The Influence of Reward on Cognitive Empathy

Thesis submitted for the degree of Doctor of Philosophy

Kristian Hjalmer Tempelmans Plat

June 2020

The School of Psychology and Clinical Language Sciences

Declaration of original authorship

I confirm that this is my own work, and the use of all materials from other sources has been properly and fully acknowledged.

Kristian Hjalmer Tempelmans Plat (PhD candidate)

Abstract

We mimic people more if we like them. This observation suggests that certain component processes of empathy can be modulated by reward. Crucially, this reward-driven modulation of mimicry is decreased in people with high autistic symptoms. Autism has also been associated with impairments in other, more cognitive, components of empathy. However, it remains ill-investigated whether reward similarly modulates these cognitive component processes of empathy and, in turn, whether this modulation is altered in individuals with high autistic symptoms.

The current thesis aims to address this gap in the literature. Commonly used tasks requiring judgments about others' mental states were used to index different processes related to cognitive empathy. In separate experiments, this was done for Theory of Mind (Chapter 2), perspective taking (Chapter 3), and empathic accuracy (Chapter 4). The impact of reward on task performance was tested through manipulating the reward value associated with the 'other'. The analyses show that reward value associated with the other positively affects Theory of Mind performance, but only when the participant is required to make an explicit judgment about the other's belief. Additionally, there is some evidence that reward aids visual perspective taking. Lastly, reward benefits making inferences about how positive or negative the other feels.

However, the results show that the reward-driven modulation of cognitive empathy-related processes is not related to autistic traits. Interestingly, the various employed task measures of cognitive empathy were not related to each other (Chapter 5). Taken together, the analyses in this thesis indicate that reward

improves performance on some cognitive empathy tasks, likely mediated by the employment of underlying cognitive capacities that are unaffected in those with high autistic traits.

Contributions

Below, the reader finds an overview of my contributions and the contributions of others to the studies discussed in the experimental chapters.

Chapter 2

Mouse-tracking task study

- Kristian Tempelmans Plat: design, programming, collection, analysis
- Anthony Haffey: programming

Sandbox task study

- Simon Brett: design, programming, collection, analysis
- Kristian Tempelmans Plat: analysis

Chapter 3

Director task study

- Garret O'Connell: design, programming
- Kristian Tempelmans Plat: design, programming, collection, analysis

Dots task study

- Kristian Tempelmans Plat: design, programming, analysis
- Kelsey Britton, Gemma Bullard, Rosie Harrison, Georgia Matthews: collection

Chapter 4

Video rating task study

- Kristian Tempelmans Plat: design, programming, collection, analysis

- Shivani Arora, Christina Chan, Zaneta Filipczak, Jemima Holloway, Eleanor Hounds, Rosie Howard, Nadine Loefflad, Rose Narey: collection

Chapter 5

Task battery study

- Anthony Haffey: design, programming
- Aoife O'Mahoney, Safi Johal-Ayub: collection
- Kristian Tempelmans Plat: design, programming, analysis

Acknowledgements

Doing a PhD abroad is an extraordinary thing. You find yourself writing to a professor and all of a sudden you commit yourself to a project in the UK for 3+ years. Adapting to a new environment and culture comes in addition to the challenges inherent to a PhD project. I know *I* sang the second-year blues with passion. The past few years have been turbulent, challenging, but also exhilarating - a veritable rollercoaster with ups and downs I bet familiar to others pursuing a PhD degree.

My time at the University of Reading has given me a lot, but most of all a family - both in the academic and the common sense of the word. First of all, I would like to take this opportunity to thank everyone at BhismaLab. Lunch sessions, social gatherings and dinners, they were all memorable, and I regret not tagging along to more events than I have. In particular I would like to thank Nick Thompson, fellow bibliophile, who was always my first port of call for books, brainstorming and banter. Also, Anthony Haffey, most esteemed member of BhismaLab, whose coding wizardry and general wisdom were invaluable, far beyond the realms of Collector. Without him, I would never have handed in my thesis - quite literally.

Outside of BhismaLab, I would like to thank all members of the MeMo lab. I am still quite flattered that my name suggestion for the lab group was picked. A special word of thanks goes out to Nico Biagi, who criticised my peculiar methods in Italian cuisine (and justly!), but whom I have come to regard as a brother.

My move to the UK also resulted in finding a different type of family. I thank Elise for always being there for me, even when malfunctioning MATLAB scripts kept me from coming home on time for dinner. Running into you in the SPCLS corridors made a huge difference. I also thank Nigel ('Dad'), who provided support when I was in dire need.

I would also like to thank everyone who put up with listening to my poetry. My time in Reading was certainly a fecund source of inspiration. I even considered writing these acknowledgements in rhyme, but that would have put undesirable constraints on phrasing my utmost gratitude to my supervisors: Professor Kou Murayama, who taught me more about statistics than a course could ever do and, of course, Professor Bhisma Chakrabarti. From insomnia because the allocation of my scholarship was yet undecided, to countless meetings that always re-instilled me with motivation, to making me grow as a researcher, Bhisma showed his support in many ways. Infinite thanks.

Lastly, one big thank you to my friends and family in the Netherlands. In particular: Niels, for being the big brother I look up to; Mum and Dad for being indefatigable sources of support and love. Thank you for always being there for me during my PhD (and otherwise), regardless of distance, and for generally making me believe in myself.

I can definitely say that this experience has been most valuable, and if one thing is for certain, it is that I will never forget my time in Reading. For all of those not specifically mentioned here: thank you too.

Table of contents

Declaration of original authorship **3**

Abstract **4**

Contributions **6**

Acknowledgements **8**

Table of contents **10**

List of figures **16**

List of tables **19**

1 Introduction **21**

1.1 Component processes of empathy 22

1.1.1 Affective and cognitive empathy processes 22

1.1.2 Empathy and the autism spectrum 32

1.2 The influence of reward on empathy-related processes 38

1.2.1 Manipulating reward value 38

1.2.2 Reward and affective empathy 41

1.2.3 Reward and cognitive empathy 43

2 Reward and belief tracking **47**

2.1 Introduction 48

2.2 Aims 53

2.3 Predictions 55

2.4 Mouse-tracking study 55

2.4.1 Methods 55

2.4.1.1 Participants 55

2.4.1.2 Stimuli 56

2.4.1.2.1 Evaluative conditioning task 56

2.4.1.2.2 Implicit belief tracking task (mouse-tracking) 57

2.4.1.3 Procedure 57

2.4.1.3.1 Evaluative conditioning task 57

2.4.1.3.2 Implicit belief tracking task (mouse-tracking) 60

2.4.1.3.3 Trait measure 63

2.4.1.4 Data analysis 64

2.4.2 Results 65

2.4.2.1 Data pre-processing 65

2.4.2.2 Implicit belief tracking (replication analysis) 67
2.4.2.3 The effect of reward 67
2.4.2.4 Individual differences in the reward effect.. 68
2.4.3 Brief discussion.. 70
2.5 Sandbox study 71
2.5.1 Methods 71
2.5.1.1 Participants 71
2.5.1.2 Stimuli 71
2.5.1.2.1 Evaluative conditioning task 71
2.5.1.2.2 Explicit belief tracking task (Sandbox) 72
2.5.1.3 Procedure.. 72
2.5.1.3.1 Evaluative conditioning task 72
2.5.1.3.2 Explicit belief tracking task (Sandbox) 73
2.5.1.3.3 Trait measure... 76
2.5.1.3.4 Recall task 76
2.5.1.4 Data analysis... 76
2.5.2 Results... 77
2.5.2.1 Data pre-processing 77
2.5.2.2 Explicit belief tracking and the reward effect.................................... 78
2.5.2.3 Individual differences in the reward effect.. 80
2.5.3 Brief discussion.. 81
2.6 Discussion.. 84
2.7 Conclusion 88

