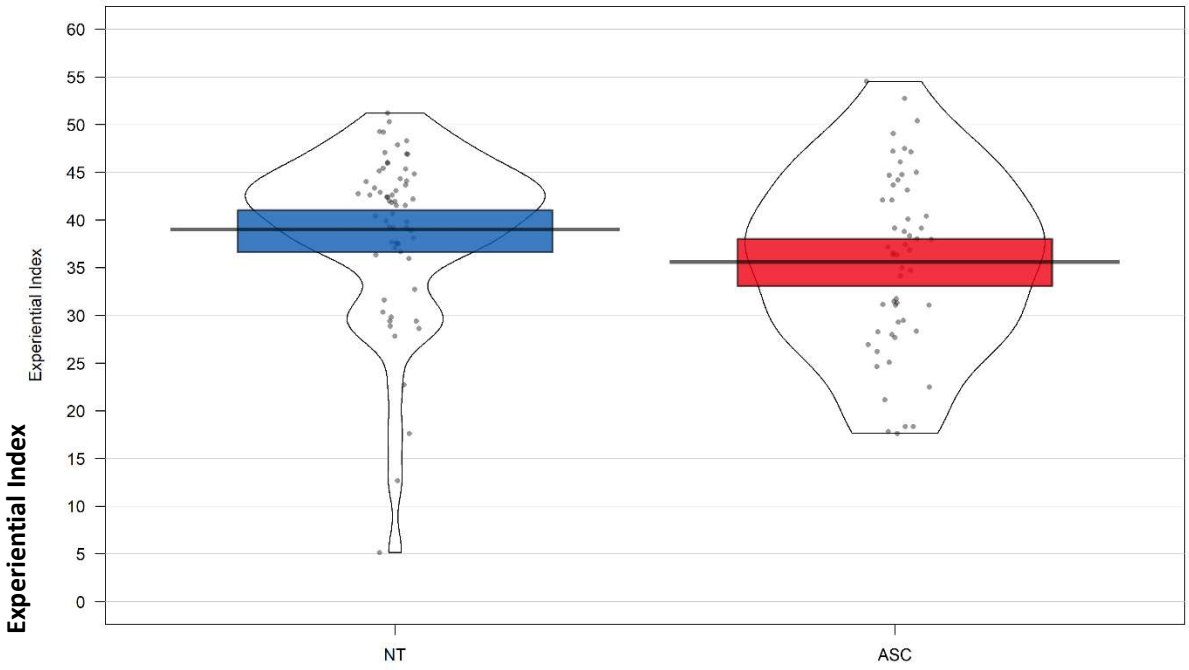


Figure 3.1. Mean experiential index score in each diagnostic group. Shaded area represents 95% confidence interval.

b) What are the cognitive predictors of scene construction ability?

We ran correlation analyses to test the relations between experiential index and (i) AQ, (ii) animations, (iii) TAS-20, and (iv) ADOS total (ASC group only) across the whole group of participants (see Table 2). In addition, since the categorical analysis observed a significant main effect of diagnosis on experiential index, we ran separate exploratory correlations for the autistic and neurotypical groups.

Table 2. Matrix displaying correlations between experiential index and cognitive predictors in the total sample, and in the autistic and neurotypical (NT) groups separately.

	Experiential index	AQ	Animations	ADOS
TOTAL SAMPLE				
AQ (N = 117)	-.233*	-		
Animations (N = 54)	.399**	-.134	-	-

TAS-20 (N = 71)	-.292*	.690**	.084	-
AUTISTIC GROUP				
AQ (N = 55)	-.041	-		
Animations (N = 32)	.373*	-.009	-	
TAS-20 (N = 40)	-.220	.589**	.173	-.064
ADOS (N = 51)	-.242	-.061	-.255	-
NEUROTYPICAL GROUP				
AQ (N = 62)	-.298*	-		
Animations (N = 22)	.353	.191	-	-
TAS-20 (N = 40)	-.312	.468**	.408	-

*p <.05; **p <.01; *** p <.001.

In the full sample analysis, experiential index was positively correlated with animations score, $r(54)$ =.40, p =.003, BF^{10} =13.00, and negatively correlated with AQ, $r(117)$ = -.23, p =.01, BF^{10} =1.68, and TAS-20, $r(71)$ = -.29, p = .01, BF^{10} = 2.97. Participants who scored higher on the experiential index scored higher on the animations task, and lower on the AQ and TAS-20 scales.

In the autistic group analysis, there was a significant positive correlation between experiential index and animations score, $r(32)$ =.37, p =.04, BF^{10} = 1.82; participants who scored higher on the experiential index also scored higher on the animations task.

In the neurotypical group analysis, there was a significant negative correlation between experiential index and AQ, $r(62)$ = -.30, p =.02, BF^{10} = 2.38; participants who scored higher on the experiential index scored lower on the AQ scale.

Next, a multiple regression analysis was conducted across the whole group of participants[11], to test whether experiential index score was predicted by individual performance on the animations task, or AQ and TAS-20 scales (these three variables were included because they were significantly correlated with experiential index). Additionally, the diagnostic group was included in the regression

[11] Note that some tests have missing data, so results are based on participants who have data for all variables (N = 49).

model. This regression analysis revealed that performance on the animations and TAS-20 were significant predictors of participants' experiential index, but AQ and diagnosis did not significantly predict the experiential index once these effects were accounted for (see Table 3).

Table 3. Multiple Regression Analysis Results Summary.

	SE B	*β*	95% CI *B*	*p*
TOTAL SAMPLE: Experiential index; R^2 Adj = .20, $F(4, 45) = 4.12$, $p = .006$**				
Animations	.657	.446	.786, 3.435	.002***
TAS-20	.145	-.474	-.634, -.048	.024*
AQ	.252	.172	-.337, .680	.501
Diagnosis	3.743	-.057	-8.736, 6.351	.751

1. ***a) How does autism diagnosis influence scene construction of sensory experiences?***

To investigate the effect of diagnosis and sensory modality on sensory descriptions, a 2 (Diagnosis: ASC/NT) × 5 (Sensory Modality: sound/taste/touch/smell/sight) mixed ANOVA was conducted on the frequency of sensory references, with repeated measures on the last factor (see Figure 3.2). This revealed a significant main effect of sensory modality, $F(4, 464) = 99.24$, $p < .001$, $\eta_p^2 = .46$, $BF^{10} = 1.00$. A series of paired samples t-tests revealed that the frequency of sensory references for each sensory modality followed a pattern of sight > sound > touch = smell > taste. Sight was referenced significantly more that sound, $t(117) = -6.94$, $p < .001$, $d = -.64$. Sound was referenced significantly more than touch, $t(117) = -4.51$, $p < .001$, $d = -.42$, and smell, $t(117) = 6.95$, $p < .001$, $d = -.64$. Reference to smell and touch did not significantly differ in frequency, $t(117) = -1.94$, $p = .054$, $d = -.18$, $BF^{10} = 0.63$. Taste was referenced significantly less than both smell, $t(117) = 12.56$, $p < .001$, $d = -1.16$, and touch, $t(117) = 12.56$, $p < .001$, $d = 1.16$. There was no main effect of diagnosis, $F(1,116) < .01$, $p = .99$, $\eta_p^2 < .001$, $BF^{10} = < .001$, and no diagnosis x sensory modality interaction, $F(1,116) = 1.2$, $p = .31$, $\eta_p^2 = .01$, $BF^{10} = < .001$.

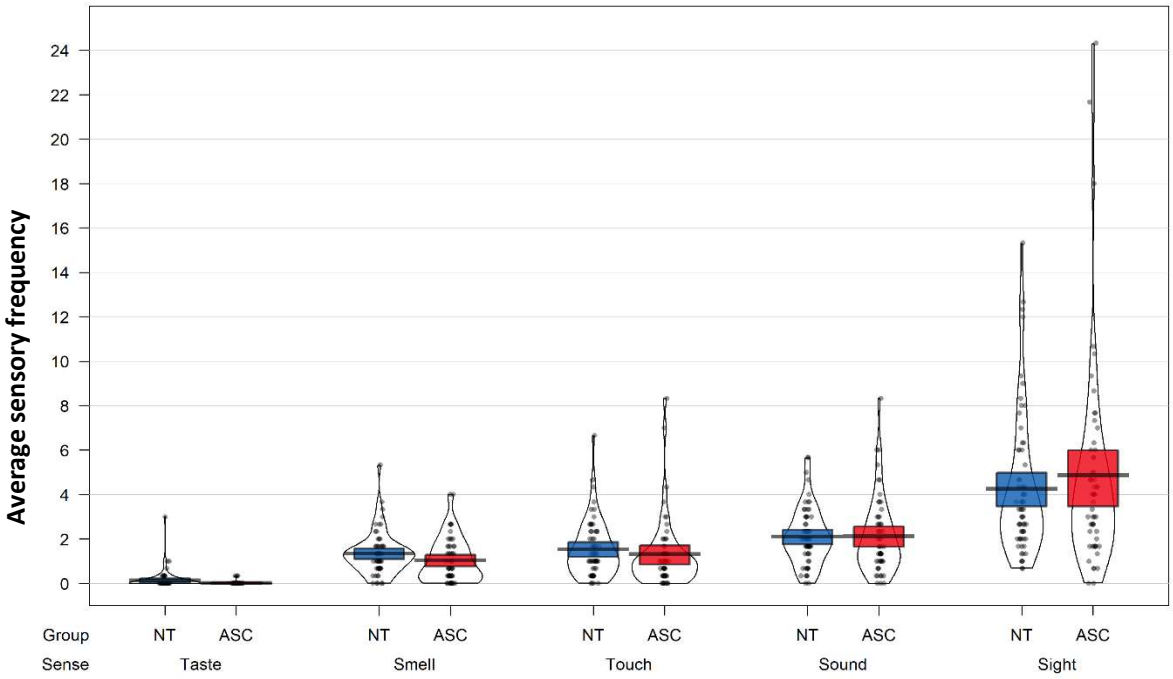


Figure 3.2. Mean frequency of sensory descriptions for each sensory modality and diagnostic group. Shaded area represents 95% confidence interval.

2. *b) What are the cognitive predictors of sensory scene construction?*

Correlation analyses investigating the relations between sensory experience (summed frequency across all five senses) and (i) experiential index, (ii) AQ, (iii) animations, and (iv) TAS-20, was conducted across the whole group of participants since group did not modulate effects in the categorical analysis (see Table 4).

Analysis revealed a significant positive correlation between frequency of sensory experience and experiential index, $r(118) = .63$, $p < .001$, $BF^{10} = 19.88$, and animations score, $r(54) = .39$, $p = .003$, $BF^{10} = 11.67$, whereby participants with a higher frequency of sensory references also scored higher on the experiential index and the animations task.

Table 4. Matrix displaying correlations between sensory experience and cognitive predictors in the total sample.

	Sensory experience	Experiential index	AQ	Animations
TOTAL SAMPLE				
Experiential index (N=118)	.629**			
AQ (N= 117)	-.026	-.233*		
Animations (N= 54)	.394**	.399**	-.134	
TAS-20 (N= 71)	-.062	-.292*	.690**	.084

*p <.05; **p <.01; ***p <.001.

3. a) How does autism diagnosis influence self-referential scene construction?

To investigate the effect of diagnosis on self-reference during scene construction, a one-way between subjects (Diagnosis: ASC/NT) ANOVA was conducted on frequency of first-person pronoun use (see Figure 3.3). There was no significant main effect of diagnosis, $F(1, 114) = 0.11$, $p = .74$, $\eta_p^2 = .001$, $BF^{10} = 0.21$, thus participant groups did not differ in their likelihood of self-reference[12].

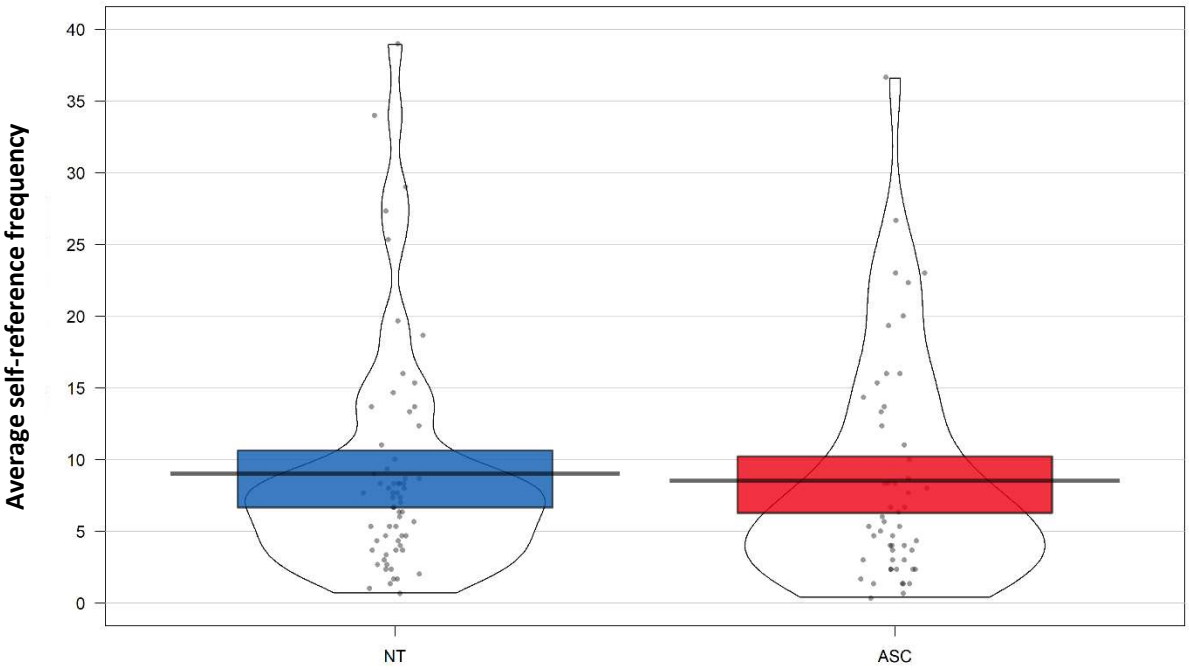


Figure 3.3. Mean Self-Reference Frequency in each diagnostic group. Shaded area represents 95% confidence interval.

[12] The main effect of diagnosis was also non-significant if the two outlier participants were included in the analysis $F(1, 116) = 0.08$, $p = .77$, $\eta_p^2 = .001$, $BF^{10} = 0.28$).

3. b) What are the cognitive predictors of self-referential scene construction?

Since the previous categorical analysis observed no significant main effect of diagnosis, the subsequent correlation analyses tested associations between self-reference and (i) experiential index, (ii) sensory experience, (iii) AQ, (iv) animations, and (v) TAS-20, across the whole group of participants (see Table 5).

Self-reference was positively correlated with experiential index, $r(116) = .54$, $p < .001$, $BF^{10} = >100$, and sensory experience, $r(116) = .33$, $p < .001$, $BF^{10} = >100$, whereby participants who used more frequent first-person pronouns scored higher on the experiential index and described more sensory experiences.

Table 5. Matrix displaying correlations between self-reference and cognitive predictors in the total sample.

	Self-Reference	Experiential index	Sensory Experience	AQ	Animations
TOTAL SAMPLE					
Experiential index (N = 116)	.361**	-			
Sensory Experience (N = 116)	.329**	.629**	-		
AQ (N = 115)	-.017	-.233*	-.026	-	
Animations (N = 54)	.116	.399**	.394**	-.134	-
TAS-20 (N = 71)	-.058	-.292*	-.062	.690**	.084

DISCUSSION

The current study aimed to investigate scene construction ability, and the cognitive predictors of scene construction ability (including ToM, autistic traits, and emotions) in a larger sample of autistic and neurotypical adults than previous studies. Hassabis et al.'s (2007) scene construction task was

used, in which participants were asked to vividly imagine and describe fictitious scenes; descriptions were also coded for frequency of first-person pronouns and sensory experiences.

Scene Construction

Scene construction ability was significantly better – characterised by descriptions that were greater in quality (i.e., more specific, more episodic, more coherent) – in neurotypical than autistic participants. Furthermore, scene construction ability was negatively associated with autistic traits (AQ): participants with a higher experiential index had fewer autistic traits than participants with a lower experiential index. This pattern replicates that reported in previous studies that have used smaller participant samples (Black et al., 2018; Lind et al., 2014b), thus increasing our confidence in the reliability of these findings. According to Hassabis et al. (2007), episodic simulation of atemporal, fictitious scenes which are not self-relevant does not require self-projection. This suggests that episodic simulation ability in autistic adults cannot be attributed solely to difficulties in self-projection, and that impairments in scene construction might be the critical factor (Lind et al., 2014b).

Impaired scene construction in autism might be caused by impaired hippocampally-mediated *relational binding* (Bowler et al., 2011) or "weak central coherence" (WCC; Frith, 1989; Frith & Happé, 1994), which causes diminished multisensory integration of phenomenological details (e.g., Waterhouse et al., 1996; Foss-Feig et al., 2010). Consequently, this might impair the ability to flexibly retrieve this multisensory information as an integrated whole required to mentally construct a coherent scene in mind (Lind et al., 2014b; Mullally & Maguire, 2013). Like previous studies that have used the scene construction task (Lind et al., 2014b), the present study found that the spatial coherence sub-score of the experiential index was higher among neurotypical (M = 2.36; SD = 1.29) than autistic (M = 1.80; SD = 1.75) participants (p = .049), suggesting that mental representations were less coherent, and more fragmented in the autistic group. This supports the notion of diminished relational binding in ASC.

As predicted, scene construction ability was significantly associated with ToM (animations task) and alexithymia (TAS-20 scale) in the overall sample of participants: participants who had greater scene construction ability (a higher experiential index), had higher ToM ability (higher animations score) and lower alexithymia levels (lower TAS-20 score). These findings are in line with the notion that scene construction commonly underlies a series of cognitive functions – including the ability to understand one's own and others' emotions (i.e., good ToM and low alexithymia) – that involve episodic simulation (Schacter et al., 2012)[13]. For example, neuroimaging evidence has suggested that brain regions typically associated with ToM are active during scene construction, although significantly less than during self-related episodic future thinking and episodic remembering (see Hassabis, Kumaran & Maguire, 2007).

However, a further multiple regression analysis investigating the cognitive predictors of scene construction ability revealed that when diagnosis and autistic traits (AQ) were included in a multiple regression model along with ToM (animations score) and alexithymia (TAS-20 score) measures, diagnosis and autistic traits no longer predicted scene construction ability, but ToM and alexithymia both remained significant predictors. ToM is impaired in autism (Baron-Cohen, 1989; Baron-Cohen et al., 1985; Castelli, 2002; Deschrijver et al., 2016; Tager-Flusberg, 2007), and clinically significant levels of alexithymia affect approximately 50% of autistic people (Berthoz & Hill, 2005), compared to only 13% of neurotypical people (Salminen et al., 1999). This suggests that the less coherent, and more fragmented scene construction descriptions given by the autistic group might not be associated with ASC specifically – i.e. impaired relational binding – but might be related to diagnosis-defining social difficulties including alexithymia and ToM impairment (or the underlying cognitive mechanisms associated with ToM and alexithymia), which are both more prevalent in ASC than in the neurotypical population. Thus, the ability to understand others' mental states and one's

[13] Importantly, the direction of causation – i.e., whether impaired ToM and alexithymia cause impaired scene construction, *or* vice versa – cannot be determined from a correlation; therefore, this simply highlights the close link between the variables.

own emotions might be better predictors of overall scene construction ability than ASC diagnosis because they can compensate for the key impairments associated with ASC – e.g., poor understanding of social/narrative conventions, and the need to include sensory experiences and self-referential cognition.

Sensory Experiences

To investigate the relation between scene construction and perceptual processing style, the current study investigated the frequency of sensory references in each of the five key sensory modalities (sound, smell, sight, touch, taste). Across participant groups, the frequency of sensory references for each sensory modality followed a pattern of sight > sound > touch = smell > taste, showing that some sensory modalities were more frequently used to support scene construction than other sensory modalities (Viberg, 1983). Importantly, the frequency of sensory reference did not differ according to diagnosis and was not associated with autistic traits. This suggests that impaired scene construction is not influenced by atypical sensory experience.

However, it might be that the explicit prompt to include sensory details in descriptions provided sufficient task support to enable typical performance in the ASC group. Previous research found that autistic participants' episodic descriptions were similar to neurotypical participants *only* when provided with task support cueing what information to provide (i.e., pictures cueing for components: what, how when, where, who, emotions, Likert scales, perceptions [e.g., colours]; perspective [allocentric vs. egocentric] etc.) (Anger et al., 2019). Therefore, difficulty in ASC might be knowing what social/sensory information to provide (e.g., Hamilton, 2013**;** Southgate & Hamilton, 2008; Wang & Hamilton, 2012). Supporting this, in the current study, ToM (animations) was positively associated with sensory descriptions, whereby participants with higher ToM ability – and hence better understanding of social/narrative conventions – provided more descriptions of sensory experiences.

Furthermore, our results showed that the frequency of sensory reference was positively correlated with scene construction ability (experiential index), reflecting the overlap in what is captured by these measures. More interestingly, the frequency of sensory reference was positively correlated with ToM: participants with higher frequency of sensory reference were also better at inferring mental states (animations ToM task). This is in line with our suggestion that ToM might contribute to better understanding of social narrative conventions including sensory processing, which in turn supports scene construction ability.

Self-Referential Cognition

Frequency of self-reference did not differ according to diagnosis, and was not associated with autistic traits (AQ and ADOS). This challenges the notion that self-referential cognition is routinely absent/diminished in ASC, and instead appears in line with more recent evidence suggesting that self-referential cognition is intact in some tasks/cognitive domains (Lind et al., 2020; Williams et al., 2018). Intact frequency of self-reference suggests that autistic participants appropriately reconstructed egocentric representations from stored allocentric representations. Therefore, it might be that impairments are specific to episodic memory (Lind & Bowler, 2010).

Furthermore, this frequency of self-reference (egocentric perspective) was positively correlated with scene construction ability (experiential index) and sensory experience: participants with higher scene construction ability had a higher frequency of self-reference and sensory perceptions. This suggests that participants appropriately used episodic memory of past experiences to support construction of a contextually rich novel scene (Committeri et al., 2020). It is interesting to note that, contrary to our predictions, self-reference was not correlated with ToM (animations) ability, which suggests that self-referential processing does not consistently rely upon an understanding of narrative conventions.

Conclusion

Overall, our results replicate and extend previous findings (Black et al., 2018; Lind et al., 2014b) showing that scene construction ability is impaired in autistic people. This suggests that episodic simulation ability cannot be attributed solely to difficulties in self-projection in autistic adults, and that impairments in scene construction might be the critical factor (Lind et al., 2014b). However, the ability to infer others' mental states (i.e. ToM) and to recognise emotions in themselves (i.e. alexithymia) appear to be better predictors of scene construction ability than an autism diagnosis. ToM might allow autistic people to better understand social/narrative conventions – e.g., describing sensory experiences – required to construct a contextually rich scene, and hence compensate for scene construction impairments.

Chapter 4| Indistinguishable Behavioural and Neural Correlates of Perceptual Self-Other Distinction in Autistic and Neurotypical Adults

Indistinguishable Behavioural and Neural Correlates of Perceptual Self-Other Distinction in Autistic and Neurotypical Adults

Marchella Smith and Heather J Ferguson

Datasets supporting this paper can be found on the open science framework (https://osf.io/br98c/?view_only=601bdd15758e49acac86cc1ed5c42a74).

ABSTRACT

People typically show enhanced cognitive processing of self-relevant compared to other-relevant information – i.e., a self-bias. Previous research has suggested that self-bias may be impaired in autism spectrum conditions (ASC), perhaps due to difficulties with self-other distinction. However, empirical evidence for this is inconsistent, and the neural basis of processing differences remains unknown. We present two experiments that aimed to test perceptual self-bias and familiarity effects in ASC using a perceptual-association task. Participants were asked to distinguish face/label associations of the self from those of other people of differing levels of familiarity (i.e., friend *vs.* stranger). Experiment 3.1 took an individual differences approach by testing whether behavioural self-bias is associated with the number of autistic traits in a neurotypical adult sample (N = 59). Experiment 3.2 took a case-control approach by testing whether behavioural self-bias and associated ERP responses differ between neurotypical (N = 27) and autistic (N = 30) adults. Across both experiments, behavioural results showed that participants experienced both a self-bias (self > friend and stranger) and a familiarity effect (e.g., friend > stranger); neither effect was affected by the number of autistic traits or autism diagnosis. In Experiment 3.2, analysis of N1, N2, and P3 ERP components revealed a typical self-bias in both groups (self > friend and stranger), and only the autistic group evidence showed evidence of a familiarity effect (N2 friend > stranger). The findings are discussed in relation to self-other distinction ability, and the relevance of other neuropsychological and psychiatric conditions such as anxiety and alexithymia are also considered.

Self-biases in cognition have been observed to be impaired (or at least diminished) in adults and children with autism spectrum conditions (ASC) compared to neurotypical individuals in the domain of memory (Burrows et al., 2017; Grisdale et al., 2014; Henderson et al., 2009; Lombardo et al., 2007; Smith et al., submitted). However, in the perceptual domain, research suggests that self-bias effects may be intact in ASC. For example, Williams et al. (2018) used the perceptual shape association task (Sui et al., 2012) and found that autistic adults (N = 22), like neurotypical adults (N = 21), showed enhanced performance (i.e. faster RT and higher accuracy) for self-relevant matches compared to other-relevant matches. Importantly, there was no between-group difference in the speed and accuracy of responses, and unlike findings in the memory domain (e.g., Henderson et al., 2009), individual differences in the number of autistic traits were unrelated to the magnitude of self-bias effects in either this case-control sample, or in a larger sample of neurotypical participants (N = 124; Williams et al., 2018; see also Amodeo et al., 2021). Further evidence of intact self perceptual processing in ASC has been found using more complex (unfamiliar) face stimuli paired with self and other labels (Zhao et al., 2018). Together, these findings suggest that atypicalities in self-referential processing are not universal in autism.

The finding that self-relevant processing may be intact in ASC only under certain conditions might reflect Bowler et al.'s (2015) Task Support Hypothesis, which proposes that autistic individuals may be better able to compensate for their impaired self-referential cognition when task difficulty is reduced. Tasks that tap memory processes are likely to be more cognitively demanding than those that tap lower-level perception (Quesque & Brass, 2019). Another factor that has been shown to reduce the difficulty of other processing is to relate information to familiar (e.g., self or mother) rather than unfamiliar (e.g., stranger) others. In the perceptual shape association task described above (Sui et al., 2012), responses were faster and more accurate for shapes that have been associated with familiar than unfamiliar other labels. Indeed, some researchers have reported that familiar others are processed comparably to the self due to the high degree of overlap in their representations (Ketay et al., 2019). Importantly, Williams et al. (2008) replicated the familiarity

effect on perceptual performance in autistic adults, and no difference was found between autistic and neurotypical groups. In this paper, we consider that the purest index of self-other distinction involves comparing processing for self and familiar others since both they are both familiar to the self (unlike famous/unfamiliar others) and only differ on personal relevance. In contrast, the purest index of familiarity involves comparing processing for familiar and unfamiliar others (i.e., friend vs. stranger) since they are both distinct from the self and differ on familiarity.

To date, very little is known about the real-time neural correlates that underlie self- and familiarity-biases in autism, though such investigations will undoubtedly enhance our understanding of the underlying mechanisms. Finding that neural responses differ between autistic and neurotypical people, despite undiminished behavioural responding, would suggest that autistic people experience atypicalities in their early perceptual processing, but may recruit additional processes to compensate for this in their behavioural expression of self-other processing. Alternatively, finding that neural responses during self-referential tasks do not differ between autistic and neurotypical people would support the notion that self-bias and familiarity effects are undiminished in ASC from the perceptual level in the earliest moments of processing (e.g., Williams et al., 2018; Zhao et al., 2018).

In this paper, we address these questions by extending previous studies in the perceptual domain using a face-label task that distinguishes self- and other-relevant processing, and testing the extent to which effects of self-reference and familiarity are altered in autistic adults (Experiment 3.2) or neurotypical adults with a higher number of autistic traits (Experiment 3.1). In Experiment 3.2, we also complement the standard behavioural data collected in these tasks by recording temporally sensitive event-related potentials (ERPs) in response to face-name pairings to examine the real-time mechanisms that underlie self- and other-relevant processing in autism and neurotypical adults. This allowed us to test for real-time evidence of compensatory mechanisms that might enable autistic adults to mask difficulties with self-other distinction at a behavioural level.

Neural Underpinnings of Perceptual Self-Bias in ASC: Evidence from EEG

A growing body of research has examined self-other processing in neurotypical adults using ERPs (see Knyazev, 2013 for a review). This work typically manipulates the consistency of self-reference with auditory, visual or sensory experiences (e.g. own name/face pairings, Cygan et al., 2014; self/other touch, Deschrijver, Wieserma, & Brass, 2016; self/other perspective-taking, Ferguson, Brunsdon, & Bradford, 2019), and has converged on showing modulation of the P3 ERP component when processing self-relevant compared to other-relevant information, suggesting that this component indexes the distinction between self and other perspectives. There is mixed evidence on whether P3 is influenced by the familiarity of the other, with some studies showing no difference in P3 between self- and close-other names (both elicit a P3 that is larger than unfamiliar-other names; Tacikowski et al., 2014), and others showing attenuated P3 modulation for self-other conflicts that involve a less vs. more similar other (Ferguson et al., 2019). Very little research has examined the neural underpinnings of self-reference and familiarity in ASC, therefore it remains unclear whether neural indices of self and other processing differ in autism despite intact behavioural responses.

Two recent ERP studies have compared self- and other-referential processing between autistic and neurotypical people in the visual domain (Cygan et al., 2014; Nowicka et al., 2016). Among neurotypical participants, ERPs showed an enhanced P3 in response to participants' own face/name stimuli compared to both close- and unfamiliar-other relevant stimuli, as well as preferential processing (i.e. enhanced P3 amplitude) for close others compared to unfamiliar others. These patterns reflect the expected self-bias and familiarity-bias in neurotypical people. In contrast, autistic participants showed the typical P3 effect for self- relative to unfamiliar-other stimuli, but no difference in P3 amplitude between self- and close-other stimuli (i.e., both own and close-other stimuli elicited a higher P3 amplitude). These patterns suggest that at a neural level, autistic people experience an atypical self-reference effect on perception (i.e., attention is not preferentially allocated for the self- *versus* familiar-others), but an intact familiarity effect (i.e., familiar self and other stimuli are processed preferentially to unfamiliar-other stimuli).

One other study used an oddball task to compare neural responses to self-other processing between autistic and neurotypical people in the auditory domain (Nijhof et al., 2018). They found that while neurotypical adults elicited an enhanced late parietal positivity (similar to P3) to self-relevant stimuli (i.e. auditory presentation of their own name) relative to other-relevant stimuli (close- and unfamiliar names did not differ), autistic adults did not distinguish the self from either close- or unfamiliar-other stimuli on the late parietal positivity component. Both groups showed evidence of a familiarity effect on the early N1 component (i.e. higher amplitude for self and close-other *vs.* unfamiliar-other). These patterns are consistent with findings in the visual domain, showing that the neural responses to self-relevant information may be atypical in autism, but familiarity is intact (Cygan et al., 2014; Nowicka et al., 2016).

Despite showing global similarities in the biases that autistic adults experience in response to self-relevant and familiar stimuli, P3 responses to self-relevant stimuli were only enhanced relative to unfamiliar-others in the visual tasks (Cygan et al., 2014; Nowicka et al., 2016) and did not differ in Nijhof et al.'s (2018) auditory task. The discrepant findings might reflect differences in modality and task demands between studies. In Cygan et al.'s (2014) and Nowicka et al.'s (2016) visual tasks, the processing of self-relevant stimuli was task *relevant,* in that participants were required to press a key in response to seeing self-relevant stimuli. In contrast, the processing of self-relevant stimuli was task *irrelevant* in Nijhof et al.'s (2018) auditory task; participants were not required to respond to self-relevant stimuli, and attending to it would impede rather than support task performance. This suggests that autistic adults do not automatically distinguish familiar (e.g., self and familiar-other) from unfamiliar (unfamiliar-other) stimuli in the same way as neurotypical people. Providing explicit instructions to attend to specific stimuli, may prompt autistic participants to compensate for this impairment.

The possibility of different self-related attentional mechanisms in autism has been tested recently using the attentional blink paradigm. Nijhof et al. (2022) found that the attentional blink (i.e.

difficulty detecting the second of two target stimuli presented in close temporal succession) is reduced when the second target was the participant's own name compared to a close other (self-bias) or an unfamiliar other name (i.e., a familiarity effect), and did not differ in magnitude between ASC and NT adults. This suggests that the attentional self-bias is intact in ASC. This self-bias was reflected in higher N2 amplitude (alongside an increased N2 and P3 for the familiarity effect), which also did not differ between ASC and NT adults. Since previous studies that showed impaired ERP correlates of self-bias in ASC (Cygan et al., 2014; Nijhod et al., 2018; Nowicka et al., 2016) all did so at later stages of cognitive processing (late P3/late parietal positivity), it may be that early-stage self-referential cognition is intact in ASC (Nijhof et al., 2022; Nijhof & Bird, 2019; Williams et al., 2018), with impairments only appearing at later stages (Nijhof et al., 2022).

The current study

The current study used a perceptual association task (Williams et al., 2018; Zhao et al., 2018) to test whether self-biases and familiarity-biases on perception are altered in ASC by exploring whether binding of information to the self elicits a similar ERP response between autistic and neurotypical adults. Like Zhao et al. (2018), rather than using geometric shapes (e.g., Williams et al., 2018), the current study will use social stimuli – specifically, unfamiliar-faces that have been learned with a label for the self, familiar-other or unfamiliar other. Participants' own faces or faces of known familiar-others were not used to minimise the influence of bottom-up familiarity or personal salience. In Experiment 3.1, we tested whether behavioural indices of self-bias and familiarity effects are associated with the number of autistic traits in a sample of neurotypical adults. Experiment 3.2 took a case-control approach by testing whether behavioural and ERP indices of self-bias and familiarity effects differ between neurotypical and autistic adults.

In line with previous behavioural research using the perceptual association task, we predicted that the size of the self-bias and familiarity effects on behavioural measures would be unrelated to the number of autistic traits (Experiment 3.1) and would be undiminished in autistic adults compared to

matched neurotypical adults (Experiment 3.2). Furthermore, we predicted that ERP measures would show an enhanced P3 amplitude for self matches compared to both friend and stranger matches, which would also differ significantly (Self>friend>stranger). Based on the existing behavioural findings in the perceptual domain in ASC, we predicted that autistic and neurotypical groups would not differ in their neural responses, with both showing evidence of a self-bias and familiarity effect on ERP responses.

Experiment 3.1| Individual Differences Approach

METHOD

Participants

This study included a volunteer sample of 59 (14 male; 45 female) University of Kent undergraduate students, aged 18-27 (*M* = 19.22; *SD* = 1.44). An *a priori* power analysis using G*Power (Vinet & Zhedanov, 2010) revealed that 12 participants would be enough to detect a referent x trial type interaction with effect size 1.31 (Williams et al., 2017, Exp. 1) and power of .90. The larger sample size of *n*=59 allows for a more representative sample.

Participants had normal or corrected-to-normal vision and self-reported no history of psychiatric or neurological conditions. All participants provided written informed consent and following completion of the experiment, participants were fully debriefed and received partial course credits for their participation. Ethical approval was provided by University of Kent Research Ethics Committee.

Design

This was a 3 (Referent: Self/Friend/Stranger) x 2 (Trial type: match/mismatch) within-subjects design.

Materials and Procedure

Self-Referential Perceptual Matching Task: The task was created using Psychopy2 (Peirce et al., 2019). The *learning phase* (see Figure 4.1) had the purpose of teaching participants about the identity of three neutral faces. Participants were presented sequentially with three faces (taken from Ma et al., 2015) for 20 seconds each. Each face had direct gaze, and non-face features (e.g., hair; ears) were removed using Microsoft Paint/Paint 3D. Simultaneously, a label (either "You", "Friend", or "Stranger") was presented below each face, in Calibri bold white font. Participants were instructed to "remember the three face-label matches". Importantly, the identity of each face (e.g., either you, friend, or stranger), and the order that each face-label match was presented in the learning phase was randomized across participants to avoid order effects.

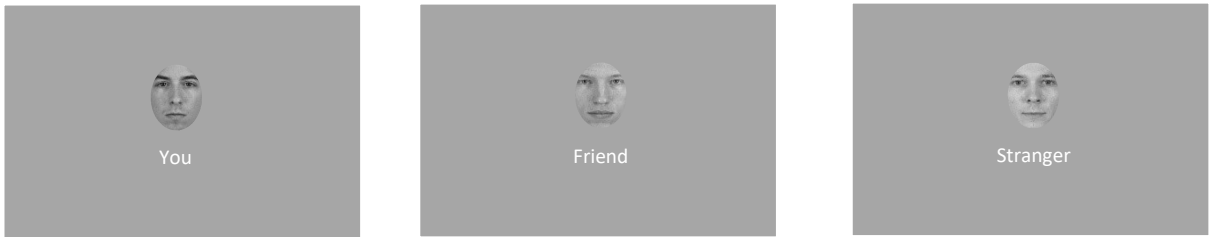


Figure 4.1. Illustration of learning phase slides. At the beginning of the experiment, each face-label match was presented for 20 seconds.

Test phase: on each trial (see Figure 4.2), there was a blank screen (500ms) followed by a centrally positioned white fixation cross (500ms), after which one of the three faces was presented (100ms) above *either* the label that was previously associated with the face – a *match* (self-associated face + "You"; friend-associated face + "Friend"; stranger-associated face + "Stranger") – or one of the labels that was associated with one of the other two faces – a *mismatch* (self-associated face + "Friend" or "Stranger"; friend-associated face + "You" or "Stranger"; stranger-associated face + "You" or "Friend"). Participants were asked to respond as quickly and accurately as possible when the face-label pair presented was a match (by pressing key "c") or a mismatch (by pressing key "m"). The response window started from the onset of the face-label pair presentation and continued for a further 1400-1800ms (randomised across trials). Performance feedback ("correct" or "incorrect")

was presented centrally on screen (400ms) after each of the 12 familiarisation trials at the start of each block. The main task included three blocks, each containing 120 experimental trials presented in a random order (i.e., 360 experimental trials in total), including 60 match (20 for each of the three possible correct combinations) and 60 mismatch (10 for each of the six possible incorrect combinations) trials. No feedback was given on performance. The dependent variables were participants' mean RT and accuracy (proportion of correct responses) on match and mismatch trials for each referent. The task took approximately 15 minutes to complete.

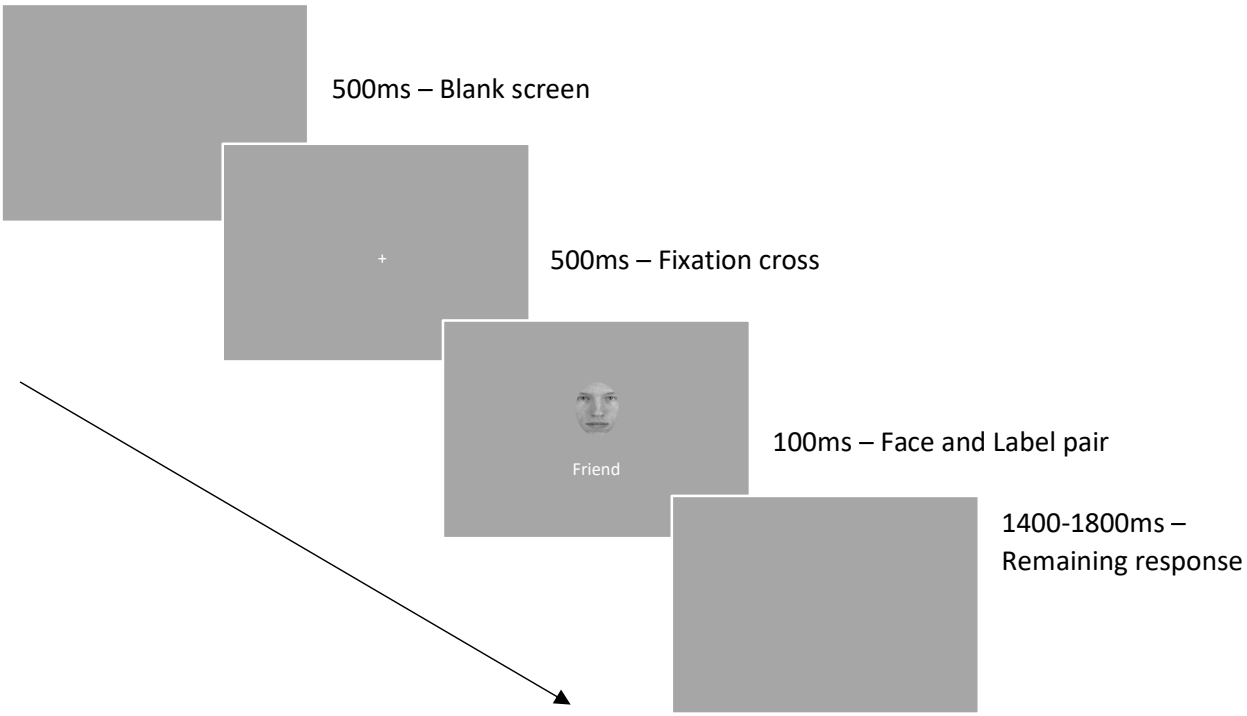


Figure 4.2. Illustration of an example trial in the test phase.

Measures of Personality and Mood

Two established measures were used to assess autistic traits (Autism-spectrum Quotient (AQ); Baron-Cohen et al., 2001), and camouflaging (Camouflaging Autistic Traits Questionnaire (CAT-Q); Hull et al., 2019); see Supplementary Materials.

Bayesian Analysis

Bayesian analysis was also used to interpret the results (conducted using JASP 0.14.1; JASP Team, 2020), which enables a more graded interpretation than only using *p* values or effect sizes. This estimates the relative strength of the alternative hypothesis over the null hypothesis, or vice versa (e.g., Dienes, 2014; Rouder et al., 2009). For the Bayesian analyses, we adopted the default Cauchy priors as recommended by Wagenmakers *et al.* (2018). Bayes factor (BF^{10}) < 1 is evidence supporting the null hypothesis (<0.33 is firm evidence), and Bayes factors > 3, >10, > 30, and >100 are firm, strong, very strong, and decisive evidence supporting the alternative hypothesis respectively.

RESULTS

Trials were excluded from RT and accuracy analysis if RT exceeded three standard deviations from the participant's overall mean or from the mean RT of the group overall (this excluded 13.98% of trials); trials were also excluded from the RT analysis if answered incorrectly. Two ANOVAs were conducted to analyse the effect of referent, and trial type on RT and accuracy using 3 (Referent: self/friend/stranger) x 2 (Trial type: match/mismatch) repeated measures ANOVAs. See Figure 4.3. Datafiles can be found on open science framework (https://osf.io/br98c/?view_only=601bdd15758e49acac86cc1ed5c42a74).

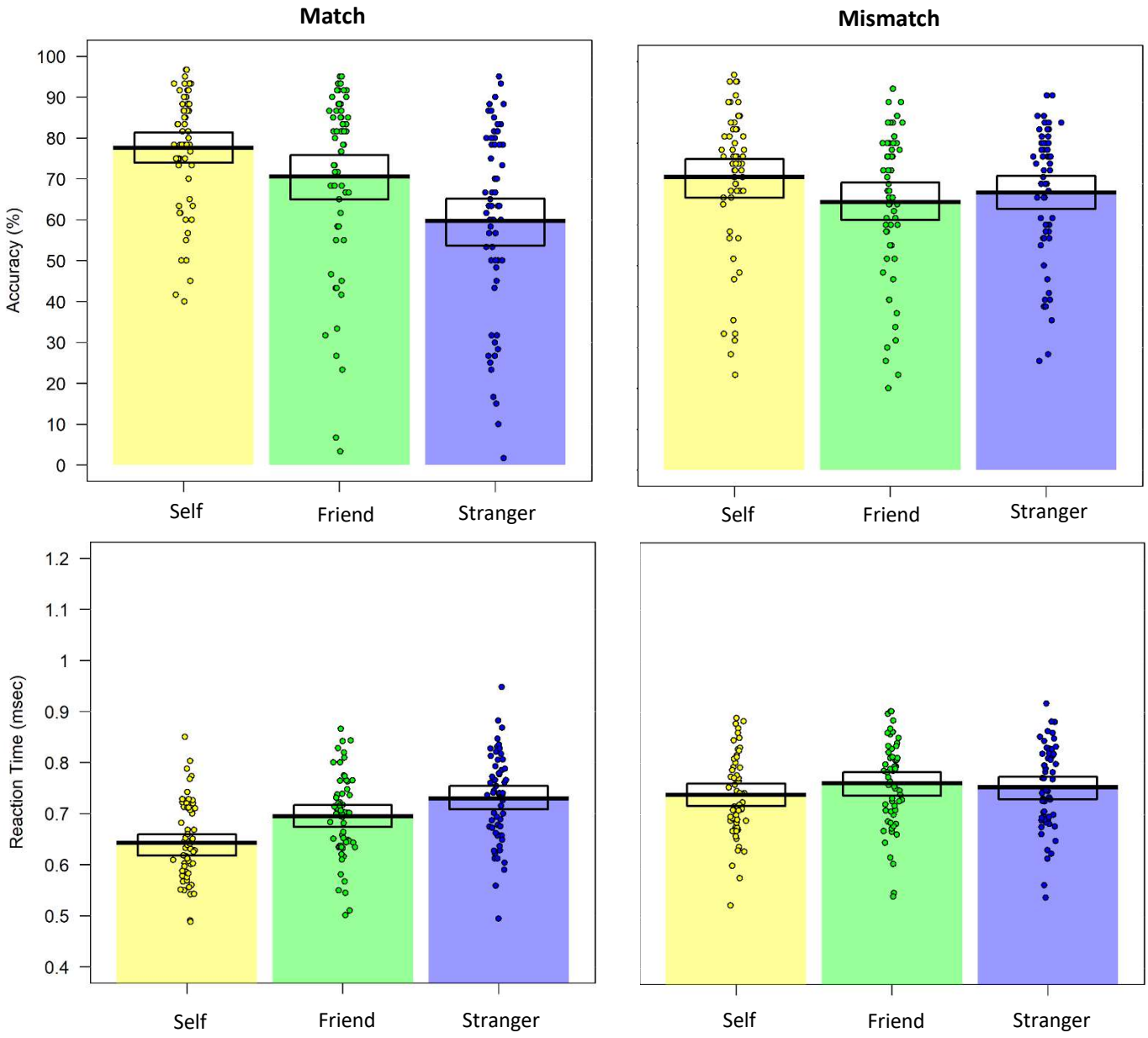


Figure 4.3. Accuracy and RT descriptive statistics for each referent in match (left panel) and mismatch (right panel) trials in Experiment 3.1. Boxed area represents 95% confidence interval.

Accuracy: There was a significant main effect of referent, $F(2, 116) = 35.21$, $p < .001$, $\eta_p^2 = .39$, $BF^{10} > 100$. Post-hoc paired samples t-tests revealed that accuracy followed a pattern of self>friend>stranger, whereby accuracy on self trials was higher than for friend, $t(58) = 4.54$, $p < .001$,

d = .59, and stranger trials, *t*(58) = 7.55, *p* <.001, *d* = .98; accuracy was also higher for friend than stranger trials, *t*(58) = 4.45, *p* <.001, *d* = .58. There was no significant main effect of trial type, *F*(1, 58) = 0.59, *p* = .45, η_p^2 = .01, BF^{10} = 0.15.

The trial type x referent interaction was significant, *F*(2, 116) = 22.45, *p* <.001, η_p^2 = .28, BF^{10} >100. Post-hoc paired samples t-tests revealed that accuracy followed a pattern of self>friend>stranger conflated over both match and mismatch trial types, with the difference between conditions as described above (all *t*s > 2.27, *p*s <.03).

RT: Analysis of RT revealed a significant main effect of trial type, *F*(1, 58) = 300.81, *p* <.001, η_p^2 = .84, BF^{10} >100, whereby RT was faster on match than mismatch trials. There was also a main effect of referent, *F*(2, 116) = 60.73.81, *p* <.001, η_p^2 = .51, BF^{10} >100. Post-hoc paired samples t-tests revealed that RT followed a pattern of self<friend<stranger, whereby RT for self trials was significantly faster than for friend, *t*(58) = 8.42, *p* <.001, *d* = 1.10, and stranger trials, *t*(58) = 10.88, *p* <.001, *d* = 1.42; RT was also significantly faster for friend than stranger trials, *t*(58) = 2.62, *p* =.01, *d* = .34.