3 Reward and perspective taking.. 89
3.1 Introduction 90
3.2 Aims.. 95
3.3 Predictions 96
3.4 Director task study... 97
3.4.1 Methods.. 97
3.4.1.1 Participants 97
3.4.1.2 Stimuli 98
3.4.1.2.1 Evaluative conditioning task 98
3.4.1.2.2 Director task... 98
3.4.1.3 Procedure.. 98
3.4.1.3.1 Evaluative conditioning task 99
3.4.1.3.2 Director task... 100

3.4.1.3.3 Trait measure 106
3.4.1.4 Data analysis 106
3.4.2 Results 109
3.4.2.1 Data pre-processing 109
3.4.2.2 Perspective taking 111
3.4.2.3 The effect of reward 111
3.4.2.4 Individual differences in the reward effect 113
3.4.2.5 Perspective taking as the difference between Experimental and Control 113
3.4.3 Brief discussion 114
3.5 Dots task study 115
3.5.1 Methods 115
3.5.1.1 Participants 115
3.5.1.2 Stimuli 116
3.5.1.2.1 Evaluative conditioning task 116
3.5.1.2.2 Dots task 116
3.5.1.3 Procedure 117
3.5.1.3.1 Evaluative conditioning task 117
3.5.1.3.2 Dots task 117
3.5.1.3.3 Trait measure 122
3.5.1.4 Data analysis 122
3.5.2 Results 124
3.5.2.1 Data pre-processing 124
3.5.2.2 Perspective taking and the reward effect 125
3.5.2.3 Individual differences in the reward effect 128
3.5.2.4 Individual differences in perspective taking 130
3.5.3 Brief discussion 130
3.6 Discussion 132
3.7 Conclusion 135

4 Reward and empathic accuracy 136
4.1 Introduction 138
4.2 Aims 141
4.3 Predictions 142
4.4 Methods 143
4.4.1 Participants 143
4.4.2 Stimuli 144

4.4.3 Procedure 144
4.4.3.1 Session 1 144
4.4.3.1.1 Video recording and self-ratings 144
4.4.3.1.2 Trait measure 148
4.4.3.1.3 Post-session 148
4.4.3.2 Session 2 148
4.4.3.2.1 Social closeness questionnaire 149
4.4.3.2.2 Other-ratings 149
4.4.4 Data analysis 150
4.5 Results 151
4.5.1 Data pre-processing 151
4.5.2 Preliminary analyses 152
4.5.3 The effect of reward 154
4.5.3.1 Empathic accuracy for valence 154
4.5.3.2 Empathic accuracy for arousal 155
4.5.4 Individual differences in the reward effect 155
4.5.4.1 Empathic accuracy for valence 155
4.5.4.2 Empathic accuracy for arousal 156
4.6 Discussion 156
4.7 Conclusion 159

5 The interrelationship between putative measures of cognitive empathy 161
5.1 Introduction 162
5.2 Aims 164
5.3 Predictions 165
5.4 Methods 165
5.4.1 Participants 165
5.4.2 Stimuli 166
5.4.2.1 Director task 166
5.4.2.2 Sandbox task 167
5.4.2.3 Reading the Mind in the Eyes task 168
5.4.2.4 Hinting task 168
5.4.2.5 Mouse-tracking task 168
5.4.2.6 Video rating task 168
5.4.3 Procedure 169
5.4.3.1 Trait measure 169
5.4.3.2 Pre-battery introduction 169

5.4.3.3 Director task 170
5.4.3.4 Sandbox task 172
5.4.3.5 Reading the Mind in the Eyes task 173
5.4.3.6 Hinting task 173
5.4.3.7 Mouse-tracking task 174
5.4.3.8 Video rating task 178
5.4.4 Data analysis 179
5.4.4.1 Director task 180
5.4.4.2 Sandbox task 181
5.4.4.3 Reading the Mind in the Eyes task 181
5.4.4.4 Hinting task 181
5.4.4.5 Mouse-tracking task 182
5.4.4.6 Video rating task 185
5.4.5 Overview 186
5.5 Results 188
5.5.1 Pre-processing 188
5.5.1.1 Director task 188
5.5.1.2 Sandbox task 190
5.5.1.3 Reading the Mind in the Eyes task 190
5.5.1.4 Hinting task 190
5.5.1.5 Mouse-tracking task 190
5.5.1.6 Video rating task 191
5.5.1.7 Correlation analysis sample 192
5.5.2 Director task 193
5.5.3 Sandbox task 193
5.5.4 Reading the Mind in the Eyes task 194
5.5.5 Hinting task 194
5.5.6 Mouse-tracking task 194
5.5.7 Video rating task 195
5.5.8 Relationships between task- and trait measures 195
5.5.8.1 Analysis of individual differences 195
5.5.8.2 Effect of demographic variables 197
5.6 Discussion 198
5.7 Conclusion 204

6 General discussion 206
6.1 The effect of reward on cognitive empathy 211

6.1.1 The effect of reward on belief tracking 212
6.1.2 The effect of reward on perspective taking.. 213
6.1.3 The effect of reward on empathic accuracy.. 215
6.1.4 Task-dependence of reward effects.. 215
6.2 Individual differences in the effect of reward... 220
6.3 The interrelationship between cognitive empathy measures........................ 222
6.4 A comparison between cognitive and affective empathy............................. 223
6.5 Potential limitations and future directions 225
6.6 Conclusions... 228

References 230

Appendices

Appendix A - Autism Spectrum Quotient items and scoring 272
Appendix B - Faces shown in implicit belief tracking study (Chapter 2.4)....... 275
Appendix C - Change-of-location scenarios used in explicit belief tracking study (Chapter 2.5) 276
Appendix D - Faces shown in the Director task study (Chapter 3.4) 281
Appendix E - Cartoon images used for characters in the Sandbox task (Chapter 5.4.2.2) 282
Appendix F - Change-of-location scenarios used in the Sandbox task (Chapter 5.4.2.2) 285
Appendix G - Photographs of the eye region used for the Reading the Mind in the Eyes test (Chapter 5.4.2.3) 289
Appendix H - Scenarios used in the Hinting task (Section 5.4.2.4) with possible correct answers 295
Appendix I - Description of the subject matter of the videos in the Video rating task (Chapter 5.4.2.6) 301
Appendix J - Differences between Wave 1 and Wave 2 for the task battery Director task 303
Appendix K - Differences between Wave 1 and Wave 2 for the task battery mouse-tracking task.. 304

List of figures

Figure 1.1 Schematic representation of processes underlying empathy 30

Figure 2.1 Schematic overview of the evaluative conditioning task and face-reward associations 58

Figure 2.2 Schematic overview of a trial in the MT task 62

Figure 2.3 Null effect of reward on implicit belief tracking 68

Figure 2.4 The relationship between autistic traits and the effect of reward on belief tracking bias 69

Figure 2.5 Schematic overview of a trial in the explicit belief tracking task 75

Figure 2.6 The effect of reward on explicit belief tracking in FB trials 80

Figure 2.7 The relationship between autistic traits and the effect of reward on false belief bias 81

Figure 3.1 Example trial shown on paper before the Director task 101

Figure 3.2 Director task trial procedure 103

Figure 3.3 Effect of reward on visual perspective taking in the Director task 112

Figure 3.4 Dots perspective taking task design 120

Figure 3.5 Mean reaction time in the Dots task conditions 127

Figure 3.6 Interaction between Perspective and Reward on mean reaction time 128

Figure 3.7 The relationship between autistic traits and the effect of reward on egocentric interference 129

Figure 4.1 Example of a video rating task trial 146

Figure 4.2 The relationship between social distance and empathic accuracy for valence 155

Figure 5.1 Schematic overview of a trial in the task battery MT task 177

Figure 5.2 Schematic overview of the task battery mouse-tracking AUC calculation procedure 183

Figure 5.3 Correlogram for correlations between task- and trait measures of the task battery 197

Figure 5.4 Distribution of capacities underlying cognitive empathy for two fictive individuals 203

List of tables

Table 3.1 Accuracy data for the Director task. 109

Table 4.1 Descriptives for empathic accuracy in the video rating task 153

Table 5.1 Overview of the task measures in the task battery study 187

Table 5.2 Accuracy descriptives per condition of the task battery mouse-tracking task 191

Table 5.3 Descriptive statistics for the correlation analysis sample 192

Table 5.4 Pearson correlations between task- and trait measures of the task battery 196

Table 5.5 Correlations between age and task battery measures 197

Table 5.6 Effect of gender on task battery 198

Table 5.7 Hypothesised contributions of cognitive capacities to task performance 202

Table 6.1 Overview of findings in this thesis 210

1 Introduction

“You never really understand a person until you consider things from his point of view, until you climb inside of his skin and walk around in it.”

~ Atticus Finch, in Harper Lee’s ‘To Kill a Mockingbird’

This chapter sets out the background for the following experimental chapters. It first introduces and delineates the component processes of empathy, employing a central dichotomy of affective and cognitive empathy. Whereas affective empathy entails mechanisms through which someone else’s mental state can be shared, cognitive empathy relies on making inferences about such states. It is explained how the processes underlying these components are altered in autism, and how impaired social motivation may underlie such alterations. This account is supported by evidence that reward generally modulates affective empathy-related processes, but not for those high in autistic traits. In contrast, it is unclear whether processes related to cognitive empathy can also be modulated by reward, and whether this modulation varies along the spectrum of autistic traits. The chapter concludes with a discussion of studies that have looked at this relationship for affective and cognitive empathy, respectively. By summarising existing methodologies to manipulate reward, it provides a framework for experimentally investigating the research questions in the following chapters.

1.1 Component processes of empathy

Understanding others (i.e. their thoughts, feelings and intentions) lies at the core of social interaction. This understanding is often referred to as empathy. Although there are subtle differences in how researchers define empathy, it is generally divided into a cognitive and an affective component (e.g. Baron-Cohen & Wheelwright, 2004; Decety & Jackson, 2004; de Waal, 2008; Leiberg & Anders, 2006). The cognitive component encompasses processes that allow the understanding of others' mental states through inferential cognitive processing, whereas the affective component entails mechanisms through which someone else's emotional state can be spontaneously shared (Thompson, Uusberg, Gross, & Chakrabarti, 2019). Although the understanding of others can be arrived at through both forms of empathy individually (Soto & Levenson, 2009), the two components are thought to be interrelated and complementary (Zurek & Scheithauer, 2017; Baron-Cohen & Wheelwright, 2004; Decety & Jackson, 2006; Dziobek et al., 2008). The following sections will further delineate affective and cognitive empathy and discuss their importance in development and social interaction.

1.1.1 Affective and cognitive empathy processes

Affective empathy relies on mechanisms that can contribute to the sharing of someone else's affective state by the perceiver. Two mechanisms that have often been regarded as subserving affective empathy are emotional contagion and motor mimicry (Batson, 2009; de Waal & Preston, 2017). The evolutionary importance of these interrelated mechanisms is underlined by their presence in other primates, but also in some other mammalian and avian species (de Waal, 2008; de Waal & Preston, 2017). Emotional contagion is an adaptive process in which the affective state perceived in others is copied by the perceiver. Consider the following: when a parent sees a dangerous animal posing a threat, they respond to this situation with fear. Consequently, this fear is communicated to the infant, who in turn becomes afraid too. According to Preston and de Waal (2002), this emotional contagion relies on the perception-action mechanism. This mechanism is responsible for generating autonomic and somatic responses ('actions') based on the perception of an emotional state in others (Adolphs, 2002; de Waal, 2008). As a result, the perceiver resonates with the perceived emotional state. In other words, infants respond to affective states in their parent by matching that state (Sagi & Hoffman, 1976; Ungerer et al., 1990; Zahn-Waxler, & Radke-Yarrow, 1982; Simner, 1971). This perception-action link allows the infant to adapt to a situation.

Motor mimicry, and facial mimicry in particular, is commonly thought of as the main mechanism subserving emotional contagion (Bavelas, Black, Lemery, & Mullett, 1986; Hatfield, Cacioppo, & Rapson, 1994; Lundqvist & Dimberg, 1995), and affective empathy in general (Dimberg, Andréasson, & Thunberg, 2011; Hermans, Putman, & van Honk, 2006; Meltzoff, 2002; Sonnby-Borgström, 2002). When an infant sees its parent's emotional facial expression, it will generally mirror that facial expression (Field, Woodson, Greenberg, & Cohen, 1982; Meltzoff & Moore, 1977; Stern, 1977; but see Catmur, Walsh, & Heyes, 2009 and Oostenbroek et al., 2016, for evidence against innate imitation in early infancy). It has been suggested that this mirroring process depends on the mirror neuron system (Gallese, 2007). This account can be retraced to the discovery of a certain type of neurons in primates. It was found that these neurons were active when a monkey performed an action, but also when it observed that same action (di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti, Fadiga, Fogassi, & Gallese, 1996). These mirror neurons were therefore thought to underlie the perception-action link in primates. Subsequently, evidence has been found for the existence of a similar neuron system in the human brain (see Rizzolatti & Craighero, 2004, for a review). Later studies then identified this mirror neuron system as important for the matching of others' emotions through mimicry (see Bastiaansen, Thioux, & Keysers, 2009, for a review).

Mimicry is crucial for normal development. For example, in addition to its basic evolutionary purpose, it is a requisite for bonding in a parent-child relationship (Preston & de Waal, 2002; Fawcett & Liszkowski, 2012) and it aids vicarious learning (Bandura, 1977; Galef, 1988; Heyes, 1993; Piaget, 1946; Tomasello, Savage-Rumbaugh, & Kruger, 1993). Mimicry is arguably the primary medium through which an infant first gets acquainted with the social world. Not surprisingly, therefore, Feldman (2007) reported that behavioural synchrony between parent and child in the first years of life predicted empathy in adolescence. In short, mimicry lays the fundament for social interaction later in life.