There was a significant trial type x referent interaction, *F*(2, 116) = 30.15, *p* <.001, η_p^2 = .34, BF^{10} >100. Post-hoc paired samples t-tests revealed that in match trials, RT followed a pattern of self<friend<stranger, with the difference between conditions as described above (all *t*s > 4.00, *p*s <.001). In mismatch trials, RT followed a pattern of self<friend=stranger, whereby RT for self mismatches were significantly faster than for friend mismatches, *t*(58) = 4.46, *p* <.001, *d* = .58, and stranger mismatches, *t*(58) = 2.97, *p* =.004, *d* = .39. Friend mismatches did not significantly differ in RT from stranger mismatches, *t*(58) = 1.50, *p* =.14, *d* = .20.

Association Analysis

We calculated a self-bias score for RT and accuracy by subtracting the RT/accuracy for self-relevant matches from the average combined RT/accuracy for familiar-other and unfamiliar-other matches. A Pearson's correlation revealed that there was no significant correlation between the size of the RT

self-bias effect and AQ, $r(59)= -.15$, $p = .25$, $BF^{10} = 0.31$, or CAT-Q, $r(59)= -.23$, $p = .69$, $BF^{10} = 0.69$. and no significant correlation between the size of the accuracy self-bias effect and AQ, $r(59)= -.004$, $p = .97$, $BF^{10} = 0.16$, or CAT-Q, $r(59)= .05$, $p = .69$, $BF^{10} = 0.18$. See Figure 4.4.

We also ran exploratory (i.e. not pre-registered) correlations on the familiarity-effect for RT and accuracy, to inform understanding of whether familiarity-effect differs from similarity-effect. This was calculated by subtracting the average combined RT/accuracy for self and familiar-other matches from the unfamiliar-other matches. A Pearson's correlation revealed that there was no significant correlation between the size of the RT familiarity effect and AQ, $r(59)= .06$, $p = .63$, $BF^{10} = 0.18$, or CAT-Q, $r(59)= .01$, $p = .96$, $BF^{10} = 0.16$, and no significant correlation between the size of the accuracy familiarity effect and AQ, $r(59)= .03$, $p = .82$, $BF^{10} = 0.17$, or CAT-Q, $r(59)= .06$, $p = .66$, $BF^{10} = 0.18$. See Figure 4.4.

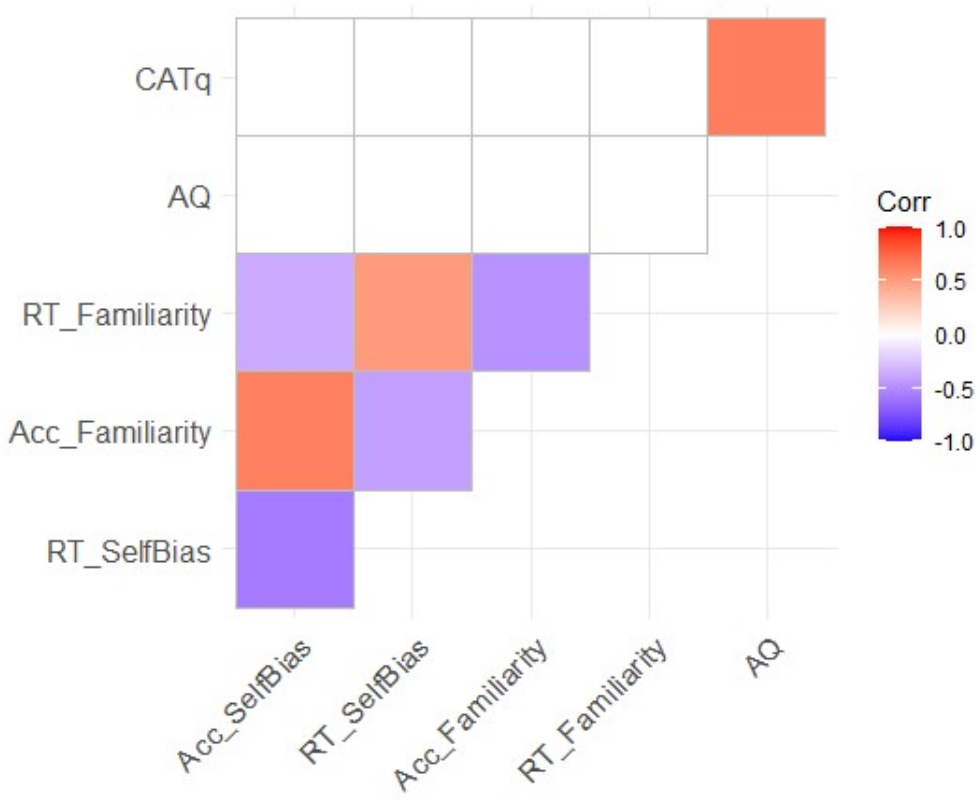


Figure 4.4. Correlation Matrix showing correlations between accuracy/RT self-bias/familiarity effect and AQ/CAT-Q.

Summary
In Experiment 3.1, we replicated both the self-bias effect and the familiarity effect on perception, using temporarily learned associations between (unfamiliar) faces and referent labels (self, friend, stranger). Participants were faster and more accurate when responding to a self-associated face compared to a friend-associated face, and even slower/less accurate when responding to a stranger-related face. Importantly, the self-bias effect was not associated with individual differences in autistic traits or camouflaging. The results are therefore consistent with our predictions and those reported for more abstract shape-label pairings (Williams et al., 2019, Experiment 3.1).

Experiment 3.2| Case-Control Approach

METHOD

Participants

Experiment 3.2 was fully pre-registered on the Open Science Framework (https://osf.io/br98c/?view_only=601bdd15758e49acac86cc1ed5c42a74). Participants were recruited from the Autism Research at Kent participant database. This included 27 NT adults with no history of psychiatric or neurological conditions (self-report) and 30 adults with a clinical diagnosis of ASC: specifically, this was either autistic disorder (N = 6), autism spectrum disorder (N = 4), Asperger's syndrome (N = 10) or not known (N = 10). We note that of the total participants who took part in this experiment, five autistic and three neurotypical were excluded from the ERP analyses due to poor quality of EEG data (resulting in a trial loss of > 40%) or technical problems with EEG recording. Thus, the final ERP sample included 25 autistic participants and 24 neurotypical participants (see Table 3 in Supplementary Materials). All participants were included in the behavioural analyses to maximise power (see Table 1).

Participant groups were matched on age, gender, and Intelligence Quotient (IQ, Wechsler Adult Intelligence Scale, WAIS-III or WAIS-IV; Wechsler, 1997; Wechsler, 2008). To assess the current autistic characteristics, autistic participants were also assessed on module 4 of the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000) by a trained, research-reliable researcher, and videos were double coded to ensure reliability of scoring (inter-rater reliability was found to be excellent with intraclass correlation of .89).

Participants provided written informed consent and were fully debriefed following completion of the experiment. They received £10 per hour of their participation, plus additional travel expenses. Ethical approval was provided by School of Psychology Research Ethics Committee at University of Kent.

Table 1. Participant characteristics in Experiment 3.2. Note: some tasks have missing data, so statistics are reported on maximal data for each task.

	Autistic (N = 30; 19 male)				Neurotypical (N = 27; 16 male)						
	Mean	SD	Range	N	Mean	SD	Range	N	*t*	*p*	*d*
Age (years)	33.46	11.91	20-59	30	38.93	14.81	19-68	27	1.54	.13	0.41
IQ: Full Scale	108.5	13.87	73-137	30	110.54	9.51	93-133	26	0.63	.53	0.17
IQ: Verbal	105.2	10.99	86-126	30	109.42	10.97	92-129	26	1.44	.16	0.39
IQ: Performance	110.63	18.9	65-143	30	110.5	11.4	89-144	26	0.03	.98	0.01
ADOS: Total	8.11	5.69	1-21	19	-	-	-	-	-	-	-
AQ	30.62	10.13	10-59	29	13.68	5.81	3-28	25	7.38	<.001	2.01
TAS-20	58.21	18.69	5-84	29	43.18	10.89	28-65	22	3.36	.002	0.05
EQ	26.38	11.91	8-55	29	46.83	11.14	26-67	24	6.41	<.001	1.77
BAI	26.62	13.35	6-62	29	11.42	7.8	0-30	26	5.08	<.001	1.37
CAT-Q	117.93	27.65	55-163	29	82.81	17.44	38-109	26	5.56	<.001	1.50

Design

This was a 2 (Group: ASC/neurotypical) x 3 (Referent: Self/Friend/Stranger) x 2 (Trial type: match/mismatch) mixed design with repeated measures on the last two factors.

Materials and Procedure

The materials and procedure for the self-referential perceptual matching task was identical to that outlined in Experiment 3.1.

Measures of Personality and Mood

Participants completed the same questionnaire measures (AQ, CAT-Q) as outlined in Experiment 3.1. In addition, participants completed three established measures to assess Theory of Mind (Animations Task; Abell et al., 2000), anxiety (Beck Anxiety Inventory (BAI); Beck et al., 1988), and emotions (Toronto Alexithymia Scale (TAS-20); Bagby et al., 1994); see Supplementary Materials.

Bayesian Analysis

Like Experiment 3.1, Bayesian analysis was also used to interpret the behavioural results (conducted using JASP 0.14.1; JASP Team, 2020).

EEG Recording and analysis

A Brain Vision Quickmap amplifier system was used to record continuous EEG activity during the self-referential perceptual matching task. A 32-channel ActiCap was used to measure EEG over midline electrodes Fz, Cz, Pz, and Oz, over the left hemisphere from electrodes Fp1, F3, F7, FC1, FC5, T7, C3, CP1, CP5, TP9, P3, P7, PO9, O1, and from the homologue electrodes over the right hemisphere. EEG data was referenced online to electrode FCz, and grounded to electrode AFz. EEG and EOG recordings were sampled at a rate of 500 Hz. Electrode impedances were kept at <25 KΩ.

Brain Vision Analyzer 2 software was used to pre-process the data before analysis. Offline, noisy or faulty electrodes were interpolated from surrounding channels (a maximum of 3 channels) and EEG data was re-referenced to the mean of right and left mastoids. EEG and EOG activity was then band-pass filtered (0.1-30 Hz, notch filter at 50Hz), and data containing blinks was corrected using a semi-automatic ocular ICA correction approach (Brain Vision Analyzer 2.1). Next, EEG data was aligned to the onset of the visual stimuli (i.e. the face and label) and segmented into epochs from -200ms to 1500ms. Semi-automatic artifact detection was used to identify and discard segments with non-ocular artifacts (drifts, channel blockings, EEG activity exceeding ± 50μV). Together, these procedures resulted in an average of 58 accepted segments per condition/participant. The remaining data was aligned to a -200-0ms baseline period, then averaged per participant and condition.

Three ERP components were identified for analysis in our pre-registration: mean amplitude in the N1 (130-180ms) (Rousselet et al., 2008), N2 (250-320ms) (Lu et al., 2015) and P3 (250-600ms) (Comerchero & Polich, 1999; Polich, 2007; Polich & Kok, 1995) time-windows. N1 is a negative-going component that peaks between 150-200ms, and earlier over anterior than posterior regions. N1 is thought to reflect an early *involuntary* shift of attention to familiar stimuli (Holler et al., 2011; Tateuchi et al., 2012), particularly those in which people have a level of expertise/high degree of familiarity (Tanaka & Curran, 2001). N2 is a negative-going component that peaks between 200-350ms primarily over anterior regions. N2 is thought to reflect conflict monitoring and/or response inhibition mechanisms (Donkers & van Boxtel, 2004; Jolicœur et al., 2006; Kranczioch et al., 2007; Nijhof et al., 2022). The P3 is a positive-going component that peaks between 250-500ms over central parietal electrode sites, and has been linked with self-referential processing and self-other distinction (see Introduction). In contrast to typical oddball effects on the P3, self-reference effects typically elicit larger P3 amplitudes for self-*compatible* conditions compared to self-*incompatible* conditions (Deschrijver et al., 2017).

Statistical analyses were run separately for midline and lateral electrode sites. Midline analyses included electrodes Fz, Cz, Pz and Oz. Lateral analyses were conducted on quadrant means involving the left anterior electrodes F3, F7, FC1, FC5; right anterior electrodes F4, F8, FC2, FC6; left posterior electrodes CP1, CP5, P3, P7, O1; and right posterior electrodes CP2, CP6, P4, P8, O2 (representing an Anterior–posterior x Hemisphere design).

Community Involvement

There was no community involvement in the design or implementation of the reported study.

RESULTS

Behavioral Analyses

Trials were excluded from further analysis if the RT exceeded three standard deviations from the participant's overall mean or from the mean RT of the diagnosis group (this excluded 14.17% of trials in the neurotypical group, and 3.65% of trials in the autistic group); trials were also excluded from the RT analysis if answered incorrectly.[14] Two ANOVAs were conducted to analyse the effect of group, referent, and trial type on RT (ANOVA 1) and accuracy (ANOVA 2) using 2 (Group: autistic/neurotypical) × 3 (Referent: self/friend/stranger) x 2 (Trial type: match/mismatch) mixed ANOVAs with repeated measures on the last two factors. See Figure 4.5.

[14] We additionally pre-registered to exclude participants from analysis if their overall accuracy fell below 60%, based on the accuracy means reported in Williams et al.'s (2018, Experiment 1) shape-label version of this task. However, performance on our face-label version of the task was clearly more difficult, reflected by lower overall accuracy in both Experiments 3.1 and 3.2, therefore applying this exclusion criteria resulted in high data loss. We therefore report the data here without applying this accuracy cutoff criteria, but report the full pre-registered analysis In Supplementary Materials, showing that results were the same when this exclusion criteria was applied.

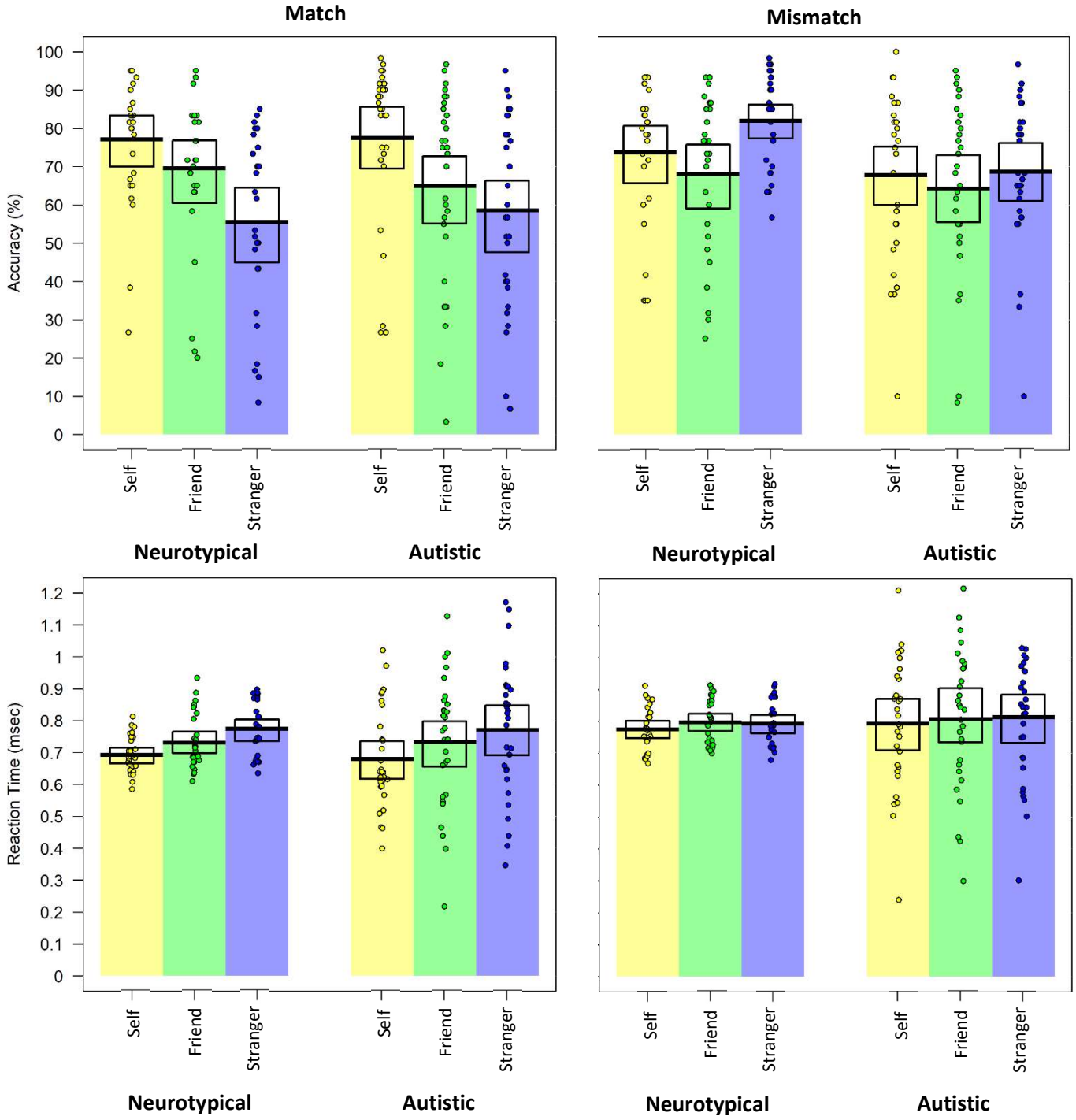


Figure 4.5. Accuracy and RT descriptive statistics in the neurotypical and ASC groups for each referent in match (left panel) and mismatch (right panel) trials in Experiment 3.2. Boxed area represents 95% confidence interval.

Accuracy: All assumptions for the current analysis were met. Analysis of accuracy revealed a significant main effect of trial type, $F(1, 55) = 6.06$, $p = .02$, $\eta_p^2 = .10$, $BF^{10} = 1.29$, whereby accuracy was higher for mismatches than matches. There was also a significant main effect of referent, $F(2, 110) = 18.38$, $p < .001$, $\eta_p^2 = .25$, $BF^{10} > 100$. Post-hoc paired samples t-tests revealed that accuracy followed a pattern of self>friend=stranger, whereby accuracy on self trials was higher than for friend, $t(56) = 4.59$, $p < .001$, $d = .61$, and stranger trials, $t(56) = 5.12$, $p < .001$, $d = .69$; however, accuracy did not differ between friend and stranger trials, $t(56) = 0.49$, $p = .63$, $d = .07$. There was no significant main effect of group, $F(1, 55) = 0.73$, $p = .40$, $\eta_p^2 = .01$, $BF^{10} = 0.43$, and no interaction between group x referent, $F(2, 110) = 0.35$, $p = .71$, $\eta_p^2 = 006$, $BF^{10} > 100$. However there were significant interactions between trial type x referent, $F(2, 110) = 36.11$, $p < .001$, $\eta_p^2 = .40$, $BF^{10} > 100$, trial type x group, $F(1, 55) = 6.32$, $p = .56$, $\eta_p^2 = .10$, $BF^{10} = 0.56$, both subsumed under a 3-way trial type x referent x group interaction, $F(2, 110) = 3.97$, $p = .02$, $\eta_p^2 = .07$, $BF^{10} > 100$.

Post-hoc tests investigated the 3-way effect by testing the group x referent interaction in each trial type separately. The group x referent interaction was not significant in match trials, $F(2, 110) = 1.12$, $p = .33$, $\eta_p^2 = .02$, $BF^{10} > 100$, but was significant in mismatch trials, $F(2, 110) = 5.68$, $p = .01$, $\eta_p^2 = .09$, $BF^{10} > 100$. Paired samples t-test revealed that in mismatch trials, accuracy in the neurotypical group followed a pattern of stranger>self>friend, whereby, accuracy was higher for stranger than self, $t(26) = 3.21$, $p = .003$, $d = .62$, and friend, $t(26) = 4.73$, $p < .001$, $d = .91$; accuracy was also higher for self than friend, $t(26) = 3.66$, $p = .001$, $d = .71$. In the autistic group, accuracy did not differ between self and friend, $t(29) = 1.90$, $p = .07$, $d = .35$, or self and stranger, $t(29) = 0.44$, $p = .66$, $d = .08$, but stranger was significantly higher than friend, $t(29) = 2.54$, $p = .02$, $d = .46$.

RT: Leven's test revealed that RT violated the homogeneity of variance assumption, so here we report significance tests with Greenhous-Geisser correction on dfs. Analysis of RT revealed a significant main effect of trial type, $F(1, 55) = 103.43$, $p < .001$, $\eta_p^2 = 0.65$, $BF^{10} > 100$, whereby matches were responded to more quickly than mismatches. There was also a significant main effect

of referent, $F(2, 110) = 31.29$, $p < .001$, $\eta_p^2 = 0.36$, $BF^{10} > 100$. Post-hoc paired samples t-tests revealed that RT followed a pattern of self<friend<stranger, whereby RT for self trials was significantly faster than for friend, $t(56) = 4.66$, $p < .001$, $d = .62$, and stranger trials, $t(56) = 7.87$, $p < .001$, $d = 1.04$; RT was also significantly faster for friend than stranger trials, $t(56) = 3.27$, $p = .002$, $d = .43$. There was no significant main effect of group, $F(1,55) = 0.03$, $p = .87$, $\eta_p^2 = .00$, $BF^{10} = 0.46$, and group did not interact with any of the other variables (Fs <1, ps >.12).

However, the referent x trial type interaction was significant, $F(2, 110) = 11.45$, $p < .001$, $\eta_p^2 = 0.17$, $BF^{10} > 100$. Post-hoc paired samples t-tests revealed that in match trials, RT followed a pattern of self<friend<stranger, with the difference between conditions as described above (all ts > 3.6, ps <.001). In mismatch trials, RT followed a pattern of self<friend=stranger, whereby, RT for self mismatches were significantly faster than for friend mismatches, $t(56) = 2.36$, $p = .02$, $d = .31$, and stranger mismatches, $t(56) = 3.48$, $p < .001$, $d = .46$. Friend mismatches did not significantly differ in RT from stranger mismatches, $t(56) = 0.17$, $p = .87$, $d = .02$.

ERP effects

Plots of the ERP waveforms in each group and condition are shown in Figure 4.6. Note that due to space constraints, only significant ERP effects that involve group or condition are presented in the text. Full statistical effects for each measure are summarised in the Supplementary Materials, and full data is available on the Open Science Framework (https://osf.io/br98c/).

Note that we report analyses for the pre-registered time windows for N1, N2 and P3, based on effects found in previous studies. However, visual inspection of grand average ERPs suggests that these time windows do not map exactly onto the time course of N1 and N2 components elicited by our face-label stimuli. Additional analyses are needed to verify effects on these components with adjusted time windows (e.g., N1 = 80-120ms and N2 = 200-260ms), however this was outside the scope of this thesis.

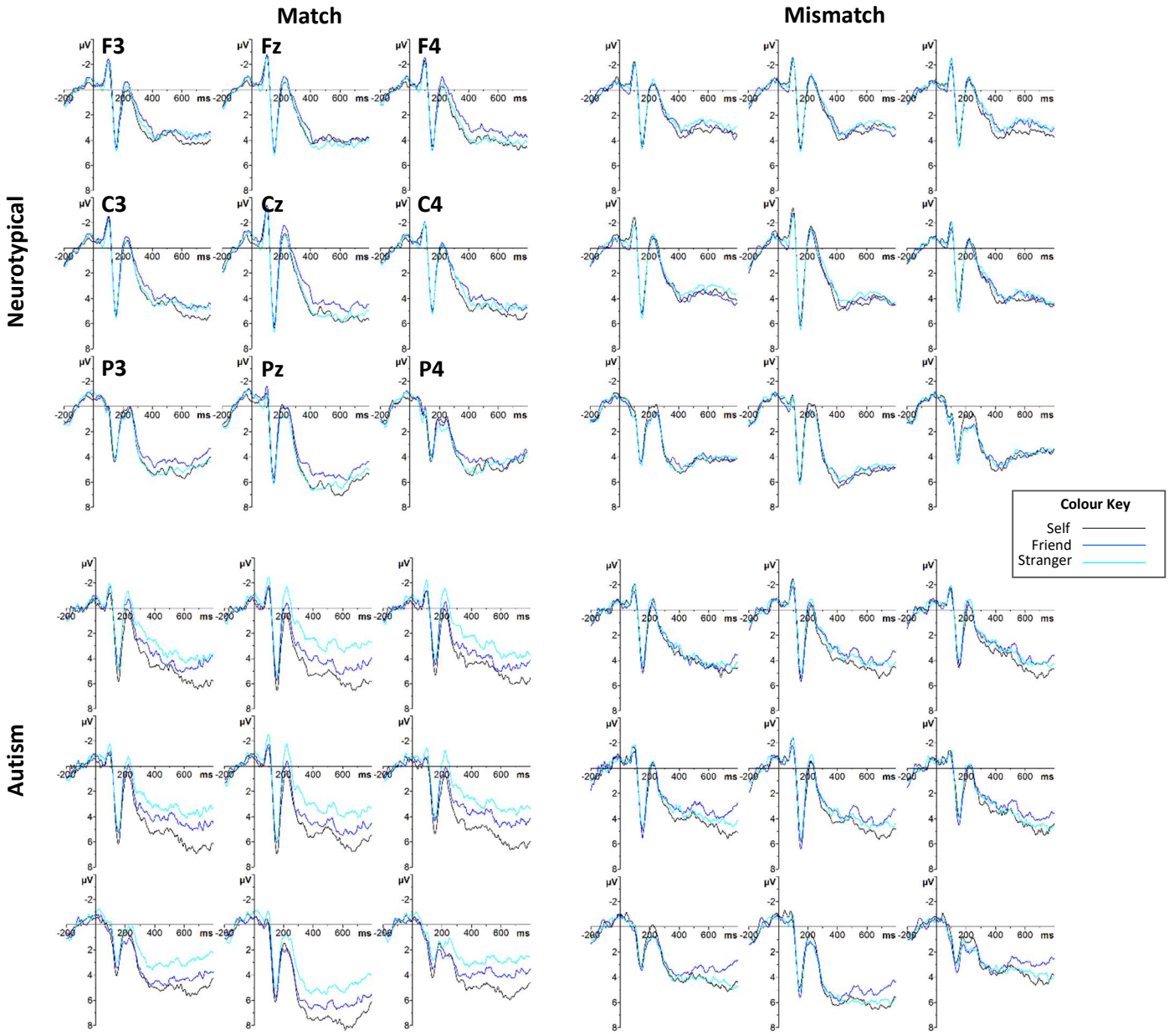


Figure 4.6. Grand average ERPs (over F3, Fz, F4, C3, Cz, C4, P3, Pz, P4) elicited by the face-label match (left panels) and mismatch (right panels) for self (black line), friend (dark blue line), and stranger (light blue line) in the neurotypical and ASC groups (see colour key).

Midline N1: Analysis of the N1 over midline electrodes revealed a trial type x referent x electrode interaction, $F(6, 282) = 4.03$, $p < .001$ $\eta_p^2 = .08$, however, follow-up analyses showed that the trial type x referent interaction was not significant in any electrodes ($Fs < 2.75$, $ps > .07$).

Lateral N1: Analysis of the N1 over lateral electrodes revealed a trial type x group x Ant-Pos interaction, F(1, 47) = 5.60, p = .02, η_p^2 = .11, however follow-up analyses showed that the trial type x group interaction was not significant over either anterior or posterior sites (*Fs* < 3.05, *p*s >.09).

There was also a trial type x referent x Ant-Pos interaction, F(2, 94) = 3.13, p = .05, η_p^2 = .06, which revealed a significant trial type x referent interaction over anterior, *F*(2, 94) = 3.65, *p* = .03, η_p^2 = .07, but not posterior, *F*(2, 94) = 2.03, *p* = .14, η_p^2 = .04, sites. In match trials, N1 anterior site amplitude was less negative-going for self (*M* = 4.06, *SD* = 3.87) than friend (*M* = 3.53*, SD* = 3.85), *t*(48) = 2.59, *p* = .01, *d* = .37, and stranger (*M* = 3.60, *SD* = 4.04), *t*(48) = 0.27, *p* = .04, *d* = .30, but there was no difference between friend and stranger, *t*(48) = 0.27, *p* = .79, *d* = .04. In mismatch trials, there was no difference in N1 anterior site amplitudes between referents (*t*s < 1.47, *ps* >.15).

Midline N2: The analysis of N2 over midline electrodes revealed a significant effect of trial type, *F*(1, 47) = 4.62, *p* = .04, η_p^2 = .09, with mismatch trials (*M* = 2.19, *SD* = 2.76) eliciting a more negative-going N2 amplitude than match trials (*M* = 2.53, *SD* = 3.08) . The significant interaction between trial type x electrode showed that this match effect was maximal over electrodes Fz and Cz.

There was also a significant referent x electrode interaction, *F*(6, 282) = 2,98, *p* = .01, η_p^2 = .06, showing that N2 amplitude was more negative-going for stranger than self, but only over Fz, *t*(48) = 3.18, *p* = .003, *d* = .45, and Cz, *t*(48) = 3.07, *p* = .003, *d* = .44, electrodes (all other electrodes *t*s < 1.83, *ps* >.07).

Lateral N2: The analysis of N2 over lateral electrodes revealed a significant effect of referent, *F*(2, 94) = 1.41, *p* = .046, η_p^2 = .06, reflecting a more negative-going N2 amplitude for stranger (*M* = 1.51, *SD* = 2.16) than self (*M* = 1.94, *SD* = 2.44), *t*(48) = 2.81, *p* = .01, *d* = .40, but no difference between self and friend (*M* = 1.67, *SD* = 2.45), *t*(48) = 1.54, *p* = .13, *d* = .22, or friend and stranger, *t*(48) = 0.95, *p* = .35, *d* = .142. The significant interaction between referent x Ant-Pos, *F*(2, 94) = 6.48, *p* = .002, η_p^2 = .12, showed that this self > stranger effect, and additionally, a more negative-going N2 amplitude

for friend (*M* = 1.62, *SD* = 2.97) than self (*M* = 2.08, *SD* = 2.80), *t*(48) = 2.39, *p* = .02, *d* = .34, was only present over anterior electrodes and not posterior electrodes.

There was a referent x group interaction, *F*(2, 94) = 3.52, *p* = .03, η_p^2 = .07. Post-hoc tests showed that in the NT group, there was no difference in N2 amplitude between referents (*ts* < 0.91, *ps* >.37). In the ASC group, stranger (*M* = 1.48, *SD* = 2.40) elicited a more negative-going N2 amplitude than friend (*M* = 2.04, *SD* = 2.19), *t*(24) = 3.44, *p* = .002, *d* = .69, and self (*M* = 2.31, *SD* = 2.89), *t*(24) = 4.00, *p* <.001, *d* = .82, but there was no difference between self and friend, *t*(24) = 1.37, *p* =.19, *d* = .27. Independent samples t-tests showed that there was no difference between groups in N2 amplitude for any of the referents (*ts* < 1.08, *ps* >.14).

Midline P3: Analysis of P3 over midline electrodes revealed a significant effect of trial type, *F*(1, 47) = 10.05, *p* = .003, η_p^2 = .18, with match trials eliciting larger amplitude (*M* = 4.00, *SD* = 2.92) than mismatch trials (*M* = 3.54, *SD* = 2.59). There was also a significant effect of referent, *F*(2, 94) = 3.54, *p* = .03, η_p^2 = .07, whereby self (*M* = 4.13, *SD* = 3.16) trials elicited larger amplitude than friend (*M* = 3.63, *SD* = 3.03), *t*(48) = 2.15, *p* = .04, *d* = .31, and stranger (*M* = 3.54, *SD* = 2.41), *t*(48) = 2.18, *p* = .03, *d* =.31, trials, but there was no difference in amplitude between friend and stranger trials, *t*(48) = .47, *p* = .64, *d* = .07. The significant interaction between referent x electrode showed that this self-bias effect on P3 was maximal over electrodes Cz and Pz.

There was also a trial type x referent x group interaction, F(2, 94) = 3.57, p = .03, η_p^2 = .07, which revealed that the referent x group interaction was only significant on match trials, *F*(2, 94) = 4.78, *p* = .01, η_p^2 = .09, and not on mismatch trials, *F*(2, 94) = .28, *p* = .75, η_p^2 = .006. Independent samples t-tests revealed that there was no between group difference in amplitude for any referent (*ts* < 1.89, s*p* >.07). Paired samples t-tests revealed that in the neurotypical group P3 amplitude did not differ between referents (ts < 1.88, ps >.07), but in the autistic group, amplitude was greater for self (*M* = 4.93, *SD* = 3.92) than for stranger (*M* = 2.83, *SD* = 2.82), *t*(24) = 2.43, *p* = .02, *d* = .49; friend did not differ from self, *t*(24) = 2.01, *p* = .06, *d* = .40, or stranger, *t*(24) = 1.73, *p* = .10, *d* = .35. See Figure 4.7.

Lateral P3: Analysis of P3 over lateral electrodes revealed a significant effect of trial type $F(1, 47) = 7.36$, $p = .01$, $\eta_p^2 = .14$, with match trials eliciting larger P3 amplitude ($M = 2.98$, $SD = 2.36$) than mismatch trials ($M = 2.62$, $SD = 1.99$). There was also a significant effect of referent, $F(2, 94) = 3.71$, $p = .03$, $\eta_p^2 = .07$, whereby self ($M = 3.09$, $SD = 2.37$) trials elicited larger P3 amplitude than friend ($M = 2.69$, $SD = 2.36$), $t(48) = 2.12$, $p = .04$, $d = .30$, and stranger ($M = 2.62$, $SD = 2.03$), $t(48) = 2.46$, $p = .02$, $d = .35$, trials, but there was no difference in amplitude between friend and stranger trials, $t(48) = .39$, $p = .70$, $d = .06$. See Figure 4.7.

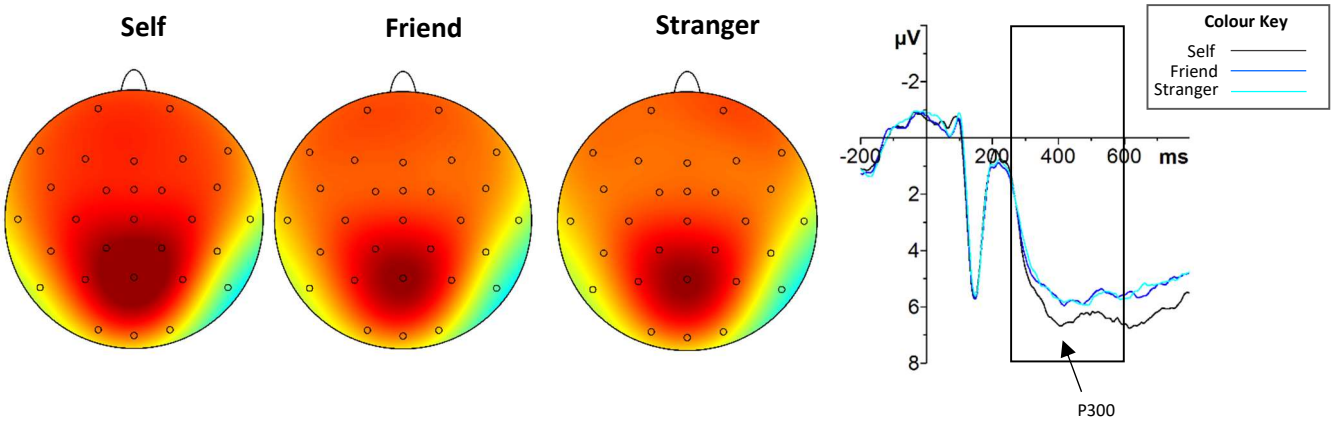


Figure 4.7. Topographic maps of the P3 component averaged over match and mismatch trials across all participants, separately for the self, friend, and stranger conditions. Grand average ERP waveform over Pz electrode, showing P3 amplitude for each referent (see colour key).

Association Analyses

The association between each of the self-bias and familiarity effect measures (i.e. RT, accuracy, and ERP amplitudes) and individual difference measures of autistic traits, ToM, personality and mood was tested using a series of Pearson's correlations, as in Experiment 3.1. ERP measures were calculated for match trials only using the mean amplitude of N1, N2 or P3 over midline electrodes. Resulting statistical details are shown in Table 2. None of the individual measures correlated with self-bias or familiarity effect on RT, accuracy, N2 or P3, and only TAS-20 showed a small positive correlation with the self-bias effect on the N1 ERP component.

Table 2. Matrix displaying correlations between each self-bias and familiarity effect measure and cognitive predictors in the total sample.

	Self-Bias			Familiarity Effect			Self-Bias			Familiarity Effect		
	N	RT	Accuracy	RT	Accuracy	N	N1 amplitude	N2 amplitude	P3 amplitude	N1 amplitude	N2 amplitude	P3 amplitude
AQ	54	0.12	-0.10	0.05	0.22	46	0.11	0.05	-0.07	-0.05	0.09	0.18
TAS-20	51	-0.08	0.03	0.04	0.11	44	0.33*	-0.06	-0.20	-0.18	0.18	0.28
CAT-Q	55	0.04	0.01	0.02	0.11	48	0.14	-0.04	-0.16	-0.07	0.21	0.28
EQ	53	-0.14	-0.09	-0.06	-0.05	46	-0.27	0.14	0.20	0.26	-0.14	-0.23
BAI	55	-0.07	-0.07	0.06	0.17	48	0.15	0.07	-0.05	-0.22	-0.09	-0.03
ADOS (ASC group only)	19	0.16	-0.28	0.02	0.15	16	0.13	-0.18	-0.08	-0.27	-0.02	-0.15

DISCUSSION

The current study aimed to further test whether perceptual self-bias and familiarity effects are undiminished in ASC (Williams et al., 2018; Zhao et al., 2018). Using a perceptual-association task (Sui et al., 2012), participants were asked to distinguish face/label associations relating to the self *vs.* friend (familiar-other) *vs.* stranger (unfamiliar-other). Experiment 3.1 took an individual differences approach by testing whether behavioural indices of self-bias and familiarity are associated with the number of autistic traits in neurotypical adults (N = 59). Experiment 3.2 took a case-control approach by testing whether behavioural and neural indices (measured using ERPs) of self-bias and familiarity differ between neurotypical (N = 27) and autistic (N = 30) adults.

Across both experiments, behavioural results showed that participants experienced both a self-bias and a familiarity effect. This was characterised by faster RT and higher accuracy for self-relevant stimuli compared to familiar-other (e.g., Friend) and unfamiliar-other (e.g., Stranger), and faster RT and higher accuracy for familiar-other stimuli compared to unfamiliar-other stimuli (this familiarity effect was only significant on RTs in Experiment 3.2). Importantly, the number of autistic traits was unrelated to the magnitude of self-bias and familiarity effects on RT or accuracy (in Experiments 3.1 and 3.2) and there was no between-group difference in the speed and accuracy of responses (Experiment 3.2).

The findings largely support previous research finding intact behavioural self-bias, and hence self-other distinction, in the perceptual domain in ASC (Zhao et al., 2018; Williams et al., 2018), but are inconsistent with evidence from other domains, particularly memory, where self-bias is often impaired (Henderson et al., 2009; Lombardo et al., 2007; Burrows et al., 2017; Grisdale et al., 2014; Smith et al., submitted; but see Lind et al., 2020). This domain-level difference in self-reference impairments might reflect reduced task difficulty in some tasks – e.g., perceptual tasks which only require a first-order representation of self, as opposed to trait memory tasks which require a second-order representation of self (Williams et al., 2018). This reduced difficulty might allow

autistic people to use alternative strategies, distinct from those used by neurotypical people to compensate for their impaired self-other distinction (e.g., Bowler et al., 2015; Nijhof & Bird, 2019). However, differences in the order of self representation are not sufficient to explain differences between tasks, since intact processing in ASC has been reported in some tasks that involve lower cognitive demands. Our studies therefore aimed to examine whether there is any evidence that autistic people compensate for general difficulties with self-other distinction by recruiting additional cognitive resources that enable them to mask difficulties at a behavioural level.

Neural Underpinnings of Perceptual Self-Bias in ASC: Evidence from EEG

In the N1 time window, in match trials only, over anterior lateral regions, there was a *self-bias* – whereby amplitude was more negative-going in the friend and stranger trials compared to the self. N1 is thought to reflect an early *involuntary* shift of attention to familiar stimuli (Höller et al., 2011; Tateuchi et al., 2012; Haider et al., 1964; Luck et al., 2000), particularly those in which people have a high degree of familiarity/expertise (Tanaka & Curran, 2001). Therefore, the findings are in line with previous research in neurotypical people showing a distinction in N1 amplitudes in response to seeing familiar faces (Caharel et al., 2002; Marzi & Viggiano, 2007), and hearing familiar sounds (Kirmse et al., 2009), and names (Nijhof et al., 2018), compared to when seeing or hearing unfamiliar stimuli.

Crucially, this self-bias in the N1 time window was observed in both the neurotypical and autistic group, and the magnitude did not differ between groups. The typical *self-bias* in ASC at this early stage supports previous research which found that, like neurotypical participants, autistic participants showed greater ERP amplitude in the N1 time window when *hearing* one's own name vs. the name of an unfamiliar-other (Nijhof et al., 2018); no research has found this typical early modulation of N1 reflecting a *self-bias* in ASC for *seeing* self vs. other relevant stimuli, or indeed in any other modality. Therefore, this provides strong evidence for spontaneous self-bias in ASC, even for temporarily learned associations between an unfamiliar face and self/other labels. Since these

self-bias effects are apparent in the earliest moments of visual processing (within 200ms of onset), this suggests that autistic adults are sensitive to these self-associations very quickly.

In the N2 time window, there was an overall *self-bias* over the midline – whereby amplitude differed between self and stranger trials (self did not differ from friend, which would have provided the strongest evidence of a self-bias). Furthermore, over lateral regions, in the autistic group *only,* there was some evidence of a self-bias (self > stranger) and a familiarity effect (friend > stranger). Since N2 is thought to reflect conflict monitoring and/or response inhibition mechanisms (Donkers & van Boxtel, 2004; Jolicœur et al., 2006; Kranczioch et al., 2007; Nijhof et al., 2022), we interpret this as reflecting the monitoring of conflicts between the self and others of differing familiarity in our task, and/or the cognitive control required to inhibit incorrect responses. The group differences might suggest that different types of conflict influence processing in each group. The observed self-bias is in line with Nijhof et al.'s (2022) research finding that the reduced attentional blink (i.e. difficulty detecting the second of two target stimuli presented in close temporal succession) when the second target was the participant's own name compared to an unfamiliar person's name, or a familiar person's name, was associated with increased N2 amplitudes in both ASC and NT participants.

In the later P3 time window, there was a *self-bias* – whereby P3 amplitude was greater in the self compared to the friend and stranger trials (which did not differ from each other). Previous research has shown modulation of the P3 ERP component in neurotypical people when processing self-relevant compared to other-relevant information, suggesting that this component indexes the distinction between self and other perspectives (e.g. own name/face pairings, Cygan et al., 2014; self/other touch, Deschrijver, Wieserma, & Brass, 2017; self/other perspective-taking, Ferguson, Brunsdon, & Bradford, 2018). This self-bias in the P3 ERP component was observed on match trials in both neurotypical and autistic participants, and did not differ in magnitude between groups.

Is Self-Bias intact in ASC?

The difference between processing of self and familiar-other (i.e. self-bias) is thought to be the purest index of self-other distinction because both the self and close-other are familiar to the self (unlike self vs. unfamiliar-others), therefore any difference in perception between the two is thought to result from personal relevance rather than familiarity. It is interesting to note that across both behavioural and neural measures, we found strong evidence to support the effect of a self-bias on processing in both groups, despite the use of complex face stimuli that did not genuinely match the self-other label in any condition. That is, while face-label associations were clearly learned in this task, they also presented a major processing conflict in self and familiar other conditions since the depicted face was not of the self or a friend (i.e. the label matched the learned face, but conflicted with a known face). The fact that self-bias and familiarity effects were detected here using these unfamiliar stimuli demonstrates the strength of the associations that participants were able to encode and respond to.

Previous research suggested that at the early N2 stage of processing, *self-bias* is intact in ASC (Nijhof et al., 2022), however, at the later P3 stage, there is a *familiarity effect*– whereby familiar-other stimuli (both self and familiar-others) are processed preferentially compared to unfamiliar-other (e.g., stranger) stimuli – but there is no *self-bia*s effect specifically – in that self-relevant stimuli are not processed preferentially than familiar-other stimuli (Cygan et al., 2014; Nowicka et al., 2016; Nijhof et al., 2022; Nijhof et al., 2018). These findings suggested that seemingly intact behavioural indices of perceptual self-bias at the behavioural level in the associations task (e.g., Williams et al., 2018; Zhao et al., 2018) might be supported by alternative underlying mechanisms.

In the current study (which also required only first-order self-representation) there was no between group difference in the magnitude of the *self*-bias effect at N1 or P3, and indeed, the influence of self-bias was evident within the earliest moments of processing (within 200ms of stimuli onset). The findings do not provide support for compensation in autism, and suggest that *self-bias* in the perceptual matching task is intact in ASC (Williams et al., 2017); although there is some evidence

that different types of conflict might influence processing (N2 lateral). It is unlikely that lack of between-group difference in behavioural and neural self-bias is caused by low sample size. The sample size in the current study (EEG: NT = 24, ASC = 25) was comparable to or exceeded that used by Nijhof et al. (2018) (NT = 21, ASC = 21), Cygan et al. (2014) (NT = 23, ASC = 23) and Nowicka et al.'s (2016) (NT = 15, ASC = 15), which suggests that we were sufficiently powered to detect a between group difference.

However, methodological differences between studies might account for some discrepancy in findings. Most importantly, previous tasks (Cygan et al., 2014; Nowicka et al., 2016; Nijhof et al., 2018) used familiar stimuli that were personally salient (e.g., self/friend faces/names). Therefore, the minimal evidence (only in the autistic group on the N2 component) of a familiarity effect on ERP responses contrasts with previous research, but is likely to be due to different processing that is activated by unfamiliar faces/names compared to familiar faces/names. In previous tasks, the self, familiar- and unfamiliar-other stimuli are likely to have had: 1) different levels of exposure (e.g., self face/name stimuli might have been exposed to the self more than familiar- and unfamiliar-other face/name stimuli), and consequently 2) different pre-established status (e.g., conditioning to respond to own name more than the name of our friend). In contrast, in the current study, participants' own faces and faces of familiar-others were not included to minimise any exposure/pre-established status (i.e., familiarity) from influencing the results. Rather, "tagging" new stimuli to the *concept* of self/friend/stranger tests the effectiveness of this initial tagging, and not the strength of pre-established (i.e. already tagged) associations. Therefore, it might be that the initial tagging of stimuli to the self is intact in ASC, but the strengthening of these associations over time is diminished, regardless of whether the task relies on first- or second-order representation of self. Furthermore, in the current study, participants were told which stimuli to tag to each referent, and hence which stimuli are most important; under real world conditions, they must make these judgements themselves. Therefore, difficulty in ASC might be deciphering which information to

tag/up-regulate consolidation of. As such, difficulty in ASC might be more apparent in tasks using familiar stimuli because this relies on associations that they have previously formed themselves.

Individual variance and comorbid factors that influence Self-Other distinction

Concluding *categorically* that self-awareness and hence self-other distinction is intact/typical or impaired/atypical in *any person* might be too simplistic since self-awareness is not necessarily fixed – rather, it can be a transient state. For example, in neurotypical people, self-other distinction – and hence presumably self-bias – is associated with/modified by a range of more transient neuropsychological and psychiatric conditions including depression, anxiety, stress, and post-traumatic stress disorder (Eddy, 2022; Lanius et al., 2020). Therefore, variations in self-bias effects between studies might be explained by individual differences in the presence/severity of co-morbid factors.

In the current study, we conducted exploratory correlation analyses to examine how individuals' self-reported alexithymia and anxiety levels impacted self-other distinction in the perceptual matching task. No significant correlation was observed between behavioural self-bias and anxiety (BAI) or alexithymia (TAS-20) levels. However, the self-bias effect on N1 was positively associated with alexithymia: participants reporting greater difficulty identifying and recognising their own emotions, had greater self-bias on N1. Alexithymia is associated with diminished self-other distinction (Saito et al., 2016), therefore greater early attention allocation for self vs. other in people with higher alexithymia levels might appear contradictory, but perhaps it reflects a heightened focus on the self in this group. This is the first study to associate ERP self-bias responses with alexithymia, therefore further research is required to determine the validity of these findings/conclusions.

Importantly, the association between self-bias (at the neural level) and alexithymia, highlights the importance of controlling for co-morbid factors. Since neuropsychological, and psychiatric conditions – including alexithymia – are more prevalent in ASC than in the general population (Salminen et al., 1999), previous research finding impaired self-bias in ASC might have been confounded by between

group differences in co-morbid conditions. If alexithymia (and anxiety) is a transient state (as opposed to autistic traits in someone with an autism diagnosis), then self-bias might be *temporarily* diminished; but support to reduce alexithymia (and anxiety) symptoms might enable typical self-other distinction. Since alexithymia (and anxiety) did not significantly differ between groups in the current study, it is unlikely that these co-morbid factors had a confounding effect, which increases our confidence in the validity of the findings.