Affective empathy and its constituent processes remain important into adulthood. For instance, Chartrand and Bargh (1999) found that mimicking others' motor behaviour results in liking between adult interaction partners. For this reason, mimicry was nicknamed "social glue" by these authors. Lakin, Jefferis, Cheng, and Chartrand (2003) consequently suggested that, whilst not denying its evolutionary importance, mimicry foremostly serves a social function in adults. Being able to resonate with others' emotions results in feelings of compassion and empathic concern (Bernhardt & Singer, 2012). Therefore, the expression of affective empathy has beneficial effects on relationships (Boele et al., 2019; Verhofstadt et al., 2016).

If affective empathy represents mechanisms underlying the *matching* of someone else's feelings, cognitive empathy can be thought of as the *reasoning* about these feelings. Unsurprisingly, because of their reliance on reasoning, cognitive empathy processes emerge later in development than those related to affective empathy (Hoffman, 1977; Preston & de Waal, 2002; Gallese, 2003; Decety & Jackson, 2004; Knafo et al., 2009). Similar to affective empathy, cognitive empathy can be further differentiated into component processes. Aan het Rot and Hogenelst (2014) suggested that, to investigate cognitive empathy, one can look at Theory of Mind, perspective taking and empathic accuracy (see below for definitions of these processes). The remainder of this section covers literature examining these respective processes.

According to Apperly (2012), Theory of Mind, or sometimes 'mindreading' (Apperly, 2011), can be thought of in three ways. First, it is the conceptual understanding that others have a mind, an 'internal world', of their own. This understanding gives rise to explanations, predictions and justifications of behaviour that are combined into a single 'theory' about the individual who lies on the receiving end of the mindreading. Second, Theory of Mind can also be understood as a set of cognitive processes. Making Theory of Mind inferences, holding them in mind, and applying them to social interaction relies on various executive processes, such as working memory and inhibitory control (Apperly, 2011; Davis & Pratt, 1995; Carlson, Moses, & Breton, 2002; Gordon & Olson, 1998; Mutter, Alcorn, & Welsh, 2006). Third, Theory of Mind can be regarded as an ability or propensity that can differ between individuals. It is especially this third approach that allows the investigation of what factors play a role in influencing Theory of Mind inferences.

A common way to measure Theory of Mind is through the use of so-called false belief paradigms. The flagship test for this purpose was devised by Wimmer and Perner (1983). In this task, designed for children, a character places an object in one container. After this character leaves the scene, the object is subsequently moved to a different container by a second character. When the first character returns, the child is asked where that character will look for the object. At this point in the story, the child has a true belief about the object's location, whereas the first character has a false belief. If the child subsequently indicates the object's original container, he/she is said to exhibit an understanding of the character's belief, i.e. Theory of Mind. This task was further popularised as the 'Sally-Anne' test by Baron-Cohen, Leslie, and Frith (1985), who corroborated Wimmer and Perner's findings (1983) that children develop this understanding around the age of four. However, it has been argued that success on the Sally-Anne test relies on executive functioning and language ability, two abilities which are greatly developed around the age that children start to pass these tasks (e.g. Astington & Baird, 2005; Pellicano, 2007; Sabbagh, Moses, & Shiverick, 2006). It could therefore be that Theory of Mind understanding is present earlier in development, but that this hypothesis cannot be tested with the Sally-Anne test. To address this issue, a number of studies used looking behaviour rather than a verbal test question to assess where the infant expected characters to search for the object. Using this adaptation of the Sally-Anne test, it was shown that infants as young as 13 months show some degree of false-belief understanding (Onishi & Baillargeon, 2005; Surian, Caldi, & Sperber, 2007; Southgate, Senju & Csibra, 2007; but see Perner & Ruffman, 2005, for arguments against this interpretation).

Studies like these led to the formulation of the two-systems theory of belief understanding (Apperly & Butterfill, 2009). This theory holds that there is (1) an early-developing system that allows for cognitively efficient, automatic, yet inflexible tracking of others' beliefs and (2) a later-developing system that enables belief tracking that can be directed by cognitive resources (i.e. is flexible), but is therefore cognitively demanding. As other authors have done (e.g. Schuwerk, Vuori, & Sodian, 2015; Sodian, Schuwerk, & Kristen, 2015), the remainder of this thesis will refer to these two systems as 'implicit' and 'explicit' Theory of Mind, respectively. According to Sodian and colleagues (2015), implicit Theory of Mind bypasses deliberate reflection on others' mental states and is therefore independent of verbal abilities. On the other hand, verbal abilities are important for guiding explicit Theory of Mind judgments. Therefore, to reach an adult-like explicit Theory of Mind, developments in language play an important role. Evidence for implicit Theory of Mind reasoning has subsequently been found in adults too (van der Wel, Knoblich, & Sebanz, 2014; Low & Watts, 2013; Edwards & Low, 2017; Schneider, Bayliss, Becker, & Dux, 2012a; Schneider, Lam, Bayliss, & Dux, 2012b; but see the latter for the assertion that the efficient system taps into executive resources as well). In short, the expression of Theory of Mind can be either implicit or explicit, across development.

The second process related to cognitive empathy discussed here is perspective taking. This ability is closely related to Theory of Mind, as both require the understanding that someone else can have a different representation of the world compared to oneself (Pearson, Ropar, & Hamilton, 2013). Perspective taking refers to the ability to ‘see things through the eyes of someone else’. To investigate this cognitive empathy-related process, research often focuses on visual perspective taking.

Along the same vein as for Theory of Mind, visual perspective taking can be divided into two abilities, commonly referred to as level-1 and level-2 visual perspective taking (e.g. Flavell, 1977; Flavell, Abrahams Everett, Croft, & Flavell, 1981a; Apperly & Butterfill, 2009). Level-1 perspective taking is concerned with *whether* someone else can see an object, whereas level-2 perspective taking pertains to *how* objects appear to that other person’s view. The leading theory is that level-2 perspective taking is cognitively demanding, whereas level-1 perspective taking can be automatic (Qureshi, Apperly, & Samson, 2010; Samson, Apperly, Braithwaite, Andrews, & Bodley Scott, 2010; Surtees & Apperly, 2012; but see Westra, 2016; Elekes, Varga, & Király, 2016, 2017, for a contrasting account). Level-2 perspective taking appears later in development than level-1 perspective taking (Flavell et al., 1981a; Flavell, Flavell, Green, & Wilcox, 1981b). Both types of perspective taking, and paradigms tapping into these respective abilities, have in common that they involve a discrepancy between the participant’s and someone else’s perspective (e.g. Samson et al., 2010; Keysar, Barr, Balin, & Brauner, 2000). To respond correctly, participants have to take the other’s perspective while suppressing one’s own (usually privileged) perspective. Taking someone else’s perspective can then contribute to the understanding of that person's mental state (Hamilton, Brindley, & Frith, 2009).

Interestingly, Samson and colleagues (2010) showed that, just as one's own perspective can interfere with responding from someone else's perspective ('egocentric interference'), someone else's perspective can interfere with responding from one's own ('altercentric interference'). This finding suggests that others' perspectives are computed automatically, even when they are not relevant to the task at hand. In other words, the perspectives of others hold such value that they are not easily ignored.

Although Theory of Mind and perspective taking are crucial to arrive at an understanding of others, the employment of these abilities by the perceiver does not in itself guarantee that the person on the other end (the perceived) is understood. Ultimately, the understanding of others relies on the ability to accurately infer their thoughts and feelings, which has been termed 'empathic accuracy' (Ickes, Stinson, Bissonette, & Garcia, 1990; Ickes, 1997). Studies that look at empathic accuracy often involve a comparison between what the perceived person thinks/feels over time and the inferences made thereof by the perceiver (e.g. Stinson & Ickes, 1992; Zaki, Bolger, & Ochsner, 2008; Martin-Key, Brown, & Fairchild, 2017; aan het Rot & Hogenelst, 2014; Blanke, Rauers, & Riediger, 2015). A closer match represents more accurate inferences on the part of the perceiver. These inferences rely on the combination and integration of information from various sources, such as observed cues, knowledge and reasoning (Ickes, 1997). These types of information can be obtained by engaging in cognitive empathy processes (i.e. Theory of Mind and perspective taking), affective empathy processes, or the interpretation of the situational context

(Davis, 1994; Ickes, 1997; Blanke et al., 2015). Therefore, although empathic accuracy has sometimes been equated with cognitive empathy (e.g. Zaki et al., 2008; Blanke & Riediger, 2019), the remainder of this thesis operationally defines empathic accuracy as the integration of information sources to arrive at mental state inferences (see Figure 1.1 for a schematic representation of the processes underlying empathy).

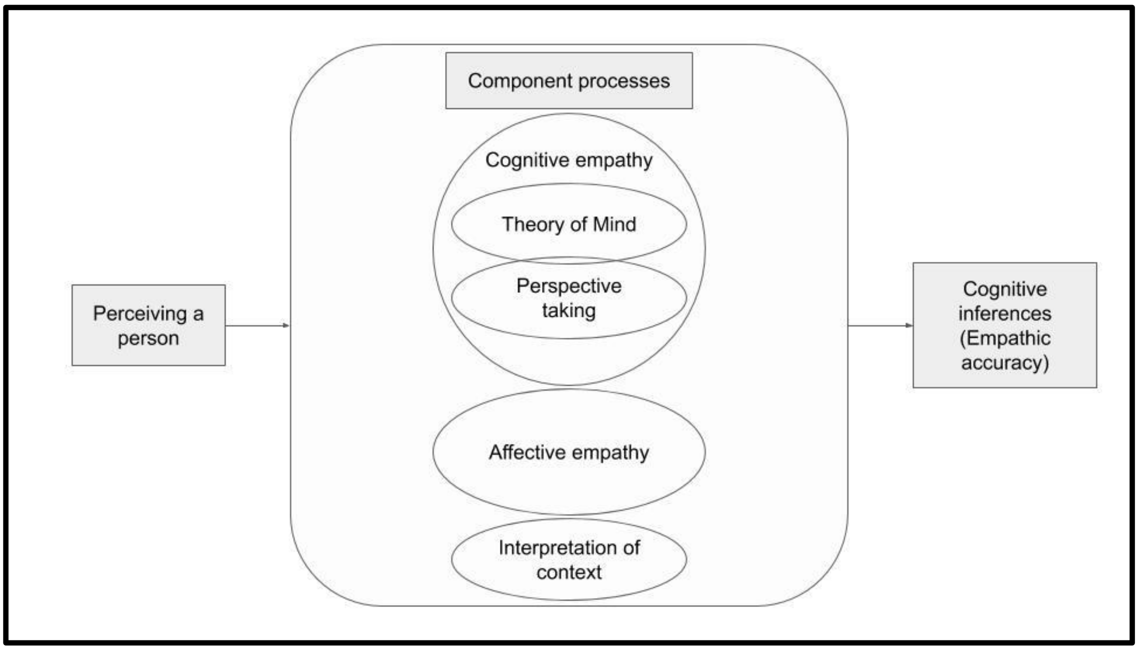


Figure 1.1 Schematic representation of processes underlying empathy. This model represents the processes through which the perceiver makes inferences about the thoughts and emotions of a perceived other. When this person is perceived, the perceiver first has to understand that the other person can experience mental states of their own. This understanding relies on a Theory of Mind, and is commonly reflected in the accuracy of belief tracking. In addition, the perceiver has to take the other person's perspective instead of their own. These conceptually related cognitive processes contribute to, but are not sufficient to, making accurate inferences. The accuracy of these inferences relies

on the integration of the above cognitive processes with affective empathy processes and an interpretation of the context. Affective empathy processes include mimicry and emotional contagion, while the context provides cues (e.g. bodily and auditory cues, the content of communicated messages) and can activate episodic and semantic memory. To make accurate inferences about others' thoughts and feelings, this context needs to be taken into account. Empathic accuracy therefore represents the integration of Theory of Mind reasoning, perspective taking, affective empathy processes, and an interpretation of the context. This latter process is beyond the scope of this thesis, as it does not relate directly to the perceived other person.

As the above discussion of cognitive empathy processes illustrates, there are a variety of ways in which cognitive empathy can be investigated. In addition, the literature covered so far has shown a clear distinction between processes underlying cognitive and affective empathy, respectively. This distinction has been further supported by studies examining the neural systems underlying the two empathy components. Cognitive empathy processes have been found to rely on areas such as the temporo-parietal junction (Saxe & Kanwisher, 2003; Saxe & Wexler, 2005) and ventromedial prefrontal regions (Eslinger, 1998; Shamay-Tsoory, Tomer, Berger, & Aharon-Peretz, 2003). Affective empathy processes, on the other hand, are associated with activation in e.g. the inferior frontal gyrus, which has been regarded as the human mirror neuron system (Dapretto et al., 2006). This differentiation was further corroborated by a lesion study in which a double dissociation between the associated neural systems was found (Shamay-Tsoory, Aharon-Peretz, & Perry, 2009). Work by Kanske, Böckler, Trautwein, and

Singer (2015) confirmed this pattern in non-lesioned individuals. They first specified two neural systems based on the existing literature. They subsequently found that activation in the neural system underlying cognitive empathy was related to performance on a cognitive empathy task but not an affective empathy task, and vice versa. Moreover, this pattern was observed for both task-related neural activation and resting-state functional connectivity. In short, the evidence suggests that the distinction of cognitive versus affective empathy is a meaningful one.

On the other hand, cognitive and affective empathy should not be regarded as completely independent constructs, as they both contribute to the understanding of others. For example, some studies found a positive relationship between cognitive and affective empathy (Verhofstadt et al., 2016; Jolliffe and Farrington, 2006; but see Kanske, Böckler, Trautwein, Parianen Lesemann, & Singer, 2016). Additionally, in a meta-analysis, Boele and colleagues (2019) showed that both cognitive and affective empathy have a positive influence on relationship quality. Altogether, it is clear that both affective and cognitive empathy contribute to social interaction. This distinction needs to be upheld in research, whilst keeping in mind that they are interrelated.