Conclusion

Overall, the findings replicated previous research finding intact behavioural self-bias and familiarity effects in the perceptual association task in autistic adults (Williams et al., 2018; Zhao et al., 2018). Analysis of the ERP indices revealed the real-time processes that underlie these effects. Clear evidence of a self-bias was revealed in both groups, with self trials eliciting larger N1 and P3 components than familiar ('friend') trials. This pattern shows that autistic adults can encode novel associations between faces and self/other labels and are influenced by learned self-biases in the earliest moments of processing (i.e within 200ms). Moreover, the intact behavioural perceptual self-bias appears to be driven by the same underlying real-time mechanisms between groups. The familiarity effect was largely absent on ERP measures, though the ASC group elicited some modulation of N2 reflecting familiarity conflict. Methodological factors, such as familiarity of the stimuli, are likely to account for the small neural response to familiarity and discrepancies with previous studies. Future research should examine the influence of co-morbid factors which may influence self-bias and familiarity effects in autistic and neurotypical people.

Chapter 5| Autistic adults experience greater malleability of the self after simulating other people than neurotypical adults

Autistic adults experience greater malleability of the self after simulating other people than neurotypical adults

Marchella Smith and Heather J Ferguson

School of Psychology, Keynes College, University of Kent, Canterbury CT2 7NP, UK

Corresponding author

Professor Heather Ferguson (h.ferguson@kent.ac.uk)

Acknowledgements

The datasets supporting this article are available on the Open Science Framework (https://osf.io/n6muv/?view_only=67123f8b300f4286be9f6e5cd7107d9a).

ABSTRACT

Simulating the knowledge state of other people can change self-knowledge, causing one to incorporate information about the simulated person into the self (termed, simulation induced malleability). In this paper, we examine the nature of this simulation induced malleability in autistic adults, who are known to have atypical self-referential cognition. In addition, we test whether self-reflection might also change other-knowledge, causing one to incorporate information about the self into the other (termed, self-reflection induced malleability), and whether this differs between neurotypical and autistic adults. In a neurotypical adult sample, we found evidence of both simulation induced malleability (Experiment 4.1; N = 163) and self-reflection induced malleability (Experiment 4.2; N = 162). Furthermore, the size of simulation induced malleability was positively associated with the number of autistic traits, but only when the simulated other was considered as the "average person" and not a "similar-other". In two case-control experiments, there was greater simulation induced malleability in autistic than neurotypical adults (Experiment 4.3; N = 50) irrespective of similarity of the target, but no self-reflection induced malleability was observed in either diagnostic group (Experiment 4.4; N = 49). The findings are discussed in relation to self-awareness in autism.

Keywords: autism; simulation; self-reflection; memory malleability

Theory of mind (ToM) – also known as mindreading – refers to the ability to attribute mental states such as thoughts, feelings and intentions to other people. According to the "simulation" account, mindreading requires us to imagine our own mental states as though we are in a similar position to the target of mental state attribution (e.g., Goldman, 2006; Mitchell et al., 2006). This allows us to simulate (i.e. predict) the other's mental state based on direct first-person access to our own (Goldman, 2006). For example, when trying to understand how your friend is feeling about being bullied, you might recall *how you felt when you were bullied* previously (e.g., *I felt sad)*. This mental state is then taken "offline" so that it can be attributed to the subject of simulation (Goldman, 2010).

However, this strategy is only useful if the target of simulation is sufficiently similar to the self, and hence one's own experience is considered applicable to the other person. For example, it is reasonable to assume that our friend, who has similar characteristics to ourselves, would feel similarly about events. However, if the other person is dissimilar to the self, then we cannot be certain that they would respond in the same way as ourselves. Indeed, neuroimaging evidence typically finds a linear effect of familiarity/similarity in activation of regions associated with self-referential mental state processing. The medial prefrontal cortex is preferentially recruited for self-referential cognition but is also activated (to a lesser extent) for similar-other processing, and less so again for dissimilar-other processing (Kelley et al., 2002). This suggests that mentalising about similar-others relies on self-referential processing more than mentalising about dissimilar-others (Denny et al., 2012; Mobbs et al., 2009; Jenkins et al., 2008; Rilling et al., 2008).

Simulation Induced Malleability of Self-Representation

Recently, research has investigated whether simulating other people might have *consequences for the self* (Meyer et al., 2019; Rubin-McGregor et al., 2022). It might be that information considered about the other person changes the self-knowledge that is simultaneously retrieved. Thus, knowledge about the self might be malleable.

Once a memory has been consolidated, it is thought to be relatively stable, in that it does not modify in content. However, during retrieval of a consolidated memory trace, memories become malleable again – similar to when the memory trace was being consolidated. In this malleable state, memories are susceptible to modification from new information. For example, in a classic misinformation effect study (Loftus, 2005) participants were shown several images depicting an event – e.g., a car stopping before a stop sign. They were then given post-event descriptions of the event, one of which involved misinformation that the car stopped at a yield sign. When subsequently asked what they saw in the original event, participants who were given misinformation were more likely to report having seen a yield sign than those who were given not given misinformation. This susceptibility to incorporating new information into a retrieved memory is termed retroactive interference, which occurs because of memory reconsolidation.

According to Neisser's (1995) social psychological conception of self, our self-concept develops through our direct interactions with other people – i.e. interpersonal self-awareness. During social interaction people learn new information about themselves which can change their self-knowledge/concept. For example, if people are described as submissive by their conversation partner and are not given the chance to deny the assessment, they will subsequently change their self-concept to align with the description (Swann & Hill, 1982). Similarly, if people are asked leading questions that imply introversion or extroversion, trait self-ratings show that they subsequently shift their self-concept to align with the leading questions (Fazio et al., 1981). Therefore, external factors that provide new information about the self can cause the malleable self to reconsolidate, incorporating the newly considered information.

Interestingly, people can generate new information themselves, and this impacts on self-representation. Meyer et al. (2019) proposed that simulating other people – which involves retrieving self-information and simultaneously considering information about the other – reshapes self-knowledge; i.e. simulation induced malleability (SIM). This proposal was based on the premise

that if pre-existing semantic and episodic knowledge is susceptible to modification (i.e. post-event information becomes incorporated into pre-existing knowledge) and simulation involves considering knowledge about the self, then simulating other people should change self-knowledge to incorporate information about the simulated person (Meyer et al., 2019). As such, simulating other-knowledge (e.g. *My friend likes dogs a lot*) causes the self to be perceived as more like the simulated other (i.e., *I like dogs a lot*) than pre-simulation (i.e., *I like dogs a bit*).

Empirical research has provided support for this proposal. In a series of studies, Meyer et al. (2019) asked participants to first recall personal information (either trait knowledge, or episodic memory, e.g., "remember a time when you heard a dog barking; how did you feel?"), then simulate other people in similar contexts (e.g., "how would the average person [or, your friend] feel when they heard a dog barking?"), before finally recalling the personal information again (e.g., remember the time when you heard a dog barking; how did you feel?"). They found that simulating other people changed self-knowledge, whereby the self became more like the simulated other, and this effect persisted 24 hours after simulation (see also Rubin-McGregor et al., 2022). These effects show that self-knowledge is malleable and susceptible to misinformation effects, similar to other forms of semantic and episodic knowledge.

Moreover, Meyer et al. (2019) found that SIM was amplified for similar-others (similar-friend) compared to dissimilar-others (the average person). Greater SIM for similar-others is consistent with the simulation account's proposal that a greater amount of self-knowledge is recruited for simulation of similar-others, causing more self-knowledge to become malleable and susceptible to change. Thus, SIM might serve an evolutionary function. The social comparison theory (Festinger, 1954) proposes that people have an innate tendency to perceive themselves (i.e., our self-concept) in terms of how they compare to other people. Importantly, this tendency is more pronounced when the other person is similar to the self (e.g., Meyer et al., 2019): there is more pressure to attain uniformity with friends who we hence want to conform and perhaps affiliate with,

more so than for dissimilar-others who we might have less desire conform and affiliate with (see also Markus & Kunda, 1986).

Simulation Induced Malleability of Self-Representation in Autism

People with autism spectrum conditions (ASC)[15] have diminished social motivation, characterised by reduced interest in orienting to social cues compared to neurotypical people (Chevallier et al., 2012; Harrop et al., 2018; Lawrence et al., 2020). This suggests that they may have a reduced desire to attain uniformity with other people, and therefore that the self-knowledge/concept may be less susceptible to others' influence among autistic people than neurotypical people (e.g., Meyer et al., 2019). However, no research has yet empirically tested malleability of self-representation in autism.

Self-awareness atypicality is a key characteristic of ASC (Lombardo & Baron-Cohen, 2010), which is often characterised by difficulty adapting one's self-concept appropriately to different environments/contexts (Kanner, 1943; Lyons & Fitzgeral, 2013). ASC is diagnosed based on two key areas of behavioural impairments: social communication and interaction, and the tendency for restrictive and repetitive behaviours and interests (Diagnostic and Statistical Manual of Mental Disorders; American Psychiatric Association, 2013). These two areas are strongly connected and associated with self-awareness atypicality.

Much research has found diminished self-other distinction in ASC. For example, Lombardo et al. (2010) found that when asked to make reflective judgments about the physical and mental state characteristics of the self (self-representation) and a familiar-other (mentalizing), the ventral mPFC responded equivalently for both self- and other-representation in autistic males, unlike neurotypical males who showed the expected preferential recruitment of ventral mPFC for self-representations (Lombardo et al., 2010; see also Kennedy & Courchesne, 2008). This was characterised by diminished

[15] We acknowledge the ongoing debates about the terminology used to describe autism, and therefore use a mixture of commonly endorsed terms to reflect the fact that "there is no single way of describing autism that is universally accepted and preferred" (Kenny et al., 2016).

activation in the self-condition, but not the other condition relative to neurotypical participants (see also Di Martino et al., 2009; Bird & Viding, 2014). This is consistent with findings at the behavioural level, whereby unlike neurotypical people, autistic people show diminished or absent preferential processing of self-relevant information (Grisdale et al., 2014; Henderson et al., 2009; Lombardo et al., 2010; Smith et al., submitted; c.f., Lind et al., 2019; Williams et al., 2018; Wuyun et al., 2020).

Diminished self-other distinction in ASC suggests that autistic people may be more likely to attain uniformity with others – i.e. a greater tendency for SIM – compared to neurotypical people. This contradicts existing proposals based on diminished social motivation and an inflexible self-concept in ASC (Geurts et al., 2009; Kana et al., 2011; Minshew et al., 1997; Sanders et al., 2008), but is consistent with observations that autistic people experience difficulties maintaining a constant self-concept (Kanner, 1943; Lyons & Fitzgeral, 2013). Diminished self-awareness might mean that one's self concept is less strongly consolidated in memory, as evidenced by weaker episodic memory in ASC (Bowler et al., 2000; Gaigg et al., 2015; Lind, 2010). Consequently, the diminished self-other distinction in ASC might result from an overly malleable/flexible self-concept which is more susceptible to modulation in light of new information – e.g., simulation of other people.

Self-Reflection Induced Malleability of Representation of Other People

According to the social comparison theory (Festinger, 1954), one way that evaluators might attain uniformity/reduce divergence, is by changing their own views to be more like the other (e.g., Meyer et al., 2019). However, an alternative route is by attempting to shift their representation of the other person to be more like the self. This can be achieved by persuading the other person to shift their views, *or* by shifting one's own representation of the other person. While empirical research has examined the malleability of neurotypical self-representations following simulation of others (Meyer et al., 2019; Rubin-McGregor et al., 2022), to the best of our knowledge, malleability of one's own representation of other people has not been investigated in either neurotypical or autistic people.

We examine these two alternative explanations for diminished self-other distinction in ASC, testing whether other people are perceived to be more like the self than they actually are. ASC is characterised by impaired ToM, involving a diminished ability to understand and conceptualise other people's mental states (Baron-Cohen, 1989; Castelli et al., 2002; Frith & Happé, 1994). This might reduce appreciation of self-other differences (e.g., Uddin et al., 2008). Therefore, rather than understanding the full extent/complexity of a person, autistic people might put others into over-general categories which are easier to conceptualise. Consequently, rather than holding separate representations for self and other, autistic people might presume the characteristics of other people are like their own, and hence hold only one general representation for all social beings. Thus, in contrast to an account of diminished self-other distinction based on a more flexible self-concept in ASC, this alternative explanation proposes that diminished self-other distinction in ASC might be caused by over-attribution of the egocentric perspective during mental state attribution.

The Current Study

The current study advances Meyer et al.'s (2019) work by 1) replicating the SIM effect in a large sample of neurotypical adults and testing whether autistic people display SIM to a comparable extent as neurotypical people, and 2) adapting the paradigm to test whether a new kind of simulation (i.e. self-reflection) reshapes knowledge of others to become more like the self (i.e. self-reflection induced malleability; SRIM) and whether these effects differ between neurotypical and autistic adults. If self-reflection involves considering knowledge about other people (e.g., as a reference point), then self-reflection should change knowledge about others by incorporating knowledge about the self. As such, self-reflection should cause other people to be perceived as more like the self than pre-self-reflection.

We present four experiments. Experiments 4.1 and 4.2 take an individual differences approach by considering whether SIM (Experiment 4.1) and SRIM (Experiment 4.2) are associated with the number of autistic traits in neurotypical adults. Experiments 4.3 and 4.4 were pre-

registered, and took a case-control approach by considering whether level of SIM (Experiment 4.3) and SRIM (Experiment 4.4) differs between neurotypical and autistic adults. To measure SIM (Experiments 4.1 and 4.3) we replicated Meyer et al. (2019), by asking participants to: first recall personal information (e.g., "remember a time when you heard a baby crying; how did you feel?"), then simulate other people in similar contexts (e.g., "how would [the average person/your friend] feel when hearing a baby crying?"), and finally recall the personal information again (e.g., remember the time when you heard a baby crying; how did you feel?"). We predicted that participants would display SIM (i.e. perceiving the self as more like the simulated other on the second recall), and that this would be more pronounced for similar-others than for the average person. To measure SRIM (Experiments 4.2 and 4.4), the presented order of events was changed, such that participants first rated other people, then rated the self, and finally recalled ratings for other people again. We predicted that participants would display SRIM (i.e. perceiving the other as more like the self on the second recall), and that this would be more pronounced for similar-others than for the average person. Furthermore, we aim to test the two opposing predictions by considering whether 1) effects of simulation and/or self-reflection would be smaller/absent in people with more autistic traits because of their inflexible self-representation or reduced social motivation, or 2) effects of simulation and/or self-reflection would be greater in people with more autistic traits because they already have atypical self-other distinction and more flexible processing between self and other.

Experiment 4.1|Simulation Induced Malleability, Individual Differences Approach

METHOD

Participants

This study was carried out on a sample of 172 psychology undergraduate students at the University of Kent. Nine participants failed to pass the attention check on catch trials (see below), so were

excluded from the analyses. This gave a final sample of 163 participants (24 male and 138 female, 1 non-binary) aged between 18-43 years old (*M* = 19.47*, SD* = 2.77). An *a priori* power analysis using G*Power (Vinet & Zhedanov, 2010) revealed that 152 participants would be enough to detect a target x time interaction with effect size 0.23 (Meyer et al., 2019, Exp. 1) and power of .80.

All participants were native or fluent English speakers and had normal or corrected to normal vision. Participants gave written informed consent, and following completion of the experiment they were fully debriefed in writing and rewarded for their time with course credits. Completion of the experiment took approximately 30 minutes. Ethical approval was provided by the School of Psychology's Research Ethics Committee at the University of Kent.

Design, Measures and Procedure

This was a 2 (Target: similar-other/average person) x 2 (Time: baseline/post-simulation) repeated measures design. Participants completed the following tasks, in a fixed order, online via Qualtrics:

Simulation Task

Self Phase (Part 1): First, participants were prompted to recall 42 unique experiences from their past, 21 of which were positive (e.g., 'remember a time that you received good news'), and 21 were negative (e.g., 'remember a time that you received bad news'), presented in a random order across participants. For each experience, participants were asked to judge how they felt during this event, using a sliding marker on a continuous scale ranging from *extremely negative* to *extremely positive* (see Figure 5.1). To prevent participants from attempting to remember the value that they input on this or any subsequent rating, the scale did not show numerical values or markers associated with each increment.

Target Phase (Part 2): Immediately after, participants were asked to judge how two other targets would feel during a subset of these memory events (e.g., 'how would [target] feel receiving good news?') using the same sliding scale as the self phase. This included 1) a similar other person (chosen

by the participant as a friend whom they feel has similar personality, temperament, beliefs, values, major likes, and dislikes as themselves, and 2) someone described as the "average person". From the 42 experiences that participants recalled in the self phase, 14 experiences (7 positive; 7 negative) were presented to be judged in relation to the similar-other, 14 experiences (7 positive; 7 negative) were presented to be judged in relation to the average person, and 14 (7 positive; 7 negative) were discarded (i.e., three sets of experiences). The set of 14 experiences shown for each target or discarded was counterbalanced across participants, such that that each set was used for each variable equally. Target order (similar-other/familiar-other) was also counterbalanced across participants.

Post-Simulation Self Phase (Part 3): Immediately after, they were presented with the original prompts from the self phase again. They were asked to recall the same experiences that they had done previously, and judge how they felt during each experience again, using the same sliding scale (*post-simulation self*). This enables measurement of how simulation of the target changes the self. Completion time: 10 minutes.

Catch questions were presented in each self phase (x 2) and in each similar-other (x 1) and average person (x 1) phase, whereby the typical question was replaced with the instruction to "slide the marker all the way to the left [right]". This gave a total of 6 catch questions. Participants were required to answer at least 4/6 catch questions correctly for their data to be included in the analyses.

Perceived similarity measure: Finally, participants were asked to judge how similar they felt to 1) the similar-other, and to 2) the average person, in terms of their personality, temperaments, beliefs, values, and major likes and dislikes, on a scale ranging from 0 (not at all similar) to 100 (the same).

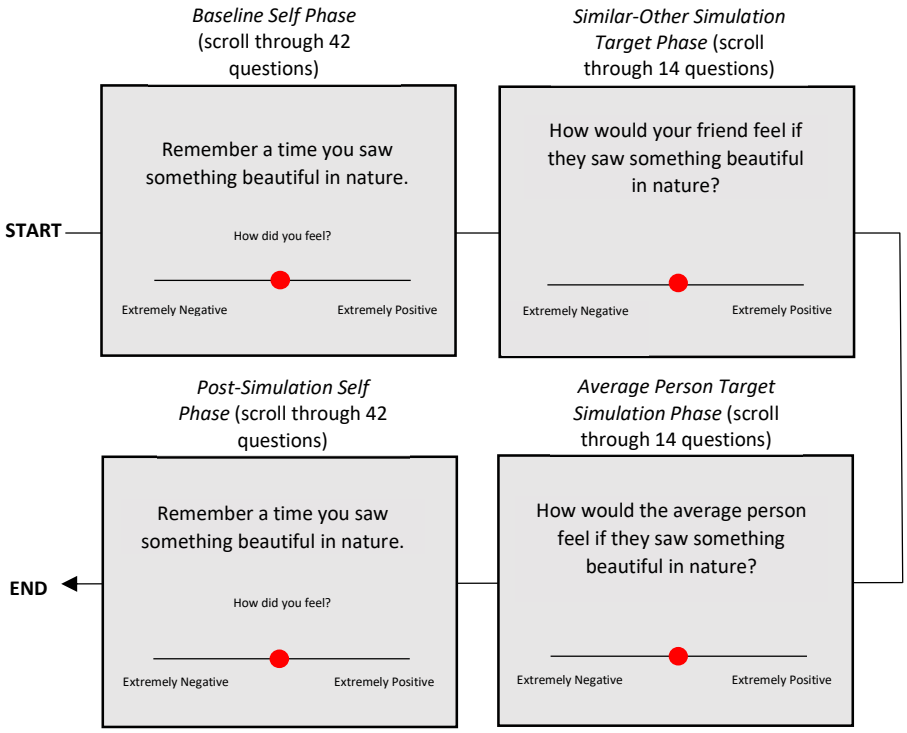


Figure 5,1. Schematic diagram of Simulation Task. Participants were first asked to rate 42 personal memories; next rate a subset (14 questions) of these memory events in relation to one of the targets, and then a different subset (14 questions) in relation to the second target (the order of which was counterbalanced across participants); and last, rate their 42 personal memories again. Participants could complete this on a laptop/desktop device, or a mobile/tablet device as shown.

Autism-spectrum Quotient

Participants completed the Autism-spectrum Quotient (AQ; Baron-Cohen et al., 2001). This is a standardized 50-item self-report questionnaire used to measure people's number of autistic traits. Participants were asked the extent to which they agree (definitely agree, slightly agree, slightly disagree, definitely disagree) with each statement about the self (e.g., "I find social situations easy"). The dependent variable is the total AQ score (ranging from 0-50), with higher scores indicating a greater number of ASC traits (Woodbury-Smith et al., 2005). Good test-retest reliability scores have been observed (range between r = .70 and r = .95; Broadbent et al., 2013). In the current sample, the mean AQ score was 18.65 (SD = 7.16; ranging from 5-41).

Toronto Alexithymia Scale

Participants completed the Toronto Alexithymia Scale (TAS-20; Bagby et al., 1994). This is a standardised 20-item self-report questionnaire which intends to measure one's level of alexithymia. Participants are asked the extent to which they agree (ranging from 1: "strongly disagree" to 5: "strongly agree") with each statement relating to understanding their own emotions (e.g., "I am often confused about what emotion I am feeling"). The dependent variable is the total TAS-20 score, which can range from 20-100, with higher scores indicating greater levels of alexithymia. Good test-retest reliability has been observed (r = .77; Bagby et al., 1994). In the current sample, the mean TAS-20 score was 51.68 (SD = 11.93; ranging from 25-84).

Statistical Analysis Procedure

For each memory prompt, two similarity values were computed: i) *baseline similarity* refers to the absolute difference between baseline self-rating and target rating (similar-other/average person separately, i.e. baseline self *minus* target rating); and ii) *post-simulation similarity* refers to the absolute difference between post-simulation self-rating and target rating (similar-other/average person separately). These absolute differences for each prompt were then averaged to produce variables representing baseline similarity to the 1) similar-other, and 2) average person; and post-simulation similarity to the 3) similar-other, and 4) average person. Lower similarity scores reflect greater perceived similarity between the self and the target.

Bayesian Analysis

Bayesian analysis was also used to interpret the results (conducted using JASP 0.14.1; JASP Team, 2020), which enables a more graded interpretation than only using p values or effect sizes. This estimates the relative strength of the alternative hypothesis over the null hypothesis, or vice versa (e.g., Dienes, 2014; Rouder et al., 2009). For the Bayesian analyses, we adopted the default Cauchy priors as recommended by Wagenmakers *et al.* (2018). Bayes factor (BF^{10}) < 1 is evidence

supporting the null hypothesis (<0.33 is firm evidence), and Bayes factors > 3, >10, > 30, and >100 are firm, strong, very strong, and decisive evidence supporting the alternative hypothesis respectively.

The datasets supporting this article are available on the Open Science Framework (https://osf.io/n6muv/?view_only=67123f8b300f4286be9f6e5cd7107d9a).

RESULTS

Alpha was set to .05 for all analyses. A paired samples t-test was computed to verify that the Target manipulation worked (i.e. that participants rated themselves to be more similar to their nominated similar other than to the average person; Meyer et al, 2019; Ames, 2004; Benoit et al., 2010; Mitchel et al., 2006; Tamir & Mitchell, 2013) in the *perceived similarity measure*. This confirmed that participants perceived the similar-other to be more similar to the self (M = 77.80, SD = 11.16) than the average person (M = 52.94, SD = 18.06), $t(162) = 19.60$, $p < .001$, $d = 1.54$.

Analysis of Variance: To investigate the effect target and time on similarity to the target, a 2 (Target: similar-other/average person) x 2 (Time: baseline/post-simulation) repeated measures ANOVA was conducted on similarity values (see Figure 5.2). There was a significant main effect of time, $F(1, 162) = 88.99$, $p < .001$, $\eta_p^2 = 0.33$, $BF^{10} > 100$, whereby people perceived themselves to be more similar to the target post-simulation (M = 20.95, SD = 7.54) than at baseline (M = 24.67, SD = 8.20). There was no significant main effect of target, $F(1, 162) = 0.00$, $p = .99$, $\eta_p^2 = 0.00$, $BF^{10} = .09$, and no target x time interaction, $F(1, 162) = 2.75$, $p = .10$, $\eta_p^2 = 0.02$, $BF^{10} > 100$.

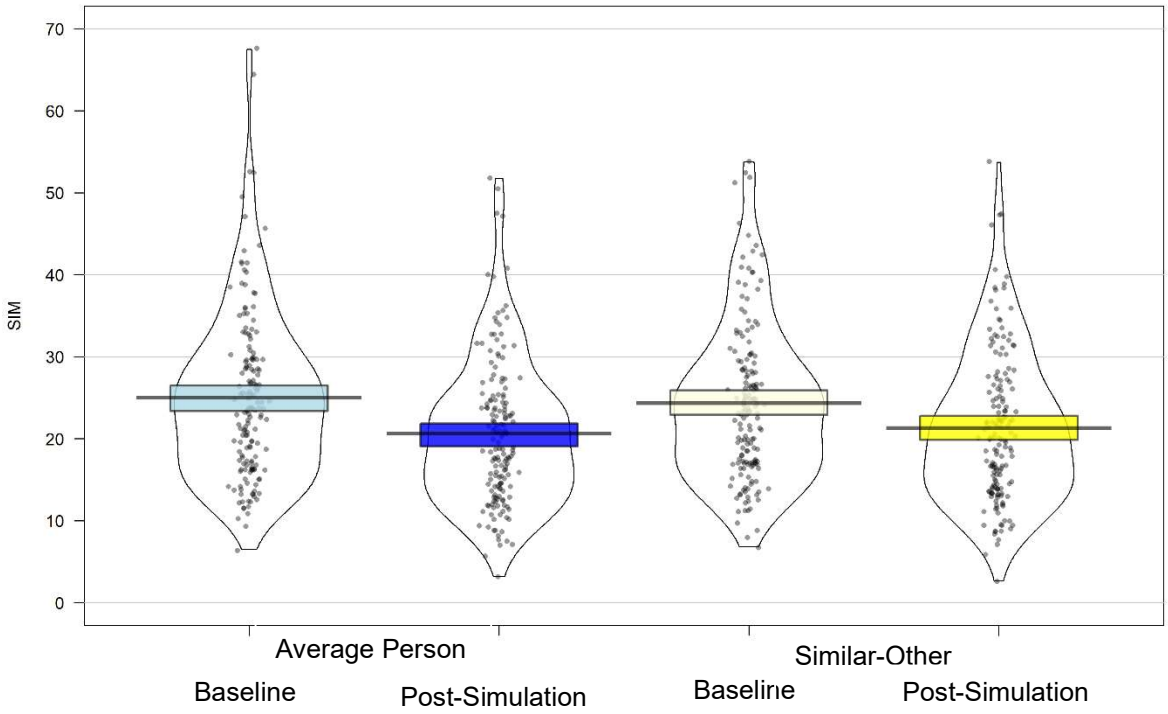


Figure 5.2. Mean similarity to target in the in each target x time condition. Shaded area represents 95% confidence interval.

Association Analyses: A series of bivariate correlation analyses were conducted to examine whether individual differences in AQ and/or TAS-20 are associated with overall SIM (i.e. baseline similarity to the target minus post-simulation similarity to the target). Because Meyer et al. (2019) observed different SIM effects in similar-other and average person targets, correlations were run separately for each Target (similar-other/average person)[16]. These correlations revealed that AQ score was not associated with similar-other SIM, $r(163) = .04$, $p = .64$, $BF^{10} = .11$, but was positively associated with average person SIM, $r(163) = .20$, $p = .01$, $BF^{10} = 2.26$; participants who had higher average person SIM, had higher AQ (See Figure 5.3). TAS-20 was not associated with similar-other SIM, $r(163) = .01$, $p = .93$, $BF^{10} = .10$, or average person SIM, $r(163) = .13$, $p = .09$, $BF^{10} = .41$.

[16] When considering SIM for both targets combined, these correlations revealed that SIM was not significantly correlated with AQ score, $r(163) = .15$, $p = .06$, $BF^{10} = .59$, or TAS_20, $r(163) = .09$, $p = .27$, $BF^{10} = 0.18$.

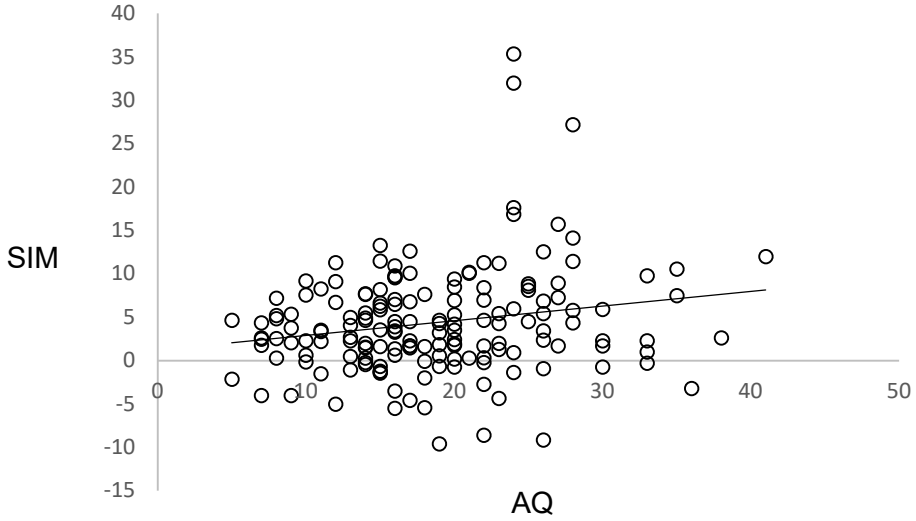


Figure 5.3. Scatter plot depicting significant positive correlation between AQ and average person SIM. Line depicts line of best fit.

DISCUSSION

Experiment 4.1 aimed to replicate previous findings (Meyer et al., 2019; Rubin-McGregor et al., 2022) that simulating other people changes the self to become more similar to the other, and that the magnitude of this simulation induced malleability (SIM) is greater when simulating similar-other people than the average person. We partially replicated these previous findings. We found a main effect of time, whereby participants perceived themselves to be more similar to the target post-simulation than at baseline. However, in contrast to Meyer et al., this was not modulated by target; the magnitude of the effect did not differ between the similar-other and the average person.

The second aim of Experiment 4.1 was to investigate the relationship between SIM and autistic traits. Analyses revealed a significant positive association between autistic traits (AQ) and SIM in the average person condition: participants with a higher number of autistic traits had higher SIM. In contrast, autistic traits were not associated with similar-other SIM. Therefore, the current findings suggest that in a case-control design, autistic adults would show greater SIM compared to neurotypical

adults in the average person condition, but comparable SIM in the similar other condition. We test this prediction in Experiment 4.3.

Experiment 4.2| Self-Reflection Induced Malleability, Individual Differences Approach

METHOD

Participants

This study was carried out on a new sample of 167 psychology undergraduate students at University of Kent. Six participants were excluded due to an experimenter error in the task, and five participants failed to pass the attention checks on catch trials (see below), so were excluded from subsequent analyses. This gave a final sample of 162 participants (19 male, 140 female and 3 non-binary) aged between 18-36 (*M* = 19.33, *SD* = 2.30). Like Experiment 4.1, an *a priori* power analysis using G*Power (Vinet & Zhedanov, 2010) revealed that 152 participants would be enough to detect a target x time interaction with effect size 0.23 (Meyer et al., 2019, Exp. 1) and power of .80.

All participants were native or fluent English speakers and had normal or corrected to normal vision. Participants gave written informed consent, and following the completion of the experiment they were fully debriefed in writing and rewarded course credits for their time. Completion of the experiment took approximately 30 minutes. Ethical approval was provided by the School of Psychology's Research Ethics Committee at the University of Kent.

Design, Measures and Procedure

This was a 2 (Target: similar-other/average person) x 2 (Time: baseline/post-self-reflection) repeated measures design. Participants completed the following tasks, in a fixed order, online via Qualtrics:

Self-Reflection Task

Participants completed the same Self and Target phases as Experiment 4.1, however the order of presentation was changed (see Figure 5.4). Participants first completed the (baseline) Target phase (average-person = 14 experiences; and similar-other = 14 experiences; target order randomised), followed by the Self phase (42 experiences). Immediately after, they were presented with the original prompts from the Target phase and were asked to rate how the targets would feel during each experience again (post-self-reflection target phase). Here, the subset of memory events allocated to each target was the same as the baseline target phase, and the order in which participant completed the similar-other and average person phases was randomised. This enabled measurement of how self-reflection changes representation of the target. Completion time: 10 minutes.

Catch questions were presented in each self-phase (x 2) and in each similar-other (x 1) and average person (x 1) phase, whereby the typical question was replaced with the instruction to "slide the marker all the way to the left [right]". This gave a total of 6 catch questions. Participants were required to answer at least 4/6 catch questions correctly for their data to be included in the analyses.

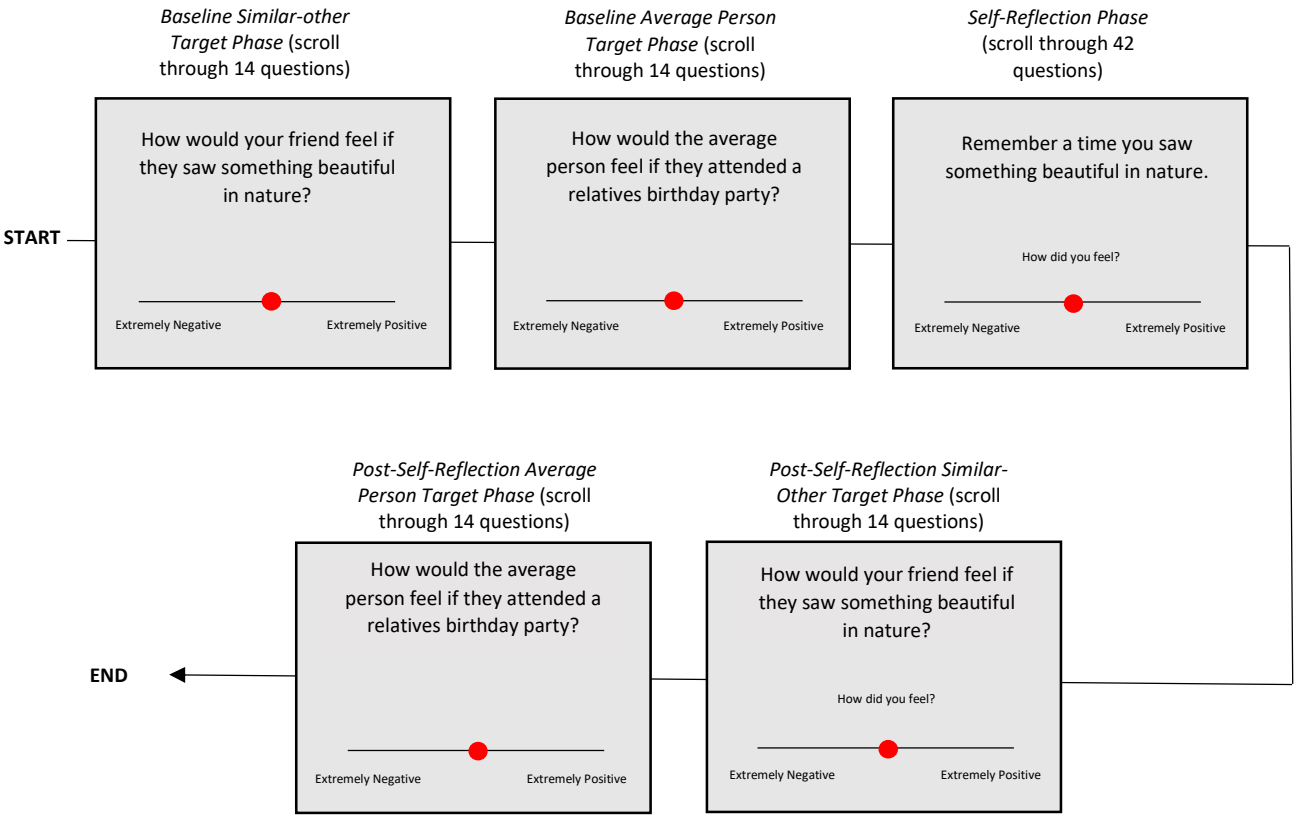


Figure 5.4. Schematic diagram of Self-Reflection Task. Participants were first asked to rate a subset (14 questions) of memory events in relation to one of the targets, and then a different subset (14 questions) in relation to the second target (the order of which was counterbalanced across participants); next rate 42 personal memories (all of those events that were previously rated in relation to the targets were included here); and finally, rate the subsets of questions in relation to the each of the targets (14 + 14 questions; the order of which was counterbalanced across participants) again. Participants could complete this on a laptop/desktop device, or a mobile/tablet device (as shown in Figure 5.1).

Perceived similarity measure: Finally, participants were asked to judge how similar they felt to 1) the similar-other, and to 2) the average person in terms of their personality, temperaments, beliefs, values, and major likes and dislikes, on a scale ranging from 0 (not at all similar) to 100 (the same).

Bayesian Analysis

Like Experiment 4.1, Bayesian analysis was also used to interpret the results (conducted using JASP 0.14.1; JASP Team, 2020).

Autistic traits

As in Experiment 4.1, participants completed the Autism-spectrum Quotient (AQ; Baron-Cohen et al., 2001): AQ scores ranged between 5–34 (M = 19.30, SD = 6.01), and the Toronto Alexithymia Scale (TAS-20; Bagby et al., 1994): TAS-20 scores ranged between 22–78 (M = 51.90, SD = 11.41).

Statistical Analysis Procedure

Similar to Experiment 4.1, for each memory prompt, two similarity values were computed: i) *baseline similarity* refers to the absolute difference between self rating and baseline target rating (similar-other/average person separately; baseline self *minus* target rating); and ii) *post-self-reflection similarity* refers to the absolute difference between the self rating and post-self-reflection target rating (similar-other/average person separately). These absolute differences for each prompt were then averaged to produce variables representing baseline similarity to the 1) similar-other, and 2) average person; and post-self-reflection similarity to the 3) similar-other, and 4) average person. Lower similarity scores reflect greater perceived similarity between the self and the target.

The datasets supporting this article are available on the Open Science Framework (https://osf.io/n6muv/?view_only=67123f8b300f4286be9f6e5cd7107d9a).

RESULTS

Alpha was set to .05 for all analyses. A paired samples t-test on the *perceived similarity measure* confirmed that participants perceived the similar-other to be more similar to the self (M = 77.84, SD = 12.51) compared to the average person (M = 52.81, SD = 16.86), $t(161) = 20.30$, $p < .001$, $d = 1.60$.

Analysis of Variance: To investigate the effect of target and time on similarity to the target, a 2 (Target: similar-other/average person) x 2 (Time: baseline/post-self-reflection) repeated measures ANOVA was conducted on similarity values (see Figure 5.5). There was a significant main effect of target, $F(1, 161) = 9.53$, $p = .002$, $\eta_p^2 = 0.06$, $BF^{10} > 100$, whereby people perceived themselves to be more similar to the similar-other ($M = 26.44$, $SD = 9.94$) than to the average person ($M = 28.88$, $SD = 9.94$). There was no significant main effect of time, $F(1, 161) = 0.21$, $p = .65$, $\eta_p^2 = 0.001$, $BF^{10} = .16$, but the target x time interaction was significant, $F(1, 61) = 67.14$, $p < .001$, $\eta_p^2 = 0.29$, $BF^{10} = 17.48$.

Simple main effects revealed that whereas the similar-other was rated as more similar to the self post-self-reflection than at baseline, $t(161) = 5.11$, $p < .001$, $d = .40$, the average person was more similar to the self at baseline than post-self-reflection, $t(161) = 5.84$, $p < .001$, $d = .46$. Furthermore, at baseline, similarity to the similar-other and the average person did not differ, $t(161) = 1.28$, $p = .20$, $d = .10$, however post-self-reflection, the similar-other was more similar to the self than the average person, $t(161) = 6.66$, $p < .001$, $d = .52$.

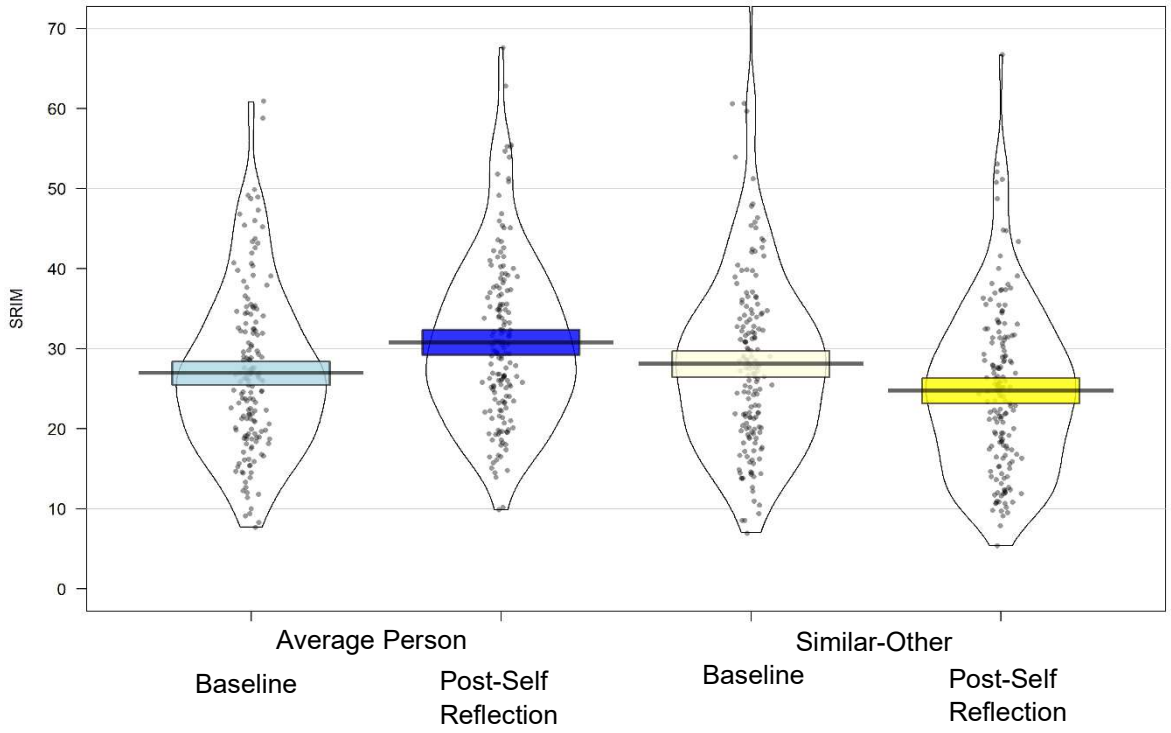


Figure 5.5. Mean similarity to target in the in each target x time condition. Shaded area represents 95% confidence interval.

Association Analyses: A series of bivariate correlation analyses were conducted, separately for each Target (similar-other/average person), to examine whether individual differences in AQ and/or TAS-20 are associated with overall SRIM (calculated as baseline similarity to the target minus post-self-reflection similarity to the target). Correlations revealed that AQ score was not associated with similar-other SRIM, $r(162) = .06$, $p = .48$, $BF^{10} = .13$, or average person SRIM, $r(162) = .04$, $p = .62$, $BF^{10} = .11$. Furthermore, TAS-20 was not associated with similar-other SRIM $r(162) = .06$, $p = .46$, $BF^{10} = .13$, and average person SRIM, $r(162) = .07$, $p = .35$, $BF^{10} = .15$.

DISCUSSION

Experiment 4.2 aimed to extend previous findings (Meyer et al., 2019; Rubin-McGregor et al., 2022) by considering whether self-reflection changes one's perception of other people to become more like the self, and furthermore, whether the magnitude of this self-reflection induced malleability (SRIM) is greater when rating people who are similar to the self compared to the average person. Our results partially supported this prediction. The similar-other was rated as being more like the self post-self-reflection than at baseline, suggesting that self-reflection changed one's perception of the similar-other to become more like the self. However, the average person was rated as being more like the self at baseline than post-self-reflection. This suggests that self-reflection changed one's perception of the other differently depending on level of similarity to the self. It is possible that self-reflection caused participants to become more aware of the similarities between the self and similar-other, which increased the overlap between representations of the self and similar other. In contrast, self reflection caused participants to become more aware of the differences between the self and the average person, which increased the distinction between representations of the self and an average (unfamiliar) other (Ferguson et al., 2018; Savitsky et al., 2011; Simpson & Todd, 2017).

The second aim of Experiment 4.2 was to investigate the relationship between SRIM and autistic traits. Analyses found no significant association between autistic traits and SRIM in ether the similar-other or average person. Therefore, the current findings suggest that in a case-control design, autistic adults would show comparable SRIM to matched neurotypical adults (in opposite directions for similar-other and average person). We test this prediction in Experiment 4.4.

Experiments 4.3 and 4.4|Case-Control Approach

METHOD

Experiments 4.3 and 4.4 were fully pre-registered on the Open Science Framework (https://osf.io/n6muv/?view_only=67123f8b300f4286be9f6e5cd7107d9a).

Based on previous research (Meyer et al. 2019), in the simulation task (Experiment 4.3), it was hypothesised that participants would rate themselves to be more like the target post-simulation than at baseline. Additionally, in line with results from Experiment 4.1, we did not expect this SIM effect to be modulated by target or to differ overall between diagnostic groups.

Based on Experiment 4.2, in the self-reflection task (Experiment 4.4), it was hypothesised that there would be a target x time interaction. We expected the similar-other target would become more like the self post-self-reflection than at baseline, but the average person would become more different to the self post-self-reflection than at baseline. Furthermore, whilst at baseline, there would be no difference between similarity to the similar-other and average person, post-self-reflection, the similar-other would be more like the self than the average person. Additionally, in line with results from Experiment 4.2, we did not expect this SRIM effect to differ between diagnostic groups.

For both experiments, we also tested the contribution of individual differences in personality traits/cognitive abilities to SIM/SRIM. Based on the findings from Experiments 4.1 and 4.2, we predicted that there would be no association between SIM/SRIM and AQ or TAS-20. However, we predicted positive relationships between SIM/SRIM and i) CAT-Q, ii) BAI, iii) EQ, iv) animations and v)

scene construction, whereby people with higher SIM/SRIM would have higher CAT-Q (indicating more camouflaging), BAI (indicating higher anxiety), EQ (indicating higher empathy), animations (indicating better TOM) and scene construction (indicating better scene construction).

Participants

Participants were primarily recruited from the Autism Research Kent participant database, and some additional participants were recruited from other University of Kent databases. In total, Experiment 4.3 included 24 neurotypical adults with no history of psychiatric or neurological conditions (Self-report) and 26 adults with a clinical diagnosis of ASC: specifically, this was either autistic disorder (N = 6), autism spectrum disorder (N = 4), Asperger's syndrome (N = 10) or not known (N = 6). Experiment 4.4 included 25 neurotypical adults with no history of psychiatric or neurological conditions (Self-report) and 25 adults with a clinical diagnosis of ASC: specifically, this was either autistic disorder (N = 5), autism spectrum disorder (N = 4), Asperger's syndrome (N = 9) or not known (N = 7). Official diagnostic information was checked.

To assess the current autistic characteristics, autistic participants were also assessed on module 4 of the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000) by a trained, research-reliable researcher, and videos were double coded to ensure reliability of scoring (inter-rater reliability was found to be excellent with intraclass correlation of .89).

All participants had normal or corrected to normal hearing and vision. The two participant groups were matched on age, gender, and Intelligence Quotient (IQ, Wechsler Adult Intelligence Scale, WAIS-III or WAIS-IV; Wechsler, 1997; Wechsler, 2008), but were not matched on Autism spectrum Quotient (AQ; Baron-Cohen et al., 2001). See Table 1 (Exp. 3) and 2 (Exp. 4) for participant demographics.

Participants provided written informed consent and they were fully debriefed following completion of the experiment. They received £10 per hour of their participation, plus additional

travel expenses. Ethical approval was provided by the School of Psychology's Research Ethics Committee at the University of Kent.

Table 1. Participant demographics in Experiment 4.3.