1.1.2 Empathy and the autism spectrum

Based on the above, it can be stated that empathy is of paramount importance both in development and typical social interaction. This importance is further emphasised by what happens when an individual has empathy deficits. The condition that has been studied in this context the most is autism spectrum disorder (ASD), to which dysfunctional empathy is central (Blair, 2005). Although it is an umbrella term ranging between high-functioning and non-verbal, there are commonalities across ASD. The defining characteristics are (1) deficits in social communication and interaction, and (2) restricted, repetitive behaviour or activities (American Psychiatric Association, 2013). Worldwide, the prevalence of ASD is about 62 per 10,000 people, although this number is possibly an underestimation, as awareness of the condition is limited in some countries (Elsabbagh et al., 2012a). Typically, ASD can be diagnosed in a child around the age of two-three years (Blair, 2005; Charman & Baird, 2002), although the first signals of abnormal social attention can be observed in infants as young as six months (Elsabbagh et al., 2012b). Throughout development, individuals with ASD experience difficulties in orienting towards social stimuli (Osterling & Dawson, 1994; Swettenham et al., 1998), initiating and following joint attention (Colombi et al., 2009; Leekam, López, & Moore, 2000; Leekam & Ramsden, 2006; Naber et al., 2008; Sullivan et al., 2007), facial information processing (Osterling, Dawson, & Munson, 2002; Osterling & Dawson, 1994; Deruelle, Rondan, Gepner, & Tardif, 2004; Rosset et al., 2008), imitating motor behaviour (Colombi et al., 2009; Rogers, Hepburn, Stackhouse, & Wehner, 2003), and responding to emotional cues (Golan, Baron-Cohen, & Hill, 2006b). However, impaired social functioning remains throughout life (e.g. Neuhaus, Beauchaine, & Bernier, 2010). For this reason, it is important to elucidate the mechanisms underlying empathy dysfunction in ASD.

Previous research on the nature of empathy deficits in ASD has yielded mixed results. Studies that looked at cognitive empathy component processes in ASD have generally showed an impairment. Individuals with ASD exhibit difficulties with Theory of Mind (Jones et al., 2017; Demurie, De Corel, & De

Roeyers, 2011; Burnside, Wright, & Poulin-Dubois, 2017; Begeer, Bernstein, van Wijhe, Scheeren, & Koot, 2012), although there is some evidence that these difficulties are more evident in implicit than explicit Theory of Mind (Begeer et al., 2016; Schuwerk et al., 2015). In addition, an impairment in perspective taking has often been reported in ASD. However, whereas studies have found this impairment in level-2 perspective taking (e.g. Hamilton et al., 2009), the majority of the evidence shows that level-1 perspective taking is intact in ASD (Santiesteban, Shah, White, Bird, & Heyes, 2015; Begeer, Malle, Nieuwland, & Keysar, 2010; see Pearson et al., 2013, for a review on both levels of perspective taking). Lastly, an impairment has been found for empathic accuracy (Demurie et al., 2011; Ponnet, Buysse, Roeyers, & De Clercq, 2008; Ponnet, Buysse, Roeyers, & de Corte, 2005; Bartz et al., 2010). Studies that combined measures of affective and cognitive empathy have generally found impairments in cognitive, but not affective, empathy (Jones, Happé, Gilbert, Burnett, & Viding, 2010; Dziobek et al., 2008). However, when affective empathy is operationalised as facial mimicry, an impairment emerges (McIntosh, Reichmann-Decker, Winkielman, & Wilbarger, 2006; Oberman, Winkielman, & Ramachandran, 2009; Stel, van den Heuvel, & Smeets, 2008). Hence, the literature suggests that the complexity of the pattern of empathy deficits in ASD transcends the primary dichotomy.

Traditionally, theories of autism focus on dysfunction to explain empathy deficits. One dominant theory states that there is a dysfunction of the human mirror neuron system (e.g. Iacoboni et al., 2005; Williams, Whiten, Suddendorf, & Perrett, 2001; see Iacoboni & Dapretto, 2006, for a supporting review). This account holds that mirror neurons fire in response to actions, their underlying intentions, and emotions. These activation patterns then contribute to a representation of the other's mental states in the perceiver. In autism, its proponents argue, these mirror neurons do not become active in response to observed actions and emotions (e.g. Altschuler et al., 2000). However, more recent studies on ASD found normal activation levels in the classic mirror neuron system in response to action observation (e.g. Marsh & Hamilton, 2011; Hamilton et al., 2007; Dinstein et al., 2010). Therefore, the general mirror neuron dysfunction account of ASD was not supported.

An alternative account of empathy deficits in ASD advocates that there is reduced propensity rather than ability (Keysers & Gazzola, 2014; Chevallier, Kohls, Troiani, Brodkin, & Schultz, 2012). In other words, although the ability to empathise is not necessarily impaired in ASD, there is less motivation to express empathy. This interpretation has come to be known as the social motivation theory of autism (Chevallier et al., 2012). Studies supporting this theory have found that people with ASD show an impairment in spontaneous, but not intentional facial mimicry (Gillespie, McCleery, & Oberman, 2014; McIntosh et al., 2006). In addition, Senju, Southgate, White, and Frith (2009) showed that people with ASD are impaired in spontaneously attributing mental states to others. The question is therefore why, in ASD, the ability for empathy does not necessarily translate to its expression.

One possibility to bridge this gap between propensity and ability lies in social reward. Social reward refers to incentives such as smiles and compliments, whereas non-social rewards encompass money or food. In typically developed individuals, social incentives are inherently rewarding and thus sought after and appreciated. However, this drive is diminished in individuals with ASD

(Dawson, Webb, & McPartland, 2005; Schultz, 2005). Kohls and colleagues (2011) argue that, because of this lack of social motivation, children with ASD do not orient towards socio-emotional input, resulting in ill-developed social skills. The social motivation theory has been tested with a variety of behavioural and neuro-imaging studies. It has been found that the learning of social rewards, but not non-social rewards, is impaired in ASD (Lin, Rangel, & Adolphs, 2012). On the other hand, several studies have shown atypical processing of both social and non-social rewards in the brain in ASD (Dichter, Richey, Rittenberg, Sabatino, & Bodfish, 2012; Kohls et al., 2011, 2013; Clements et al., 2018; Richey et al., 2014). This observation suggests a general reward processing impairment that is not exclusive to social reward. Nevertheless, only atypical processing of social rewards predicts autistic symptom severity (Dichter et al., 2012). In addition, in a review, Kohls, Chevallier, Troiani, and Schultz (2012) argued that reward processing atypicality is greater for the drive to obtain reward ('wanting') than for reward consumption ('liking'), with the atypicality being more pronounced for social rewards. Together, these studies suggest that especially the processing of social rewards is impaired in ASD.

Regardless of the specific nature of the reward processing impairment, it is more interesting how reward subsequently affects behaviour in ASD compared to typically developed individuals. For instance, Panasiti, Puzzo, and Chakrabarti (2016) showed that autistic traits do not decrease the ability to learn social reward, but they do modulate the beneficial effect that learned social reward has on prosocial behaviour. Similarly, Haffey, Press, O'Connell, and Chakrabarti (2013) showed that autistic traits modulate the mimicry of stimuli associated with reward, but only for social stimuli. In short, although the exact nature of reward dysfunction in ASD remains elusive, it is clear that it affects behaviour.