	Autistic (n = 26; 17 male)			Neurotypical (n = 24; 14 male)					
(Autistic, Neurotypical)	Mean (SD)	Range	N	Mean (SD)	Range	N	*t*	*p*	*d*
Age (years)	35.84 (11.09)	20.17-59.88	26	41.12 (14.27)	19.98-68.90	24	1.47	.149	0.42
IQ: Full Scale	107.12 (17.29)	73-137	26	111.78 (9.15)	97-132	23	1.16	.253	0.33
IQ: Verbal 6, 23)	104.35 (12.63)	78-126	26	110.17 (10.89)	92-129	23	1.72	.092	0.49
IQ: Performance	108.46 (22.30)	65-143	26	112.00 (11.24)	89-144	23	0.69	.495	0.20
ADOS: Total	7.91 (5.62)	1-21	20	-	-	-	-	-	-
AQ	30.77 (7.18)	19-45	26	15.14 (5.96)	4-28	22	8.12	<.001	2.35
TAS-20	59.64 (15.65)	5-81	25	45.90 (10.64)	31-65	19	3.29	.002	1.00
Animations	4.67 (2.54)	0-8	12	5.75 (1.83)	3-8	8	1.04	.31	0.47
EQ	38.71 (17.38)	15-69	24	49.41 (9.22)	29-66	22	2.57	.01	0.76
BAI	23.35 (12.39)	3-54	26	10.54 (7.00)	0-24	24	4.45	<.001	1.26
CAT-Q	114.58 (26.72)	55-162	26	85.83 (14.73)	64-109	24	4.66	<.001	1.32
Scene Construction	35.89 (6.72)	17.80-46.67	19	41.59 (5.69)	30.33-46.87	7	1.99	.06	0.88

Table 2. Participant demographics in Experiment 4.4.

	Autistic (n = 25; 18 male)			Neurotypical (n = 25; 15 male)					
(Autistic, Neurotypical)	Mean (SD)	Range	N	Mean (SD)	Range	N	*t*	*p*	*d*
Age (years)	34.61 (10.23)	20.17-55.42	25	40.51 (14.29)	19.98-68.90	25	1.68	.100	0.48
IQ: Full Scale	108.16 (16.82)	73-137	25	111.75 (8.95)	97-133	24	0.93	.359	0.27
IQ: Verbal	104.88 (15.35)	78-151	25	110.46 (10.72)	92-129	24	1.47	.148	0.42
IQ: Performance	110.00 (20.97)	65-143	25	111.79 (11.04)	89-144	24	0.37	.712	0.11
ADOS: Total	7.95 (5.67)	1-21	20	-	-	-	-	-	-
AQ	30.68 (7.29)	19-45	25	14.61 (6.34)	3-28	23	8.12	<.001	2.35
TAS-20	61.46 (11.04)	33-81	24	45.00 (11.10)	28-65	20	4.91	<.001	1.49
Animations	5.36 (2.06)	2-8	11	5.75 (1.83)	3-8	8	0.42	.339	.20
EQ	37.91 (17.53)	15-69	22	50.30 (9.98)	29-70	23	2.93	.005	0.87
BAI	23.88 (12.08)	3-54	25	10.72 (6.91)	0-24	25	4.73	<.001	1.34

CAT-Q	116.32 (24.12)	55-162	25	83.92 (17.31)	38-109	25	5.46	<.001	1.54
Scene Construction	35.82 (6.99)	17.80-46.67	16	41.59 (5.69)	30.33-46.87	7	1.92	.07	0.87

Materials and Procedure

Participants completed the *Simulation* and *Self-reflection tasks*, as outlined in Experiments 4.1 and 4.2 respectively[17]. The order in which participants complete the two tasks was counterbalanced across participants, such that half the participants completed the simulation task in session 1, then the self-reflection task in session 2; and half the participants completed the self-reflection task in session 1, then the simulation task in session 2.

Measures of Personality and Mood

See Tables 1 and 2 for participant demographics on each of the measures detailed below.

Toronto Alexithymia Scale (TAS-20; Bagby et al., 1994): See Experiment 4.1 for full description.

Autism-spectrum Quotient (AQ; Baron-Cohen et al., 2001): See Experiment 4.1 for full description.

Animations Task (Abell et al., 2000): This is a measure of Theory of Mind ability. Participants were shown four silent animations (34-45 seconds each) of a large red triangle and a small blue triangle that are either: (i) moving randomly; (ii) moving in a goal directed fashion (e.g., fighting); or (iii) moving interactively with implied intentions (coaxing, tricking). Participants were asked "What was happening in this animation?". Condition three is designed to elicit descriptions of the mental states of the triangles – i.e., requires mentalising. Prior to the test trials, participants completed one practice trial. Verbal descriptions are recorded and coded. Scores range from 0-2 for each animation, with higher scores indicating more accurate attribution of the triangle's mental states (e.g., desire and intentions).

Empathising Quotient (Baron-Cohen & Wheelwright, 2004): This is a standardized 60-item self-report questionnaire which intends to measure empathy. Participants were asked the extent to which they agree (definitely agree, slightly agree, slightly disagree, definitely disagree) with each statement

[17] In Experiment 3 and 4, the *same* set of 14 experiences were always rated in relation to the similar-other and the average person – i.e., this was not randomised. However, the order in which the target phases were presented was counterbalanced.

relating to the self (e.g., “I can easily tell if someone else wants to enter a conversation”). The dependent variable is the total empathy quotient (EQ) score.

Beck Anxiety Inventory (BAI; Beck et al., 1988): This is a standardized 21-item self-report questionnaire which intends to measure one’s level of anxiety. Participants were asked the extent to which symptoms of anxiety (e.g., “Unable to relax”) affected them during the last month on a four-point scale (“not at all”; “mildly - but it didn’t bother me much”; “moderately – it wasn’t pleasant at times”; “severely – it bothered me a lot”). The dependent variable is the total BAI score which can range from 0 (low anxiety) to 63 (high anxiety).

Camouflaging Autistic Traits Questionnaire (CAT-Q; Hull et al., 2019): This is a standardized 25-item self-report questionnaire which intends to measure the frequency of camouflaging behaviours. Participants were asked the extent to which they agree [ranging from “strongly disagree” (1) to “strongly disagree” (7)] with each statement relating to their own social interactions (e.g., “In social situations, I feel like I’m performing rather than being myself”). The dependent variable is the total CAT-Q score which can range from 25-175. Acceptable internal consistency ($\alpha = .94$) and preliminary test-retest reliability ($r = .77$) has been observed (Hull et al., 2019).

The *scene construction task* (Hassabis, Kumaran, & Maguire, 2007): This measures the ability to mentally construct a visual scene. The experimenter asked the participant to close their eyes and vividly imagine three ordinary fictitious scenes (e.g., “Imagine you’re lying on beach in a tropical bay”). They were then given two to three minutes to “Describe what you can see, hear, smell and feel in as much (multimodal) detail as possible”. Participant ratings: After each description, participants were asked to complete a questionnaire rating their imagined scenes for perceived salience/vividness (salience of the imager: 1, not vivid, to 5, very vivid), and for sense of presence/difficulty (how hard was the trial: 1, easy, to 5, hard). Spatial Coherence Index: Participants were given a post-description questionnaire consisting of 12 statements, each providing a possible

qualitative description of the newly constructed imagined scene/episode. Participants were asked to tick any of the statements that they think accurately applied to their imagined scene/episode.

Bayesian Analysis

Like the previous experiments, Bayesian analysis was used to interpret the results (conducted using JASP 0.14.1; JASP Team, 2020).

The datasets supporting this article are available on the Open Science Framework (https://osf.io/n6muv/?view_only=67123f8b300f4286be9f6e5cd7107d9a).

Community Involvement

There was no community involvement in the design or implementation of the reported study.

RESULTS

Experiment 4.3: Simulation Task

A paired samples t-test on the perceived similarity measure confirmed that overall[18], participants perceived the similar-other to be more similar to the self (M = 74.54, SD = 13.29) than the average person (45.80, SD = 22.22), $t(49) = 9.15$, $p = <.001$, $d = 1.29$.

To investigate the effect of diagnosis, target, and time on similarity to the target, a 2 (Diagnosis: ASC/NT) x 2 (Target: similar-other/average person) x 2 (Time: baseline/post-simulation) mixed ANOVA was conducted on similarity values, with repeated measures on the last two factors (see Figure 5.6). There was a significant main effect of time, $F(1, 48) = 33.42$, $p = <.001$, $\eta_p^2 = 0.41$, $BF^{10} >100$, whereby people perceived themselves to be more similar to the target post-simulation ($M = 22.17$, $SD = 8.94$) than at baseline ($M = 27.70$, $SD = 10.53$). There was also a significant main effect of diagnosis, $F(1, 48) = 6.72$, $p = .01$, $\eta_p^2 = 0.12$, $BF^{10} = 4.12$, whereby the neurotypical group perceived themselves to be more similar to the target (M = 21.64, SD = 4.78) than the ASC group

[18] This was also found in each diagnostic group separately. NT: $t(23) = 5.19$, $p = <.001$, $d = 1.060$ ($M_{\text{similar-other}} = 74.79$, $SD_{\text{similar-other}} = 13.11$; $M_{\text{average Person}} = 55.54$, $SD_{\text{average Person}} = 15.71$); ASC: $t(25) = 8.58$, $p = <.001$, $d = 1.682$ ($M_{\text{similar-other}} = 74.31$, $SD_{\text{similar-other}} = 13.71$; $M_{\text{average Person}} = 36.81$, $SD_{\text{average Person}} = 23.78$).

(27.98, SD = 11.06). There was no significant main effect of target, $F(1, 48) = 0.001$, $p = .98$, $\eta_p^2 = 0.00$, $BF^{10} = .15$, and no interactions between target x diagnosis, $F(1, 48) = 0.37$, $p = .55$, $\eta_p^2 = 0.01$, $BF^{10} = .67$, target x time, $F(1, 48) = 0.44$, $p = .51$, $\eta_p^2 = 0.01$, $BF^{10} > 100$, or target x time x diagnosis, $F(1, 48) = 0.40$, $p = .53$, $\eta_p^2 = 0.01$, $BF^{10} > 100$.

However, there was a time x diagnosis interaction, $F(1, 48) = 5.57$, $p = .02$, $\eta_p^2 = 0.10$, $BF^{10} > 100$. Simple main effects revealed that the neurotypical group perceived themselves to be more similar to the target post-simulation ($M = 20.03$, $SD = 4.71$) than at baseline ($M = 23.25$, $SD = 5.41$), $t(23) = 4.61$, $p = <.001$, $d = .94$. The ASC group also perceived themselves to be more similar to the target post-simulation ($M = 24.15$, $SD = 11.31$) than at baseline ($M = 31.81$, $SD = 12.40$), $t(25) = 4.54$, $p <.001$, $d = .89$. Furthermore, whilst in the baseline condition, the neurotypical group ($M = 23.25$, $SD = 5.41$) perceived themselves to be more similar to the target than the ASC group ($M = 31.81$; SD 12.40), $t(48) = 3.12$, $p = .003$, $d = .88$; in the post-simulation condition, the neurotypical ($M = 20.03$, $SD = 4.71$) and ASC ($M = 24.15$, $SD = 11.30$) groups did not differ in perceived similarity to the target, $t(48) = 1.65$, $p = .11$, $d = 4.68$.

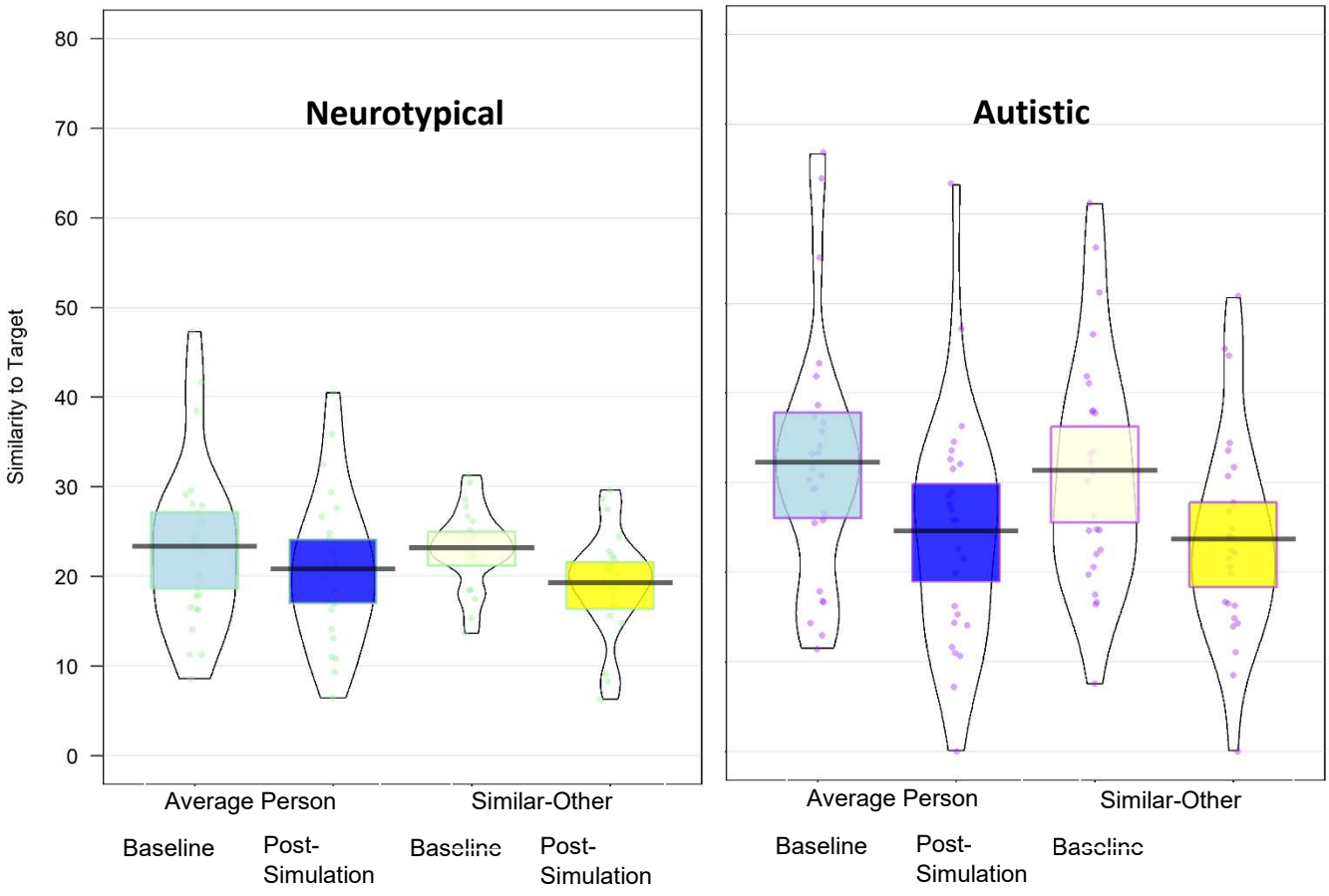


Figure 5.6. Mean similarity to target in each target x time condition in neurotypical (left/green) and ASC (right/pink) groups. Shaded area represents 95% confidence interval.

Association Analyses

Correlation analyses investigated the relations between SIM (in each target condition) and (i) AQ (ii) animations, (iii) BAI, (iv) TAS-20, (v) CAT-Q, (vi) EQ, and (vii) scene construction across the whole group of participants. See Table 2 and Figure 5.7.

Table 2. Matrix displaying correlations between SIM and cognitive predictors in the total sample.

	Similar-Other SIM	Average Person SIM	AQ	animations	BAI	TAS-20	CAT-Q	EQ Post-Simulation	Imagination
AQ (48)	.273	.097	-						
Animations (20)	-.549*	-.673**	.059	-					
BAI (n= 50)	.076	-.081	.080	-.063*	-				
TAS-20 (n= 44)	-.026	-.147	.605**	-.441	.102	-			
CAT-Q (n = 50)	.130	-.093	.546**	.270	-.033	.495**	-		
EQ (n = 46)	-.254	-.156	-.377*	-.155	-.053	-.322*	-.053	-	
Imagination (26)	-.464*	-.108	-.407*	.234	-.352	-.381	-.384	.262	-

*p <0.05; **P < 0.01; *** p < 0.001.

Animations was negatively associated with similar-other SIM, $r(20) = -.55$, $p = .01$, $BF^{10} = 5.19$, and average person SIM, $r(20) = -.67$, $p = .001$, $BF^{10} = 37.88$. Participants with higher SIM had

lower animations score. Imagination was also negatively associated with similar-other SIM, $r(26) = -.46$, $p = .02$, $BF^{10} = 3.64$. Participants with higher similar-other SIM had lower imagination ability. There were no significant correlations between similar-other SIM or average person SIM and any other variables.

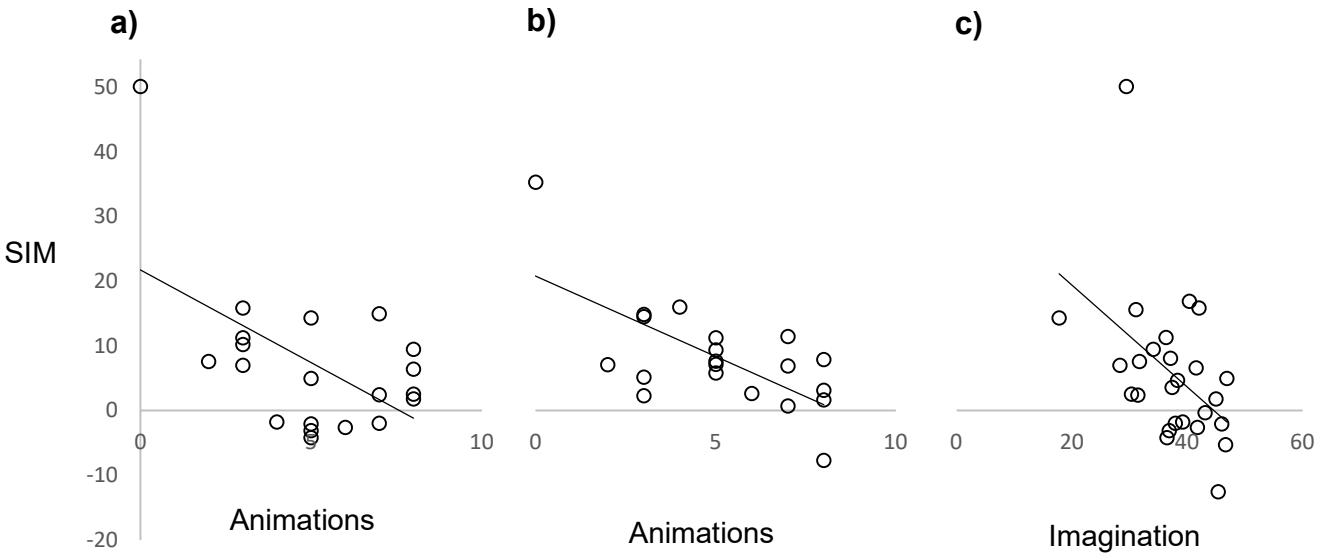


Figure 5.7. Scatter plots depicting significant correlation between animations and similar-other SIM (a), and average person SIM (b), and between imagination and similar-other SIM (c). Lines depict lines of best fit.

Experiment 4.4: Self-Reflection Task

A paired samples t-test on the perceived similarity measure confirmed that overall[19], participants perceived the similar-other to be more similar to the self (M = 71.28, SD = 16.21) than the average person (M = 44.14, SD = 20.33), $t(49) = 8.89$, $p = <.001$, $d = 1.26$.

To investigate the effect of diagnosis, target and time on similarity to the target, a 2 (Diagnosis: ASC/NT AQ) x 2 (Target: similar-other/average person) x 2 (Time: baseline/post-self-reflection) mixed ANOVA was conducted on similarity values, with repeated measures on the last

[19] This was also found in each diagnostic group separately. NT: $t(24) = 6.52$, $p = <.001$, $d = 1.30$ ($M_{\text{similar-other}} = 73.04$, $SD_{\text{similar-other}} = 14.32$; $M_{\text{average Person}} = 55.64$, $SD_{\text{average Person}} = 13.95$); ASC: $t(23) = 7.69$, $p = <.001$, $d = 1.50$ ($M_{\text{similar-other}} = 69.52$, $SD_{\text{similar-other}} = 18.02$; $M_{\text{average Person}} = 32.64$, $SD_{\text{average Person}} = 19.33$).

two factors (see Figure 5.8). There were no significant main effects of diagnosis, $F(1, 48) = 0.90$, $p = .35$, $\eta_p^2 = 0.02$, $BF^{10} = .50$, target, $F(1, 48) = 0.12$, $p = .73$, $\eta_p^2 = 0.002$, $BF^{10} = .17$, or time, $F(1, 48) = 0.04$, $p = .85$, $\eta_p^2 = <0.001$, $BF^{10} = .15$. There was also no significant interactions between diagnosis x target, $F(1, 48) = 1.00$, $p = .32$, $\eta_p^2 = 0.02$, $BF^{10} = .08$, diagnosis x time, $F(1, 48) = 0.75$, $p = .39$, $\eta_p^2 = 0.02$, $BF^{10} = .08$, or diagnosis x target x time, $F(1, 48) = 0.24$, $p = .63$, $\eta_p^2 = 0.01$, $BF^{10} = .01$.

There was however a significant target x time interaction, $F(1, 48) = 4.25$, $p = .045$, $\eta_p^2 = 0.81$, $BF^{10} = .02$. However, post-hoc tests found that participants perceived no difference in similarity to the similar-other at baseline *versus* post-self-reflection, $t(49) = 1.64$, $p = .11$, $d = .23$, and no difference in similarity to the average person at baseline *versus* post-self-reflection, $t(49) = 1.21$, $p = .23$, $d = .17$.

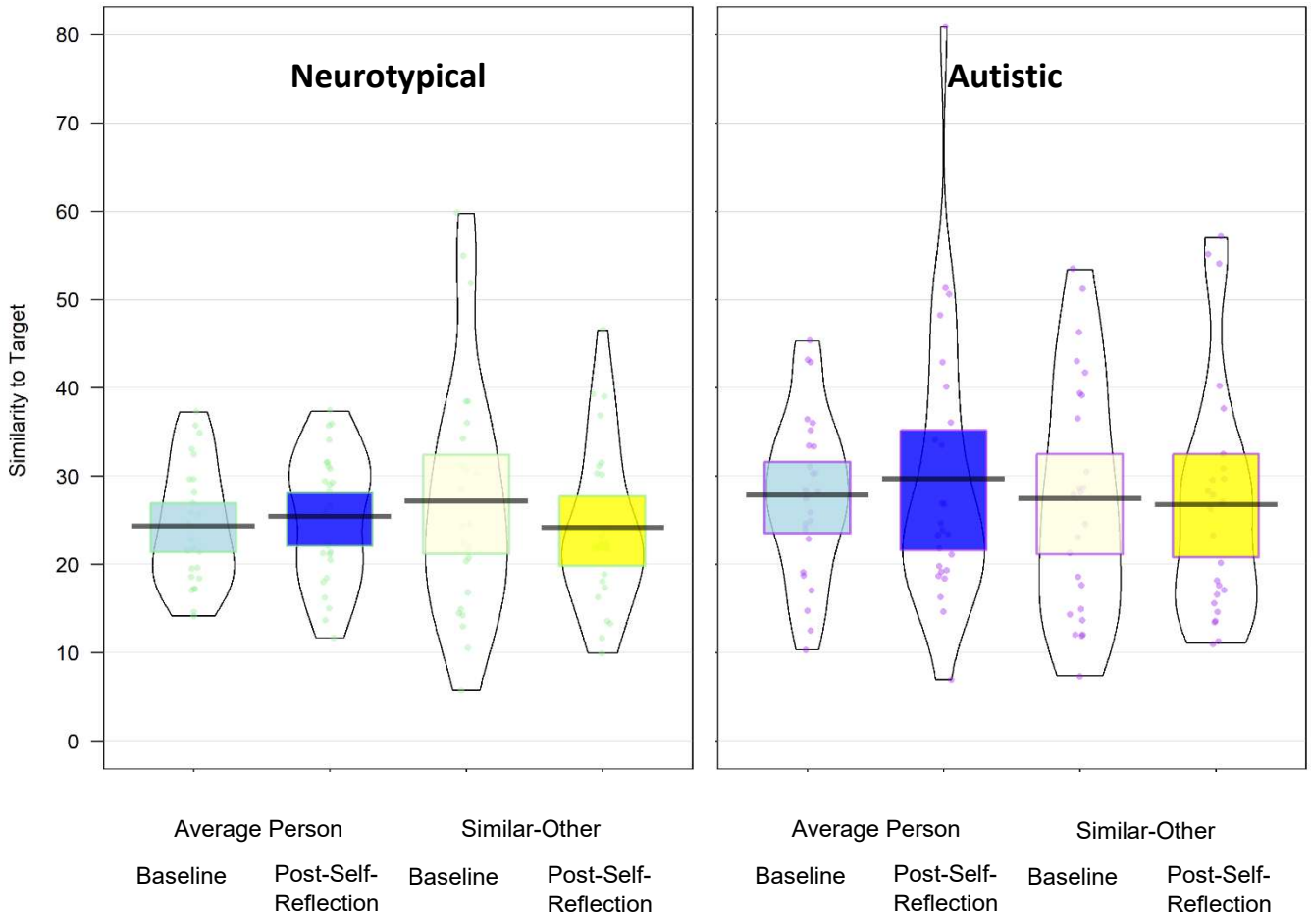


Figure 5.8. Mean similarity to target in each target x time condition in neurotypical (left/green) and autistic (right/pink) groups. Shaded area represents 95% confidence interval.

Association Analyses

Correlation analyses investigated the relations between SRIM (in each target condition) and (i) AQ (ii) animations, (iii) BAI, (iv) TAS-20, (v) CAT-Q, (vi) EQ, and (vii) scene construction across the whole group of participants. See Table 3.

Table 3. Matrix displaying correlations between SIM and cognitive predictors in the total sample.

	Similar-Other SRIM	Average Person SRIM	AQ	animations	BAI	TAS-20	CAT-Q	EQ	Imagination
AQ (48)	-.072	.074	-						
Animations (20)	.232	.331	.318	-					
BAI (n= 50)	-.253	-.003	.464**	-.230	-				
TAS-20 (n= 44)	-.052	.077	.739**	.478*	.408**	-			
CAT-Q (n = 50)	.085	.090	.633**	.211	.230	.495**	-		
EQ (n = 46)	.106	.029	-.413**	-.331	-.261	-.419*	-.169	-	
Imagination (26)	-.273	.056	-.376	.200	-.215	-.407	-.422	.214	-

*p <0.05; **P < 0.01; *** p < 0.001.

There were no significant correlations between similar-other SRIM or average person SRIM and any variables.

DISCUSSION

Experiment 4.3

Replicating Experiment 4.1, and partly replicating Meyer et al. (2019), neurotypical adults displayed SIM, whereby they perceived themselves to be more like the target post-simulation than at baseline. As in Experiment 4.1, this did not differ according to target. This suggests that considering information about the other person (irrespective of target) changed self-knowledge such that it incorporated information about the simulated target. We predicted no between-group difference in a case-control design. In line with this, Experiment 4.3 found that autistic participants also displayed SIM. However, autistic participants perceived themselves to be less like the target than neurotypical participants did at baseline, but not at post-simulation (whereby there was no group difference). This suggests that the ASC group experienced more SIM than the neurotypical group. There was no significant association between size of SIM and autistic traits.

Furthermore, across the whole sample, SIM was negatively associated with ToM and positively associated with imagination: participants who displayed greater average-person SIM and similar-other SIM had lower ToM ability; and participants with greater similar-other SIM also had lower imagination ability.

Experiment 4.4

Findings from Experiment 4.2 suggested that self-reflection changes one's perception of the targets – i.e., SRIM – differently depending on level of similarity to the self. Importantly, the lack of association between SRIM and autistic traits suggested that SRIM is not influenced by autistic tendencies in neurotypical people; therefore we predicted that SRIM would be intact in ASC. In contrast to Experiment 4.2, Experiment 4.4 found no significant SRIM in neurotypical or autistic participants irrespective of target. The contradictory findings might reflect the significantly larger sample size in Experiment 4.2 (N = 162) than Experiment 4.4 (N = 50), and therefore suggests that

SRIM effects are small and/or sensitive to unknown experimental design factors. The lack of SRIM in Experiment 4.4 (at least in the neurotypical group) might therefore reflect insufficient power required to detect the same target x time interaction as was observed in Experiment 4.2.

Based on the findings in Experiment 4.2, we intended to compare the size of SRIM in autistic participants with the SRIM that we expected to observe in neurotypical participants. However, since there was no significant SRIM in the neurotypical group (Experiment 4.4), we cannot infer whether the lack of SRIM in the autistic group reflects typical/atypical SRIM (or simply low power here). Future research is required to determine the reliability and validity of the SRIM observed in neurotypical adults (Experiment 4.2) before we can meaningfully interpret SRIM in a larger sample of autistic adults.

GENERAL DISCUSSION

Across four experiments, the current study aimed to investigate malleability of self- and other-representations in neurotypical and autistic adults. Experiments 4.1 and 4.2 took an individual differences approach by testing whether SIM (Experiment 4.1) and SRIM (Experiment 4.2) are associated with the number of autistic traits in a large sample of neurotypical adults. Experiments 4.3 and 4.4 took a pre-registered case-control approach by testing whether level of SIM (Experiment 4.3) and SRIM (Experiment 4.4) differs between neurotypical and autistic adults. To measure SIM, in Experiments 4.1 and 4.3, we replicated Meyer et al. (2019). Participants were asked to: 1) recall personal information (e.g., “remember a time when you heard a dog barking; how did you feel?”), 2) simulate other people in similar contexts (e.g., “how would [the average person/your friend] feel when hearing a dog barking?”), and 3) recall the personal information again (e.g., “remember the time when you heard a dog barking; how did you feel?”). To measure SRIM, in Experiments 4.2 and 4.4, we changed the order of events, such that participants first rated other people, then the self, then other people again.

Experiment 4.1 replicated Meyer et al. (2019) by showing that neurotypical adults perceived themselves to be more like the target post-simulation than at baseline. This shows that simulating other people changes the perceived self-concept to be more like the target (i.e. SIM). SIM was positively correlated with autistic traits only in the average person condition, which suggests that in a case-control design, relative to neurotypical participants, autistic participants would show higher SIM in the average person condition. Experiment 4.3 found that both groups perceived themselves to be more like the target post-simulation than at baseline, and that this effect was greater in autistic than neurotypical participants. However, in contrast to our predictions based on Meyer et al. (2019) and Experiment 4.1, SIM was not modulated by similarity to the target, and was not correlated with autistic traits. Experiment 4.2 showed SRIM for the first time using this paradigm, which differed according to the target. Neurotypical adults perceived the similar-other to be more like the self post-self-reflection than at baseline, but they perceived the average person to be more like the self at baseline than post-self-reflection. This SRIM was not modulated by autistic traits in either target condition, suggesting that there would be no group difference in a case-control design. Experiment 4.4 found no significant SRIM irrespective of target, diagnosis, or their interaction.

Simulation Induced Malleability of Self-Representation

The SIM observed in neurotypical participants (Experiment 4.1 and 4.3) *partly* replicates previous research that has used the same task in neurotypical adults (Meyer et al., 2019; Rubin-McGregor et al., 2022). This supports the proposal that pre-existing episodic knowledge is susceptible to modification whereby post-event information becomes incorporated into pre-existing knowledge. More specifically in this case, simulation of other people – which involves retrieving self-information and simultaneously considering information about the other person – can change self-knowledge, causing it to incorporate information about the simulated person (i.e., *My friend likes dogs a lot*). As such, simulation causes the self to be perceived as more like the simulated other (i.e., *I like dogs a*

lot) than pre-simulation (i.e., *I like dogs a bit*). This motivation or tendency to achieve uniformity with others might serve an evolutionary function by facilitating social cohesion (Festinger, 1954).

However, the finding that in neurotypical participants (Experiment 4.1 and 3) SIM was *not* modulated by similarity to the target (i.e. similar-other vs. average person), contradicts previous research which found that SIM was greater when simulating a similar-other person than the average person (Meyer et al., 2019; Rubin-McGregor et al., 2022). According to the simulation account, similar-others recruit self-referential processing to a greater extent than dissimilar-others (Denny et al., 2012; Mobbs et al., 2009; Jenkins et al., 2008; Rilling et al., 2008). Previous research has shown that the degree of similarity to the self influences neural activation when mentalising about another person (Mitchell et al., 2006; Pfeifer et al., 2009), which suggests that people refer to their own perspective to understand how a similar person might be seeing, feeling, or thinking. Greater SIM was expected for similar-others since simulation relies more on self-knowledge which in turn causes more self-knowledge to become malleable and susceptible to change. This was not supported by the current data, which highlights a high degree of individual variation in similarity effects, and suggests that the effects reported by Meyer et al. (2019) may not be as universal as originally thought.

One possible explanation for why SIM was not modulated by target in our study is that simulation of similar-others did not recruit self-knowledge to any greater extent than simulation of the average person. If so, then simulation would cause a comparable amount self-knowledge to become labile and reconsolidated for both targets. This possibility is unlikely, since in both experiments, participants reported feeling more like the similar-other than the average person in the manipulation check questions. A more likely alternative is that the small effect size (η_p^2 = 0.05) reported by Meyer et al. (2019) and the non-significant target x time interaction in our Experiments 4.1 ad 4.3 reflects an effect of similarity-to-self that is sensitive to experimental context, and that a large sample size is required to control for high variability between individuals. Meyer et al.'s Experiment 4.1 included a comparable sample size (*N* = 185) to our Experiment 4.1 (*N*=163) but was

significantly larger than our Experiment 4.3 (*N*=50). It is likely that effects of similarity are subject to greater individual variance than the overall SIM effect.

Simulation Induced Malleability of Self-Representation in Autism

The more rigid thinking style that is characteristic of ASC might suggest that the self-concept would be less likely to become more like simulated others (e.g., Meyer et al., 2019; Rubin-McGregor et al., 2022). If so, then this might contribute to the apparent diminished social motivation in autistic people (Chevallier et al., 2012; Harrop et al., 2018; Lawrence et al., 2020): a self-concept that is not malleable is less able to attain uniformity and hence social cohesion with other people. However, contradicting this proposal, the finding in Experiment 4.3, that like neurotypical participants, autistic participants also perceived the target to be more like the self post-simulation than at baseline, suggests that participants with ASC did demonstrate SIM. Importantly, the finding that autistic participants perceived themselves to be less like the target than neurotypical participants did at baseline, but did not differ at post-simulation, suggests that the autistic group in fact displayed *more* SIM than the neurotypical group. This suggests that the self-concept is *more* malleable and susceptible to modification in ASC (Kanner, 1943; Lyons & Fitzgeral, 2013). This challenges the diminished social motivation proposal, since autistic people show greater tendency to attempt to attain uniformity with other people, which might facilitate social cohesion (Festinger, 1954).

A greater tendency to attain uniformity and hence social cohesion might reflect autistic people's attempts to camouflage. Camouflaging refers to compensatory/masking strategies which reduce the appearance of one's autistic traits, and hence enables autistic people to better fit in with the current social environment (Hull et al., 2020). This might include pretending to be interested in topics that others are interested in (Hull et al., 2017; Hull, et al., 2020; Livingston & Happé, 2017). Since there was a lack of association between SIM and self-reported frequency of camouflaging tendencies (CAT-Q) in Experiment 4.3, it might appear that SIM does not support/reflect camouflaging. However, tendency to attain uniformity (i.e., SIM) was negatively associated with ToM

and imagination: participants who showed higher SIM had poorer ToM and imagination ability. This might reflect a greater need to camouflage in those with lower ToM and imagination ability, and hence presumably more difficulty fitting in.

The greater SIM and enhanced self *malleability* in ASC might mean that autistic people's self concept is less strongly consolidated in memory and so may shift more flexibly than is observed in neurotypical people. This is consistent with clinical observations and self-reports that autistic people have difficulty maintaining a constant self-concept and a tendency to take on different identities over time (Kanner, 1943; Lyons & Fitzgeral, 2013). This can be explained by atypicalities in autobiographical memory. Autobiographical information – involving both projection into one's past and future – enables construction and updating of one's self-identity whilst also maintaining a sense of being a coherent person through time (Tulving, 1985; Conway, 2005; Bluck & Alea, 2002). Autistic people have impaired autobiographical memory, characterised by reduced retrieval, specificity, coherence, and phenomenological richness of episodic memories (Anger et al., 2019; Bowler et al., 2000; Brown et al., 2012; Bruck et al., 2007; Gaigg et al., 2015; Lind, 2010), and a gradual forgetting of personal knowledge (Bon et al., 2012). Thus, reduced ability or tendency to reflect on and rehearse past experiences might cause diminished consolidation of episodic information into semantic knowledge (Bon et al., 2012). Consequently, the less consolidated memories of one's self-identity might be *more malleable* in ASC, thus causing greater SIM.

Another possibility is that greater SIM in ASC reflects use of *more self-knowledge* for simulation of other people. Research suggests that autistic people have diminished ability to distinguish self-representation from other representations (Lombardo et al., 2010; Nijhof et al., 2018; Simantov et al., 2021). If self-other distinction is diminished, then autistic people will erroneously recruit too much self-knowledge for simulation of others, and so will not rate the other distinctly from the self. Consequently, recruitment of *more self-knowledge* would make m*ore self-knowledge malleable* and susceptible to modulation. However, since autistic people perceived

greater self-other distinction at baseline than neurotypical participants, they should recruit self-knowledge *less* during simulation (Denny et al., 2012; Mobbs et al., 2009; Rilling et al., 2008; Jenkins et al., 2008). This question of *what* the greater SIM reflects – i.e., *more self-knowledge* used, or same amount but *more malleable* self-knowledge – cannot be directly answered by the current data but is an important avenue for future research.

Implications, Limitations, and Future Directions

The results suggest that autistic adults demonstrate greater SIM than neurotypical adults, thus it would be important to investigate whether this change in self-knowledge/concept is short-term or long-term. Rubin-McGregor et al. (2021) proposed that SIM might reflect at least two possible outcomes for the self. Work on memory malleability predicts that simulation alters the content of self-knowledge (as discussed earlier in this paper), and thus SIM should cause long-term change. However, if SIM changes only what is active in *working self-concept* – whereby different aspects of the self (e.g., someone might have a conception of themselves as an employee and a mother) can be differentially active depending on context (e.g., at work, attributes of the employee-self are heightened, whereby the person considers themselves to be more professional and less caring than they would be when around their children; Hinkley & Anderson, 1996; Markus & Wurf, 1987) – then it should induce only short-term change (Rubin-McGregor et al., 2022). In neurotypical adults, Rubin-McGregor et al. (2021) found that SIM induced long-term change in the self, thus supporting the first proposal. However, duration of change from SIM has not been investigated in ASC; since self-awareness is atypical – and perhaps diminished – in autistic people, it is possible that SIM does not have the same long-term effect in autistic people as it does in neurotypical people.

SIM provides insight into how the self develops over time. Changing over time can have positive impacts, such as self-improvement and learning how to appropriately adapt to the demands and expectations of different environments. However, too much changing of the self might also lead to poorer mental health, because it violates the self-consistency standard (Keyes & Ryff, 2000).

Indeed, autistic people often self-report that greater camouflaging predicts exhaustion, isolation and loss of identity and acceptance of self (Bradley et al., 2021). Therefore, encouraging increased acceptance of the self in autistic people is important for mental health.

Finally, the lack of a clear and consistent SRIM effect in Experiments 4.2 and 4.4 suggests that SRIM effects are small and/or sensitive to individual variance between participants and unknown experimental design factors. Additionally, while the lack of association between SRIM and autistic traits in Experiment 4.2 suggested that SRIM would be intact in autistic people this pattern could not be verified in Experiment 4.4 since no significant SRIM was found (likely due to the lower sample size). As such, we cannot conclude whether the lack of SRIM in the autistic participants reflects typical/atypical performance. Future research is needed to recruit larger sample sizes that provide sufficient power to replicate the pattern observed here in Experiment 4.2 and to test for and compare SRIM in neurotypical and autistic participants.

Conclusion

We have presented four experiments that investigated malleability of self and other representations in neurotypical and autistic adults. In Experiment 4.1, SIM was observed in a large neurotypical sample. This SIM did not differ according to similarity to the target, but was associated with autistic traits only in the average person condition. Experiment 4.3 found greater SIM (irrespective of target) in autistic than neurotypical adults, characterised by greater perceived similarity to the target in autistic than neurotypical participants at baseline but not post-simulation. This challenges the *inflexible* self-concept explanation, and instead supports the overly *flexible* self explanation – whereby autistic people have a less strongly consolidated self-concept, which causes an overly malleable self (Kanner, 1943; Lyons & Fitzgeral, 2013). This might enable them to camouflage in attempt to appear more like neurotypical people, but might also have consequences for mental health. In Experiment 4.2, SRIM was observed in a large neurotypical sample, which differed according to target. SRIM was not modulated by autistic traits in either target condition. However,

since Experiment 4.4 found no significant SRIM in neurotypical participants for either target, we cannot conclude whether the lack of SRIM in the autistic participants reflects typical/atypical performance. Future research is required to determine the reliability and validity of SRIM.

Chapter 6| Motor Self-Other Distinction in Neurotypical and Autistic Adults

Motor Self-Other Distinction in Neurotypical and Autistic Adults

Marchella Smith and Heather J Ferguson

Datasets supporting this paper can be found on the open science framework (https://osf.io/4utsg/?view_only=6bf6a10287864a8cab5d81ae50a12cdb).

ABSTRACT

Previous research has produced inconsistent findings regarding whether imitation-inhibition is typical in autism spectrum conditions. Some research has suggested that the identity of the referent might account for the discrepancy between studies (Oberman et al., 2007; Bird et al., 2007). Therefore, using a stimulus response compatibility paradigm, the current study measured participants ability to distinguish their own motor repertoire from those displayed by other people of differing levels of *similarity* – specifically, the self vs. unfamiliar adult (similar-other) vs. unfamiliar child (dissimilar-other). Experiment 5.1 took an individual differences approach by testing whether behavioural self-bias is associated with the number of autistic traits that neurotypical adults have (N = 59). Experiment 5.2 took a case-control approach by testing whether behavioural self-bias and associated mu/alpha (8-13Hz) and beta (13-35Hz) desynchronisation – which are thought to reflect an index of mirror neuron activity – differs between neurotypical (N = 27) and autistic (N = 30) adults. Typical behavioural imitative and spatial compatibility effects were observed, which did not differ between neurotypical and autistic participants (Experiment 5.2), and were not associated with differences in autistic traits (Experiment 5.1). Associated mu/alpha (8-13Hz) and beta (13-35Hz) desynchronisation (Experiment 5.2) did not suggest consistent differences between groups, but might suggest top-down control deficits in autism. However, there were no effects of referent which might be explained by methodological factors.

The mirror neuron system (MNS) refers to neurons in the inferior frontal gyrus (IFG) and inferior parietal lobule (IPL) that activate both when people perform an action (e.g., emotional facial expression or hand movement), and when they observe a similar action being performed by others (Rizzolatti & Sinigaglia, 2016). The MNS functions by neural connections that map observed actions from one's higher order visual areas (medial temporal gyrus (MTG) and superior temporal sulcus (STS) via the inferior parietal cortex and premotor cortex) to similar actions in our own motor system (visuo-motor/self-other mapping; Hamilton, 2015). The common coding theory (Prinz, 1997) explains that action perception (i.e. watching someone else's actions) and action execution (of our own planned actions) rely on a 'shared representational system' (a common code; Brass & Heyes, 2005) that develops from repeated exposure to – and hence learning of – associations between a motor command (e.g., twisting door handle) and its perceptual consequences (e.g., door opens; Catmur et al., 2009; Cook et al., 2014; Heyes, 2010). Importantly however, the common code does not specify who caused the corresponding motor-representations (i.e., self or other). Consequently, if one fails to distinguish between internally generated (self) and externally triggered (observed) motor representations, then observing another's action would cause overt imitative behaviour – whereby one executes the motor plan of the externally triggered motor representation (Chartrand & Bargh, 1999; Brass & Wiersema, 2021). Thus, existence of shared representations also makes it necessary to distinguish these representations (Brass & Wiersema, 2021). In this paper we investigate people's ability to distinguish their own motor repertoire from those displayed by other people of differing levels of *similarity* – specifically, the self vs. similar-other *vs.* dissimilar-other – and whether this ability differs in people with autism spectrum conditions.

The functional control of the MNS/shared representations has been studied extensively using the stimulus response compatibility (SRC) paradigm (Brass et al., 2000). In this task, an imperative stimulus (e.g., a number or coloured cue) indicates that participants should lift either their index or middle finger as quickly and accurately as possible, whilst simultaneously watching a task-irrelevant on-screen hand move either congruently (i.e., same finger) or incongruently (i.e.,

opposite finger) with one's own pre-specified action. Congruent trials do not require participants to distinguish their own intended actions from the referents': participants' response is a 'quasi-imitative' reaction (Brass et al., 2000). People typically show higher reaction time (RT) and lower accuracy when responding during observation of incongruent movements relative to congruent movements (Brass et al., 2000). This reflects people's tendency to automatically imitate other people's actions, and hence the greater cognitive resources required (thus leading to slower RT and lower accuracy) to inhibit automatic imitation of an incongruent action in favour of the desired goal (pre-specified action). Therefore, the 'interference effect' (difference in RT and accuracy between congruent and incongruent conditions) is a measure of the control of automatic imitation and thus is thought to be an index of self-other distinction. Importantly, SRC tasks are considered a measure of automatic rather than intentional imitation because movement compatibility is task-irrelevant in that participants are neither instructed nor intend to imitate, and in half of the trials (incongruent trials) imitation hinders task performance (Bird et al., 2007).

Neuroimaging during the SRC paradigm has identified the anterior fronto-medial cortex (aFMC) and the TPJ as being associated with self-other control (Brass et al., 2009; Spengler et al., 2009; Marsh et al., 2016). Furthermore, this self-other control at the action level relies on the same brain regions – posterior TPJ – as for self-other control at the mental state level (Decety & Lamm, 2006; Spengler et al., 2009). For example, Santiesteban et al. (2012) found that excitatory transcranial direct stimulation (tDCS) of the rTPJ enhanced self- and inhibited other-representations in the SRC paradigm (i.e., enhanced performance), but also enhanced other- and inhibited self-representations in a perspective-taking task (i.e., enhance performance; see also evidence from repetitive transcranial magnetic stimulation studies; Costa et al., 2008; Young et al., 2010). This pattern suggests that enhanced TPJ activity is associated with up-regulation of the control of self and other representations which might reflect a domain-specific low-level process supporting a range of socio-cognitive tasks (Sowden & Catmur, 2015).

Mirror neuron activity has also been investigated using EEG (Arnstein et al., 2011; Cannon et al., 2014; Oberman et al., 2008). At rest, neurons in the sensorimotor cortex fire in synchrony; however, during action execution and action observation, they become desynchronised, reflecting cortical activity. This causes desynchronisation of the mu rhythm – an EEG oscillation between 8-13Hz – over sensorimotor areas, which is thought to be an index of mirror neuron activity (Oberman et al., 2008; Perry & Bentin, 2009). There is debate about the distinction between mu rhythm and alpha, since they are composed of the same frequency bands. In this investigation (Experiment 5.2), we distinguished between mu and alpha EEG activity based on their topography: mu originates from central areas overlying the sensorimotor cortex and alpha originates from occipital areas (as recommended by Hobson & Bishop, 2017). Additionally, since mu rhythm appears to have spectral peaks at both 10Hz and 20Hz, we also consider beta oscillations (13-35Hz) as an index of MNS activity (as in Brunsdon et al., 2019, 2020).

Shared Representations in Autism and Similarity to Referent

The broken mirror theory (J.H.G. Williams et al., 2001) claims that a dysfunctional mirror system is the primary social difficulty in autism spectrum conditions, which might first cause reduced imitation, and later cause difficulties understanding other people's intentions. However, several studies have investigated automatic imitation in autistic people and found it to be typical (Forbes et al., 2017; Gowen et al., 2008; Press et al., 2010; Sowden et al., 2016; Grecucci et al., 2013) or even enhanced relative to neurotypical people (i.e., greater hyper-imitation/reduced imitation inhibition). For example, Spengler et al. (2010) found that autistic adults demonstrated greater increase in RT and reduction in accuracy during incongruent vs. congruent trials compared to neurotypical adults (Spengler et al., 2010; see also Bird et al., 2007; Schunke et al., 2016). This enhanced interference effect is consistent with clinical features of echolalia and echopraxia in ASC (Grossi et al., 2013), and suggests that impairment is not specific to MNS self-other matching (e.g., Williams et al., 2001), but rather in the top-down modulation or control of MNS (Spengler et al., 2010). Specifically, Spengler et

al. (2010) argue that autistic people have difficulty distinguishing self- from other-representations, which means they are slower to inhibit automatic imitation, and so display hyper- rather than hypo-imitation. Furthermore, in ASC, the interference effect was found to be associated with (i) poorer reciprocal social interaction abilities (Sowden et al., 2016), (ii) decreased behavioural ToM task performance (Spengler et al., 2010), and (iii) reduced medial prefrontal cortex (mPFC) and TPJ (mentalizing regions) activity during this ToM task (Spengler et al., 2010).

One factor that might cause differences in the ability to distinguish the self from others is the identity of the other person (i.e. the referent). The mapping of observed actions to similar actions in our own motor repertoire is thought to be flexible in that it allows some generalisation to comparable stimuli (Hamilton, 2015). Nevertheless, if people use these internal models to simulate other's actions, then the strength of self-other matching – and hence subsequent automatic imitative response – should depend on the *similarity* between an observed (referent's) and stored (own) motor actions (Shanks & Darby, 1998). As such, adults should display a self- and similarity-bias whereby the interference effect follows a pattern of self > similar-other (adult) > dissimilar-other (child), reflecting greater difficulty inhibiting automatic imitation and hence self-other distinction for people who are more similar to the self (Chaminade & Cheng, 2009; Gowen & Poliakoff, 2012; Press, 2011).