If reward can drive the modulation of mimicry and prosocial behaviour, the possibility arises that a similar relationship applies to empathy-related processes in general. Similarly, it could be that ASD alters this relationship domain-wide. Therefore, it is important to investigate the direct effect of reward on cognitive empathy, and how ASD may alter this relationship. This investigation can be done in two ways. The majority of studies above looking into affective versus cognitive empathy deficits recruited typically developed individuals and individuals with ASD. However, to achieve greater granularity in autistic symptoms, other studies often use a measure of autistic traits. The most common questionnaire used for this purpose is the Autism Spectrum Questionnaire (AQ), for which a score of 32 (range 0-50) or higher is indicative of clinical severity (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001b). Several studies have shown that AQ scores are on average higher for those with a diagnosis than those without (e.g. Baron-Cohen et al., 2001b; Ruzich et al., 2015). In addition, it has been shown that high AQ scores in those without a diagnosis reflect similar latent traits and similar brain activation anomalies as in ASD (Murray, Booth, McKenzie, Kuenssberg, & O'Donnell, 2014; Focquaert & Vanneste, 2015). On the other hand, other studies have provided evidence against the notion that autistic traits lie on a continuum in the general population or are predictive of a ASD diagnosis (Koolschijn, Geurts, van der Leij, & Scholte, 2015; Ashwood et al., 2016). This counterevidence suggests that the phenotypic heterogeneity observed in autism makes it difficult to capture autistic traits in a questionnaire. Altogether, it can be argued that, although a direct generalisability from findings in the non-clinical population to the ASD population is debated, the use of the AQ can nevertheless result in useful insights and inform autism research.

Most studies looking into the autism-reward relationship for empathy have used this measurement of autistic traits instead of recruiting participants with ASD. The following sections will cover these studies, starting with a section on how reward is commonly operationalised. They will then discuss existing evidence on the effect of reward on affective empathy, and explain why this question is currently not sufficiently answered for cognitive empathy.

1.2 The influence of reward on empathy-related processes

1.2.1 Manipulating reward value

Typically, paradigms tapping into processes underlying cognitive empathy require observers to make a judgment about the thoughts or feelings of a target person. In contrast, in affective empathy paradigms, it is typically measured to what extent the emotional state of the target person resonates in the observer. In both these types of paradigms, the reward value associated with the target person can be manipulated. The rationale is that if reward can have an influence on empathy-related processes, this manipulation should have observable effects on the measure of interest.

In the same way that there is a diversity of these measures of empathy, there are several ways in which reward value can be operationalised and manipulated. For the purpose of this thesis, reward value is defined as ‘any positive association with a social target (i.e. a person)'. These associations include subjective evaluations of closeness (e.g. Stinson & Ickes, 1992), similarity (e.g. Majdandžić, Amashaufer, Hummer, Windischberger, & Lamm, 2016), likability (e.g. Marsh et al., 2010), but also associations of reward with a target learned through conditioning (e.g. Sims, van Reekum, Johnstone, & Chakrabarti, 2012). Various studies have examined the effect of reward by comparing friends against strangers (e.g. Stinson & Ickes, 1992; Savitsky, Keysar, Epley, Carter, & Swanson, 2011). Although this friend versus stranger difference yields an ecologically valid proxy of reward, the variegated nature of relationships forms a challenge for methodological control. In contrast, using a conditioning manipulation in experiments allows for more control of the reward value associated with others.

Several studies suggest that these associations ultimately reflect a common positive 'currency' in the brain. Two candidates for the representation of social value are the ventromedial prefrontal cortex (vMPFC) and medial orbitofrontal cortex (mOFC), which are two adjacent brain regions (Murray, Debbané, Fox, Bzdok, & Eickhoff, 2015; Behrens, Hunt, Woolrich, & Rushworth, 2008; Ruff & Fehr, 2014). These brain regions have been found to underlie the representation of reward, social value information and approach/avoidance (Blair, 2008; Schmitz & Johnson, 2007; Bzdok et al., 2013; Ruff & Fehr, 2014; Rolls, 2015). Taken together, these findings suggest that these regions form a general neural system to represent subjective hedonic value (Klein, Shepherd, & Platt, 2009; Kühn et al., 2010). It has been found that activation in these regions correlates with feelings of social closeness when one is mimicked by another person (Kühn et al., 2010). Another study suggested that closeness is represented on a topological continuum in the medial prefrontal cortex, with distant others evoking neural activation more dorsally than close others (Murray, Schaer, & Debbané, 2012). Relatedly, several studies have shown a close relationship between ventral striatum activation and social closeness (Kawamichi et al., 2016; Fareri, Niznikiewicz, Lee, & Delgado, 2012; see Laurita, Hazan, & Spreng, 2019, for a review of brain regions involved in the representation of close others). Together, these studies suggest that brain regions typically associated with reward processing are also involved in the representation of the value of social stimuli.

Interestingly, some studies examining empathy-related processes have also found a link between different subjective evaluations of others. For example, in a review, Lakin and colleagues (2003) stated that there is a bidirectional positive relationship between nonconscious mimicry and ratings of affiliation and liking. Specifically, being mimicked by someone increases perceived affiliation with and liking of that person, which in turn results in more mimicry of *their* expressions. Neufeld and Chakrabarti (2006) confirmed that being mimicked results in a higher reward value assigned to the mimicker, especially if the mimickee is high in empathic traits. This observation was further corroborated by the finding that mimickers and mimickees not only report increased bonding, liking and interaction smoothness, but also become more similar in their emotional experiences (Stel & Vonk, 2010; McIntosh, 2006). In addition, it has been found that taking the perspective of someone results in increased perceived similarity, liking, a reduction of stereotypes, and more positive attitudes towards that person (Erle & Topolinski, 2017; Davis, Conklin, Smith, & Luce, 1996; Vorauer & Sasaki, 2014; Todd, Bodenhausen, Richeson, & Galinsky, 2011;

Galinsky & Moskowitz, 2000; Batson et al., 1997). Although these studies do not provide evidence for a direct link between such evaluations, they suggest that social interaction results in a general increase in positivity towards others. This suggestion resonates with Trope and Liberman's proposal (2010) that social closeness is directly related to positivity. In short, through showing links between the different quantifications of positive evaluation, the literature provides support for the operational definition of reward employed in this thesis.

The following sections will examine the influence of reward on empathy-related processes. Originally, empathy-related tasks typically yielded a pass/fail measure (e.g. the Sally-Anne test; Wimmer & Perner, 1983; Baron-Cohen et al., 1985). However, the emergence of novel paradigms has made it possible to quantify the effect of reward in a continuous rather than binary way. Hence, subtle differences along the AQ continuum can be detected. For each study reporting an effect of reward, it will be explained in what way reward value was manipulated, and whether autistic traits modulated the reward effect.

1.2.2 Reward and affective empathy

Reports on the effect of reward on affective empathy-related processes have generally been congruent. McIntosh (2006) demonstrated in two experiments that people mimic someone more if the observed person is liked. In the first experiment, a confederate acted likable and friendly, or cold and judgmental, towards a participant. The participant then watched a video in which the confederate sometimes displayed emotional facial expressions. Participants in the condition with the likable versus dislikable confederate showed more spontaneous facial mimicry. In the second experiment, in which the likable/dislikable confederate was replaced by a friend or stranger, observers mimicked friends more than they did strangers. This study did not examine further modulation by autistic traits.