Oberman et al. (2008) tested effects of other *familiarity* on imitation and found that in neurotypical children, mu suppression (i.e., MNS activity) is equal during observation of hand actions performed by the self, a familiar-other (e.g., parent or sibling) and an unfamiliar-other (i.e. stranger) person (self = familiar = unfamiliar). This implies that self-other distinction demands are equal between the three conditions. In contrast, among autistic children, mu suppression followed a pattern of self = familiar-other > unfamiliar-other. Since greater mu suppression puts greater demands on cognitive self-other distinction, the difficulty of distinguishing self-other should also follow a pattern of self = familiar > unfamiliar other. Importantly, mu suppression did not differ

between the neurotypical and autistic groups for the self and familiar-other conditions, but was significantly diminished in the unfamiliar-other condition for autistic children (which did not differ from baseline). Oberman et al.'s (2008) findings indicate a group difference in the perception of familiarity to targets. Neurotypical people might consider all people (both familiar and unfamiliar others) to be "like me" (i.e., self-relevant) and therefore as socially relevant referents; in contrast, people with ASC might only consider familiar-others (including themselves) as socially relevant "like me" referents (Bird et al., 2007). If so, then self-other distinction demands may be reduced in ASC compared to neurotypical participants under unfamiliar-other conditions.

In a related study, Press et al. (2005) tested the effect of other *similarity* in the SRC paradigm. Participants performed a pre-specified action (open or close their right hand) during simultaneous observation of hand actions performed by *either* a similar-other (e.g., human hand) or dissimilar-other (e.g., non-human: robot, wooden hand, lights). Neurotypical participants showed an automatic imitation effect whereby they responded faster on compatible than incompatible trials – in *both* human and robot conditions, supporting the notion that self-other mapping can be flexibly generalised to dissimilar others. However, this effect was greater during observation of the human (~30 ms; Brass et al., 2001) than the non-human (~15 ms; Brass et al., 2001) hand – i.e., an "animacy bias" in both RT (e.g., Bird et al., 2007; Gowen et al., 2016; Press et al., 2005, 2006, 2007) and kinematic similarity (e.g., Chaminade et al., 2005; Cook et al., 2014; Gowen et al., 2008; Kilner et al., 2003). This animacy effect is thought to be a direct consequence of increased MNS activity in response to observing similar (e.g., human) compared to dissimilar (e.g., non-human) ideomotor representations (Tai et al., 2004; Chaminade et al., 2005; Cook et al., 2014; Sowden & Shah, 2014; Gowen et al., 2008; Kilner et al., 2003).

If self-other distinction is diminished in ASC, then we would expect both hyper-imitation (e.g., Deschrijver et al., 2017; Spengler et al., 2010) and a diminished similarity bias. However, empirical data contradicts this prediction. Bird et al., (2007) found that automatic imitation was

modulated by similarity to the other in both neurotypical and autistic participants; both groups experienced greater interference for human than non-human (robot) referents. In fact, this animacy bias was larger in the ASC group than the neurotypical group in the human condition (replicating previous studies using the SRC paradigm, Deschrijver et al., 2017; Spengler et al., 2010) but did not differ between groups in the non-human condition. This suggests self-other distinction is intact in ASC if the referent is dissimilar to the self (e.g., a robot).

The current study

The current study will investigate self-other distinction for shared representations in relation to *similarity* of the target. To the best of our knowledge the effect of similarity of the referent involving *human stimuli* of differing levels of *similarity* to the self (which make equal ToM demands) as opposed to a human (familiar) and non-human (e.g., robot) comparison (e.g., Press et al., 2005; Bird et al., 2007) has not been investigated using the SRC paradigm. In our experiments, we establish different levels of similarity between the self and other by manipulating whether the depicted other is an adult or child. Age of an avatar has been used to manipulate similarity to self in a perspective-taking task. Ferguson et al. (2018) found that neurotypical adults show reduced interference from a child avatar's perspective compared to an adult avatar's perspective. This effect of similarity was evident on ERP measures as a reduced consistency effect for the child versus adult avatar on the P3 and late frontal slow wave components. This suggests that adults do not automatically compute a dissimilar other's (child) perspective to the same extent that they do for similar others (adults). This suggests that distinguishing one's own mental state from that of a similar other is more difficult than from a dissimilar other since the former has a stronger shared representation with the self.

The current experiments adapt the SRC paradigm to examine effects of *similarity* to the self on imitation – self vs. adult (similar other) vs. child (dissimilar other). Using only human referents allows us to equate ToM demands between referents. We aim to investigate whether the effect of *similarity* of the referent on imitative compatibility is typical in ASC. Experiment 5.1 took an

individual differences approach by testing whether behavioural responses to self- and similarity-bias is associated with the number of autistic traits in a neurotypical adult sample. Experiment 5.2 took a case-control approach by testing whether the behavioural self- and similarity-bias differs between neurotypical and autistic adults. In Experiment 5.2, we supplemented the behavioural measures with EEG to examine associated mu/alpha (8-13Hz) and beta (13-35Hz) desynchronisation – which are thought to index mirror neuron activity.

We expected to replicate basic effects of imitation and imitative inhibition (Brass et al., 2000), with participants showing longer RTs, lower accuracy, and increased mu/alpha and beta desynchronisation when the cued response key is incompatible with the observed finger movement compared to when it is compatible. We also predicted that this imitation compatibility effect would be smaller – characterised by faster RTs, higher accuracy, and greater mu/alpha and beta desynchronisation – when the referent is more similar to the self (self > similar-other [adult] > dissimilar-other [child]) – i.e. a similarity effect. Moreover, we predicted that in Experiment 5.1, the magnitude of this imitative compatibility effect on RT and accuracy would not correlate with autistic traits; and in Experiment 5.2 the magnitude of the imitative compatibility and similarity effects on RT, accuracy and mu/alpha and beta desynchronisation would not differ between neurotypical and autistic participants.

Experiment 5.1| Individual Differences Approach

METHOD

Participants

Experiment 5.1 included a volunteer sample of 54 (12 male; 42 female) University of Kent undergraduate students, aged 18-27 (*M* = 19.22; *SD* = 1.49). An *a priori* power analysis using G*Power (Vinet & Zhedanov, 2010) revealed that 12 participants would be enough to detect a referent x imitative compatibility interaction with effect size 1.00 (Bird et al., 2007) and power of .90. The larger sample size of *n*=54 allows for a more representative sample.

Participants had normal or corrected-to-normal vision and self-reported no history of autism, psychiatric or neurological conditions. All participants provided written informed consent and following completion of the experiment, were fully debriefed and received partial course credits for their participation. Ethical approval was provided by School of Psychology Research Ethics Committee at University of Kent.

Design

This was a 3 (Referent: Self/Adult/Child) x 2 (Imitative compatibility: compatible/incompatible) repeated measures design.

Materials and Stimuli

Imitation-inhibition task

A stimulus-response compatibility paradigm (Brass, Bekkering, Wohlschlager, & Prinz, 2000) was used to measure the effects of referent on imitative compatibility (see Figure 6.1). At the start of each trial, participants were required to hold down keys “n” and “m” with their right index and middle fingers respectively, in order for the trial to start. Next, there was a 500ms blank screen followed by a 500ms fixation cross in the centre of the screen. Then an image of a referent was shown with direct gaze and resting hand position for a variable interval (randomly varies between 800ms and 1200ms) to prevent anticipatory responding. After this time, the referent’s hand image changed to induce apparent motion, showing either the index or middle finger lifting. This was made up of three static images (presented for 34ms each) indicating the progression from a resting hand position to lifted finger position (e.g., two intermediate index finger lift position images, and one maximum index finger lift position image). 80ms after the onset of the first intermediate movement, either an orange or a purple dot (response cue) appeared in the centre between the index and middle finger, signalling that the participant should either lift their own index (orange circle) or middle (purple circle) finger as quickly and as accurately as possible.

Therefore, there were two main types of trials: (i) *compatible trials*, in which the on-screen finger movement was consistent with the participant's finger lift cued by the coloured dot (e.g., the onscreen hand lifts index finger and an orange dot is presented, or the onscreen hand lifts middle finger and the purple circle is presented); and (ii) *incompatible trials*, in which the on-screen finger movement was inconsistent with the participant's finger lift cued by the coloured dot (e.g., the onscreen hand lifts middle finger and the orange dot is presented, or the onscreen hand lifts index finger and the purple dot is presented). Furthermore, the face behind the onscreen hand either depicted the self, a similar-other (adult stranger, gender-matched with the participant) or a dissimilar-other (child stranger, gender-matched with the participant).

Throughout the task, participants were required to hold down keys "n" and "m" with their right index and middle finger respectively. The key release when participants lifted their finger indicated their response. Reaction times were recorded from the onset of the response cue (purple/orange dot) until the key release (response). After the cue onset, the maximum finger lift position remained on screen for a further 500ms, and then the screen cleared until the key response or the end of the response window (1.92 seconds after cue onset).

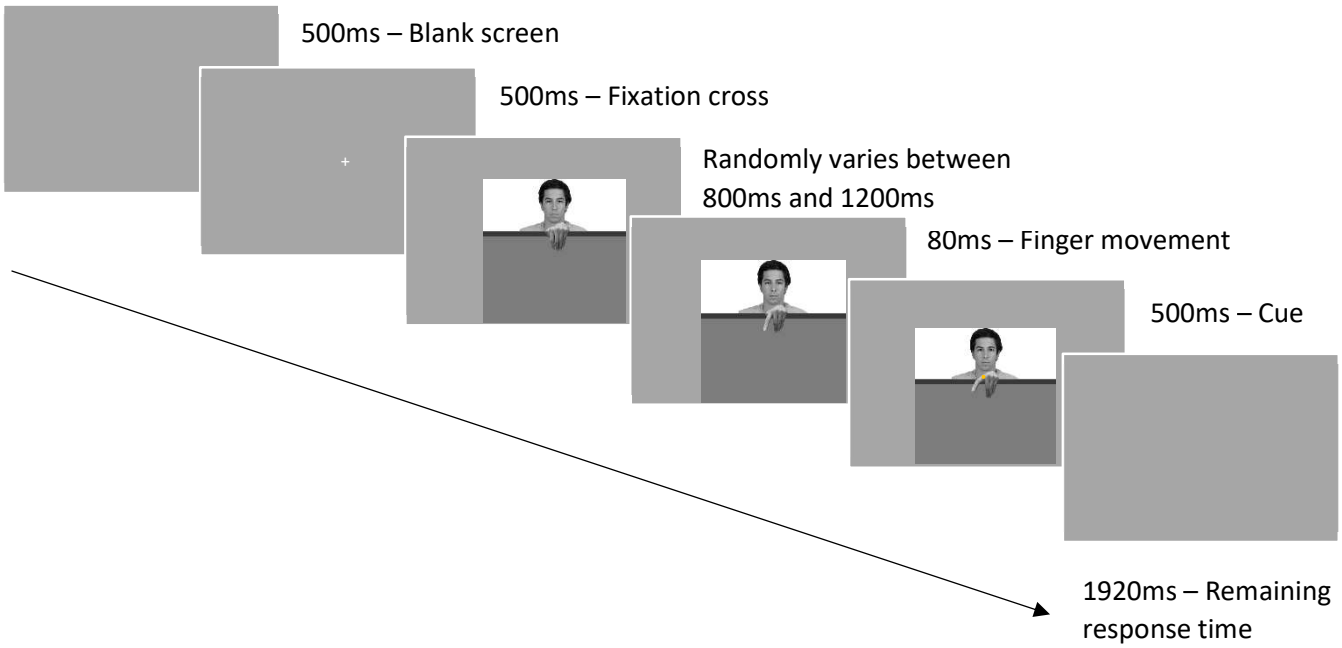


Figure 6.1. Illustration of an example trial.

Participants completed a total of 576 trials in a random order, which consisted of 96 trials per cell of the 3 (Referent: Self/Adult/Child) x 2 (Imitative compatibility: compatible/incompatible) design. This was split into six blocks of 96 trials. The task took approximately 40 minutes to complete.

Measures of Personality and Mood

Two established measures were used to assess autistic traits (Autism-spectrum Quotient (AQ); Baron-Cohen et al., 2001), and camouflaging (Camouflaging Autistic Traits Questionnaire (CAT_Q); Hull et al., 2019); see Supplementary Materials.

Bayesian Analysis

Bayesian analysis was also used to interpret the results (conducted using JASP 0.14.1; JASP Team, 2020), which enables a more graded interpretation than only using *p* values or effect sizes. This estimates the relative strength of the alternative hypothesis over the null hypothesis, or vice versa (e.g., Dienes, 2014; Rouder et al., 2009). For the Bayesian analyses, we adopted the default Cauchy priors as recommended by Wagenmakers *et al.* (2018). Bayes factor (BF^{10}) < 1 is evidence supporting the null hypothesis (<0.33 is firm evidence), and Bayes factors > 3, >10, > 30, and >100 are firm, strong, very strong, and decisive evidence supporting the alternative hypothesis respectively.

RESULTS

Trials were excluded from the RT analysis if they were answered incorrectly or if the RT exceeded three standard deviations from the participant's overall mean or from the mean RT of the group overall (this excluded 10.81% of trials). Two ANOVAs were conducted to analyse the effect of referent and imitative compatibility on RT and accuracy using 3 (Referent: Self/Adult/Child) x 2 (Imitative compatibility: compatible/incompatible) repeated measures ANOVAs. See Figures 6.2 and 6.3.

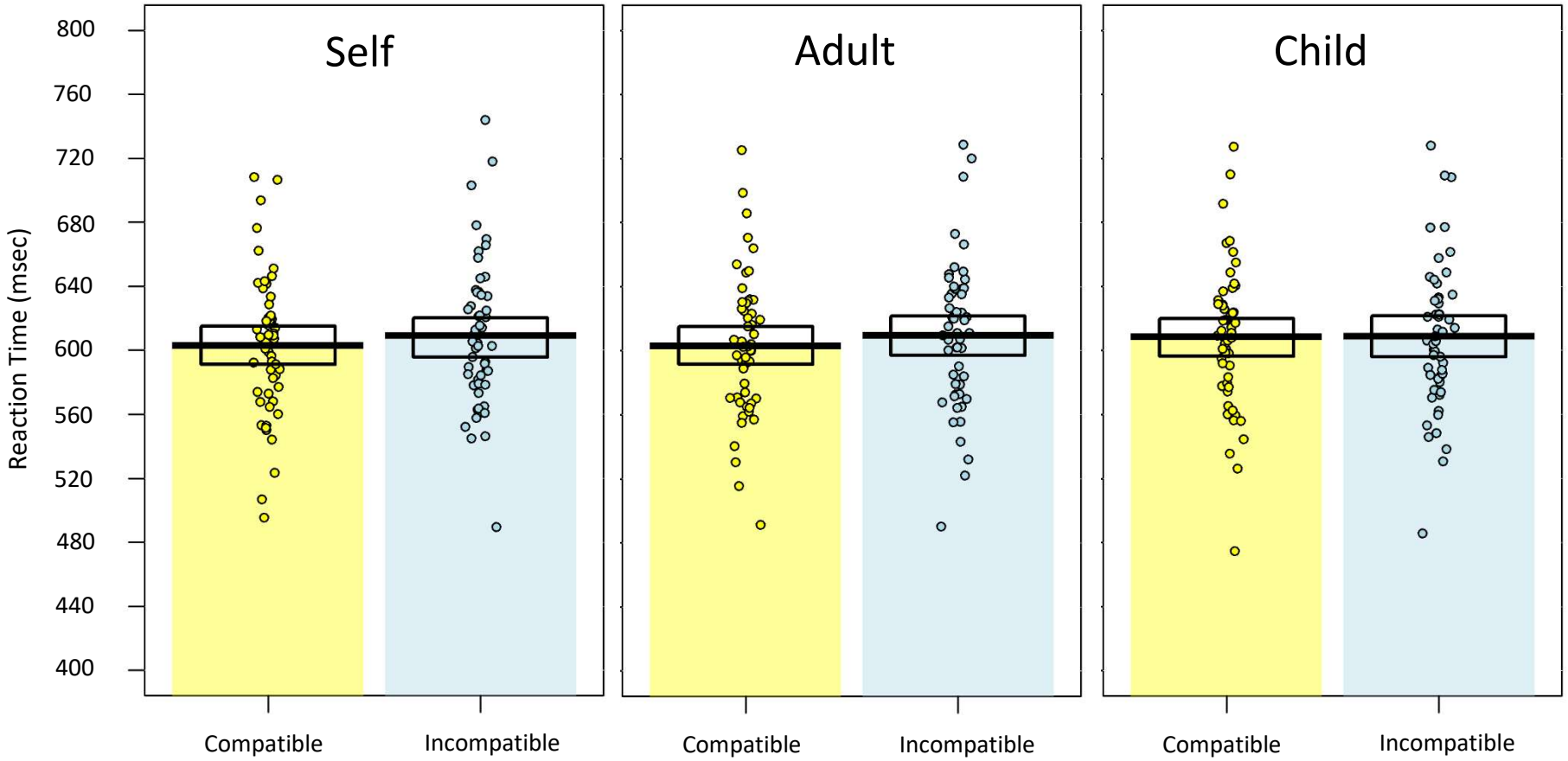


Figure 6.2. Plots depicting mean reaction time (msec) for each condition in Experiment 5.1. Boxed area represents 95% confidence interval.

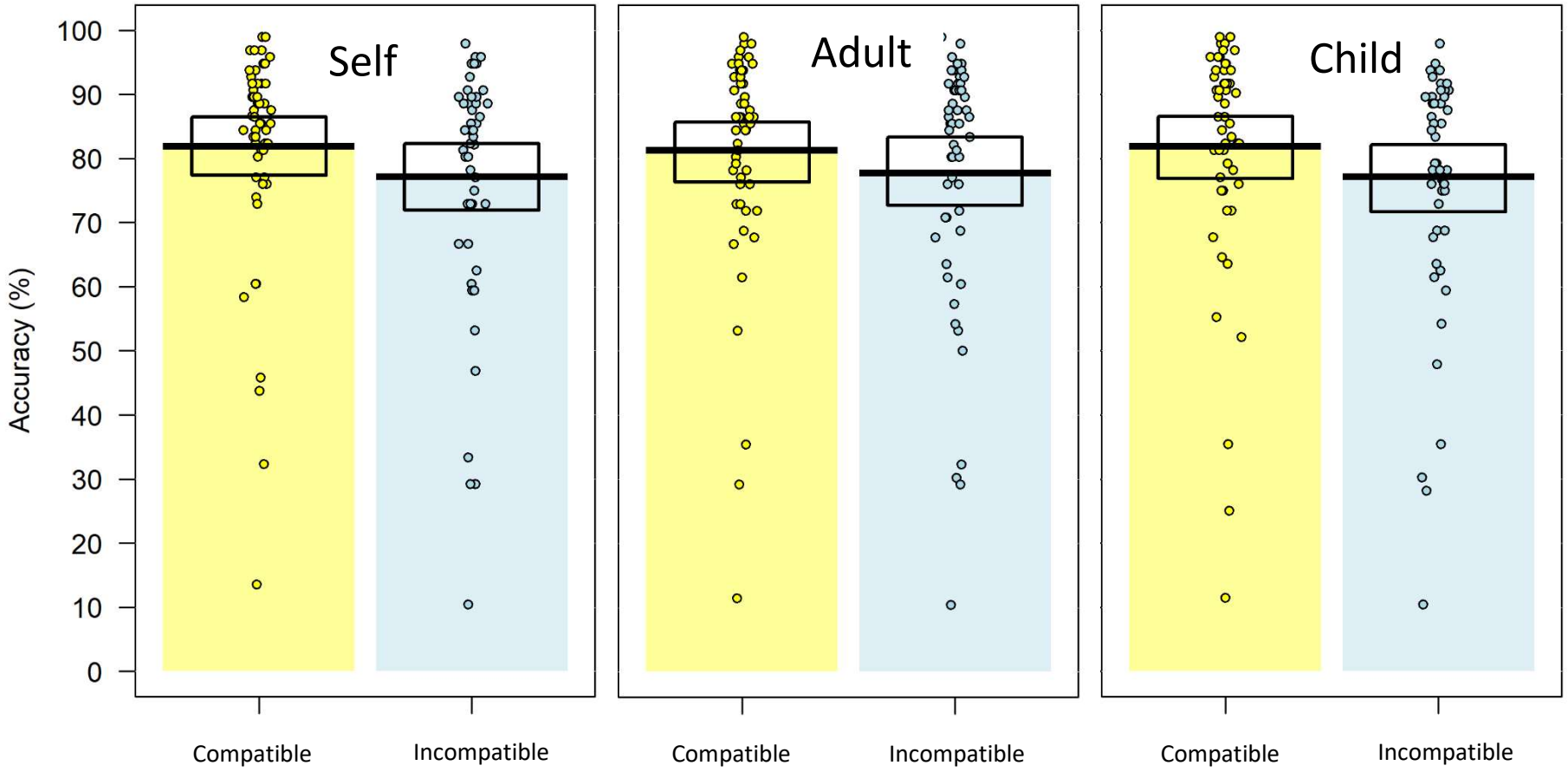


Figure 6.3. Plots depicting mean accuracy (%) for each condition in Experiment 5.1. Boxed area represents 95% confidence interval.

RT: Analysis revealed no significant main effect of imitative compatibility, $F(1, 53) = 3.82$, $p = .06$, $\eta_p^2 = 0.07$, $BF^{10} = 1.00$, or referent, $F(2, 106) = 2.36$, $p = .10$, $\eta_p^2 = 0.04$, $BF^{10} = .11$, and no significant imitative compatibility x referent interaction $F(2, 106) = 3.05$, $p = .05$, $\eta_p^2 = 0.05$, $BF^{10} = 1.20$.

ACCURACY: There was a significant main effect of imitative compatibility, $F(1, 53) = 22.05$, $p < .001$, $\eta_p^2 = 0.29$, $BF^{10} > 100$, whereby accuracy was higher for imitatively compatible than imitatively incompatible trials. There were no other significant main or interaction effects (all Fs < 0.89, ps > .42).

Association Analyses

Compatibility scores were calculated as the mean RT/accuracy on incompatible trials *minus* the mean RT/accuracy on compatible trials. A series of Pearson's correlations revealed that the size of the RT imitative compatibility effect was not significantly correlated with AQ, $r(54) = .01$, $p = .94$, $BF^{10} = 0.17$, or CAT-Q, $r(54) = -.01$, $p = .92$, $BF^{10} = 0.17$. Furthermore, the size of the accuracy imitative compatibility effect was not significantly correlated with AQ, $r(54) = .11$, $p = .43$, $BF^{10} = 0.23$, or CAT-Q, $r(54) = .11$, $p = .44$, $BF^{10} = 0.23$.

Summary

In Experiment 5.1, we found an imitative compatibility effect when using the SRC paradigm. Participants showed better accuracy and RT (higher accuracy and lower RT) in compatible than incompatible trials. Moreover, imitative compatibility was not associated with autistic traits, which supports our prediction that in a case-control sample, imitative compatibility would not differ between neurotypical and autistic participants.

Experiment 5.2|Case-Control Approach

METHOD

Participants

Experiment 5.2 was fully pre-registered on the Open Science Framework (https://osf.io/4utsg/?view_only=6bf6a10287864a8cab5d81ae50a12cdb). Participants were recruited from the Autism Research at Kent participant database. This included 27 NT adults with no history of autism, psychiatric or neurological conditions (self-report) and 31 adults with a clinical diagnosis of ASC: specifically, this was either autistic disorder (N = 6), autism spectrum disorder (N = 4), Asperger's syndrome (N = 10) or not known (N = 11). We note that of the total participants who took part in this experiment, four autistic and three neurotypical were excluded from the ERP analyses due to poor quality of EEG data (resulting in a trial loss of > 40%) or technical problems with EEG recording. Thus, the final ERP sample included 27 autistic participants and 24 neurotypical participants (see Table 1 in Supplementary Materials). All participants were included in the behavioural analyses to maximise power.

Participant groups were matched on age, gender, and Intelligence Quotient (IQ, Wechsler Adult Intelligence Scale, WAIS-II; Wechsler, 1997; Wechsler, 2008) (See Table 1). To assess current autistic characteristics, autistic participants were also assessed on module 4 of the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000) by a trained, research-reliable researcher, and videos were double coded to ensure reliability of scoring (inter-rater reliability was found to be excellent with intraclass correlation of .89).

Participants provided written informed consent and were fully debriefed following completion of the experiment. They received £10 per hour of their participation, plus additional travel expenses. Ethical approval was provided by School of Psychology Research Ethics Committee at University of Kent.

Table 1. Participant characteristics in Experiment 5.2. Note: some tasks have missing data, so statistics are reported on maximal data for each task.

	Autistic (N = 31; 20 male)				Neurotypical (N = 27; 16 male)						
	Mean	SD	Range	N	Mean	SD	Range	N	*t*	*p*	*d*
Age (years)	33.64	11.75	20-59	31	38.93	14.81	19-68	27	1.52	.14	0.40
IQ: Full Scale	108.94	13.85	73-137	31	110.54	9.51	93-133	26	0.50	.62	0.13
IQ: Verbal	105.68	11.13	86-126	31	109.42	10.97	92-129	26	1.27	.21	0.34
IQ: Performance	110.87	18.63	65-143	31	110.5	11.4	89-144	26	0.09	.93	0.02
ADOS: Total	8.2	5.55	1-21	20	-	-	-	-	-	-	-
AQ	30.19	10.56	10-59	31	13.68	5.81	2-38	25	7.00	<.001	1.88
TAS-20	57.71	18.71	5-84	31	43.18	10.89	28-65	22	3.27	.002	0.91
EQ	26.17	11.77	8-55	30	46.83	11.14	26-67	24	6.57	<.001	1.80
BAI	26.13	13.38	6-62	30	11.42	7.8	0-30	26	4.92	<.001	1.32
CAT-Q	117.83	27.17	55-163	30	82.81	17.44	38-109	26	5.64	<.001	1.51

Design

This was a 2 (Group: ASC/NT) × 3 (Referent: Self/Adult/Child) x 2 (Imitative compatibility: compatible/incompatible) mixed design with repeated measures on the last two factors.

Materials and Procedure

The materials and procedure for the motor imitation/inhibition task was identical to that outlined in Experiment 5.1.

Measures of Personality and Mood

Participants completed the same two questionnaire measures (AQ and CAT-Q) as outlined in Experiment 5.1. In addition, participants completed two established measures to assess anxiety (Beck Anxiety Inventory (BAI); Beck et al., 1988), and emotions (Toronto Alexithymia Scale (TAS-20); Bagby et al., 1994); see Supplementary Materials.

Bayesian Analysis

Like Experiment 5.1, Bayesian analysis was also used to interpret the behavioral results (conducted using JASP 0.14.1; JASP Team, 2020).

EEG recording and analysis

A Brain Vision Quickmap amplifier system was used to record continuous EEG activity during the motor imitation task. A 32-channel ActiCap was used to measure EEG over midline electrodes Fz, Cz, Pz, and Oz, over the left hemisphere from electrodes Fp1, F3, F7, FC1, FC5, T7, C3, CP1, CP5, TP9, P3, P7, PO9, O1, and from the homologue electrodes over the right hemisphere. EEG data was referenced online to electrode FCz, and grounded to electrode AFz. EEG and EOG recordings were sampled at a rate of 500 Hz. Electrode impedances were kept at <25 KΩ.

Brain Vision Analyzer 2 software was used to pre-process the data before analysis. Offline, noisy or faulty electrodes were interpolated from surrounding channels (a maximum of 3 channels) and EEG data was re-referenced to the mean of right and left mastoids. EEG and EOG activity was then band-pass filtered (0.1-40 Hz, notch filter at 50Hz), and data containing blinks were corrected using a semi-automatic ocular ICA correction approach.

Next, EEG data was aligned to the onset of the referent's finger movement and segmented into two consecutive epochs: a baseline period from -600ms to 0ms (i.e. showing the face and resting hand), and a test period from 0ms to 600ms (i.e. showing the full finger lift movement and cue onset). Semi-automatic artifact detection software was used to identify and discard segments with non-ocular artifacts (drifts, channel blockings, EEG activity exceeding ± 50μV). Together, these procedures resulted in an average of 70 accepted segments per condition/participant.

A fast fourier transformation (with 10% Hanning window) was applied to retained segments, and power was averaged in the mu/alpha (8-13 Hz) and beta (13-35 Hz) ranges for each participant and condition over the sensorimotor (C3, Cz, C4) and occipital (O1, Oz, O2) electrodes. This allowed us to test whether changes in mu/alpha and beta desynchronization over central sites are distinct from occipital sites. The dependent variable was the percentage change in mu/alpha and beta power for each condition during the 600ms 'test' period from finger movement, which was calculated relative to the 600ms pre-movement baseline "reference" period using the formula: (reference-test/reference) x 100. Negative values indicate alpha and beta synchronisation; positive values indicate desynchronisation.

Community Involvement

There was no community involvement in the design or implementation of the reported study.

RESULTS

Behavioral Analyses

As in Experiment 5.1, trials were excluded from RT analysis if incorrectly answered or if the RT exceeded three standard deviations from the participant's overall mean or from the mean RT of the diagnosis group (this excluded 7.27% of trials in the neurotypical group, and 2.73% of trials in the autistic group). Two ANOVAs were conducted to analyse the effect of group, referent and imitative compatibility on RT and accuracy using 2 (Group: ASC/NT) × 3 (Referent: Self/Adult/Child) x 2 (Imitative compatibility: compatible/incompatible) mixed ANOVAs with repeated measures on the last two factors. See Figures 6.4 and 6.5.

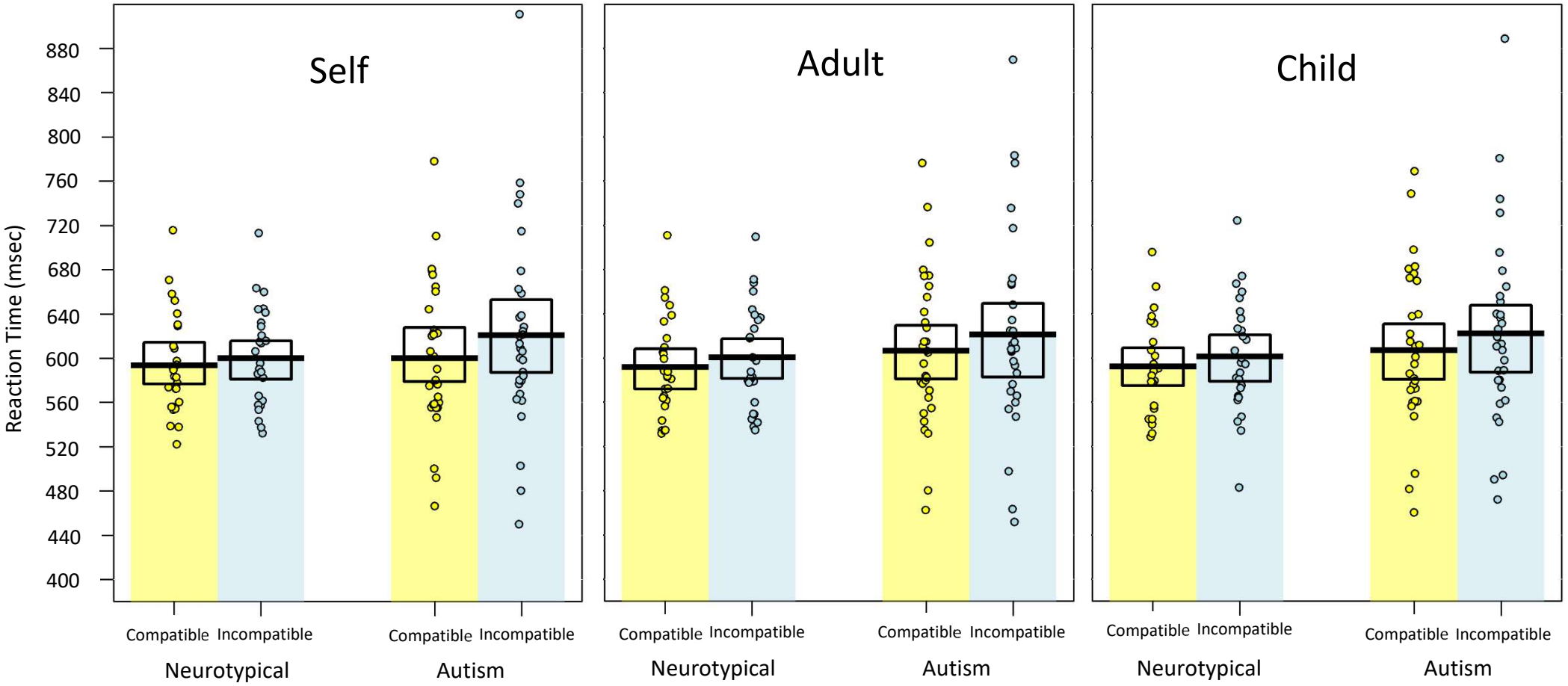


Figure 6.4. Plots depicting mean reaction time (msec) for each condition in Experiment 2. Boxed area represents 95% confidence interval.

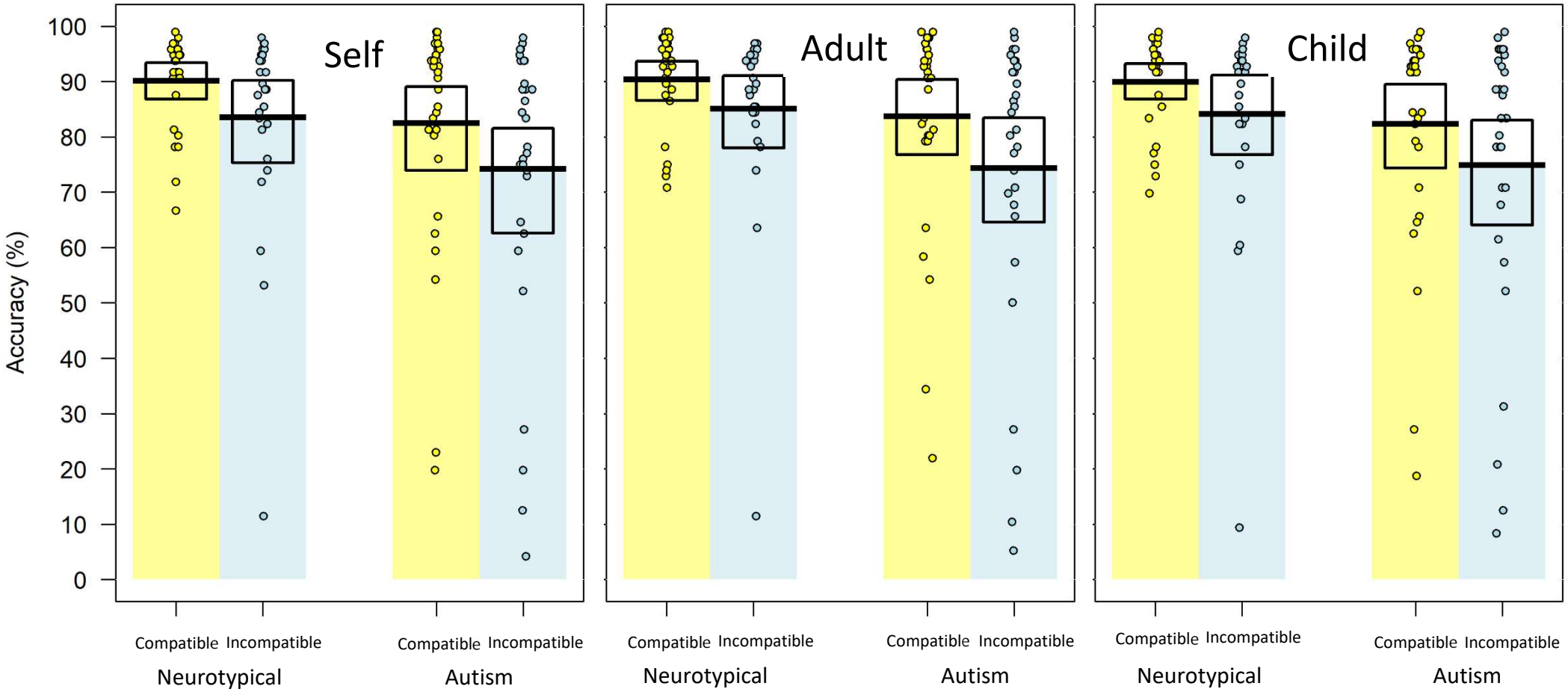


Figure 6.5. Plots depicting mean accuracy (%) for each condition in Experiment 2. Boxed area represents 95% confidence interval.

RT: Analysis revealed a significant main effect imitative compatibility, $F(1, 56) = 9.69$, $p = .003$, $\eta_p^2 = 0.15$, $BF^{10} > 100$, whereby RT was faster for imitatively compatible than imitatively incompatible trials. None of the remaining main effects or interactions reached significance (all Fs < 1.18, ps > .28).

ACCURACY: Analysis revealed a significant main effect of imitative compatibility, $F(1, 56) = 17.37$, $p < .001$, $\eta_p^2 = 0.24$, $BF^{10} > 100$, whereby accuracy was higher for imitatively compatible than imitatively incompatible trials. The remaining main effects and interactions were non-significant (all Fs < 3.24, ps > .08).

Association Analyses

Pearson's correlation analyses revealed that the size of the imitative compatibility effect (RT and accuracy) was not significantly associated with either AQ, TAS-20, CAT-Q, or BAI (all r < .24, ps > .21).

EEG Analyses

Power changes in each frequency band, group and condition are shown in Figures 6.6 and 6.7. Note that due to space constraints, only significant ERP effects that involve group or condition are presented in the text. Full statistical effects for each measure are summarised in the Supplementary Materials, and full data is available on the Open Science Framework (https://osf.io/4utsg/). See Figures 6.6 and 6.7.

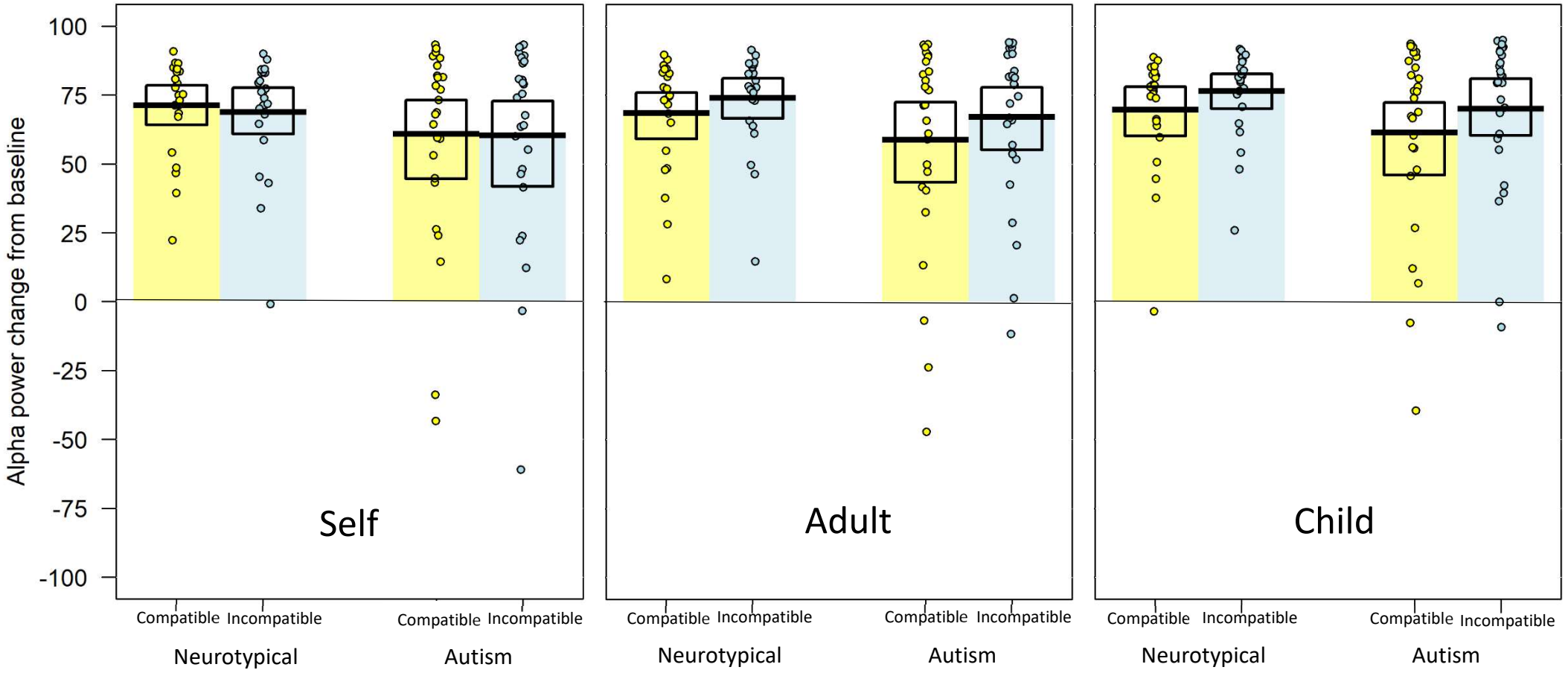


Figure 6.6. Plots depicting alpha power change from baseline over central site for each condition in Experiment 5.2. Boxed area represents 95% confidence interval.

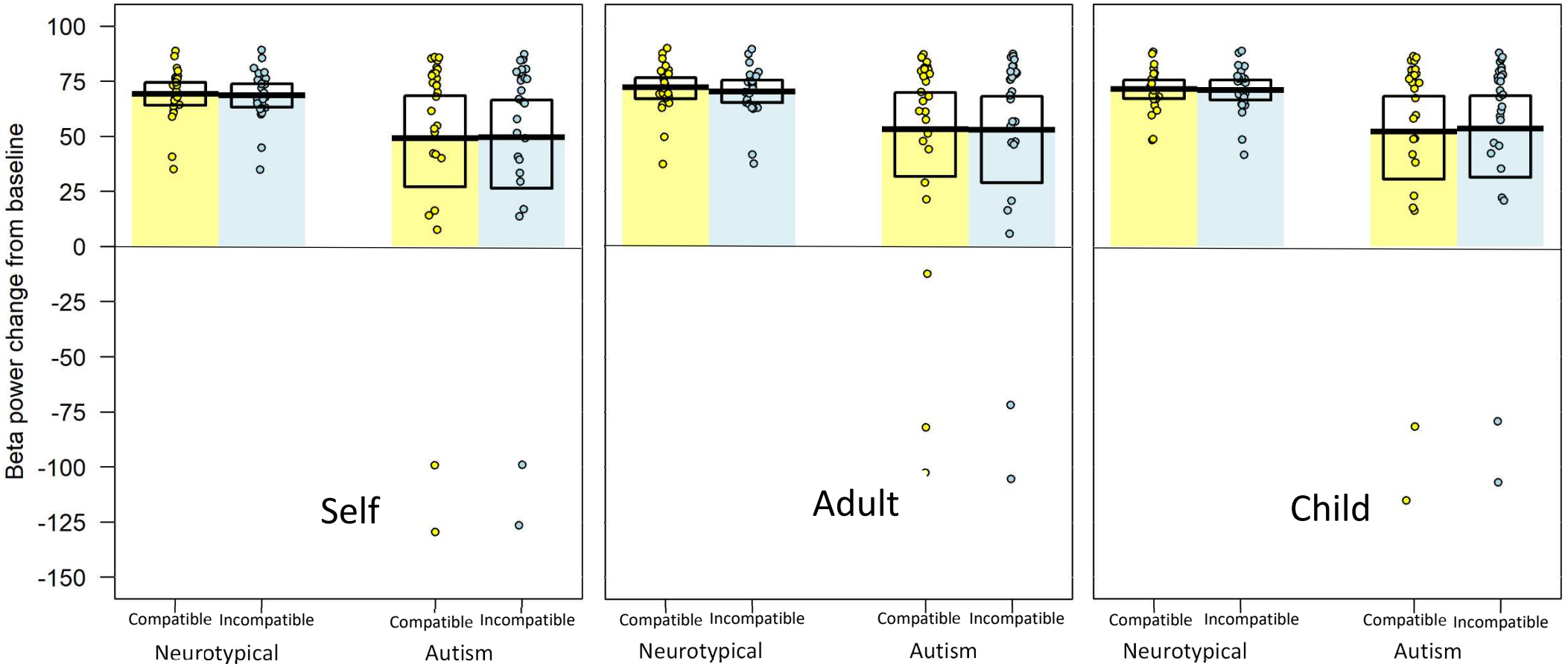


Figure 6.7. Plots depicting beta power change from baseline over central site for each condition in Experiment 5.2. Boxed area represents 95% confidence interval.

Mu/Alpha (8-13Hz): Analysis revealed a significant main effect of imitative compatibility, $F(1, 49) = 37.45$, $p < .001$, $\eta_p^2 = 0.43$, showing greater mu/alpha desynchronisation during incompatible (M = 51.12%) than compatible (M = 43.60%) trials. There were no main or interaction effects involving group, however, numerous 2-way interactions reached significance (site x referent, site x imitative compatibility, referent x imitative compatibility), which were subsumed under a significant 3-way interaction between site x referent x imitative compatibility, $F(2, 98) = 10.38$, $p < .001$, $\eta_p^2 = 0.18$. This 3-way interaction was broken down by testing the referent x imitative compatibility interaction over each site separately.

Over the central site, the referent x imitative interaction was significant, $F(2, 100) = 26.69$, $p < .001$, $\eta_p^2 = 0.35$. Paired samples t-tests revealed that in self trials, mu/alpha desynchronisation did not significantly differ according to imitative compatibility, t(50) = 1.43, $p = .16$, d = .20; in adult trials, mu/alpha desynchronisation was greater during imitatively incompatible (M = 70.31%) than compatible (M = 63.35%) trials, t(50) = 6.00, p <.001, d = .84; in child trials, mu/alpha desynchronisation was also greater during imitatively incompatible (M = 73.05%) than compatible (M = 65.24%) trials, t(50) = 7.61, p <.001, d = 1.07. Further paired sample t-tests revealed that the size of the imitative compatibility effect (imitative incompatible minus imitatively compatible) significantly differed between the self (M = -1.36%) and adult (M = 6.96%) trials, t(50) = 5.62, p <.001, d = 0.79, and self and child (M = 7.82%) trials, t(50) = 5.88, p <.001, d = 0.78; there was no significant difference in the size of the imitation compatibility effect between the adult and child trials, t(50) = 0.91, p = .37, d = 0.13.

Over the occipital site, the referent x imitative compatibility interaction was also significant, $F(2, 100) = 22.37$, $p < .001$, $\eta_p^2 = 0.31$. Paired samples t-tests revealed that in self trials, mu/alpha desynchronisation was greater during imitatively compatible (M = 28.97%) than incompatible (M = 25.62%) trials, t(50) = 2.42, p = .02, d = .34; in adult trials, mu/alpha desynchronisation was greater during imitatively incompatible (M = 36.10%) than compatible (M = 18.92%) trials, t(50) = 6.07, p

<.001, d = .85, and in child trials, mu/alpha desynchronisation was also greater during imitatively incompatible (M = 35.68%) than compatible (M = 17.68) trials, t(50) = 4.62, p <.001, d = .65. Further paired sample t-tests revealed that the size of the imitative compatibility effect (imitative incompatible minus imitatively compatible) significantly differed between the self (M = -3.36%) and adult (M = 17.18%) trials, t(50) = 6.48, p <.001, d = 0.91, and self and child (M = 18.00%) trials, t(50) = 5.00, p <.001, d = 0.65; there was no significant difference in the size of the imitation compatibility effect between the adult and child trials, t(50) = 0.29, p = .78, d = 0.04.

Beta (13-35Hz): Analysis revealed a significant main effect of group, $F(1, 49) = 4.46$, $p = .04$, $\eta_p^2 = 0.08$, showing greater power in the neurotypical (M = 52.59%) than the ASC (M = 13.04%) participants. There was a significant referent x imitative compatibility effect, which was subsumed under a significant 3-way interaction between site x referent x imitative compatibility, $F(2, 98) = 41.21$, $p < .001$, $\eta_p^2 = 0.46$. The 3-way interaction was broken down by looking at the referent x imitative compatibility interaction over each site separately.

Over the central site, the referent x imitative compatibility interaction was not significant, $F(2, 100) = 2.00$, $p = .14$, $\eta_p^2 = 0.04$. Over the occipital site, the referent x imitative compatibility interaction was significant, $F(2, 100) = 35.02$, $p < .001$, $\eta_p^2 = 0.41$. Paired samples t-tests revealed that in self trials, beta desynchronisation was greater during imitatively incompatible (M = 4.54) than compatible (M = -4.76) trials, t(50) = 2.18, p =.03, d = .31; there was no difference in beta desynchronisation according to imitative compatibility in the adult, t(50) = 0.82, p =.42, d = 0.11, or child, t(50) = 0.67, p = .51, d = .09, trials.

GENERAL DISCUSSION

The current study aimed to investigate self-bias at the motor level in ASC. Using an SRC paradigm, this measured participants ability to distinguish their own motor repertoire from those displayed by

other people of differing levels of *similarity* – specifically, the self vs. unfamiliar adult (similar-other) vs. unfamiliar child (dissimilar-other). Experiment 5.1 took an individual differences approach by testing whether behavioural self-bias is associated with the number of autistic traits that neurotypical adults have (N = 59). Experiment 5.2 took a case-control approach by testing whether behavioural self-bias and associated mu/alpha (8-13Hz) and beta (13-35Hz) desynchronisation – which are thought to reflect an index of mirror neuron activity – differs between neurotypical (N = 27) and autistic (N = 30) adults.

Analysis of behavioral performance found a main effect of imitative compatibility, whereby accuracy (Experiments 5.1 and 5.2) and RT (Experiment 5.2 only) were better (higher accuracy and faster RT) in compatible than incompatible trials. In Experiment 5.2, this was also reflected in greater mu/alpha desynchronisation during imitatively incompatible than compatible trials (adult and child conditions only), but no imitative x referent effect on beta power. This mu/alpha desynchronisation imitative compatibility effect reinforces the use of EEG measures to indicate sensorimotor activity reflecting imitation.

This imitative compatibility effect reflects people's tendency to execute the motor plan of an externally triggered (i.e. observed) motor representation (Brass & Wiersema, 2021; Chartrand & Bargh, 1999), and hence the greater cognitive resources required (thus leading to slower RT and lower accuracy) to inhibit mirror neuron activity and subsequent automatic imitation of an incongruent action in favour of the desired goal (pre-specified action). Importantly, since the size of the behavioral imitative compatibility effect did not differ between diagnostic groups (Experiment 5.2), or differ according to autistic traits (Experiments 5.1 and 5.2), this suggests that the control of automatic imitation and hence self-other distinction in this task is intact in ASC. This is supported by the EEG evidence which found that the effect of imitative compatibility on alpha and beta power did not differ between groups. Comparable imitation effects on power between groups suggests normal mu suppression – indicative of intact MNS activity (challenging the MNS dysfunction hypothesis,

Williams et al., 2001) – during the observation of imitatively congruent/incongruent hand actions (however see later discussion about main effect of group in beta desynchronisation), and hence comparable self-other distinction demands between groups.

Intact imitative compatibility replicates some previous research which has found typical automatic imitation in autistic people (Forbes et al., 2017; Gowen et al., 2008; Press et al., 2010; Sowden et al., 2016; Grecucci et al., 2013), but is inconsistent with others which have found enhanced automatic imitation relative to neurotypical people (i.e., greater hyper-imitation/reduced imitation inhibition; Spengler et al., 2010). Some research has suggested that the identity of the referent might account for the discrepancy between studies (Bird et al., 2007; Oberman et al., 2008). Therefore, the current study considered the possibility that differences in the ability to distinguish the self from others might be caused by the *similarity* of the other referent. If people use shared representations to simulate other's actions, then the strength of self-other matching – and hence subsequent automatic imitative response – should depend on the *similarity* between observed (referent's) and stored (own) motor actions (Shanks & Darby, 1998). As such, people should display a self-bias whereby interference effect follows a pattern of self > similar-other (adult) > dissimilar-other (child), reflecting greater difficulty inhibiting automatic imitation and hence self-other distinction for people more similar to the self (Chaminade & Cheng, 2009; Gowen & Poliakoff, 2012; Press, 2011). However, challenging our prediction the current study found that, at the behavioural level (Experiments 5.1 and 5.2) imitative compatibility effects were not modulated by identity of the referent. This might either suggest that 1) referent similarity does not modulate imitative compatibility, or 2) the task is not sensitive to differences between referents.

In relation to the first possibility, evidence of modulation of imitation by group membership/similarity has not been consistently found in previous research. On the one hand Marsh et al. (2016) found a compatibility effect in both spatial and imitative components; however only spatial compatibility was modulated by group membership/similarity. Furthermore, several

studies using the SRC paradigm have investigated how other social factors (e.g., power and status, Farmer et al., 2016) might influence compatibility effects, and have often found no effect of referent, or have not dissociated between imitative and spatial components (e.g., Farmer et al., 2016, Exp. 1-3), thus making results hard to interpret. On the other hand, some research has supported the similarity hypothesis: in tasks that dissociated between spatial and imitative effects, automatic imitation has been found to be enhanced for human vs. non-human referents (Bird et al., 2007; Press et al., 2005, 2006, 2007), and Northern-European-looking vs. Mediterranean-looking referents in a participant sample that was assumed to be mainly of European decent (Trilla et al., 2020). To the best of our knowledge, these are the only studies to have supported the suggestion that *similarity* of the referent modulates automatic imitation.

The latter studies challenge the current findings (particularly in the neurotypical participants – Experiments 5.1 and 5.2). As discussed in the introduction, the human vs. non-human task (Press et al., 2005, 2006, 2007; Bird et al., 2007) might be confounded by animacy bias. However, the discrepancy with Trilla et al.'s (2020) study is harder to understand. The key difference between the studies is that Trilla et al.'s (2020) included a gaze factor, whereby the referent either directed or averted gaze to the participant prior to demonstrating the finger movement. This gaze factor might encourage processing/attention to the referent's identity. The current study provided no incentive to attend to the referent's identity, so attention might have been focused on the hand area (where the cue was also positioned) which was inducive of task performance. Consequently, neglecting the identity of the referent might have caused the lack of imitation modulation. Indeed, EEG (Experiment 5.2) found that there was no difference in mu/alpha or beta desynchronisation according to the main effect of referent. This is in line with previous research finding that the *familiarity* of the referent does not modulate mu/alpha desynchronisation in neurotypical people (Oberman et al., 2008), however, we expected to find an effect of *similarity* of the referent (Press et al., 2005). Furthermore, when considering the modulation of imitative compatibility by referent, there was no imitative compatibility effect in the self condition, but in both the adult and child

conditions, greater power desynchronisation (mu/alpha, but not beta) was observed for imitatively incompatible than compatible trials; the size of this compatibility effect did not differ between adult and child conditions (but the self significantly differed from both the adult and child). Therefore, the EEG findings suggest that similarity of the referent was not consistently processed/detected.

Importantly, the inconsistent support for the effect of referent similarity on imitative compatibility in the neurotypical comparison group, means that no conclusions can be made about the lack of referent x imitative compatibility effect in the autistic group – i.e. we cannot conclude that imitation is not modulated by similarity of the referent in autistic participants. Methodological factors such as the exclusion of a gaze factor, which might have prevented the expected referent effect in neurotypical participants might have similarly prevented a referent effect in autistic participants. Therefore, future research is needed to determine the conditions in which the SRC task is sensitive to referent *similarity* effects, and referent effects in general, in neurotypical participants, before comparisons and subsequent conclusions can be made about these effects in autistic participants.

Finally, there was an overall/main effect of group on beta-band activity, whereby beta power was lower in the autistic than the neurotypical group across conditions. This decreased beta power (and negative correlation with autistic traits in autistic participants) has previously been attributed to top-down control deficits in ASC (Moliadze et al., 2020). As such, perhaps overall desynchronisation (i.e. across a broad temporal window of cognitive processing) isn't the best measure to capture the online/automatic imitation responses between conditions/groups. A time-frequency approach (i.e. that tests how power changes over time, synced to the movement onset) might be better for future reanalysis of this data.

Conclusion

Overall, we found that behavioural imitative compatibility effect – characterised by faster RT and higher accuracy for compatible than incompatible trials – did not differ between neurotypical and autistic participants (Experiment 5.2), and was not associated with differences in autistic traits (Experiment 5.1). Associated mu/alpha (8-13Hz) and beta (13-35Hz) desynchronisation (Experiment 5.2) – which are thought to reflect an index of mirror neuron activity – did not suggest consistent differences between groups (there were no between group differences in Mu/alpha, and some small hints of differences on beta), but might reflect top-down control deficits in ASC. Furthermore, we considered whether similarity of the referent might account for the discrepancy in previous research findings regarding imitation-inhibition ability in ASC. However, we failed to observe any effect of referent similarity in behavioural performance (Experiments 5.1 and 5.2) and observed only weak evidence of this in EEG data. Methodological factors might have caused referent identity to not be consistently processed/detected. Importantly, the lack of referent effect in neurotypical participants, means that no comparison and subsequent conclusion can be made about the lack of referent effect in autistic participants. Future research is needed to determine the conditions in which referent characteristic modulates automatic imitation in the SRC paradigm.

Chapter 7| Discussion

In this thesis, we have tested the finding that the self has a significant influence on human cognition across different processing domains (memory, perception, attention, action). For example, research often finds *self-biases* in cognition – the phenomena whereby people show enhanced cognitive processing for information processed in relation to the self, than for information processed in relation to someone else, or in other ways. Given that the self influences cognition, it follows that any individual with an impoverished self-concept should have diminished self-bias capacity/profile. Potentially, this is highly relevant for our understanding of autism spectrum conditions (ASC). ASC is characterised by atypical development of both the understanding of self (meta-cognition) and of other people (mindreading; Frith & De Vignemont, 2005; Lind et al., 2018; White et al., 2014; Williams, 2008). One possible reason for this might be that self-other distinction is impaired – e.g., people are processing self-referential information similarly to other-referential information. If so, then there should be both impaired understanding of the "duality of self" – the appropriate context-dependent realization that one is similar to, yet different from others (i.e. impaired ToM and/or self-awareness; Brewer, 1991) – *and* impaired self-bias (Lombardo & Baron-Cohen, 2010). However, empirical evidence for impaired self-bias in ASC is inconsistent, and importantly, as of yet, no theory or explanation has been put forward that fully accounts for the discrepancy.

The Introduction (Chapter 1) reviewed previous research relevant to this topic. On the one hand, research using the trait-memory paradigm previously found impaired SRE in autistic children, and adolescents/adults (Henderson et al., 2009; Burrows et al., 2017; Lombardo et al., 2007). Across these three early studies of self-bias, the mean weighted effect size (Cohen's d = 0.65) for the difference in size of the SRE between autistic and neurotypical participants, suggested a moderate-to-large reduction of this effect in ASC (Lind et al., 2020). This impaired self-referential cognition in the trait-memory paradigm was taken as evidence of at least a diminished – if not "absent" – sense of self in ASC (Scott, 2004): if psychological self-awareness and associated extended self-concept is diminished, then the self should not structure memory in the same way that it does in most neurotypical people. On the other hand, later research found no difference in the magnitude of the

self-bias between neurotypical and autistic adults in a perceptual association task (Williams et al., 2018; Zhao et al., 2018). This led researchers to suggest that the discrepant findings regarding self-awareness ability might reflect an intact first-order representation of self (I-self; James, 1890), which is likely to be required in the perceptual association task, but an impaired second-order representation of self (Me-self; James, 1890), which is likely to be required in the trait-memory paradigm (Williams et al., 2018). This proposal has been supported by recent evidence finding typical ERP associates of self-other distinction at early (N2 ERP component; Nijhof et al., 2022) but not later (P3 ERP component; Cygan et al., 2014; Nijhof et al., 2018; Nowicka et al., 2016) stages of processing. If so, then in an incidental encoding version of the self-referential trait memory paradigm (Cunningham et al., 2014; Turk et al., 2008), requiring only first-order representation of self, self-bias should be intact in ASC. However, Lind et al. (2020) found that adults and children with ASC showed an intact SRE in *both* an incidental encoding version of the trait memory paradigm *and* when replicating the classic trait-memory paradigm, requiring second-order representation of self. Participant group demographics were matched to those of Henderson et al. (2009) and Lombardo et al. (2007) to allow for direct comparison of results, meaning that basic differences in the samples cannot account for the different patterns of results.

There is a growing amount of evidence suggesting that self-bias might in fact be intact in ASC (e.g., Nijhof et al., 2022; Nijhof & Bird, 2019; Lind et al., 2020; Williams et al., 2018), contrary to earlier conclusions (e.g., Lombardo et al., 2010; Henderson et al., 2009; Burrows et al., 2017; Grisdale et al., 2014). This is supported by the introduction (Chapter 1), which analysed data from six papers (including 10 experiments/conditions) and revealed no significant difference in magnitude of self-bias between neurotypical and autistic people based on the current evidence in the field. It might be that this area of research has been affected by the "file drawer effect" (e.g., Lind et al. 2020), and subsequently, premature conclusions have been made based on limited evidence. Indeed, the introduction meta-analysis (Chapter 1) demonstrated how little research has addressed these questions directly. Therefore, across five experimental chapters (a total of 11 experiments; six case

control studies and five studies testing individual differences in autistic traits), this thesis investigated typical and atypical development of self-referential cognition by analysing evidence of self-biases in neurotypical adults and adults with autism spectrum conditions (ASC)[20]. This aimed to gain a better understanding self-awareness/self-referential cognition in ASC, and furthermore, use evidence from ASC as a model for understanding self-awareness/self-referential cognition in the neurotypical population. We used behavioural tasks that replicate the classic paradigms/effects (e.g. the ownership effect in Chapter 2, and novel behavioural tasks (e.g. Chapter 3 and 5). Additionally for two experiments, we used ERP/EEG to measure the online mechanisms underlying self-bias, which allowed us to test online evidence of compensatory mechanisms that might enable autistic adults to mask difficulties with self-other distinction at a behavioural level.

Given the limited previous research in this area, the six experiments reported with case-control samples in the current thesis make a strong contribution to the literature. When these six experiments were added to the meta-analysis in Chapter 1, which then included 12 studies (including 18 experiments/conditions), a standard random effects analysis with these additional studies, estimated an overall effect size for the between-group difference of 0.04 (SE = 0.11; with 95% confidence limits from -0.18 to 0.26), $z = -.33$, $p = 0.74$. Therefore, even after including a further six experiments (and 8 conditions) from the current thesis, the pattern remained unchanged, in that the effect size for the between-group difference in self-bias was not significant. This supports the conclusion that self-processing is not universally impaired in ASC. See Figure 7.1.

This Discussion chapter first summarises the key findings from each experimental chapter, then considers the theoretical and clinical implications of these effects, reflects on potential limitations of the work, and finally provides suggestions for future research.

[20] Importantly, we note that since there is such heterogeneity in people with an autism diagnosis, any case-control experiment cannot reliably determine that *all* people with an autism diagnosis would perform the same way on any given task. Factors including age, gender, verbal ability, intelligence, personality factors etc. can influence differential performance among people with an autism diagnosis.

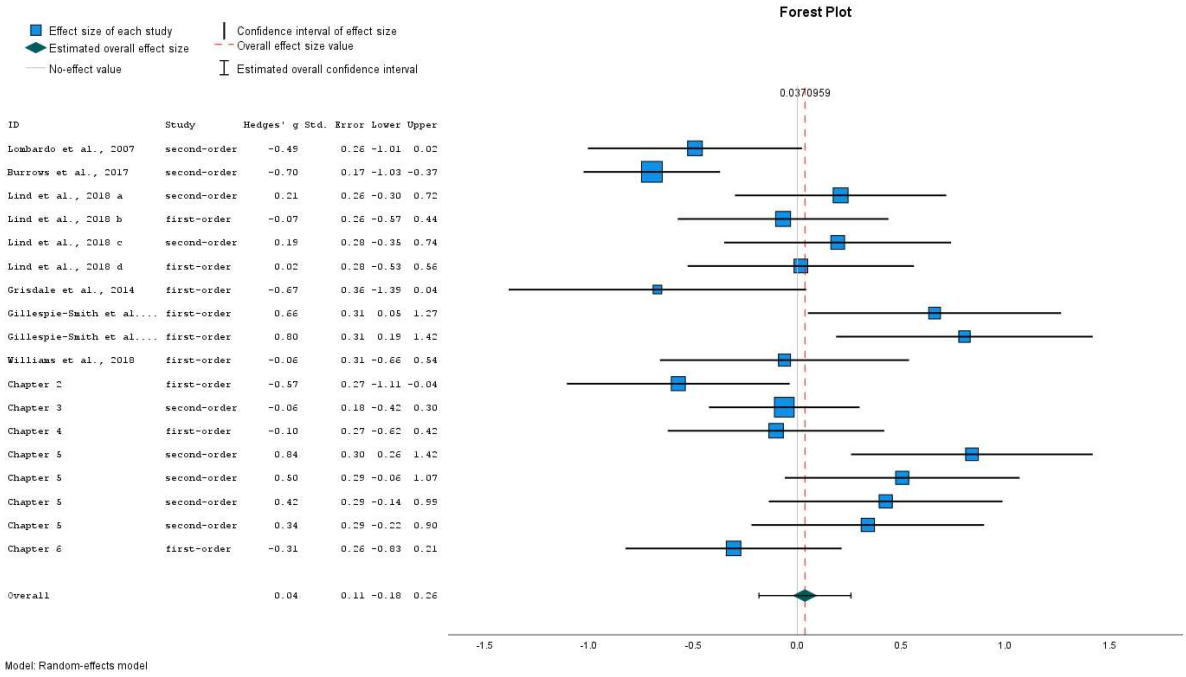


Figure 7.1. Forest plot depicting between-group difference in self-bias including previous research and the six case-control experiments from the current thesis. Effect sizes for each experiment are marked (blue square) along the x-axis (whiskers depict 95% confidence intervals). The size of each blue mark is proportional to the weight that each study has in determining the RE model combined effect size estimate (green diamond shape, whiskers depict 95% confidence interval for the estimate). Statistics are displayed numerically on the left side of the plot (effect size, SE, and confidence intervals). Vertical grey line depicts null-effect size value, and vertical red dashed line depicts observed effect size value.

Summary of Results from Empirical Chapters

Monitoring for self-relevant information is thought to be a perpetual goal of the working self (as suggested by the default state which involves self-oriented thought), thus making self-relevant information more easily retrievable. This suggests that a SRE should be observed not only when information is scaffolded by autobiographical knowledge – i.e., as in the self-referential trait memory

paradigm – but for any information encoded in association with the self (see Buck et al., 2002). It is not straightforward to test this using the classic self-referential trait memory paradigm, because scaffolding by autobiographical knowledge is likely triggered by participants making self-relevant judgements to determine the personal significance of traits. However, in the Ownership Task (Cunningham et al., 2008), self-ownership is completely arbitrary and imagined, so encoding should not elicit autobiographical self-knowledge to any greater extent than encoding of other-owned items. Therefore, the ownership task allows us to test self-bias whilst controlling for difference in scaffolding by autobiographical knowledge between conditions.

Chapter 2 (mental-state level) investigated whether the self-bias in *memory* is impaired in autistic relative to neurotypical adults, by testing the ownership effect (Cunningham et al., 2008) – whereby neurotypical people typically display a memory advantage for items encoded as being (notionally) owned by the self (i.e., self-referentially) relative to those encoded as being (notionally) owned by other people (Amodeo et al., 2021; Nijhof et al., 2020; Rogers et al., 1977; Symons & Johnson, 1997). This shows that self-bias is not limited to encoding contexts where incoming stimuli are associated with self-knowledge. Previous research using the ownership task has found that the ownership effect decreases in neurotypical adults as the number of autistic traits increases, and is absent in autistic adults (Grisdale et al., 2014). This was consistent with findings from other self-bias tasks (trait-memory paradigm) published at the time (Henderson et al., 2009; Lombardo et al., 2007; Burrows et al., 2017). However, more recent research has found an intact ownership effect in autistic children (Gillespie-Smith et al., 2018; Wuyun et al., 2020). It is hard to reconcile these findings in terms of developmental differences (i.e. that psychological self-awareness is intact in childhood but becomes impaired in adulthood). Therefore, the experiments presented in Chapter 2 aimed to determine the reliability of previous findings: we replicated the basic design and procedure, but improved on previous studies that have examined this question (e.g., Grisdale et al., 2014; Gillespie-Smith et al., 2018; Wuyun et al., 2020) by using well-powered sample sizes, pre-registered plans and more sensitive (Bayesian) statistical methods.

The key finding was that whereas neurotypical participants (N = 28) showed a moderate-to-large and statistically significant ownership effect in Experiment 1.2, autistic (N = 28) participants showed a small and statistically non-significant effect (i.e. comparable memory for items encoded in relation to the self and in relation to the experimenter). Memory for other-owned items did not differ between groups, but memory for self-owned items was better in neurotypical than autistic participants – thus suggesting diminished enhancement of self-referential encoding in ASC. Importantly however, Chi-squared analysis revealed that there was no categorical difference between groups in the number of people who did/did not display an ownership effect, which suggests that although the size of the ownership effect was smaller in ASC (as shown by our continuous variable analysis), autistic participants were no less likely than neurotypical participants to display an ownership effect, thus suggesting that the ownership effect is not definitively impaired in autism. Furthermore, the ownership effect was unrelated to the number of autistic traits in either the case-control sample (Experiment 1.2: N = 56) or in a larger sample of neurotypical adults (Experiment 1.1: N = 100). These patterns contrast with the original finding reported by Grisdale et al. (2014) who reported a moderate and statistically significant association between the size of the ownership effect and autistic traits (Exp. 1), and that the ownership effect was more likely in neurotypical (100% *vs.* 71% in the current study) than autistic (44% *vs.* 65% in the current study) participants (Exp. 2). Nevertheless, the findings add weight to the suggestion that this effect is reliably reduced among people with a clinical diagnosis of ASC and suggests that discrepant findings across studies (e.g., Grisdale et al., 2014; Gillespie-Smith et al., 2018; Wuyun et al., 2020, Experiment 1.2) are likely to reflect differences in task format (i.e. two participants *vs.* participant and experimenter; or 6 vs. 74 study items) and sample characteristics (see Chapter 2 Discussion for an analysis of these differences). Therefore, this impaired self-bias in the memory domain suggests that psychological self-awarness might be impaired in ASC.

Episodic memory and semantic knowledge, are thought to provide the building blocks and conceptual framework, respectively, for episodic simulation (M. A. Conway et al., 2016, 2019) – the

ability to mentally simulate or imagine events (Schacter et al., 2012). This is a key component of episodic memory (Maguire, 2001; Rugg et al., 2002; Svoboda et al., 2006)**,** episodic future thinking (Addis et al., 2007; Atance & O'Neill, 2001; Szpunar et al., 2007)**,** spatial navigation (Burgess et al., 2002)**,** theory of mind (Holmes & Frith, 2006) and 'mind-wandering' (Mason et al., 2007). Therefore, in ASC, atypical influence of the self on memory (e.g., Chapter 2) might influence differences in episodic simulation.

Chapter 3 (mental-state level) investigated *scene construction* – the ability to mentally generate (i.e., imagine) and maintain a coherent spatial scene/event in mind (critical for episodic simulation) – using a task in which participants were asked to vividly imagine and describe fictitious scenes (Hassabis et al., 2007). Results showed that scene construction ability was diminished (i.e., lower in quality) in autistic (N = 55) compared to neurotypical (N=62) adults, and was negatively associated with autistic traits. This pattern replicates that reported in previous studies that have used smaller participant samples (Black et al., 2018; Lind et al., 2014b), thus increasing our confidence in the reliability of these findings. This suggests that episodic simulation ability cannot be attributed solely to difficulties in self-projection (the ability to shift one's current state of self into the past or future to mentally simulate an alternative perspective – i.e., mental time travel of the self, or autonoetic consciousness; Tulving, 2005) in autistic adults, and that impairments in scene construction might be the critical factor (Lind et al., 2014b). However, the ability to infer others' mental states (i.e. Theory of Mind; ToM) and to recognise emotions in oneself (i.e. alexithymia) were found to be better predictors of scene construction ability than an autism diagnosis. ToM might allow autistic people to better understand social/narrative conventions – e.g., describing sensory experiences (the frequency of which followed the same pattern in both groups: sight > sound > touch = smell > taste, but importantly, was positively associated with ToM) – required to construct a contextually rich scene, and hence compensate for scene construction impairments.

Chapter 3 also considered whether between-group differences in scene construction ability are influenced by differences in self-referential cognition – specifically, the use of egocentric perspective in scene construction, indexed by first-person pronoun use (e.g., *I, me, my, mine, we, our*). This is important because self-referential cognition is likely to have important implications for contextually rich scene construction (Bicanski & Burgess, 2018; Burgess et al., 2001; Byrne et al., 2007). Furthermore, if self-referential cognition is diminished in ASC (e.g., Chapter 2), then we would expect autistic people's scene constructions to show atypical reliance on first-person perspective. Nevertheless, ASC diagnosis did not influence the frequency of self-reference, thus suggesting that autistic people appropriately distinguished between self and other people's perspectives to support self-referential scene construction. Since scene construction was diminished in ASC, it might be that this first-person perspective was not supportive of contextually rich scene construction to the same extent that it is in neurotypical people. This is in line with the findings in Chapter 2, whereby autistic people were no less likely to show a self-bias, however this self-bias was weaker; and hence in relation to Chapter 3, perhaps less supportive of scene construction. However, this suggestion is purely speculative; the scene construction task was not sensitive/designed to investigate the effectiveness of self-referential cognition on scene construction, and therefore this warrants further investigation.

Unlike findings in the memory domain (e.g., Chapter 2), in the perceptual domain, previous research has found intact behavioural self-bias in ASC when using *unfamiliar/*neutral stimuli (Williams et al., 2018; Zhao et al., 2018). This led research to conclude that the discrepant findings regarding self-awareness ability might reflect different levels of self-awareness required by different tasks (Williams et al., 2018; Nijhof et al., 2019). However, event related potential (ERP) studies have found impaired self-bias in ASC at the P3 ERP component when seeing/hearing *familiar* self vs. other relevant stimuli (Nijhof et al., 2022; Cygan et al., 2014; Nowicka et al., 2016). This suggests that seemingly intact perceptual self-bias at the behavioural level (e.g., Williams et al., 2018; Zhao et al., 2018) might be supported by alternative compensatory underlying mechanisms. Importantly

however, unlike the behavioural tasks (Williams et al., 2018; Zhao et al., 2018), the ERP tasks have used familiar stimuli (own/familiar-other's face/name), and so might be confounded by familiarity effects and personal salience of the stimuli.

Chapter 4 (perceptual level) presented two experiments that aimed to test findings of undiminished perceptual self-bias in ASC when using unfamiliar stimuli (Williams et al., 2018; Zhao et al., 2018). Using a perceptual-association task (Sui et al., 2012), participants were asked to distinguish *novel/unfamiliar* face/label associations of the self from those of other people of differing levels of *familiarity* (e.g., self vs. friend (familiar-other) vs. stranger (unfamiliar-other)). Both in a case-control sample (pre-registered Experiment 3.2: N = 57), and in a sample of neurotypical adults (Experiment 3.1: N = 59), behavioural results showed that participants experienced both a self-bias and a familiarity effect. This was characterised by faster RT and higher accuracy for self-relevant stimuli compared to familiar-other (e.g., Friend) and unfamiliar-other (e.g., Stranger) stimuli, and faster RT and higher accuracy for familiar-other stimuli compared to unfamiliar-other stimuli (this familiarity effect was only significant on RTs in Experiment 3.2). Importantly, the number of autistic traits was unrelated to the magnitude of self-bias and familiarity effects on RT or accuracy (in Experiments 3.1 and 3.2) and there was no between-group difference in the speed and accuracy of self-bias and familiarity effects (Experiment 3.2). Analysis of the ERP indices revealed the real-time processes that underlie these effects. Clear evidence of a self-bias was revealed in both groups, with self trials eliciting larger N1 and P3 components than familiar ('friend') and unfamiliar ('stranger') other trials. This pattern shows that autistic adults can encode novel associations between faces and self/other labels, and are influenced by learned self-biases in the earliest moments of processing (i.e. within 200ms). Moreover, the intact behavioural perceptual self-bias appears to be driven by the same underlying real-time mechanisms between groups. Furthermore, the familiarity effect was largely absent on ERP measures, though the ASC group elicited some modulation of N2 reflecting familiarity conflict.

It is interesting to note that across both behavioural and neural measures, we found strong evidence to support the effect of a self-bias on processing in both groups, despite the use of complex face stimuli that did not genuinely match the self-other label in any condition. That is, while face-label associations were clearly learned in this task, they also presented a major processing conflict in self and familiar other conditions since the depicted face was not of the self or a friend (i.e. the label matched the learned face, but conflicted with a known face). The fact that self-bias and familiarity effects were detected here using these unfamiliar stimuli demonstrates the strength of the associations that participants were able to encode and respond to.

Taken together with previous research finding impaired perceptual self-bias at P3 component when using familiar stimuli (e.g., Nijhof et al., 2022; Cygan et al., 2014; Nowicka et al., 2016), the results in Chapter 4 suggest that the initial binding of stimuli to the self might result in a typical self-bias (e.g., Williams et al., 2018; Zhao et al., 2018), suggesting that personal relevance/salience is detected (e.g., P3 in the current study). Impairment in self-other distinction might only become apparent (at least at the neural level) when testing already consolidated associations (e.g., familiar-stimuli; Nijhof et al., 2022; Cygan et al., 2014; Nowicka et al., 2016), but not novel associations (e.g., unfamiliar stimuli), regardless of whether the task relies on first- or second-order representations of self.

If consolidated associations are diminished in ASC, then we might expect atypical malleability of self and/or other representations in autistic people. **Chapter 5 (mental-state level)** investigated the malleability of representations using a task in which simulating the knowledge state of other people has been found to change self-knowledge, causing one to incorporate information about the simulated person into the self (termed, simulation induced malleability). In addition, we tested whether self-reflection might also change other-knowledge, causing one to incorporate information about the self into the other (termed, self-reflection induced malleability). In two neurotypical adult samples, we found evidence of both simulation induced malleability (Experiment

4.1; N = 163) and self-reflection induced malleability (Experiment 4.2; N = 162). Furthermore, the size of simulation induced malleability was positively associated with the number of autistic traits, but only when the simulated other was considered as the "average person" and not a "similar-other". Furthermore, in two pre-registered case-control experiments, there was greater simulation induced malleability in autistic than neurotypical adults (Experiment 4.3; N = 50) irrespective of similarity of the target, but no self-reflection induced malleability was observed in either diagnosis group (Experiment 4.4; N = 49). The findings that simulation induced malleability is greater in ASC challenges the inflexible self-concept explanation, and instead supports the overly flexible self explanation – whereby autistic people have a less strongly consolidated self-concept, which causes an overly malleable self (Kanner, 1943; Lyons & Fitzgeral, 2013). Consequently, greater simulation induced malleability causes greater reduction in self-other distinction.

The mirror neuron system refers to neurons in the inferior frontal gyrus (IFG) and inferior parietal lobule (IPL) that activate both when people perform an action (e.g., emotional facial expression or hand movement) and when they observe a similar action being performed by others – a shared representation (Rizzolatti & Sinigaglia, 2016). Consequently, if one fails to distinguish between internally generated (self) and externally triggered (observed) motor representations, then observing another's action would cause overt imitative behaviour – whereby one executes the motor plan of the externally triggered motor representation (Chartrand & Bargh, 1999; Brass & Wiersema, 2021). Thus, existence of shared representations also makes it necessary to distinguish these representations (Brass & Wiersema, 2021).

Chapter 6 (action level) measured this motor self-other distinction using a stimulus response compatibility (SRC) paradigm (Brass et al., 2000): an imperative stimulus (e.g., a number or coloured cue) presented on screen indicates that participants should lift either their index or middle finger as quickly and accurately as possible, whilst simultaneously watching a task-irrelevant on-screen hand move either congruently (i.e., same finger) or incongruently (i.e., opposite finger) with

one's own pre-specified action. Previous research has found that the imitative compatibility effect – characterised by slower RTs and lower accuracy during incongruent relative to congruent trials, reflecting the greater cognitive resources required to inhibit automatic imitation of an incongruent action in favour of the desired goal (pre-specified action) – is greater in autistic than neurotypical adults (Spengler et al., 2010; see also Bird et al., 2007; Schunke et al., 2016). This suggests that top-down modulation or control of the mirror neuron system (Spengler et al., 2010) might be impaired in ASC, causing difficulty distinguishing motor representation of self- from other.

However, evidence for this enhanced imitative compatibility effect in ASC is inconsistent (Forbes et al., 2017; Gowen et al., 2008; Press et al., 2010; Sowden et al., 2016; Grecucci et al., 2013; Spengler et al., 2010; Bird et al., 2007; Schunke et al., 2016). One factor that might cause differences in the ability to distinguish the self from others is the identity of the other person (i.e. the referent) (Oberman et al., 2007; Bird et al., 2007). Therefore, Chapter 6 measured participants' ability to distinguish their own motor repertoire from those displayed by other people of differing levels of *similarity* – specifically, the self vs. unfamiliar adult (similar-other) vs. unfamiliar child (dissimilar-other). We found a typical imitative compatibility effect in behavioural performance, which was not associated with the number of autistic traits (Experiments 5.1 and 5.2), and did not differ between neurotypical and autistic participants (Experiment 5.2). Furthermore, associated mu/alpha (8-13Hz) and beta (13-35Hz) desynchronisation – which are thought to reflect an index of mirror neuron activity – also did not reveal consistent differences between groups (Experiment 5.2). However, the key manipulation here, the referent effect, did not show any effect in behavioural performance (Experiment 5.1 and 5.2), and the imitative compatibility effect in mu/alpha and beta did not differ between referents (Experiment 5.2). Methodological factors might have caused referent identity to not be consistently processed/detected (e.g. speeded visual presentation of the stimulus, or size of the avatar on-screen). Importantly, the lack of referent effect in neurotypical participants, means that no comparison and subsequent conclusion can be made about the lack of referent effect in

autistic participants. Further research is needed to determine the conditions under which referent characteristic modulates automatic imitation in the SRC paradigm.

Integrating the current findings with existing literature

Self-other distinction is thought to be dependent on two key neural networks. First, the mirror neuron system shows that action perception (i.e. watching someone else's actions) and action execution (of our own planned actions) rely on a 'shared representational system' (a common code; Brass & Heyes, 2005) that develops during infancy from repeated exposure to – and hence learning of – associations between a motor command (e.g., twisting door handle) and its perceptual consequences (e.g., door opens; Catmur et al., 2009; Cook et al., 2014; Heyes, 2010). Importantly, greater *similarity* between an observed (referent's) and stored (own) motor action (Shanks & Darby, 1998) determines greater strength of mirror neuron self-other matching, and hence lower self-other distinction. The extent of mirroring is thought to support the understanding of action goals/embodied mental states when they are *directly observable* (i.e. do not rely on attribution of unobservable mental state). The MNS is *automatically* activated before the second self-other distinction neural network – the mentalising system. The mentalising system includes the medial prefrontal cortex, precuneus, and of particular importance, the right temporal parietal junction (rTPJ) (Decety & Lamm, 2006; Spengler et al., 2009; Brass et al., 2009; Spengler et al., 2009; Marsh et al., 2016). This supports more *conscious* reasoning about mental states, and the understanding that these might differ from one's own. This is required when visual cues to mental states might be misleading or *not directly observable* (e.g., verbal tasks; understanding false beliefs, deception). Furthermore, the mentalising system is thought to exert control over mirror neuron system shared representations in the process of self-other distinction depending on task demands (Southgate & Hamilton, 2008).

Self-bias provides an index of self-other distinction, as it enables us to test the extent to which the self is processed differentially compared to others. The summary of empirical chapters

demonstrates no support for the suggestion that self-other distinction of the mirror neuron system is impaired in ASC (Chapter 6). However, self-bias was impaired in some tasks (e.g., Chapter 2: Ownership) but not others (e.g., Chapter 4: Perception), and differences between studies might be accounted for by impaired self-other distinction of the mentalising system. The following discussion will evaluate the findings at the perceptual and mental state levels in relation to previous research and theory; this aims to make inferences about functioning of the underlying mechanisms driving the self-bias, and attempts to reconcile previous research. This then considers how/why atypicalities in self-other distinction might occur, and considers the relevance of self-other distinction in the neurotypical population. Finally, this considers the nature of an ASC diagnosis (McPartland, 2019), and what this means for the current research findings.

Perceptual Level

Evidence from the perceptual level, suggests that the ability to bind novel information to the self (i.e. self-referential binding) results in a typical self-bias/preferential processing. According to the Self Attention Network (SAN) (Humphreys & Sui, 2016; Conway & Pleydell-Pearce, 2000), perception of self-relevant stimuli (e.g., self-cues, notionally owned items, traits encoded in relation to the self) triggers a response from the vmPFC. The vmPFC then cascades information that stimuli are self-relevant to the left posterior superior temporal sulcus (LpSTS) attentional system, so that it is primed to respond to self-relevant stimuli. This subsequently influences attentional biases. In Chapter 4 (Experiment 3.2), the underlying real time mechanisms appear to suggest that the ease with which autistic people experience a self-bias in binding is comparable to neurotypical people, and hence is not driven by compensatory mechanisms. This might suggest intact functioning of the SAN in ASC.

Importantly, when the current findings are taken together with previous research, this emphasises that the further consolidation/elaboration of these associations might be atypical in ASC. One of the earliest and strongest predictors of an ASC diagnosis is a failure to orient (i.e., attentional bias) to one's own name during infancy (Elmose, 2016; OSTERLING et al., 2002), thus

suggesting *impaired* binding and/or a preferential orienting to the self. This is associated with impaired P3 ERP component amplitude in autistic adults (Nijhof et al., 2018). Source localization analysis found that this group difference was associated with increased right temporal parietal junction (rTPJ) activity during perception of self-relevant names in neurotypical adults, which was diminished in autistic adults (Nijhof et al., 2018). In contrast in the ASC group, enhanced right inferior frontal gyrus (rIFG) activity was associated with hearing one's own name than when hearing the name of a close-other (Nijhof et al., 2018). Therefore, whilst adults with ASC can clearly distinguish their own name from that of close others, this was not associated with activation in an area typically associated with self-other distinction (rTPJ), but rather was associated with activation in an area typically involved in bottom-up redirecting of attention to external salient stimuli (rIFG; Hampshire et al., 2010; Nijhof et al., 2018). Therefore, Nijhof et al. (2018) speculated that this bottom-up orienting of attention to own name might reflect a compensatory mechanism: individuals with ASC might have learnt from experience that their own name is an important social stimulus (more so than other names) and that attending to it has adaptive value.

Therefore, binding of *novel* stimuli to the self (e.g., Sui et al., 2012; Cunningham & Turk, 2017) does not necessarily mean that it becomes incorporated into the participants self-image in either autistic or neurotypical people (Payne et al., 2017). If the information is not genuinely related to the goals of the working self then these temporary "extensions of self" are not sufficient to interfere with performance when the association is not task relevant (Lui & Sui, 2016). The Social Motivation Theory (Chevallier et al., 2012; Harrop et al., 2018; Lawrence et al., 2020) proposes that some of the social and non-social characteristics of ASC result from a lack of social motivation and subsequent reduced attention toward social stimuli (e.g., faces), and increased attention towards non-social stimuli (e.g., electronics, airplanes, clocks). Indeed, Chapter 2 outlined previous research which found that autistic people elicit a more economically rational strategy in their preferences: they base evaluation more on *what* the stimuli (e.g., object) is (i.e. its material qualities), rather than *who* it is related to (Hartley & Fisher, 2018; Hartley et al., 2020). Therefore, different *preferences*

between neurotypical and autistic people, might cause differences in what information is prioritised/considered goal relevant (as evidenced when stimuli are familiar, and they should form (or have already formed) these associations themselves).

Mental-State Level

The findings from the perceptual level have caused us to question what the classic memory SRE tasks are measuring. In the trait memory paradigm and ownership task (e.g., Chapter 2), a self-bias (e.g., self > other) has been thought to reflect an index of psychological (if considering personality traits/object ownership) and/or physical (if considering physical characteristics/appearance) self-awareness. In other words, enhanced memory for items encoded in relation to the self vs. someone else is thought to reflect our greater *awareness* of our own, relative to other people's characteristics. However, a self-bias in these tasks might not be *dependent* on a second-order representation of self (e.g., "am I *kind*?"). The words (trait-memory paradigm) and objects (ownership task) that are processed in relation to the self vs. other are not necessarily associated with the person that it is being encoded in relation to. As such, the self-bias is driven by automatic up-regulation of attention to *anything* that participants are told to consider in relation to their own stored traits. Therefore, like the perceptual matching task (see Chapter 4), this might only measure the *initial binding* of information to the self, and not *awareness* of one's own vs. others' characteristics. Therefore, whilst the tasks (particularly the trait-memory paradigm) ask participants to take a second-order representation of self, it is unclear whether diminished awareness of self vs. other characteristics would cause impaired self-bias in these tasks (and in the perceptual matching task). Impairment might become more apparent if the tasks are designed to more specifically measure one's already consolidated awareness of self vs. other characteristics; in which case awareness of characteristics (of self and/or other) might be less accurate.

Indeed, Chapter 5 suggested that stored self-knowledge is more susceptible to simulation induced malleability in ASC, which suggests that self-knowledge is less consolidated and less

accurate. Importantly, since simulation induced malleability caused the self to become more, rather than less, similar to the other, this might cause diminished self-other distinction at subsequent stages (e.g., Chapter 2; Henderson et al., 2009; Burrows et al., 2017; Lombardo et al., 2009; Grisdale et al., 2014)[21]. These suggestions are speculative, and so future research is needed to determine the validity of these conclusions.

Chapter 5 indicated a mechanism by which the extent of self-other distinction might become diminished in ASC. Since the autistic participants perceived themselves to be less like the target than neurotypical participants did at baseline, but did not differ at post-simulation, this suggests that 1) the autistic group displayed *more* simulation induced malleability than the neurotypical group, and 2) the extent of self-other distinction is not diminished initially, but might become diminished over time through simulation. Greater simulation induced malleability might be reflective of impaired self-awareness, which might mean that one's self concept is less strongly consolidated in memory (Bowler et al., 2000; Gaigg et al., 2015; Lind, 2010). Autobiographical information – involving projection into one's past and future – enables construction and updating of one's self-identity whilst also maintaining a sense of being a coherent person through time (Tulving, 1985; Conway, 2005; Bluck & Alea, 2002). Autistic people have impaired autobiographical memory (Anger et al., 2019; Bowler et al., 2000; Brown et al., 2012; Bruck et al., 2007; Gaigg et al., 2015; Lind, 2010), and a gradual forgetting of personal knowledge (Bon et al., 2012). Thus, reduced ability or tendency to

[21] Evidence of simulation induced malleability (Chapter 5) might help to explain the discrepancy in previous research using the trait-memory paradigm (Henderson et al., 2009; Burrows et al., 2017; Lombardo et al., 2009; Lind et al., 2020). As discussed in the introduction, these each used slightly different methodology. First, these studies differed in that they *either* presented referent conditions mixed *within* each block (Lind et al., 2020), or in *separate* blocks (Henderson et al., 2009; Lombardo et al., 2007; Burrows et al., 2017). Therefore, future work could also investigate the impact that blocking arrangement has on simulation induced malleability and self-bias. Second, the finding that the size of simulation induced malleability was positively associated with the number of autistic traits, but *only* when the simulated other was considered as the "average person" and not a "similar-other" (Chapter 5; Experiment 4.1), suggests that autistic people might experience differences in SIM relative to neurotypical people (e.g., Chapter 5, Experiment 4.2) but not for all types of referents. Therefore, since the previous studies using the trait memory paradigm used different types of "other" referents (famous/well known people vs. fictional characters; Chapter 5 only investigated "average person" and "similar-other" referents), future work could investigate the impact that this has on SIM and self-bias.

reflect on and rehearse past experiences might cause diminished consolidation of episodic information into semantic knowledge (Bon et al., 2012). Consequently, the less consolidated memories of one's self-identity might be *more malleable* in ASC, thus causing greater simulation induced malleability.

Furthermore, malleability of self might be exacerbated by external factors, such as the need to camouflage. Camouflaging relates to a set of compensatory or masking strategies that reduce the appearance of one's autistic characteristics during social interactions, despite persisting deficit(s) at a cognitive (e.g., social cognition) and/or neurobiological level (Hull et al., 2017; Hull et al., 2020; Livingston & Happé, 2017). The aim of camouflaging is to better fit in with the social environment by presenting a more socially competent persona (Hull et al., 2020), and so may have adaptive benefits. Therefore, greater malleability of self – which reduces self-other distinction – might develop over time to support camouflaging. Importantly, neurotypical people are also susceptible to camouflaging, however, the need to do this might be more pronounced in ASC (hence the greater simulation induced malleability[22]) due to the greater dissonance between their own personality/behaviour, and what the (mainly neurotypical) population expect[23].

Therefore, it appears that self-awareness is more malleable and hence less accurate in ASC (Scott, 2004). The subsequent diminished self-other distinction would cause difficulty defining oneself, which is likely to impact self-referential cognition. Specifically, if autistic people have a tendency to shift the self to be more like other people, then the SAN might be primed to respond to/detect information that is genuinely self-relevant, but perhaps also information that is incorporated from other people. Evidence of diminished self-bias in ASC in some tasks suggests that this might be the case (particularly in more naturalistic tasks/real-world conditions where people choose which information is associated with the self). Consequently, the self-concept would be inaccurately

[22] Note that in Chapter 5 frequency of self-reported camouflaging tendencies was not correlated with SIM.
[23] Also note that the notion of camouflaging suggests that it Is possible for self-awareness to be entirely intact in some people, and for them to only behave in a way that might indicate that it is impaired —i.e. to camouflage.

updated (e.g., erroneous information is bound to the self). Future research is needed to investigate how this changed perception of self, influences social cognition: for example, whether autistic people shift their behaviour – whether that be short-term or long-term – to fit with this new perception of self.

Importance of diminished self-other distinction and its significance to mental wellbeing

Self-other distinction is not fixed, but rather, the degree of appropriate self-other distinction varies according to the specific social situation, and how helpful it might be to resonate or remain detached from another's mental state. For example, on the one hand mentalising *might* benefit from high self-other distinction when it allows one to simultaneously hold two opposing perspectives in mind (e.g., the self and the other). Low self-other distinction would cause misattribution of mental states when these differ to one's own (e.g., false belief tasks). On the other hand, affective empathy *might* benefit from low self-other distinction which allows one to feel another's emotional state (emotional contagion), and hence supports the ability to respond appropriately to the other's needs. Importantly, determining the extent to which self-other distinction might be beneficial in each case is subjective, and dependent on various more subtle social factors – including the identity of the other person, the motivations of each social partner, the environment in which they are in, etc. For example, in some cases excessive resonance with others experiencing negative emotion can cause increased personal distress; and continuation of this is aversive (Eddy, 2022).

Since the degree of required/beneficial self-other distinction is not fixed, constant monitoring of the social environment is required to determine the appropriate extent of self-other distinction for any given context. In the case of ASC, poor mentalising ability and reduced social experience might cause difficulty learning the appropriate extent of self-other distinction based on previous experience and current context. The 'predictive coding' theory of ASC (van Boxtel & Lu, 2013; Van de Cruys et al., 2014) proposes that meta-learning abilities cause difficulty distinguishing

between important and less important prediction errors, which disrupts the ability to contextualise incoming information and make predictions based on previous experience. This diminished reliance on previous experience may hence impair the ability to adaptively tailor self-other distinction to the many variations of social situations. Therefore, the impairment in ASC might not be a global diminishment of self-other distinction, but rather, impaired top-down control of self-other distinction, causing both higher (e.g., baseline condition in Chapter 5) and lower (e.g., post-simulation condition in Chapter 5, also reduced self-bias in Chapter 2) self-other distinction relative to neurotypical people, depending on the task.

Self-other distinction is clearly important for mental wellbeing, and this becomes more apparent when considering the secondary effects of diminished self-other distinction. For example, in the case of increased personal distress from excessive resonance with others, people may subsequently develop cognitive strategies to restore self-other boundaries. This may include attending less to their own internal states to reduce the salience of the distress of self and/or others. However, this may cause secondary effects: emotional blunting (flat affect) and dissociation might amount to, or *appear* to amount to, alexithymia – whereby people have difficulty identifying and expressing/understanding their own emotions – and/or low empathic concern due to avoidance of internal self and/or other states (Eddy, 2022). Therefore, there is irony in that often, secondary symptoms that develop over time as an adaptive strategy because of diminished self-other distinction (whether that be because of an underlying impairment, e.g., perhaps in the case of ASC; or because of external/environmental conditions that do not allow sufficient self-other distinction), reflect greater resonance with others' mental states, but might *appear* to suggest hypo-mentalising (Eddy, 2022).

Limitations

Diagnostic instruments used to diagnose autism aim to reduce the discrepancy between clinicians, and hence increase the reliability of diagnosis. However, this reduces the specificity of criteria, which

can cause them to be over-inclusive (Mottron, 2021) and susceptible to confounding ASC with many other temporary conditions (which might be environmentally driven); most of which are affected by autistic "traits" such as impairment/atypicalities in social communication and reciprocity (Mottron 2021). Therefore, a diagnostic criteria that has evolved to include people more distant from the specific phenotype has contributed to an increase in prevalence of autism diagnosis (Fombonne, 2018; Mottron, 2021). This is important to consider because the greater heterogeneity might cause more type 2 errors in research (Mottron, 2021), whereby the autistic group do not differ in their performance on a cognitive task compared to the neurotypical comparison group, despite the autism phenotype being characterised by impairments in the specified cognitive ability. Indeed, meta-analyses have indicated that effect sizes for between-group differences in case-control designs investigating neurocognitive constructs has decreased by up to 80% (Rødgaard et al., 2019).

Therefore, in the current thesis, it is possible that lack of between group difference in some experiments reflects a type 2 error. All case-control experiments included high-functioning participants with an average or above average IQ, sufficient language abilities to complete the tasks, and many of them did not score in the autism range on the Autism Diagnostic Observation Schedule (Lord et al., 2000). Therefore, their performance is considered reflective of, but perhaps not representative of, all people with an autism diagnosis; approximately 50% of whom have below average IQ and language impairments (e.g., Bertrand et al., 2001; Charman et al., 2011; Fombonne, 2009). Therefore, impairments might be more apparent in those who are more reflective of the autism phenotype, which includes lower-functioning participants. However, it is difficult to conduct cognitive research with lower-functioning participants, since the tasks rely heavily on cognitive control and language abilities; and furthermore, sensitivities around wearing EEG caps might be more significant in lower-functioning people. To minimise the risk of type 2 errors, *a priori* power analyses based on previous research (including similar samples as ours) were conducted to predict the minimum sample size required to detect significant between group differences in task performance. In the experiments in the current thesis, sample sizes were close to, if not exceeded

that suggested by the *a priori* power analysis; however, given the high variability in the ASC population, even larger samples would increase confidence in the validity of the results. However, recruiting sufficient sample sizes is an ongoing challenge in autism research.

The existence of a broader phenotype has led to a dimensional approach, whereby clinical autism is considered the extreme of a continuous distribution of traits in the general population. As such, it is thought to be possible to learn about ASC by relating social-cognitive task performance to the number of ASC traits possessed by people in the general population. This is why we aimed to complement our case control comparison studies with studies that took an individual differences approach in Chapters 2, 4, 5, and 6 (however, we note the importance of checking the validity of these conclusions against case-control samples). Chapters 4 and 6 observed no significant correlations between autistic traits (AQ) and task performance, which might be supportive of the subsequent case-control experiments which also found no between-group difference. However, Chapters 2 and 5 found between group differences in the case-control experiments, despite the individual differences approach showing either no significant or less of an effect, respectively.

There are at least three possible explanations for this discrepancy. First, there is an emerging view that there may be differences in the underlying basis of autistic traits in diagnosed and undiagnosed individuals (Mottron, 2021), and there are findings that suggest distinct phenotypes can be observed (e.g., Chawner et al., 2021). It is therefore theoretically possible that no reliable association between autistic traits and social-cognitive task performance exists in the general population, but that individuals with a clinical diagnosis of ASC nonetheless show a significantly different effect than neurotypical people. Second, experimental designs are developed to, and hence become popular if, they produce robust/replicable effects – i.e., effects that are low in between-subject variance. However, low variance between-subjects makes these tasks unreliable/less effective for detecting individual differences (i.e. a correlational approach) which are *dependent* on this between-subject variance (Hedge et al., 2018). Therefore, it is also possible that some or all the

tasks in this thesis were unsuitable for an individual differences approach. Third, it is possible that the AQ is not a reliable or valid measure of autistic traits. Psychometric evaluations of the AQ have suggested that some factors existing in the scale are uncorrelated, or negatively correlated (Austin, 2005; Kloosterman et al., 2011). If so, then total AQ scores may not be meaningfully interpreted (English et al., 2020), and so any correlation or lack of correlation with task performance similarly lacks validity. To minimise the potential of these issues, *a priori* power analyses based on previous research were conducted to predict the minimum sample size required to detect significant correlations between autistic traits and task performance based on previous research that included similar samples and measures as ours.

Future Directions

Importantly, whilst the current thesis has added significantly to the limited previous research investigating self-referential cognition in ASC, future work, including replication attempts, is necessary to check the reliability and validity of our conclusions. Ideas for future research have been discussed throughout this thesis, in previous chapters. The more research there is on any topic, the more confident we can be about any conclusions that are made. One important next step is to compare self-bias across different cognitive domains, using the same/similar task designs, which would allow us to control for different variables/task demands across the different tasks used here. Research suggests that the magnitude of self-biases across cognitive domains are not correlated, meaning that they are likely to rely on independent mechanisms (Amodeo et al., 2021; Nijhof et al., 2020). So far, research has investigated self-bias in different cognitive domains, but using tasks that vary in other task demands. Therefore, it is hard to determine whether differences in task performance between tasks are caused by different cognitive domains tested (e.g., memory vs. perception), or whether it is caused by other extraneous factors (e.g., first vs. second order representation of self).

Furthermore, whilst the current thesis investigated self- compared to other-referential cognition for others of differing levels of similarity and familiarity to the self, future work should investigate the self-other neurotypical/autism status match/mismatch. Research has suggested that autistic people show better social understanding of other autistic compared to neurotypical people; similarly, neurotypical people show better social understanding of other neurotypical compared to autistic people (Cai et al., 2022; Crompton et al., 2020). In the current thesis, the "other" person might have been assumed to be neurotypical. It might be that autistic people show a more typical self-bias/self-other distinction ability when the other person is autistic or thought to be autistic.

Conclusion

In this thesis, 11 experiments were conducted testing self-referential processing across multiple cognitive domains and measures of performance. Overall, the findings from the current thesis, taken together with previous research, suggest that the underlying ability to bind information to the self results in a typical self-bias in ASC, but that differences between autistic and neurotypical people might become apparent at later stages. Different preferences, which might be particularly driven by differences in social motivation between autistic and neurotypical people, may cause differences in what people consider important, and hence choose to up-regulate attention of. Furthermore, perhaps there is not any difference in the ability to experience a self-bias in perceptual vs. memory tasks (but see Amodeo et al., 2021; Nijhof et al., 2020), but discrepant performance in ASC in previous research might have been driven by the familiarity of the stimuli. Impairment might be more apparent if the tasks are designed to measure one's already consolidated awareness of self vs. others characteristics – i.e. perhaps by using familiar stimuli. Indeed, the findings that self-knowledge is more susceptible to simulation induced malleability in ASC, suggests that self-knowledge is more malleable. Importantly, since simulation induced malleability caused the self to become more, rather than less, similar to the other, this might cause diminished self-other distinction at subsequent stages. This greater simulation induced malleability in ASC might reflect a

consequence of diminished consolidation of self-knowledge, hence causing impaired self-awareness; but both might also be exacerbated by the need to camouflage.

Impaired self-other distinction is likely to impact various social-cognitive abilities in ASC, including impaired ToM and alexithymia. Importantly, since the degree of appropriate self-other distinction varies according to the specific social situation, impaired ToM and social experience in ASC might cause difficulty monitoring the social environment, and hence impair the ability to adaptively tailor self-other distinction to the many variations of social situations. Therefore, the impairment in ASC might not reflect a globally diminished self-other distinction, but rather, impaired top-down control of self-other distinction, which causes both higher and lower self-other distinction relative to neurotypical people, depending on the task.

REFERENCES

Addis, D. R., Wong, A. T., & Schacter, D. L. (2007). Remembering the past and imagining the future: Common and distinct neural substrates during event construction and elaboration. *Neuropsychologia*, *45*(7), 1363–1377. https://doi.org/10.1016/j.neuropsychologia.2006.10.016

Amodeo, L., Wiersema, J. R., Brass, M., & Nijhof, A. D. (2021). A comparison of self-bias measures across cognitive domains. *BMC Psychology*, *9*(1), 132. https://doi.org/10.1186/s40359-021-00639-x

Argo, J. J., Dahl, D. W., & Morales, A. C. (2008). Positive Consumer Contagion: Responses to Attractive Others in a Retail Context. *Journal of Marketing Research*, *45*(6), 690–701. https://doi.org/10.1509/jmkr.45.6.690

Arnstein, D., Cui, F., Keysers, C., Maurits, N. M., & Gazzola, V. (2011). μ-Suppression during Action Observation and Execution Correlates with BOLD in Dorsal Premotor, Inferior Parietal, and SI Cortices. *The Journal of Neuroscience*, *31*(40), 14243–14249. https://doi.org/10.1523/JNEUROSCI.0963-11.2011

Atance, C. M., & O'Neill, D. K. (2001). Episodic future thinking. *Trends in Cognitive Sciences*, *5*(12), 533–539. https://doi.org/10.1016/S1364-6613(00)01804-0

Atance, C. M., & O'Neill, D. K. (2005). The emergence of episodic future thinking in humans. *Learning and Motivation*, *36*(2), 126–144. https://doi.org/10.1016/j.lmot.2005.02.003

Austin, E. J. (2005). Personality correlates of the broader autism phenotype as assessed by the Autism Spectrum Quotient (AQ). *Personality and Individual Differences*, *38*(2), 451–460. https://doi.org/10.1016/j.paid.2004.04.022

Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia scale—I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, *38*(1), 23–32. https://doi.org/10.1016/0022-3999(94)90005-1

Baranek, G. T., David, F. J., Poe, M. D., Stone, W. L., & Watson, L. R. (2006). Sensory Experiences Questionnaire: discriminating sensory features in young children with autism, developmental delays, and typical development. *Journal of Child Psychology and Psychiatry*, *47*(6), 591–601. https://doi.org/10.1111/j.1469-7610.2005.01546.x

Bardi, L., & Brass, M. (2016). The Control of Shared Representations and Social Cognition. In *Shared Representations* (pp. 151–170). Cambridge University Press. https://doi.org/10.1017/CBO9781107279353.009

Baron-Cohen, S. (1989). The Autistic Child's Theory of Mind: a Case of Specific Developmental Delay. *Journal of Child Psychology and Psychiatry*, *30*(2), 285–297. https://doi.org/10.1111/j.1469-7610.1989.tb00241.x

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind" ? *Cognition*, *21*(1), 37–46. https://doi.org/10.1016/0010-0277(85)90022-8

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5–17. https://doi.org/10.1023/a:1005653411471

Barzy, M., Ferguson, H. J., & Williams, D. M. (2020). Perspective influences eye movements during real-life conversation: Mentalising about self versus others in autism. *Autism*, *24*(8), 2153–2165. https://doi.org/10.1177/1362361320936820

Beggan, J. K. (1992). On the social nature of nonsocial perception: The mere ownership effect. *Journal of Personality and Social Psychology*, *62*(2), 229–237. https://doi.org/10.1037/0022-3514.62.2.229

Belk, R. W. (1988). Possessions and the Extended Self. *Journal of Consumer Research*, *15*(2), 139. https://doi.org/10.1086/209154

Benoit, R. G., Gilbert, S. J., Volle, E., & Burgess, P. W. (2010). When I think about me and simulate you: Medial rostral prefrontal cortex and self-referential processes. *NeuroImage*, *50*(3), 1340–1349. https://doi.org/10.1016/j.neuroimage.2009.12.091

Bertrand, J., Mars, A., Boyle, C., Bove, F., Yeargin-Allsopp, M., & Decoufle, P. (2001). Prevalence of Autism in a United States Population: The Brick Township, New Jersey, Investigation. *Pediatrics*, *108*(5), 1155–1161. https://doi.org/10.1542/peds.108.5.1155

Beversdorf, D. Q., Smith, B. W., Crucian, G. P., Anderson, J. M., Keillor, J. M., Barrett, A. M., Hughes, J. D., Felopulos, G. J., Bauman, M. L., Nadeau, S. E., & Heilman, K. M. (2000). Increased discrimination of "false memories" in autism spectrum disorder. *Proceedings of the National Academy of Sciences*, *97*(15), 8734–8737. https://doi.org/10.1073/pnas.97.15.8734

Bicanski, A., & Burgess, N. (2018). A neural-level model of spatial memory and imagery. *ELife*, *7*. https://doi.org/10.7554/eLife.33752

Bird, G., Leighton, J., Press, C., & Heyes, C. (2007). Intact automatic imitation of human and robot actions in autism spectrum disorders. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1628), 3027–3031. https://doi.org/10.1098/rspb.2007.1019

Bird, G., & Viding, E. (2014). The self to other model of empathy: Providing a new framework for understanding empathy impairments in psychopathy, autism, and alexithymia. *Neuroscience & Biobehavioral Reviews*, *47*, 520–532. https://doi.org/10.1016/j.neubiorev.2014.09.021

Black, J., Williams, D., & Ferguson, H. J. (2018). Imagining counterfactual worlds in autism spectrum disorder. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(9), 1444–1463. https://doi.org/10.1037/xlm0000500

Bluck, S. (2003). Autobiographical memory: Exploring its functions in everyday life. *Memory*, *11*(2), 113–123. https://doi.org/10.1080/741938206

Bon, L., Baleyte, J.-M., Piolino, P., Desgranges, B., Eustache, F., & Guillery-Girard, B. (2013). Growing

Up with Asperger's Syndrome: Developmental Trajectory of Autobiographical Memory. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00605

Boucher, J., & Warrington, E. K. (1976). Memory deficits in early infatile autism: : Some similarities to the amnesic syndrome. *British Journal of Psychology*, *67*(1), 73–87. https://doi.org/10.1111/j.2044-8295.1976.tb01499.x

Bower, G. H., & Gilligan, S. G. (1979). Remembering information related to one's self. *Journal of Research in Personality*, *13*(4), 420–432. https://doi.org/10.1016/0092-6566(79)90005-9

Bowler, D., Gaigg, S., & Lind, S. (2011). Memory in autism: binding, self and brain. In *Researching the Autism Spectrum* (pp. 316–346). Cambridge University Press. https://doi.org/10.1017/CBO9780511973918.013

Bowler, D. M., Gaigg, S. B., & Gardiner, J. M. (2010). Multiple List Learning in Adults with Autism Spectrum Disorder: Parallels with Frontal Lobe Damage or Further Evidence of Diminished Relational Processing? *Journal of Autism and Developmental Disorders*, *40*(2), 179–187. https://doi.org/10.1007/s10803-009-0845-x

Bowler, D. M., Gaigg, S. B., & Gardiner, J. M. (2014). Binding of Multiple Features in Memory by High-Functioning Adults with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *44*(9), 2355–2362. https://doi.org/10.1007/s10803-014-2105-y

Bowler, D. M., Gaigg, S. B., & Gardiner, J. M. (2015). Brief Report: The Role of Task Support in the Spatial and Temporal Source Memory of Adults with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *45*(8), 2613–2617. https://doi.org/10.1007/s10803-015-2378-9

Bowler, D. M., Gardiner, J. M., & Grice, S. J. (2000). Episodic memory and remembering in adults with Asperger syndrome. *Journal of Autism and Developmental Disorders*, *30*(4), 295–304. https://doi.org/10.1023/a:1005575216176

Bradley, L., Shaw, R., Baron-Cohen, S., & Cassidy, S. (2021). Autistic Adults' Experiences of Camouflaging and Its Perceived Impact on Mental Health. *Autism in Adulthood*, *3*(4), 320–329. https://doi.org/10.1089/aut.2020.0071

Brass, M., Bekkering, H., Wohlschläger, A., & Prinz, W. (2000). Compatibility between Observed and Executed Finger Movements: Comparing Symbolic, Spatial, and Imitative Cues. *Brain and Cognition*, *44*(2), 124–143. https://doi.org/10.1006/brcg.2000.1225

Brass, M., & Heyes, C. (2005). Imitation: is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Sciences*, *9*(10), 489–495. https://doi.org/10.1016/j.tics.2005.08.007

Brass, M., Ruby, P., & Spengler, S. (2009). Inhibition of imitative behaviour and social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1528), 2359–2367. https://doi.org/10.1098/rstb.2009.0066

Brass, M., & Wiersema, J. R. (2021). Self-Other Distinction and Social Cognition in ASD. In *Encyclopedia of Autism Spectrum Disorders* (pp. 4159–4163). Springer International Publishing. https://doi.org/10.1007/978-3-319-91280-6_102194

Brewer, M. B. (1991). The Social Self: On Being the Same and Different at the Same Time. *Personality and Social Psychology Bulletin*, *17*(5), 475–482. https://doi.org/10.1177/0146167291175001

Broadbent, J., Galic, I., & Stokes, M. A. (2013). Validation of Autism Spectrum Quotient Adult Version in an Australian Sample. *Autism Research and Treatment*, *2013*, 1–7. https://doi.org/10.1155/2013/984205

Brunsdon, V. E. A., Bradford, E. E. F., & Ferguson, H. J. (2019). Sensorimotor mu rhythm during action observation changes across the lifespan independently from social cognitive processes. *Developmental Cognitive Neuroscience*, *38*, 100659. https://doi.org/10.1016/j.dcn.2019.100659

Brunsdon, V. E. A., Bradford, E. E. F., Smith, L., & Ferguson, H. J. (2020). Short-term physical training enhances mirror system activation to action observation. *Social Neuroscience*, *15*(1), 98–107. https://doi.org/10.1080/17470919.2019.1660708

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in Cognitive Sciences*, *11*(2), 49–57. https://doi.org/10.1016/j.tics.2006.11.004

Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The Human Hippocampus and Spatial and Episodic Memory. *Neuron*, *35*(4), 625–641. https://doi.org/10.1016/S0896-6273(02)00830-9

Burgess, N., Maguire, E. A., Spiers, H. J., & O'Keefe, J. (2001). A Temporoparietal and Prefrontal Network for Retrieving the Spatial Context of Lifelike Events. *NeuroImage*, *14*(2), 439–453. https://doi.org/10.1006/nimg.2001.0806

Burrows, C. A., Usher, L. V., Mundy, P. C., & Henderson, H. A. (2017). The salience of the self: Self-referential processing and internalizing problems in children and adolescents with autism spectrum disorder. *Autism Research*, *10*(5), 949–960. https://doi.org/10.1002/aur.1727

Byrne, P., Becker, S., & Burgess, N. (2007). Remembering the past and imagining the future: A neural model of spatial memory and imagery. *Psychological Review*, *114*(2), 340–375. https://doi.org/10.1037/0033-295X.114.2.340

Caharel, S., Poiroux, S., Bernard, C., Thibaut, F., Lalonde, R., & Rebai, M. (2002). ERPs associated with familiarity and degree of familiarity during face recognition. *The International Journal of Neuroscience*, *112*(12), 1499–1512. https://doi.org/10.1080/00207450290158368

Cannon, E. N., Yoo, K. H., Vanderwert, R. E., Ferrari, P. F., Woodward, A. L., & Fox, N. A. (2014). Action Experience, More than Observation, Influences Mu Rhythm Desynchronization. *PLoS ONE*, *9*(3), e92002. https://doi.org/10.1371/journal.pone.0092002

Carpenter, K. L., Williams, D. M., & Nicholson, T. (2019). Putting Your Money Where Your Mouth is: Examining Metacognition in ASD Using Post-decision Wagering. *Journal of Autism and*

Developmental Disorders, *49*(10), 4268–4279. https://doi.org/10.1007/s10803-019-04118-6

Carrier, J. G., & Dittmar, H. (1994). The Social Psychology of Material Possessions: To Have Is To Be. *Man*, *29*(3), 735. https://doi.org/10.2307/2804376

Carruthers, P. (2009). How we know our own minds: The relationship between mindreading and metacognition. *Behavioral and Brain Sciences*, *32*(2), 121–138. https://doi.org/10.1017/S0140525X09000545

Castelli, F. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain*, *125*(8), 1839–1849. https://doi.org/10.1093/brain/awf189

Catmur, C., Walsh, V., & Heyes, C. (2009). Associative sequence learning: the role of experience in the development of imitation and the mirror system. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1528), 2369–2380. https://doi.org/10.1098/rstb.2009.0048

Chaminade, T., & Cheng, G. (2009). Social cognitive neuroscience and humanoid robotics. *Journal of Physiology-Paris*, *103*(3–5), 286–295. https://doi.org/10.1016/j.jphysparis.2009.08.011

Chaminade, T., Franklin, D. W., Oztop, E., & Cheng, G. (2005). Motor interference between humans and humanoid robots: Effect of biological and artificial motion. *Proceedings of 2005 4th IEEE International Conference on Development and Learning*. https://doi.org/10.1109/DEVLRN.2005.1490951

Charman, T., Pickles, A., Simonoff, E., Chandler, S., Loucas, T., & Baird, G. (2011). IQ in children with autism spectrum disorders: data from the Special Needs and Autism Project (SNAP). *Psychological Medicine*, *41*(3), 619–627. https://doi.org/10.1017/S0033291710000991

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*(6), 893–910. https://doi.org/10.1037/0022-3514.76.6.893

Chawner, S. J. R. A., Doherty, J. L., Anney, R. J. L., Antshel, K. M., Bearden, C. E., Bernier, R., Chung, W. K., Clements, C. C., Curran, S. R., Cuturilo, G., Fiksinski, A. M., Gallagher, L., Goin-Kochel, R. P., Gur, R. E., Hanson, E., Jacquemont, S., Kates, W. R., Kushan, L., Maillard, A. M., … van den Bree, M. B. M. (2021). A Genetics-First Approach to Dissecting the Heterogeneity of Autism: Phenotypic Comparison of Autism Risk Copy Number Variants. *American Journal of Psychiatry*, *178*(1), 77–86. https://doi.org/10.1176/appi.ajp.2020.20010015

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences*, *16*(4), 231–239. https://doi.org/10.1016/j.tics.2012.02.007

Ciaramelli, E., Spoglianti, S., Bertossi, E., Generali, N., Telarucci, F., Tancredi, · Raffaella, Filippo Muratori, ·, & Igliozzi, R. (2018). Construction of Past and Future Events in Children and Adolescents with ASD: Role of Self-relatedness and Relevance to Decision-Making. *Journal of Autism and Developmental Disorders*, *48*, 2995–3009. https://doi.org/10.1007/s10803-018-3577-y

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Comerchero, M. D., & Polich, J. (1999). P3a and P3b from typical auditory and visual stimuli. *Clinical Neurophysiology*, *110*(1), 24–30. https://doi.org/10.1016/S0168-5597(98)00033-1

Committeri, G., Fragueiro, A., Campanile, M. M., Lagatta, M., Burles, F., Iaria, G., Sestieri, C., & Tosoni, A. (2020). Egocentric Navigation Abilities Predict Episodic Memory Performance. *Frontiers in Human Neuroscience*, *14*. https://doi.org/10.3389/fnhum.2020.574224

Conway, J. R., Catmur, C., & Bird, G. (2019). Understanding individual differences in theory of mind via representation of minds, not mental states. *Psychonomic Bulletin & Review*, *26*(3), 798–812. https://doi.org/10.3758/s13423-018-1559-x

Conway, M. A. (2005). Memory and the self☆. *Journal of Memory and Language*, *53*(4), 594–628. https://doi.org/10.1016/j.jml.2005.08.005

Conway, M. A., Justice, L. V, & D'Argembeau, A. (2019). The self-memory system revisited: Past, present, and future. In *The organization and structure of autobiographical memory.*

Conway, M. A., & Loveday, C. (2015). Remembering, imagining, false memories & personal meanings. *Consciousness and Cognition*, *33*, 574–581. https://doi.org/10.1016/j.concog.2014.12.002

Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, *107*(2), 261–288. https://doi.org/10.1037/0033-295X.107.2.261

Conway, M. A., Pothos, E. M., & Turk, D. J. (2016). The self-relevance system? *Cognitive Neuroscience*, *7*(1–4), 20–21. https://doi.org/10.1080/17588928.2015.1075484

Cook, J., Swapp, D., Pan, X., Bianchi-Berthouze, N., & Blakemore, S.-J. (2014). Atypical interference effect of action observation in autism spectrum conditions. *Psychological Medicine*, *44*(4), 731–740. https://doi.org/10.1017/S0033291713001335

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215. https://doi.org/10.1038/nrn755

Costa, A., Torriero, S., Oliveri, M., & Caltagirone, C. (2008). Prefrontal and Temporo-Parietal Involvement in Taking Others' Perspective: TMS Evidence. *Behavioural Neurology*, *19*(1–2), 71–74. https://doi.org/10.1155/2008/694632

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*(3), 268–294. https://doi.org/10.1037/0096-3445.104.3.268

Crane, L., Pring, L., Jukes, K., & Goddard, L. (2012). Patterns of Autobiographical Memory in Adults with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *42*(10), 2100–2112. https://doi.org/10.1007/s10803-012-1459-2

Crompton, C. J., Ropar, D., Evans-Williams, C. V. M., Flynn, E. G., & Fletcher-Watson, S. (2020). Autistic peer-to-peer information transfer is highly effective. *Autism*, *24*(7), 1704–1712. https://doi.org/10.1177/1362361320919286

Cultice, J. C., Somerville, S. C., & Wellman, H. M. (1983). Preschoolers' Memory Monitoring: Feeling-of-Knowing Judgments. *Child Development*, *54*(6), 1480. https://doi.org/10.2307/1129810

Cunningham, S. J., Brady-Van den Bos, M., & Turk, D. J. (2011). Exploring the effects of ownership and choice on self-memory biases. *Memory*, *19*(5), 449–461. https://doi.org/10.1080/09658211.2011.584388

Cunningham, S. J., Brebner, J. L., Quinn, F., & Turk, D. J. (2014). The Self-Reference Effect on Memory in Early Childhood. *Child Development*, *85*(2), 808–823. https://doi.org/10.1111/cdev.12144

Cunningham, S. J., Turk, D. J., Macdonald, L. M., & Neil Macrae, C. (2008). Yours or mine? Ownership and memory. *Consciousness and Cognition*, *17*(1), 312–318. https://doi.org/10.1016/j.concog.2007.04.003

Cunningham, S. J., Vergunst, F., Macrae, C. N., & Turk, D. J. (2013). Exploring early self-referential memory effects through ownership. *British Journal of Developmental Psychology*, *31*(3), 289–301. https://doi.org/10.1111/bjdp.12005

Cygan, H. B., Tacikowski, P., Ostaszewski, P., Chojnicka, I., & Nowicka, A. (2014). Neural Correlates of Own Name and Own Face Detection in Autism Spectrum Disorder. *PLoS ONE*, *9*(1), e86020. https://doi.org/10.1371/journal.pone.0086020

Decety, J., & Lamm, C. (2006). Human Empathy Through the Lens of Social Neuroscience. *The Scientific World JOURNAL*, *6*, 1146–1163. https://doi.org/10.1100/tsw.2006.221

DeNigris, D., Morin, A., Best, C. A., Guillery-Girard, B., Vaillant, L. J., J-m, B., Anger, M., Wantzen, P., Le Vaillant, J., Malvy, J., Bon, L., Guénolé, F., Moussaoui, E., Barthelemy, C., Bonnet-Brilhault, F., Eustache, F., & Baleyte, J.-M. (2019). Citation:) Positive Effect of Visual Cuing in Episodic Memory and Episodic Future Thinking in Adolescents With Autism Spectrum Disorder. Front Positive Effect of Visual Cuing in Episodic Memory and Episodic Future Thinking in Adolescents With Autism Spectrum Disorder. *Psychol*, *10*, 1513. https://doi.org/10.3389/fpsyg.2019.01513

Denny, B. T., Kober, H., Wager, T. D., & Ochsner, K. N. (2012). A Meta-analysis of Functional Neuroimaging Studies of Self- and Other Judgments Reveals a Spatial Gradient for Mentalizing in Medial Prefrontal Cortex. *Journal of Cognitive Neuroscience*, *24*(8), 1742–1752. https://doi.org/10.1162/jocn_a_00233

Deschrijver, E., Bardi, L., Wiersema, J. R., & Brass, M. (2016). Behavioral measures of implicit theory of mind in adults with high functioning autism. *Cognitive Neuroscience*, *7*(1–4), 192–202. https://doi.org/10.1080/17588928.2015.1085375

Deschrijver, E., Wiersema, J. R., & Brass, M. (2017). Action-based touch observation in adults with high functioning autism: Can compromised self-other distinction abilities link social and sensory everyday problems? *Social Cognitive and Affective Neuroscience*, *12*(2), 273–282. https://doi.org/10.1093/scan/nsw126

Devlin, J. T., Matthews, P. M., & Rushworth, M. F. S. (2003). Semantic Processing in the Left Inferior Prefrontal Cortex: A Combined Functional Magnetic Resonance Imaging and Transcranial Magnetic Stimulation Study. *Journal of Cognitive Neuroscience*, *15*(1), 71–84. https://doi.org/10.1162/089892903321107837

Di Martino, A., Ross, K., Uddin, L. Q., Sklar, A. B., Castellanos, F. X., & Milham, M. P. (2009). Functional Brain Correlates of Social and Nonsocial Processes in Autism Spectrum Disorders: An Activation Likelihood Estimation Meta-Analysis. *Biological Psychiatry*, *65*(1), 63–74.

https://doi.org/10.1016/j.biopsych.2008.09.022

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

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

Dinulescu, S., Alvi, T., Rosenfield, D., Sunahara, C. S., Lee, J., & Tabak, B. A. (2021). Self-Referential Processing Predicts Social Cognitive Ability. *Social Psychological and Personality Science*, *12*(1), 99–107. https://doi.org/10.1177/1948550620902281

Donkers, F. C. L., & van Boxtel, G. J. M. (2004). The N2 in go/no-go tasks reflects conflict monitoring not response inhibition. *Brain and Cognition*, *56*(2), 165–176. https://doi.org/10.1016/j.bandc.2004.04.005

Duval, S., & Tweedie, R. (2000). Trim and Fill: A Simple Funnel-Plot-Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis. *Biometrics*, *56*(2), 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Eddy, C. M. (2022). The Transdiagnostic Relevance of Self-Other Distinction to Psychiatry Spans Emotional, Cognitive and Motor Domains. *Frontiers in Psychiatry*, *13*. https://doi.org/10.3389/fpsyt.2022.797952

Elmose, M. (2016). A conceptual framework for understanding characteristics of self-awareness associated with autism spectrum disorder. *Scandinavian Journal of Child and Adolescent Psychiatry and Psychology*, *4*(3), 109–114. https://doi.org/10.21307/sjcapp-2016-017

English, M. C. W., Gignac, G. E., Visser, T. A. W., Whitehouse, A. J. O., & Maybery, M. T. (2020). A comprehensive psychometric analysis of autism-spectrum quotient factor models using two large samples: Model recommendations and the influence of divergent traits on total-scale scores. *Autism Research*, *13*(1), 45–60. https://doi.org/10.1002/aur.2198

Farmer, H., Carr, E. W., Svartdal, M., Winkielman, P., & Hamilton, A. F. de C. (2016). Status and Power Do Not Modulate Automatic Imitation of Intransitive Hand Movements. *PLOS ONE*, *11*(4), e0151835. https://doi.org/10.1371/journal.pone.0151835

Farrer, C., Franck, N., Georgieff, N., Frith, C. ., Decety, J., & Jeannerod, M. (2003). Modulating the experience of agency: a positron emission tomography study. *NeuroImage*, *18*(2), 324–333. https://doi.org/10.1016/S1053-8119(02)00041-1

Fazio, R. H., Effrein, E. A., & Falender, V. J. (1981). Self-perceptions following social interaction. *Journal of Personality and Social Psychology*, *41*(2), 232–242. https://doi.org/10.1037/0022-3514.41.2.232

Fenigstein, A., Scheier, M. F., & Buss, A. H. (1975). Public and private self-consciousness: Assessment and theory. *Journal of Consulting and Clinical Psychology*, *43*(4), 522–527. https://doi.org/10.1037/h0076760

Ferguson, H. J., Brunsdon, V. E. A., & Bradford, E. E. F. (2018). Age of avatar modulates the altercentric bias in a visual perspective-taking task: ERP and behavioral evidence. *Cognitive, Affective, & Behavioral Neuroscience*, *18*(6), 1298–1319. https://doi.org/10.3758/s13415-018-0641-1

Fernández-Abascal, E. G., Cabello, R., Fernández-Berrocal, P., & Baron-Cohen, S. (2013). Test-retest reliability of the 'Reading the Mind in the Eyes' test: a one-year follow-up study. *Molecular Autism*, *4*(1), 33. https://doi.org/10.1186/2040-2392-4-33

Festinger, L. (1954). A Theory of Social Comparison Processes. *Human Relations*, *7*(2), 117–140. https://doi.org/10.1177/001872675400700202

Fombonne, E. (2009). Epidemiology of Pervasive Developmental Disorders. *Pediatric Research*, *65*(6), 591–598. https://doi.org/10.1203/PDR.0b013e31819e7203

Forbes, P. A. G., Wang, Y., & de C. Hamilton, A. F. (2017). STORMy Interactions: Gaze and the

Modulation of Mimicry in Adults on the Autism Spectrum. *Psychonomic Bulletin & Review*, *24*(2), 529–535. https://doi.org/10.3758/s13423-016-1136-0

Foss-Feig, J. H., Kwakye, L. D., Cascio, C. J., Burnette, C. P., Kadivar, H., Stone, W. L., & Wallace, M. T. (2010). An extended multisensory temporal binding window in autism spectrum disorders. *Experimental Brain Research*, *203*(2), 381–389. https://doi.org/10.1007/s00221-010-2240-4

Frith, U., & de Vignemont, F. (2005). Egocentrism, allocentrism, and Asperger syndrome. *Consciousness and Cognition*, *14*(4), 719–738. https://doi.org/10.1016/j.concog.2005.04.006

Frith, U., & Happé, F. (1994). Autism: beyond "theory of mind." *Cognition*, *50*(1–3), 115–132. https://doi.org/10.1016/0010-0277(94)90024-8

Gaigg, S. B., Bowler, D. M., Ecker, C., Calvo-Merino, B., & Murphy, D. G. (2015). Episodic Recollection Difficulties in ASD Result from Atypical Relational Encoding: Behavioral and Neural Evidence. *Autism Research*, *8*(3), 317–327. https://doi.org/10.1002/aur.1448

Geurts, H. M., Corbett, B., & Solomon, M. (2009). The paradox of cognitive flexibility in autism. *Trends in Cognitive Sciences*, *13*(2), 74–82. https://doi.org/10.1016/j.tics.2008.11.006

Giarelli, E., Wiggins, L. D., Rice, C. E., Levy, S. E., Kirby, R. S., Pinto-Martin, J., & Mandell, D. (2010). Sex differences in the evaluation and diagnosis of autism spectrum disorders among children. *Disability and Health Journal*, *3*(2), 107–116. https://doi.org/10.1016/j.dhjo.2009.07.001

Gillespie-Smith, K., Ballantyne, C., Branigan, H. P., Turk, D. J., & Cunningham, S. J. (2018). The I in autism: Severity and social functioning in autism are related to self-processing. *British Journal of Developmental Psychology*, *36*(1), 127–141. https://doi.org/10.1111/bjdp.12219

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

Goddard, L., Dritschel, B., & Howlin, P. (2014). A Preliminary Study of Gender Differences in Autobiographical Memory in Children with an Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *44*(9), 2087–2095. https://doi.org/10.1007/s10803-014-2109-7

Goldani, A. A. S., Downs, S. R., Widjaja, F., Lawton, B., & Hendren, R. L. (2014). Biomarkers in Autism. *Frontiers in Psychiatry*, *5*. https://doi.org/10.3389/fpsyt.2014.00100

Goldman, A. I. (2006). Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading. In *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*. https://doi.org/10.1093/0195138929.001.0001

Gowen, E., Bolton, E., & Poliakoff, E. (2016). Believe it or not: Moving non-biological stimuli believed to have human origin can be represented as human movement. *Cognition*, *146*, 431–438. https://doi.org/10.1016/j.cognition.2015.10.010

Gowen, E., & Poliakoff, E. (2012). How does visuomotor priming differ for biological and non-biological stimuli? A review of the evidence. In *Psychological Research*. https://doi.org/10.1007/s00426-011-0389-5

Gowen, E., Stanley, J., & Miall, R. C. (2008). Movement interference in autism-spectrum disorder. *Neuropsychologia*, *46*(4), 1060–1068. https://doi.org/10.1016/j.neuropsychologia.2007.11.004

Grecucci, A., Brambilla, P., Siugzdaite, R., Londero, D., Fabbro, F., & Rumiati, R. I. (2013). Emotional Resonance Deficits in Autistic Children. *Journal of Autism and Developmental Disorders*, *43*(3), 616–628. https://doi.org/10.1007/s10803-012-1603-z

Grisdale, E., Lind, S. E., Eacott, M. J., & Williams, D. M. (2014). Self-referential memory in autism spectrum disorder and typical development: Exploring the ownership effect. *Consciousness and Cognition*, *30*, 133–141. https://doi.org/10.1016/j.concog.2014.08.023

Grossi, D., Marcone, R., Cinquegrana, T., & Gallucci, M. (2013). On the differential nature of induced and incidental echolalia in autism. *Journal of Intellectual Disability Research*, *57*(10), 903–912.

https://doi.org/10.1111/j.1365-2788.2012.01579.x

Haider, M., Spong, P., & Lindsley, D. B. (1964). Attention, Vigilance, and Cortical Evoked-Potentials in Humans. *Science*, *145*(3628), 180–182. https://doi.org/10.1126/science.145.3628.180

Hamilton, A. F. d. C. (2015). The neurocognitive mechanisms of imitation. In *Current Opinion in Behavioral Sciences*. https://doi.org/10.1016/j.cobeha.2015.01.011

Hamilton, A. F. de C. (2013). Reflecting on the mirror neuron system in autism: A systematic review of current theories. *Developmental Cognitive Neuroscience*, *3*, 91–105. https://doi.org/10.1016/j.dcn.2012.09.008

Hampshire, A., Chamberlain, S. R., Monti, M. M., Duncan, J., & Owen, A. M. (2010). The role of the right inferior frontal gyrus: inhibition and attentional control. *NeuroImage*, *50*(3), 1313–1319. https://doi.org/10.1016/j.neuroimage.2009.12.109

Harrop, C., Jones, D., Zheng, S., Nowell, S. W., Boyd, B. A., & Sasson, N. (2018). Sex differences in social attention in autism spectrum disorder. *Autism Research*, *11*(9), 1264–1275. https://doi.org/10.1002/aur.1997

Hartley, C., & Fisher, S. (2018). Mine is better than yours: Investigating the ownership effect in children with autism spectrum disorder and typically developing children. *Cognition*, *172*, 26–36. https://doi.org/10.1016/j.cognition.2017.11.009

Hartley, C., Fisher, S., & Fletcher, N. (2020). Exploring the influence of ownership history on object valuation in typical development and autism. *Cognition*, *197*, 104187. https://doi.org/10.1016/j.cognition.2020.104187

Hassabis, D., Kumaran, D., & Maguire, E. A. (2007). Using Imagination to Understand the Neural Basis of Episodic Memory. *Journal of Neuroscience*, *27*(52), 14365–14374. https://doi.org/10.1523/JNEUROSCI.4549-07.2007

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, *50*(3), 1166–1186. https://doi.org/10.3758/s13428-017-0935-1

Hedges, L. V., Gurevitch, J., & Curtis, P. S. (1999). The meta-analysis of response ratios in experimental ecology. In *Ecology*. https://doi.org/10.1890/0012-9658(1999)080[1150:TMAORR]2.0.CO;2

Henderson, H. A., Zahka, N. E., Kojkowski, N. M., Inge, A. P., Schwartz, C. B., Hileman, C. M., Coman, D. C., & Mundy, P. C. (2009). Self-referenced memory, social cognition, and symptom presentation in autism. *Journal of Child Psychology and Psychiatry*, *50*(7), 853–861. https://doi.org/10.1111/j.1469-7610.2008.02059.x

Hess, T. M. (2014). Selective Engagement of Cognitive Resources. *Perspectives on Psychological Science*, *9*(4), 388–407. https://doi.org/10.1177/1745691614527465

Heyes, C. (2010). Where do mirror neurons come from? *Neuroscience & Biobehavioral Reviews*, *34*(4), 575–583. https://doi.org/10.1016/j.neubiorev.2009.11.007

Hilton, C. L., Harper, J. D., Kueker, R. H., Lang, A. R., Abbacchi, A. M., Todorov, A., & LaVesser, P. D. (2010). Sensory Responsiveness as a Predictor of Social Severity in Children with High Functioning Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *40*(8), 937–945. https://doi.org/10.1007/s10803-010-0944-8

Hobson, H. M., & Bishop, D. V. M. (2017). The interpretation of mu suppression as an index of mirror neuron activity: past, present and future. *Royal Society Open Science*, *4*(3), 160662. https://doi.org/10.1098/rsos.160662

Höller, Y., Kronbichler, M., Bergmann, J., Crone, J. S., Ladurner, G., & Golaszewski, S. (2011). EEG frequency analysis of responses to the own-name stimulus. *Clinical Neurophysiology*, *122*(1), 99–106. https://doi.org/10.1016/j.clinph.2010.05.029

Holmes, D. L., & Frith, U. (2006). Review of *Autism: Explaining the enigma (second edition)*. *American Journal of Forensic Psychology*.

Hull, L., Lai, M.-C., Baron-Cohen, S., Allison, C., Smith, P., Petrides, K., & Mandy, W. (2020). Gender differences in self-reported camouflaging in autistic and non-autistic adults. *Autism*, *24*(2), 352–363. https://doi.org/10.1177/1362361319864804

Hull, L., Petrides, K. V., Allison, C., Smith, P., Baron-Cohen, S., Lai, M.-C., & Mandy, W. (2017). "Putting on My Best Normal": Social Camouflaging in Adults with Autism Spectrum Conditions. *Journal of Autism and Developmental Disorders*, *47*(8), 2519–2534. https://doi.org/10.1007/s10803-017-3166-5

Humphreys, G. W., & Sui, J. (2016). Attentional control and the self: The Self-Attention Network (SAN). *Cognitive Neuroscience*, *7*(1–4), 5–17. https://doi.org/10.1080/17588928.2015.1044427

Irish, M., Addis, D. R., Hodges, J. R., & Piguet, O. (2012). Exploring the content and quality of episodic future simulations in semantic dementia. *Neuropsychologia*, *50*(14), 3488–3495. https://doi.org/10.1016/j.neuropsychologia.2012.09.012

JAMES, W. (2018). The principles of psychology. *Personality.Culture.Society*, *20*(3–4), 27–64. https://doi.org/10.30936/1606-951X-2018-20-3/4-27-64

JASP Team. (2020). JASP (Version 0.14.1). In *[Computer software].*

Jeannerod, M. (2003). The mechanism of self-recognition in humans. *Behavioural Brain Research*, *142*(1–2), 1–15. https://doi.org/10.1016/S0166-4328(02)00384-4

Jenkins, A. C., Macrae, C. N., & Mitchell, J. P. (2008). Repetition suppression of ventromedial prefrontal activity during judgments of self and others. *Proceedings of the National Academy of Sciences*, *105*(11), 4507–4512. https://doi.org/10.1073/pnas.0708785105

Jolicœur, P., Sessa, P., Dell'Acqua, R., & Robitaille, N. (2006). Attentional control and capture in the

attentional blink paradigm: Evidence from human electrophysiology. *European Journal of Cognitive Psychology*, *18*(4), 560–578. https://doi.org/10.1080/09541440500423210

Kana, R. K., Libero, L. E., & Moore, M. S. (2011). Disrupted cortical connectivity theory as an explanatory model for autism spectrum disorders. *Physics of Life Reviews*, *8*(4), 410–437. https://doi.org/10.1016/j.plrev.2011.10.001

Kanner L. (1943). Autistic disturbances of affective contact. The Nervous Child,. In *Autism and Childhood Scbicopbrenia,* .

Kelley, W. M., Macrae, C. N., Wyland, C. L., Caglar, S., Inati, S., & Heatherton, T. F. (2002). Finding the Self? An Event-Related fMRI Study. *Journal of Cognitive Neuroscience*, *14*(5), 785–794. https://doi.org/10.1162/08989290260138672

Kennedy, D. P., & Courchesne, E. (2008). Functional abnormalities of the default network during self- and other-reflection in autism. *Social Cognitive and Affective Neuroscience*, *3*(2), 177–190. https://doi.org/10.1093/scan/nsn011

Ketay, S., Beck, L. A., Riela, S., Bailey, C., & Aron, A. (2019). Seeing you in me: Preliminary evidence for perceptual overlap between self and close others. *Journal of Social and Personal Relationships*, *36*(8), 2474–2486. https://doi.org/10.1177/0265407518788702

Keyes, C. L. M., & Ryff, C. D. (2000). Subjective Change and Mental Health: A Self-Concept Theory. *Social Psychology Quarterly*, *63*(3), 264. https://doi.org/10.2307/2695873

Kilner, J. ., Paulignan, Y., & Blakemore, S. . (2003). An Interference Effect of Observed Biological Movement on Action. *Current Biology*, *13*(6), 522–525. https://doi.org/10.1016/S0960-9822(03)00165-9

Kirmse, U., Jacobsen, T., & Schröger, E. (2009). Familiarity affects environmental sound processing outside the focus of attention: An event-related potential study. *Clinical Neurophysiology*, *120*(5), 887–896. https://doi.org/10.1016/j.clinph.2009.02.159

Klein, S. B., & Loftus, J. (1988). The nature of self-referent encoding: The contributions of elaborative and organizational processes. *Journal of Personality and Social Psychology*, *55*(1), 5–11. https://doi.org/10.1037/0022-3514.55.1.5

Kloosterman, P. H., Keefer, K. V., Kelley, E. A., Summerfeldt, L. J., & Parker, J. D. A. (2011). Evaluation of the factor structure of the Autism-Spectrum Quotient. *Personality and Individual Differences*, *50*(2), 310–314. https://doi.org/10.1016/j.paid.2010.10.015

Knyazev, G. G. (2013). EEG Correlates of Self-Referential Processing. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00264

Kranczioch, C., Debener, S., Maye, A., & Engel, A. K. (2007). Temporal dynamics of access to consciousness in the attentional blink. *NeuroImage*, *37*(3), 947–955. https://doi.org/10.1016/j.neuroimage.2007.05.044

Lanius, R. A., Terpou, B. A., & McKinnon, M. C. (2020). The sense of self in the aftermath of trauma: lessons from the default mode network in posttraumatic stress disorder. *European Journal of Psychotraumatology*, *11*(1), 1807703. https://doi.org/10.1080/20008198.2020.1807703

Lawrence, K. E., Hernandez, L. M., Eilbott, J., Jack, A., Aylward, E., Gaab, N., Van Horn, J. D., Bernier, R. A., Geschwind, D. H., McPartland, J. C., Nelson, C. A., Webb, S. J., Pelphrey, K. A., Bookheimer, S. Y., Dapretto, M., Aylward, E., Bernier, R. A., Bookheimer, S. Y., Dapretto, M., … Wolf, J. (2020). Neural responsivity to social rewards in autistic female youth. *Translational Psychiatry*, *10*(1), 178. https://doi.org/10.1038/s41398-020-0824-8

Lee, D. N., & Aronson, E. (1974). Visual proprioceptive control of standing in human infants. *Perception & Psychophysics*, *15*(3), 529–532. https://doi.org/10.3758/BF03199297

Leekam, S. R., Nieto, C., Libby, S. J., Wing, L., & Gould, J. (2007). Describing the Sensory Abnormalities of Children and Adults with Autism. *Journal of Autism and Developmental Disorders*, *37*(5), 894–910. https://doi.org/10.1007/s10803-006-0218-7

Lewis, M., & Brooks-Gunn, J. (1979). The Development of Self Recognition. In *Social Cognition and the Acquisition of Self* (pp. 198–221). Springer US. https://doi.org/10.1007/978-1-4684-3566-5_8

Lewis, M., & Ramsay, D. (2004). Development of Self-Recognition, Personal Pronoun Use, and Pretend Play During the 2nd Year. *Child Development*, *75*(6), 1821–1831. https://doi.org/10.1111/j.1467-8624.2004.00819.x

Lind, S., & Bowler, D. (2008). Episodic memory and autonoetic consciousness in autistic spectrum disorders: the roles of self-awareness, representational abilities and temporal cognition. In *Memory In Autism* (pp. 166–187). Cambridge University Press. https://doi.org/10.1017/CBO9780511490101.011

Lind, S. E. (2010). Memory and the self in autism. *Autism*, *14*(5), 430–456. https://doi.org/10.1177/1362361309358700

Lind, S. E., & Bowler, D. M. (2010). Episodic memory and episodic future thinking in adults with autism. *Journal of Abnormal Psychology*, *119*(4), 896–905. https://doi.org/10.1037/a0020631

Lind, S. E., Bowler, D. M., & Raber, J. (2014). Spatial navigation, episodic memory, episodic future thinking, and theory of mind in children with autism spectrum disorder: evidence for impairments in mental simulation? *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.01411

Lind, S. E., Williams, D. M., Bowler, D. M., & Peel, A. (2014). Episodic memory and episodic future thinking impairments in high-functioning autism spectrum disorder: An underlying difficulty with scene construction or self-projection? *Neuropsychology*, *28*(1), 55–67. https://doi.org/10.1037/neu0000005

Lind, S. E., Williams, D. M., Grainger, C., & Landsiedel, J. (2018). The Self in Autism and Its Relation to Memory. In *The Wiley Handbook of Memory, Autism Spectrum Disorder, and the Law* (pp. 70–

91). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119158431.ch4

Lind, S. E., Williams, D. M., Nicholson, T., Grainger, C., & Carruthers, P. (2020). The self-reference effect on memory is not diminished in autism: Three studies of incidental and explicit self-referential recognition memory in autistic and neurotypical adults and adolescents. *Journal of Abnormal Psychology*, *129*(2), 224–236. https://doi.org/10.1037/abn0000467

Livingston, L. A., & Happé, F. (2017). Conceptualising compensation in neurodevelopmental disorders: Reflections from autism spectrum disorder. *Neuroscience & Biobehavioral Reviews*, *80*, 729–742. https://doi.org/10.1016/j.neubiorev.2017.06.005

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory: Figure 1. *Learning & Memory*, *12*(4), 361–366. https://doi.org/10.1101/lm.94705

Lombardo, M. V., Barnes, J. L., Wheelwright, S. J., & Baron-Cohen, S. (2007). Self-Referential Cognition and Empathy in Autism. *PLoS ONE*, *2*(9), e883. https://doi.org/10.1371/journal.pone.0000883

Lombardo, M. V., & Baron-Cohen, S. (2010). Unraveling the paradox of the autistic self. *WIREs Cognitive Science*, *1*(3), 393–403. https://doi.org/10.1002/wcs.45

Lombardo, M. V., Chakrabarti, B., Bullmore, E. T., Sadek, S. A., Pasco, G., Wheelwright, S. J., Suckling, J., & Baron-Cohen, S. (2010). Atypical neural self-representation in autism. *Brain*, *133*(2), 611–624. https://doi.org/10.1093/brain/awp306

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., Pickles, A., & Rutter, M. (2000). The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, *30*(3), 205–223. https://doi.org/10.1023/A:1005592401947

Lu, X., Ho, H. T., Liu, F., Wu, D., & Thompson, W. F. (2015). Intonation processing deficits of

emotional words among Mandarin Chinese speakers with congenital amusia: an ERP study. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00385

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences*, *4*(11), 432–440. https://doi.org/10.1016/S1364-6613(00)01545-X

Lyons, V., & Fitzgeral, M. (2013). Atypical Sense of Self in Autism Spectrum Disorders: A Neuro-Cognitive Perspective. In *Recent Advances in Autism Spectrum Disorders - Volume I*. InTech. https://doi.org/10.5772/53680

M., A. E., Mead, G. H., & Morris, C. W. (1935). Mind, Self and Society. From the Standpoint of a Social Behaviorist. *The Journal of Philosophy*, *32*(6), 162. https://doi.org/10.2307/2016046

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*(4), 1122–1135. https://doi.org/10.3758/s13428-014-0532-5

Macrae, C. N. (2004). Medial Prefrontal Activity Predicts Memory for Self. *Cerebral Cortex*, *14*(6), 647–654. https://doi.org/10.1093/cercor/bhh025

Maguire, E. A. (2001). Neuroimaging studies of autobiographical event memory. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *356*(1413), 1441–1451. https://doi.org/10.1098/rstb.2001.0944

Markram, K., & Markram, H. (2010). The Intense World Theory – A Unifying Theory of the Neurobiology of Autism. *Frontiers in Human Neuroscience*, *4*. https://doi.org/10.3389/fnhum.2010.00224

Markus, H., & Kunda, Z. (1986). Stability and malleability of the self-concept. *Journal of Personality and Social Psychology*, *51*(4), 858–866. https://doi.org/10.1037/0022-3514.51.4.858

Marsh, L. E., Bird, G., & Catmur, C. (2016). The imitation game: Effects of social cues on 'imitation'

are domain-general in nature. *NeuroImage*, *139*, 368–375. https://doi.org/10.1016/j.neuroimage.2016.06.050

Marzi, T., & Viggiano, M. P. (2007). Interplay between familiarity and orientation in face processing: An ERP study. *International Journal of Psychophysiology*, *65*(3), 182–192. https://doi.org/10.1016/j.ijpsycho.2007.04.003

Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., & Macrae, C. N. (2007). Wandering Minds: The Default Network and Stimulus-Independent Thought. *Science*, *315*(5810), 393–395. https://doi.org/10.1126/science.1131295

McDonnell, C. G., Valentino, K., & Diehl, J. J. (2017). A developmental psychopathology perspective on autobiographical memory in autism spectrum disorder. *Developmental Review*, *44*, 59–81. https://doi.org/10.1016/j.dr.2017.01.001

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746–748. https://doi.org/10.1038/264746a0

McPartland, J. C. (2019). Autism's existential crisis: a reflection on Livingston et al. (2018). In *Journal of Child Psychology and Psychiatry and Allied Disciplines*. https://doi.org/10.1111/jcpp.12989

Meyer, M. L., Zhao, Z., & Tamir, D. I. (2019). Simulating other people changes the self. *Journal of Experimental Psychology: General*, *148*(11), 1898–1913. https://doi.org/10.1037/xge0000565

Minshew, N. J., Goldstein, G., & Siegel, D. J. (1997). Neuropsychologic functioning in autism: Profile of a complex information processing disorder. *Journal of the International Neuropsychological Society*, *3*(4), 303–316. https://doi.org/10.1017/S1355617797003032

Mitchell, J. P. (2009). Inferences about mental states. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1309–1316. https://doi.org/10.1098/rstb.2008.0318

Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2006). Dissociable Medial Prefrontal Contributions to

Judgments of Similar and Dissimilar Others. *Neuron*, *50*(4), 655–663. https://doi.org/10.1016/j.neuron.2006.03.040

Mobbs, D., Yu, R., Meyer, M., Passamonti, L., Seymour, B., Calder, A. J., Schweizer, S., Frith, C. D., & Dalgleish, T. (2009). A Key Role for Similarity in Vicarious Reward. *Science*, *324*(5929), 900–900. https://doi.org/10.1126/science.1170539

Moliadze, V., Brodski-Guerniero, A., Schuetz, M., Siemann, J., Lyzhko, E., Schlitt, S., Kitzerow, J., Langer, A., Kaiser, J., Naumer, M. J., Wibral, M., Chan, J., Freitag, C. M., & Siniatchkin, M. (2020). Significance of Beta-Band Oscillations in Autism Spectrum Disorders During Motor Response Inhibition Tasks: A MEG Study. *Brain Topography*, *33*(3), 355–374. https://doi.org/10.1007/s10548-020-00765-6

Morin, A. (2006). Levels of consciousness and self-awareness: A comparison and integration of various neurocognitive views. *Consciousness and Cognition*, *15*(2), 358–371. https://doi.org/10.1016/j.concog.2005.09.006

Morin, A., & Everett, J. (1990). Inner speech as a mediator of self-awareness, self-consciousness, and self-knowledge: An hypothesis. *New Ideas in Psychology*, *8*(3), 337–356. https://doi.org/10.1016/0732-118X(94)90020-5

Mottron, L. (2021). A radical change in our autism research strategy is needed: Back to prototypes. *Autism Research*, *14*(10), 2213–2220. https://doi.org/10.1002/aur.2494

Navon, M., & Makovski, T. (n.d.). *Are Self-Related Items Unique? The Self-Prioritization Effect Revisited*. https://doi.org/10.31234/OSF.IO/9DZM4

Neisser, U. (1995). Criteria for an Ecological Self. In *Advances in Psychology* (pp. 17–34). https://doi.org/10.1016/S0166-4115(05)80004-4

Nichols, S., & Stich, S. P. (2003). How to Read Your Own Mind: A Cognitive Theory of Self-Consciousness. *Consciousness: New Essays*.

Nijhof, Annabel D., Dhar, M., Goris, J., Brass, M., & Wiersema, J. R. (2018). Atypical neural responding to hearing one's own name in adults with ASD. *Journal of Abnormal Psychology*, *127*(1), 129–138. https://doi.org/10.1037/abn0000329

Nijhof, Annabel D., Shapiro, K. L., Catmur, C., & Bird, G. (2020). No evidence for a common self-bias across cognitive domains. *Cognition*, *197*, 104186. https://doi.org/10.1016/j.cognition.2020.104186

Nijhof, Annabel D., von Trott zu Solz, J., Catmur, C., & Bird, G. (2022). Equivalent own name bias in autism: An EEG study of the Attentional Blink. *Cognitive, Affective, & Behavioral Neuroscience*, *22*(3), 625–639. https://doi.org/10.3758/s13415-021-00967-w

Nijhof, Annabel Dineke, & Bird, G. (2019). Self-processing in individuals with autism spectrum disorder. *Autism Research*, *12*(11), 1580–1584. https://doi.org/10.1002/aur.2200

Nowicka, A., Cygan, H. B., Tacikowski, P., Ostaszewski, P., & Kuś, R. (2016). Name recognition in autism: EEG evidence of altered patterns of brain activity and connectivity. *Molecular Autism*, *7*(1), 38. https://doi.org/10.1186/s13229-016-0102-z

Oberman, L. M., & Ramachandran, V. S. (2007). The simulating social mind: The role of the mirror neuron system and simulation in the social and communicative deficits of autism spectrum disorders. *Psychological Bulletin*, *133*(2), 310–327. https://doi.org/10.1037/0033-2909.133.2.310

Oberman, L. M., Ramachandran, V. S., & Pineda, J. A. (2008). Modulation of mu suppression in children with autism spectrum disorders in response to familiar or unfamiliar stimuli: The mirror neuron hypothesis. *Neuropsychologia*, *46*(5), 1558–1565. https://doi.org/10.1016/j.neuropsychologia.2008.01.010

OSTERLING, J. A., DAWSON, G., & MUNSON, J. A. (2002). Early recognition of 1-year-old infants with autism spectrum disorder versus mental retardation. *Development and Psychopathology*, *14*(2),

239–251. https://doi.org/10.1017/S0954579402002031

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Perry, A., & Bentin, S. (2009). Mirror activity in the human brain while observing hand movements: A comparison between EEG desynchronization in the μ-range and previous fMRI results. *Brain Research*, *1282*, 126–132. https://doi.org/10.1016/j.brainres.2009.05.059

Pfeifer, J. H., Masten, C. L., Borofsky, L. A., Dapretto, M., Fuligni, A. J., & Lieberman, M. D. (2009). Neural Correlates of Direct and Reflected Self-Appraisals in Adolescents and Adults: When Social Perspective-Taking Informs Self-Perception. *Child Development*, *80*(4), 1016–1038. https://doi.org/10.1111/j.1467-8624.2009.01314.x

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Polich, J., & Kok, A. (1995). Cognitive and biological determinants of P300: an integrative review. *Biological Psychology*, *41*(2), 103–146. https://doi.org/10.1016/0301-0511(95)05130-9

Povinelli, D. J., Landau, K. R., & Perilloux, H. K. (1996). Self-Recognition in Young Children Using Delayed versus Live Feedback: Evidence of a Developmental Asynchrony. *Child Development*, *67*(4), 1540. https://doi.org/10.2307/1131717

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

Press, C. (2011). Action observation and robotic agents: Learning and anthropomorphism. In *Neuroscience and Biobehavioral Reviews*. https://doi.org/10.1016/j.neubiorev.2011.03.004

Press, C., Bird, G., Flach, R., & Heyes, C. (2005). Robotic movement elicits automatic imitation.

Cognitive Brain Research, *25*(3), 632–640. https://doi.org/10.1016/j.cogbrainres.2005.08.020

Press, C., Gillmeister, H., & Heyes, C. (2006). Bottom-up, not top-down, modulation of imitation by human and robotic models. *European Journal of Neuroscience*, *24*(8), 2415–2419. https://doi.org/10.1111/j.1460-9568.2006.05115.x

Press, C., Gillmeister, H., & Heyes, C. (2007). Sensorimotor experience enhances automatic imitation of robotic action. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1625), 2509–2514. https://doi.org/10.1098/rspb.2007.0774

Press, C., Richardson, D., & Bird, G. (2010). Intact imitation of emotional facial actions in autism spectrum conditions. *Neuropsychologia*, *48*(11), 3291–3297. https://doi.org/10.1016/j.neuropsychologia.2010.07.012

Prinz, W. (1997). Perception and Action Planning. *European Journal of Cognitive Psychology*, *9*(2), 129–154. https://doi.org/10.1080/713752551

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, *6*(7), 576–582. https://doi.org/10.1038/nrn1706

Quesque, F., & Brass, M. (2019). The Role of the Temporoparietal Junction in Self-Other Distinction. *Brain Topography*, *32*(6), 943–955. https://doi.org/10.1007/s10548-019-00737-5

Quon, E., & Atance, C. M. (2010). A Comparison of Preschoolers' Memory, Knowledge, and Anticipation of Events. *Journal of Cognition and Development*, *11*(1), 37–60. https://doi.org/10.1080/15248370903453576

Rilling, J. K., Dagenais, J. E., Goldsmith, D. R., Glenn, A. L., & Pagnoni, G. (2008). Social cognitive neural networks during in-group and out-group interactions. *NeuroImage*, *41*(4), 1447–1461. https://doi.org/10.1016/j.neuroimage.2008.03.044

Rizzolatti, G., & Sinigaglia, C. (2016). The mirror mechanism: a basic principle of brain function.

Nature Reviews Neuroscience, *17*(12), 757–765. https://doi.org/10.1038/nrn.2016.135

Rødgaard, E.-M., Jensen, K., Vergnes, J.-N., Soulières, I., & Mottron, L. (2019). Temporal Changes in Effect Sizes of Studies Comparing Individuals With and Without Autism. *JAMA Psychiatry*, *76*(11), 1124. https://doi.org/10.1001/jamapsychiatry.2019.1956

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*(9), 677–688. https://doi.org/10.1037/0022-3514.35.9.677

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225–237. https://doi.org/10.3758/PBR.16.2.225

Rousselet, G. A., Husk, J. S., Bennett, P. J., & Sekuler, A. B. (2008). Time course and robustness of ERP object and face differences. *Journal of Vision*, *8*(12), 3–3. https://doi.org/10.1167/8.12.3

Rubin-McGregor, J., Zhao, Z., & Tamir, D. I. (2022). Simulation induces durable, extensive changes to self-knowledge. *Journal of Experimental Social Psychology*, *98*, 104229. https://doi.org/10.1016/j.jesp.2021.104229

Rugg, M. D., Otten, L. J., & Henson, R. N. A. (2002). The neural basis of episodic memory: evidence from functional neuroimaging. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *357*(1424), 1097–1110. https://doi.org/10.1098/rstb.2002.1102

Rutherford, M., McKenzie, K., Johnson, T., Catchpole, C., O'Hare, A., McClure, I., Forsyth, K., McCartney, D., & Murray, A. (2016). Gender ratio in a clinical population sample, age of diagnosis and duration of assessment in children and adults with autism spectrum disorder. *Autism*, *20*(5), 628–634. https://doi.org/10.1177/1362361315617879

Saito, N., Yokoyama, T., & Ohira, H. (2016). Self-Other Distinction Enhanced Empathic Responses in Individuals with Alexithymia. *Scientific Reports*, *6*(1), 35059. https://doi.org/10.1038/srep35059

Salminen, J. K., Saarijärvi, S., Äärelä, E., Toikka, T., & Kauhanen, J. (1999). Prevalence of alexithymia and its association with sociodemographic variables in the general population of finland. *Journal of Psychosomatic Research*, *46*(1), 75–82. https://doi.org/10.1016/S0022-3999(98)00053-1

Sanders, J., Johnson, K. A., Garavan, H., Gill, M., & Gallagher, L. (2008). A review of neuropsychological and neuroimaging research in autistic spectrum disorders: Attention, inhibition and cognitive flexibility. *Research in Autism Spectrum Disorders*, *2*(1), 1–16. https://doi.org/10.1016/j.rasd.2007.03.005

Santiesteban, I., Banissy, M. J., Catmur, C., & Bird, G. (2012). Enhancing Social Ability by Stimulating Right Temporoparietal Junction. *Current Biology*, *22*(23), 2274–2277. https://doi.org/10.1016/j.cub.2012.10.018

Saylor, M. M., Ganea, P. A., & Vázquez, M. D. (2011). What's mine is mine: twelve-month-olds use possessive pronouns to identify referents. *Developmental Science*, *14*(4), 859–864. https://doi.org/10.1111/j.1467-7687.2010.01034.x

Schacter, D. L., Addis, D. R., Hassabis, D., Martin, V. C., Spreng, R. N., & Szpunar, K. K. (2012). The Future of Memory: Remembering, Imagining, and the Brain. *Neuron*, *76*(4), 677–694. https://doi.org/10.1016/j.neuron.2012.11.001

Schunke, O., Schöttle, D., Vettorazzi, E., Brandt, V., Kahl, U., Bäumer, T., Ganos, C., David, N., Peiker, I., Engel, A. K., Brass, M., & Münchau, A. (2016). Mirror me: Imitative responses in adults with autism. *Autism*, *20*(2), 134–144. https://doi.org/10.1177/1362361315571757

Scott, F. J. (2004). Autism: Explaining the Enigma, 2nd edn. *Psychological Medicine*.

Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature*, *408*(6814), 788–788. https://doi.org/10.1038/35048669

Shanks, D. R., & Darby, R. J. (1998). Feature- and rule-based generalization in human associative

learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *24*(4), 405–415. https://doi.org/10.1037/0097-7403.24.4.405

Shanton, K., & Goldman, A. (2010). Simulation theory. In *Wiley Interdisciplinary Reviews: Cognitive Science*. https://doi.org/10.1002/wcs.33

Simantov, T., Lombardo, M., Baron-Cohen, S., & Uzefovsky, F. (2021). Self-Other Distinction. In *The Neural Basis of Mentalizing*. https://doi.org/10.1007/978-3-030-51890-5_5

Southgate, V., & de C. Hamilton, A. F. (2008). Unbroken mirrors: challenging a theory of Autism. *Trends in Cognitive Sciences*, *12*(6), 225–229. https://doi.org/10.1016/j.tics.2008.03.005

Sowden, S., & Catmur, C. (2015). The Role of the Right Temporoparietal Junction in the Control of Imitation. *Cerebral Cortex*, *25*(4), 1107–1113. https://doi.org/10.1093/cercor/bht306

Sowden, S., Koehne, S., Catmur, C., Dziobek, I., & Bird, G. (2016). Intact Automatic Imitation and Typical Spatial Compatibility in Autism Spectrum Disorder: Challenging the Broken Mirror Theory. *Autism Research*, *9*(2), 292–300. https://doi.org/10.1002/aur.1511

Sowden, S., & Shah, P. (2014). Self-other control: a candidate mechanism for social cognitive function. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00789

Spengler, S., Bird, G., & Brass, M. (2010). Hyperimitation of Actions Is Related to Reduced Understanding of Others' Minds in Autism Spectrum Conditions. *Biological Psychiatry*, *68*(12), 1148–1155. https://doi.org/10.1016/j.biopsych.2010.09.017

Spengler, S., von Cramon, D. Y., & Brass, M. (2009). Control of shared representations relies on key processes involved in mental state attribution. *Human Brain Mapping*, *30*(11), 3704–3718. https://doi.org/10.1002/hbm.20800

Stevenson, R. A., Zemtsov, R. K., & Wallace, M. T. (2012). Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions. *Journal of Experimental*

Psychology: Human Perception and Performance, *38*(6), 1517–1529. https://doi.org/10.1037/a0027339

Suddendorf, T. (2010). Episodic memory versus episodic foresight: Similarities and differences. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(1), 99–107. https://doi.org/10.1002/WCS.23

Suddendorf, T., & Corballis, M. C. (1997). Mental time travel and the evolution of the human mind. *Genetic, Social, and General Psychology Monographs*, *123*(2), 133–167. http://www.ncbi.nlm.nih.gov/pubmed/9204544

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*(5), 1105–1117. https://doi.org/10.1037/a0029792

Sui, J., & Humphreys, G. W. (2013). Self-referential processing is distinct from semantic elaboration: Evidence from long-term memory effects in a patient with amnesia and semantic impairments. *Neuropsychologia*, *51*(13), 2663–2673. https://doi.org/10.1016/j.neuropsychologia.2013.07.025

Sui, J., & Humphreys, G. W. (2015). The Integrative Self: How Self-Reference Integrates Perception and Memory. *Trends in Cognitive Sciences*, *19*(12), 719–728. https://doi.org/10.1016/j.tics.2015.08.015

Svoboda, E., McKinnon, M. C., & Levine, B. (2006). The functional neuroanatomy of autobiographical memory: A meta-analysis. *Neuropsychologia*, *44*(12), 2189–2208. https://doi.org/10.1016/j.neuropsychologia.2006.05.023

Swann, W. B., & Hill, C. A. (1982). When our identities are mistaken: Reaffirming self-conceptions through social interaction. *Journal of Personality and Social Psychology*, *43*(1), 59–66. https://doi.org/10.1037/0022-3514.43.1.59

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta-analysis.

Psychological Bulletin, *121*(3), 371–394. https://doi.org/10.1037/0033-2909.121.3.371

Szpunar, K. K., Watson, J. M., & McDermott, K. B. (2007). Neural substrates of envisioning the future. *Proceedings of the National Academy of Sciences*, *104*(2), 642–647. https://doi.org/10.1073/pnas.0610082104

Tacikowski, P., Cygan, H. B., & Nowicka, A. (2014). Neural correlates of own and close-otherâ€™s name recognition: ERP evidence. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00194

Tager-Flusberg, H. (1991). Semantic processing in the free recall of autistic children: Further evidence for a cognitive deficit. *British Journal of Developmental Psychology*, *9*(3), 417–430. https://doi.org/10.1111/j.2044-835X.1991.tb00886.x

Tager-Flusberg, H. (2007). Evaluating the Theory-of-Mind Hypothesis of Autism. *Current Directions in Psychological Science*, *16*(6), 311–315. https://doi.org/10.1111/j.1467-8721.2007.00527.x

Tager-Flusberg, H., & Joseph, R. M. (2005). How Language Facilitates the Acquisition of False-Belief Understanding in Children with Autism. In *Why Language Matters for Theory of Mind* (pp. 298–318). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195159912.003.0014

Tanaka, J. W., & Curran, T. (2001). A Neural Basis for Expert Object Recognition. *Psychological Science*, *12*(1), 43–47. https://doi.org/10.1111/1467-9280.00308

Tateuchi, T., Itoh, K., & Nakada, T. (2012). Neural mechanisms underlying the orienting response to subject's own name: An event-related potential study. *Psychophysiology*, *49*(6), 786–791. https://doi.org/10.1111/j.1469-8986.2012.01363.x

Terrett, G., Rendell, P. G., Raponi-Saunders, S., Henry, J. D., Bailey, P. E., & Altgassen, M. (2013). Episodic Future Thinking in Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *43*(11), 2558–2568. https://doi.org/10.1007/s10803-013-1806-y

Thye, M. D., Bednarz, H. M., Herringshaw, A. J., Sartin, E. B., & Kana, R. K. (2018). The impact of atypical sensory processing on social impairments in autism spectrum disorder. *Developmental Cognitive Neuroscience*, *29*, 151–167. https://doi.org/10.1016/j.dcn.2017.04.010

Toichi, M., & Kamio, Y. (2002). A STUDY OF SELF-CONCEPT IN AUTISTIC ADOLESCENTS. *Japanese Journal of Child and Adolescent Psychiatry.*

Trilla, I., Wnendt, H., & Dziobek, I. (2020). Conditional effects of gaze on automatic imitation: the role of autistic traits. *Scientific Reports*, *10*(1), 15512. https://doi.org/10.1038/s41598-020-72513-6

Trimble, M. R. (1986). The Frontal Lobes. By Donald T. Stuss and Frank D. Benson. New York: Raven Press. 1986. Pp 303. $65.50. *British Journal of Psychiatry*, *149*(3), 393–394. https://doi.org/10.1192/S000712500021386X

Tronick, E., Als, H., Adamson, L., Wise, S., & Brazelton, T. B. (1978). The Infant's Response to Entrapment between Contradictory Messages in Face-to-Face Interaction. *Journal of the American Academy of Child Psychiatry*, *17*(1), 1–13. https://doi.org/10.1016/S0002-7138(09)62273-1

Tulving, E. (2005). Episodic Memory and Autonoesis: Uniquely Human? In *The Missing Link in CognitionOrigins of self-reflective consciousness* (pp. 3–56). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195161564.003.0001

Turk, D. J., Cunningham, S. J., & Macrae, C. N. (2008). Self-memory biases in explicit and incidental encoding of trait adjectives. *Consciousness and Cognition*, *17*(3), 1040–1045. https://doi.org/10.1016/j.concog.2008.02.004

Turk, D. J., van Bussel, K., Waiter, G. D., & Macrae, C. N. (2011). Mine and Me: Exploring the Neural Basis of Object Ownership. *Journal of Cognitive Neuroscience*, *23*(11), 3657–3668. https://doi.org/10.1162/jocn_a_00042

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Uddin, L. Q., Davies, M. S., Scott, A. A., Zaidel, E., Bookheimer, S. Y., Iacoboni, M., & Dapretto, M. (2008). Neural Basis of Self and Other Representation in Autism: An fMRI Study of Self-Face Recognition. *PLoS ONE*, *3*(10), e3526. https://doi.org/10.1371/journal.pone.0003526

Uljarevic, M., & Hamilton, A. (2013). Recognition of Emotions in Autism: A Formal Meta-Analysis. *Journal of Autism and Developmental Disorders*, *43*(7), 1517–1526. https://doi.org/10.1007/s10803-012-1695-5

van Boxtel, J. J. A., & Lu, H. (2013). A predictive coding perspective on autism spectrum disorders. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00019

Van de Cruys, S., Evers, K., Van der Hallen, R., Van Eylen, L., Boets, B., De-Wit, L., & Wagemans, J. (2014). Precise minds in uncertain worlds: Predictive coding in autism. *Psychological Review*, *121*(4), 649–675. https://doi.org/10.1037/a0037665

van den Bos, M., Cunningham, S. J., Conway, M. A., & Turk, D. J. (2010). Mine to remember: The impact of ownership on recollective experience. *Quarterly Journal of Experimental Psychology*, *63*(6), 1065–1071. https://doi.org/10.1080/17470211003770938

Viberg, Å. (1983). The verbs of perception: a typological study. *Linguistics*, *21*(1). https://doi.org/10.1515/ling.1983.21.1.123

Vinet, L., & Zhedanov, A. (2010). A "missing" family of classical orthogonal polynomials. *Behavioral Research Methods*. https://doi.org/10.1088/1751-8113/44/8/085201

Vogeley, K., & Fink, G. R. (2003). Neural correlates of the first-person-perspective. *Trends in Cognitive Sciences*, *7*(1), 38–42. https://doi.org/10.1016/S1364-6613(02)00003-7

Wallace, M. T., & Stevenson, R. A. (2014). The construct of the multisensory temporal binding

window and its dysregulation in developmental disabilities. *Neuropsychologia*, *64*, 105–123. https://doi.org/10.1016/j.neuropsychologia.2014.08.005

Wang, Y., & Hamilton, A. F. de C. (2012). Social top-down response modulation (STORM): a model of the control of mimicry in social interaction. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00153

Waterhouse, L., Fein, D., & Modahl, C. (1996). Neurofunctional mechanisms in autism. *Psychological Review*, *103*(3), 457–489. https://doi.org/10.1037/0033-295X.103.3.457

Wechsler, D. (1997). Wechsler adult intelligence scale - Third Edition (WAIS-III). *San Antonio*.

Wechsler, D. (2008). Wechsler adult intelligence scale–Fourth Edition (WAIS–IV). *San Antonio, TX: NCS Pearson*.

Wenzel, A. (2017). Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5). In *The SAGE Encyclopedia of Abnormal and Clinical Psychology*. SAGE Publications, Inc. https://doi.org/10.4135/9781483365817.n433

Wheeler, M. A., Stuss, D. T., & Tulving, E. (1997). Toward a theory of episodic memory: The frontal lobes and autonoetic consciousness. *Psychological Bulletin*, *121*(3), 331–354. https://doi.org/10.1037/0033-2909.121.3.331

White, S. J., Frith, U., Rellecke, J., Al-Noor, Z., & Gilbert, S. J. (2014). Autistic adolescents show atypical activation of the brain's mentalizing system even without a prior history of mentalizing problems. *Neuropsychologia*, *56*, 17–25. https://doi.org/10.1016/j.neuropsychologia.2013.12.013

Williams, D. M., Nicholson, T., & Grainger, C. (2018). The Self-Reference Effect on Perception : Undiminished in Adults with Autism and No Relation to Autism Traits. *Autism Research*, *11*(2), 331–341. https://doi.org/10.1002/aur.1891

Williams, J.H.G., Whiten, A., Suddendorf, T., & Perrett, D. I. (2001). Imitation, mirror neurons and autism. *Neuroscience & Biobehavioral Reviews*, *25*(4), 287–295. https://doi.org/10.1016/S0149-7634(01)00014-8

Williams, Justin H.G. (2008). Self–other relations in social development and autism: multiple roles for mirror neurons and other brain bases. *Autism Research*, *1*(2), 73–90. https://doi.org/10.1002/aur.15

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*(3), 331–335. https://doi.org/10.1007/s10803-005-3300-7

Wuyun, G., Wang, J., Zhang, L., Wang, K., Yi, L., & Wu, Y. (2020). Actions Speak Louder Than Words: The Role of Action in Self-Referential Advantage in Children With Autism. *Autism Research*, *13*(5), 810–820. https://doi.org/10.1002/aur.2274

Young, L., Camprodon, J. A., Hauser, M., Pascual-Leone, A., & Saxe, R. (2010). Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proceedings of the National Academy of Sciences*, *107*(15), 6753–6758. https://doi.org/10.1073/pnas.0914826107

Zhao, S., Uono, S., Yoshimura, S., & Toichi, M. (2018). A functional but atypical self: Influence of self-relevant processing on the gaze cueing effect in autism spectrum disorder. *Autism Research*, *11*(11), 1522–1531. https://doi.org/10.1002/aur.2019

APPENDICES

Experiment 3 Supplementary Materials

Table 1. RT and accuracy descriptive statistics in the neurotypical and ASC groups, Experiment 3.1.

		RT (msec)		Accuracy (%)	
	Referent	M	SD	M	SD
Match	Self	642	78	46.4	8.8
	Friend	695	82	42.3	13.4
	Stranger	730	88	35.9	14.0
Mismatch	Self	739	80	43.0	10.6
	Friend	762	85	39.3	10.9
	Stranger	754	82	40.7	9.7

Table 2. RT and accuracy descriptive statistics in the neurotypical and ASC groups, Experiment 3.2.

		Neurotypical (N = 27)				ASC (N = 30)			
		RT (msec)		Accuracy (%)		RT (msec)		Accuracy (%)	
	Referent	M	SD	M	SD	M	SD	M	SD
Match	Self	692	58	46.3	10.0	679	16	46.5	12.5
	Friend	731	86	41.7	12.2	734	20	38.9	14.9
	Stranger	774	86	33.3	14.0	771	21	35.1	14.8
Mismatch	Self	775	68	44.2	11.2	793	20	40.7	12.8
	Friend	797	70	40.8	12.4	808	22	38.6	13.5
	Stranger	793	70	49.2	7.3	813	20	41.2	11.3

Table 3. Participant characteristics for the sub-sample who were included in ERP analyses of Experiment 3.2. Note: some tasks have missing data, so statistics are reported on maximal data for each task.

	Autistic (N = 25; 19 male)				Neurotypical (N = 24; 16 male)						
	Mean	SD	Range	N	Mean	SD	Range	N	*t*	*p*	*d*
Age (years)	32.04	10.22	20-54	25	39.35	15.37	20-69	24	1.97	.06	0.56
IQ: Full Scale	111.4	11.77	89-137	25	111.09	9.34	97-133	23	0.10	.92	0.03
IQ: Verbal 6, 23)	106.04	10.97	91-126	25	110.26	10.97	92-129	23	1.33	.19	0.39
IQ: Performance	115.16	14.68	85-143	25	110.65	11.32	89-144	23	1.18	.24	0.34
ADOS: Total	8.13	6.01	1-21	16	-	-	-	-	-	-	-
AQ	29.79	9.86	0-45	24	13.55	6.11	3-28	22	6.65	<.001	1.96
TAS-20	59.79	16.86	0-81	24	43.35	10.92	28-65	20	3.75	<.001	1.14
Animations	5.4	2.17	2-8	10	5.5	2.07	2-8	6	0.09	.93	0.05
EQ	26.33	11.22	13-55	24	47.91	11.01	26-67	22	6.57	<.001	1.94
BAI	24.42	10.82	6-49	24	11.79	8.01	0-30	24	4.60	<.001	1.33
CAT-Q	120.04	25.94	55-162	24	81.25	17.11	38-107	24	6.12	<.001	1.77

Table 4. Statistical effects over the N1 time window in Experiment 3.2. Asterisks show significance of effects, where * $p < .05$; ** $p < .01$; *** $p < .001$.

		df	*F*	***p***	η_p^2
N1 Midline	Trial type	1, 47	1.59	0.21	0.03
	Referent	2, 94	0.11	0.9	0.002
	Group	1, 47	0.05	0.82	0.001
	Electrode	3, 141	13.34	<.001	0.22
	Trial Type x Referent	2, 94	1.39	0.26	0.03
	Trial Type x Group	1, 47	0.002	0.96	0
	Trial Type x Electrode	3, 141	1.72	0.17	0.06
	Referent x Group	2, 94	0.3	0.74	0.006
	Referent x Electrode	6, 282	1.7	0.12	0.06
	Group x Electrode	3, 141	1.54	0.21	0.03
	Trial Type x Referent x Group	2, 94	0.53	0.59	0.01
	Trial Type x Referent x Electrode	6, 282	4.03	<.001	0.08
	Trial Type x Group x Electrode	3, 141	3.04	0.03	0.06
	Referent x Group x Electrode	6, 282	1.81	0.1	0.04
	Trial Type x Referent x Group x Electrode	6, 282	1.21	0.3	0.03
N1 Lateral	Trial type	1, 47	1.03	0.32	.02
	Referent	2, 94	0.29	0.75	.006
	Group	1, 47	0.15	0.71	.003
	Ant-Pos	1, 47	18.50	<.001	.28
	Hemisphere	1, 47	5.18	0.03	.10
	Trial Type x Referent	2, 94	2.80	0.07	.06
	Trial Type x Group	1, 47	0.68	0.41	.01
	Trial Type x Ant-Pos	1, 47	0.17	0.68	.004
	Trial Type x Hemisphere	1, 47	0.18	0.67	.004
	Referent x Group	2, 94	2.00	0.14	.04
	Referent x Ant-Pos	2, 94	8.81	<.001	.16
	Referent x Hemisphere	2, 94	2.11	0.13	.04

	Group x Ant-Pos	1, 47	1.33	0.26	.03
	Group x Hemisphere	1, 47	0.90	0.35	.02
	Ant-Pos x Hemisphere	1, 47	0.03	0.87	.001
	Trial Type x Referent x Group	2, 94	0.46	0.64	.01
	Trial Type x Referent x Ant-Pos	2, 94	3.13	0.05	.06
	Trial Type x Referent x Hemisphere	2, 94	1.29	0.28	.03
	Trial Type x Group x Ant-Pos	1, 47	5.60	0.02	.11
	Trial Type x Group x Hemisphere	1, 47	0.62	0.44	.01
	Trial Type x Ant-Pos x Hemisphere	1, 47	0.17	0.69	.004
	Referent x Group x Ant-Pos	2, 94	0.66	0.52	.01
	Referent x Group x Hemisphere	2, 94	0.19	0.83	.004
	Referent x Ant-Pos x Hemisphere	2, 94	0.19	0.82	.004
	Group x Ant-Pos x Hemisphere	1, 47	0.03	0.87	.001
	Trial Type x Referent x Group x Ant-Pos	2, 94	1.25	0.29	.03
	Trial Type x Group x Ant-Pos x Hemisphere	1, 47	0.09	0.77	0.002
	Referent x Group x Ant-Pos x Hemisphere	2, 94	0.69	0.51	0.01
	Trial Type x Referent x Group x Hemisphere	2, 94	0.72	0.49	0.02
	Trial Type x Referent x Ant-Pos x Hemisphere	2, 94	1.59	0.21	0.03
	Trial Type x Referent x Group x Ant-Pos x Hemisphere	2, 94	1.4	0.25	0.03
N2 Midline	Trial type	1, 47	1.59	0.21	0.03
	Referent	2, 94	0.11	0.9	0.002
	Group	1, 47	0.05	0.82	0.001
	Electrode	3, 141	13.34	<.001	0.22
	Trial Type x Referent	2, 94	1.39	0.26	0.03
	Trial Type x Group	1, 47	0.002	0.96	0
	Trial Type x Electrode	3, 141	1.72	0.17	0.06
	Referent x Group	2, 94	0.3	0.74	0.006
	Referent x Electrode	6, 282	1.7	0.12	0.06
	Group x Electrode	3, 141	1.54	0.21	0.03
	Trial Type x Referent x Group	2, 94	0.53	0.59	0.01

	Trial Type x Referent x Electrode	6, 282	4.03	<.001	0.08
	Trial Type x Group x Electrode	3, 141	3.04	0.03	0.06
	Referent x Group x Electrode	6, 282	1.81	0.1	0.04
	Trial Type x Referent x Group x Electrode	6, 282	1.21	0.3	0.03
	Trial type	1, 47	3.98	0.05	0.08
	Referent	2, 94	3.52	0.03	0.07
	Group	1, 47	0.55	0.46	0.01
	Ant-Pos	1, 47	0.001	0.97	0
	Hemisphere	1, 47	4.37	0.04	0.09
	Trial Type x Referent	2, 94	1.75	0.18	0.04
	Trial Type x Group	1, 47	0.15	0.7	0.003
	Trial Type x Ant-Pos	1, 47	1.1	0.3	0.02
	Trial Type x Hemisphere	1, 47	0.003	0.96	0
	Referent x Group	2, 94	4.13	0.02	0.08
	Referent x Ant-Pos	2, 94	6.48	0.002	0.12
	Referent x Hemisphere	2, 94	1.06	0.35	0.02
	Group x Ant-Pos	1, 47	1.52	0.23	0.03
	Group x Hemisphere	1, 47	0.89	0.35	0.02
	Ant-Pos x Hemisphere	1, 47	0.55	0.46	0.01
	Trial Type x Referent x Group	2, 94	2.99	0.06	.06.
	Trial Type x Referent x Ant-Pos	2. 94	0.74	0.48	0.02
	Trial Type x Referent x Hemisphere	2, 94	0.33	0.72	0.007
	Trial Type x Group x Ant-Pos	1, 47	1.95	0.17	0.04
	Trial Type x Group x Hemisphere	1, 47	0.02	0.88	0.001
	Trial Type x Ant-Pos x Hemisphere	1, 47	0.02	0.9	0
	Referent x Group x Ant-Pos	2, 94	0.06	0.94	0.001
	Referent x Group x Hemisphere	2, 94	0.06	0.95	0.001
	Referent x Ant-Pos x Hemisphere	2, 94	1.13	0.33	0.02
N2 Lateral	Group x Ant-Pos x Hemisphere	1, 47	0.89	0.35	0.02
	Trial Type x Referent x Group x Ant-Pos	2, 94	0.2	0.82	0.004

	Trial Type x Group x Ant-Pos x Hemisphere	1, 47	1.96	0.17	0.04
	Referent x Group x Ant-Pos x Hemisphere	2, 94	1.13	0.33	0.02
	Trial Type x Referent x Group x Hemisphere	2, 94	0.86	0.43	0.02
	Trial Type x Referent x Ant-Pos x Hemisphere	2, 94	0.05	0.95	0.03
	Trial Type x Referent x Group x Ant-Pos x Hemisphere	2, 94	1.14	0.24	0.03
P3 Midline	Trial Type	1, 47	10.05	0.003	0.19
	Referent	2, 94	3.54	0.03	0.07
	Group	1, 47	0.004	0.95	0
	Electrode	3, 141	11.91	<.001	0.2
	Trial Type x Referent	2, 94	1.18	0.31	0.03
	Trial Type x Group	1, 47	0.01	0.91	0
	Trial Type x Electrode	3, 141	2.52	0.07	0.05
	Referent x Group	2, 94	2.66	0.08	0.053
	Referent x Electrode	6, 282	3.12	0.006	0.06
	Group x Electrode	3, 141	2.87	0.04	0.06
	Trial Type x Referent x Group	2, 94	3.57	0.03	0.07
	Trial Type x Referent x Electrode	6, 282	1.33	0.24	0.03
	Trial Type x Group x Electrode	3, 141	1.69	0.17	0.04
	Referent x Group x Electrode	6, 282	0.87	0.52	0.02
	Trial Type x Referent x Group x Electrode	6, 282	1.07	0.38	0.02
P3 Lateral	Trial type	1, 47	7.36	0.009	0.14
	Referent	2, 94	3.71	0.03	0.07
	Group	1, 47	0.02	0.9	0
	Ant-Pos	1, 47	0.85	0.36	0.02
	Hemisphere	1, 47	9.84	0.003	0.17
	Trial Type x Referent	1, 94	0.85	0.43	0.02
	Trial Type x Group	1, 47	0.15	0.7	0.003
	Trial Type x Ant-Pos	1, 47	0.02	0.89	0
	Trial Type x Hemisphere	1, 47	0.05	0.82	0.001
	Referent x Group	2, 94	1.81	0.07	0.06

	Referent x Ant-Pos	2, 94	1.4	0.25	0.03
	Referent x Hemisphere	2, 94	1.91	0.15	0.04
	Group x Ant-Pos	1, 47	2.56	0.12	0.052
	Group x Hemisphere	1, 47	2.49	0.12	0.05
	Ant-Pos x Hemisphere	1, 47	5.15	0.03	0.1
	Trial Type x Referent x Group	2, 94	2.53	0.09	0.051
	Trial Type x Referent x Ant-Pos	2, 94	1.58	0.21	0.03
	Trial Type x Referent x Hemisphere	2, 94	0.49	0.61	0.01
	Trial Type x Group x Ant-Pos	1, 47	3.13	0.08	0.06
	Trial Type x Group x Hemisphere	1, 47	0.09	0.77	0.002
	Trial Type x Ant-Pos x Hemisphere	1, 47	0.004	0.95	0
	Referent x Group x Ant-Pos	2, 94	0.05	0.95	0.001
	Referent x Group x Hemisphere	2, 94	0.54	0.59	0.01
	Referent x Ant-Pos x Hemisphere	2, 94	0.09	0.92	0.002
	Group x Ant-Pos x Hemisphere	1, 47	0.03	0.86	0.001
	Trial Type x Referent x Group x Ant-Pos	2, 94	0.35	0.71	0.01
	Trial Type x Group x Ant-Pos x Hemisphere	1, 47	1.61	0.21	0.03
	Referent x Group x Ant-Pos x Hemisphere	2, 94	1.29	0.28	0.03
	Trial Type x Referent x Group x Hemisphere	2, 94	1.28	0.28	0.03
	Trial Type x Referent x Ant-Pos x Hemisphere	2, 94	0.08	0.92	0.002
	Trial Type x Referent x Group x Ant-Pos x Hemisphere	2, 94	0.89	0.41	0.02

Experiment 5 Supplementary Materials

Table 1. Participant characteristics for the sub-sample who were included in EEG analyses of Experiment 5.2. Note: some tasks have missing data, so statistics are reported on maximal data for each task.

	Autistic (n = 27; 18 male)				Neurotypical (n = 24; 14 male)						
	Mean	SD	Range	N	Mean	SD	Range	N	*t*	*p*	*d*
Age (years)	34.66	11.75	20-60	27	39.35	15.37	20-69	24	1.23	.22	0.35
IQ: Full Scale	110.22	13.59	73-137	27	111.09	9.34	97-133	23	0.26	.80	0.07
IQ: Verbal	105.37	11.16	86-129	27	110.26	10.97	92-129	23	1.56	.13	0.44
IQ: Performance	113.63	17.20	65-143	27	110.65	11.32	89-144	23	0.71	.48	0.20
ADOS: Total	7.24	5.12	1-21	17	-	-	-	0	-	-	-
AQ	29.12	10.14	0-45	26	13.55	6.11	3-28	22	6.30	<.001	1.82
TAS-20	58.54	16.96	0-81	26	43.35	10.92	28-65	20	3.48	.001	1.04
Animations	5.40	2.17	2-8	10	5.50	2.07	3-8	6	0.09	.93	0.05
EQ	27.15	12.13	13-55	26	47.91	11.01	26-67	22	6.16	<.001	1.78
BAI	25.42	11.15	6-49	26	11.79	8.01	0-30	24	4.93	<.001	1.40
CAT-Q	118.00	27.08	55-162	26	81.25	17.11	38-107	24	5.68	<.001	1.61

Table 2. RT and accuracy descriptive statistics in the neurotypical and ASC groups in Experiment 5.2.

		Neurotypical (n = 27)				ASC (n = 31)			
		RT		Accuracy		RT		Accuracy	
Referent	Compatibility	M	SD	M	SD	M	SD	M	SD
SELF	IC, SC	0.590	0.045	43.593	3.522	0.595	0.070	40.226	9.989
	II, SI	0.604	0.042	40.519	8.920	0.625	0.091	35.258	12.383
	IC, SI	0.597	0.049	42.889	5.071	0.606	0.066	39.000	9.737
	II, SC	0.596	0.052	39.667	8.880	0.616	0.089	35.968	12.711
ADULT	IC, SC	0.590	0.047	43.407	3.993	0.603	0.067	40.548	8.861
	II, SI	0.602	0.046	40.741	8.374	0.626	0.085	35.871	12.842
	IC, SI	0.593	0.044	43.333	4.624	0.610	0.073	39.806	9.565
	II, SC	0.598	0.049	40.926	7.893	0.616	0.095	35.548	12.511
CHILD	IC, SC	0.587	0.044	43.667	3.595	0.606	0.074	40.097	9.152

II, SI	0.606	0.050	40.185	9.463	0.627	0.094	35.323	12.360
IC, SI	0.598	0.042	42.704	5.261	0.608	0.070	38.903	10.326
II, SC	0.597	0.053	40.556	8.599	0.617	0.084	36.548	12.299

*IC = imitatively compatible; II = imitatively incompatible; SC = spatially compatible, SI = spatially incompatible.

Table 3. RT and accuracy descriptive statistics for each condition in Experiment 5.1.

		RT (msec)		Accuracy (%)	
Referent	Compatibility	M	SD	M	SD
SELF	IC, SC	0.599	0.043	39.778	8.438
	II, SI	0.612	0.047	36.944	9.636
	IC, SI	0.607	0.047	39.093	8.192
	II, SC	0.606	0.047	37.37	9.268
ADULT	IC, SC	0.599	0.044	39.426	8.621
	II, SI	0.614	0.048	36.889	9.528
	IC, SI	0.612	0.045	38.889	8.54
	II, SC	0.607	0.047	38.019	9.614
CHILD	IC, SC	0.605	0.045	39.278	8.786
	II, SI	0.608	0.049	36.667	9.286
	IC, SI	0.612	0.047	39.63	9.124
	II, SC	0.610	0.046	37.593	8.861

*IC = imitatively compatible; II = imitatively incompatible; SC = spatially compatible, SI = spatially incompatible.

Table 4. ERP ANOVA effects in Experiment 5.2.

		df	*F*	*p*	η_p^2
8-13Hz	Site	1, 49	39.69	<.001	0.45
	Agent	2, 98	1.1	0.34	0.02
	Imitative	1, 49	37.45	<.001	0.43
	Group	1, 49	0.79	0.38	0.02
	Group x site	1, 49	0.03	0.86	0.001
	Group x agent	2, 98	1.65	0.2	0.03
	Group x imitative	1, 49	0.11	0.75	0.002
	Site x agent	2, 98	3.34	0.04	0.06
	Site x imitative	1, 49	12.63	<.001	0.21
	Agent x imitative	2, 98	28.17	<.001	0.37
	Site x referent x group	2, 98	0.55	0.58	0.01
	Site x imitative x group	1, 49	0.61	0.44	0.01
	Referent x imitative x group	2, 98	0.03	0.97	0.001
	Site x referent x imitative	2, 98	10.38	<.001	0.18
	Site x referent x imitative x group	2, 98	0.19	0.77	0.004
13-35Hz	Site	1, 49	13	<.001	0.21
	Referent	2, 98	2.66	0.08	0.05
	Imitative	1, 49	0.44	0.51	0.01
	Group	1, 49	4.46	0.04	0.08
	Group x site	1, 49	1.76	0.19	0.35
	Group x referent	2, 98	0.9	0.41	0.02
	Group x imitative	1, 49	1.44	0.24	0.03
	Site x referent	2, 98	0.94	0.39	0.02
	Site x imitative	1, 49	0.83	0.37	0.02
	Referent x imitative	2, 98	19.16	<.001	0.28
	Site x referent x group	2, 98	1.69	0.19	0.03
	Site x imitative x group	1, 49	6.31	0.02	0.11
	Referent x imitative x group	2, 98	0.29	0.75	0.01
	Site x referent x imitative	2, 98	41.21	<.001	0.56
	Site x referent x imitative x group	2, 98	1.25	0.29	0.03