In contrast, one paradigm in particular *has* been used to investigate this modulation. Sims and colleagues (2012) adapted an evaluative conditioning paradigm (de Houwer, Thomas, & Bayens, 2001) to condition some faces with a positive reward value, and others with a negative reward value. This paradigm was followed by a task in which the participant was repeatedly presented with faces, some of which had been conditioned, whereas others were newly introduced. Participants had to respond when this face was novel, i.e. an ‘oddball’. Crucially, each face displayed an emotional expression. It was found that participants spontaneously mimicked happy facial expressions more if those faces had been conditioned with a positive reward value. However, this relationship was negatively modulated by the participant’s autistic traits. Hence, this study was the first to show a relationship between autistic traits and the reward-driven modulation of an affective empathy-related process.

Subsequent studies have used the same evaluative conditioning procedure to show a similar relationship in other related processes. In one study, Sims, Neufeld, Johnstone, and Chakrabarti (2014) used functional magnetic resonance imaging (fMRI) while participants performed the oddball task. More functional connectivity between the ventral striatum and the inferior frontal gyrus was observed when the participant saw a face conditioned with a positive versus negative reward value. These two regions are involved in reward processing and mimicry, respectively (Kawamichi et al., 2016; Fareri et al., 2012; Laurita et al., 2019; Dapretto et al., 2006). Importantly, here too, the reward effect was reduced for participants higher in autistic traits. In a related study, Haffey and colleagues (2013) demonstrated a similar modulation of the reward effect for the mimicry of social stimuli (human hands), but not non-social stimuli (robot hands). Additionally, Panasiti and colleagues (2016) showed a reduced positive effect of reward on prosocial behaviour for those higher in autistic traits. Lastly, Trilla Gros, Panasiti, and Chakrabarti (2015) employed electroencephalography (EEG) to measure mu suppression. This EEG component refers to a reduction of signal strength in the mu band (7-25 Hz), which reflects cortical motor simulation in the mirror neuron system (Arnstein, Cui, Keysers, Maurits, & Gazzola, 2011; Braadbaart, Williams, & Waiter, 2013; Pineda, 2005). The EEG signal was measured while participants performed the oddball task. The results showed more mu suppression for positively versus negatively rewarded faces, indicating more simulation. This study, however, did not quantify autistic traits. Altogether, it is clear that (1) the reward-driven modulation of affective empathy-related processes is positive and that (2) when autistic traits are measured, they further modulate this negatively.

1.2.3 Reward and cognitive empathy

The question arises whether there is also an effect of reward on cognitive empathy, and if so, whether this is similar to the effect for affective empathy, and whether a similar modulation by autistic traits can be observed. As was discussed above, affective and cognitive empathy can be dissociated from each other, yet they also contribute together to the understanding of other people. For this reason, a similar reward-driven modulation for both is possible.

Two studies looked at the effect of reward on mentalising. Mentalising refers to the process "by which an observer views a target as possessing higher cognitive faculties such as goals, intentions and desires" (Marsh et al., 2010), and is therefore largely synonymous to Theory of Mind (see also Frith & Frith, 2006; Premack & Woodruff, 1978). Kozak, Marsh, and Wegner (2006) looked at whether liking influences mentalising. Target likability was manipulated by letting the participant read short vignettes about a person with a range of either positive or negative personality characteristics. The participants were then presented with descriptions of various ambiguous actions (e.g. "Highlighting text") and given a binary choice between a mentalistic interpretation (e.g. "Aiding his memory") and a mechanistic interpretation (e.g. "Using yellow ink"). The researchers found that participants chose the mentalistic interpretation more often with likable targets. In other words, likable targets were mentalised more. Interestingly, in a follow-up study, this pattern was replicated, but found to be reduced for people with more autistic traits (Marsh et al., 2010). It can be argued that a prompted binary choice between two equally correct alternatives does not reveal to what extent a participant would ordinarily mentalise. Nevertheless, these studies provide preliminary evidence that reward influences Theory of Mind and that this is negatively modulated by autistic traits.

More studies have looked at the effect of reward on cognitive empathy, but did not measure autistic traits. In a study by Savitsky and colleagues (2011), pairs of friends and strangers performed a task that taps into level-1 perspective taking. In each pair, one participant had to follow the other participant's instructions whilst ignoring their own privileged knowledge. Surprisingly, and contrary to findings reported above, the authors found that friends were more egocentric in their responses than strangers. It was concluded that people relax their perspective taking efforts when interacting with someone who is close to them.

Stinson and Ickes (1992) looked at the effect of reward on empathic accuracy. Here too, participants were paired up as friends or strangers. They were surreptitiously videotaped while they engaged in a natural social interaction. Subsequently, they watched the video and indicated their own thoughts and feelings throughout the interaction at set points. Next, they watched the video again and tried to infer the thoughts and feelings of the other person at these set points. It was found that friends showed more empathic accuracy than strangers. Therefore, there is evidence that closeness is positively related to empathic accuracy.

As the research covered here shows, there are only a few studies that have looked into the effect of reward on processes underlying cognitive empathy. For studies that *have* looked at this reward-driven modulation, the results are mixed. In addition, no studies but one have looked at the further modulation of this effect by autistic traits. The aim of the current thesis is to address this gap in literature. Following the operationalisations suggested by aan het Rot and Hogenelst (2014), it is investigated whether reward influences respectively Theory of Mind, perspective taking and empathic accuracy. In addition, it is examined whether there is a relationship between individual differences in this potential reward-driven modulation and autistic traits. To achieve a high degree of granularity, paradigms using continuous measures are employed. To investigate how a potential effect of reward compares between cognitive empathy-related processes, the same reward manipulation is used, where possible. Additionally, a task battery study is included in which various task measures of cognitive empathy are compared to each other.

2 Reward and belief tracking

Traditionally, false belief tests have been used to determine whether an individual possesses a theory of mind. It was later shown that someone can exhibit an implicit false belief understanding, even when they fail to show the explicit counterpart. It remains unclear whether characteristics of the target of these processes influence performance. In two studies, the reward value associated with targets was manipulated. Subsequently, participants performed an explicit or implicit belief tracking task. The results showed that performance was modulated by reward on the explicit, but not the implicit task. In addition, this reward-driven modulation was not related to autistic traits. This evidence suggests that explicit belief tracking is more accurate when the target is rewarding, possibly due to improved depth of processing.

2.1 Introduction

Traditionally, Theory of Mind (ToM) is the component process of cognitive empathy that has most frequently been investigated. ToM refers to the ability to reason about others’ mental states, and to the understanding of how these mental states underlie people’s behaviour (Apperly, 2012). The reason that ToM has taken centre stage in the field of cognitive empathy research is its crucial role in development. The emergence of a ToM is usually taken as a hallmark in the development of a child, as it paves the way for social interactions with all degrees of complexity. Being able to understand that reality can present itself to others differently is crucial to understanding others (Keysar, Lin, & Barr, 2003).

A plethora of research has used various versions of one type of paradigm to test whether someone possesses this understanding. The false belief paradigm was originally devised by Wimmer and Perner (1983), for use with children. In what has become known as the Sally-Anne test, the participant is first presented with two characters. The first character (Sally) has an object (a ball). She stores the object in one container and leaves the scene. The other character (Anne) then takes the object from its container and moves it to another container. When Sally reappears, the participant is asked the crucial question: “Where will Sally look for the object?” Of course, the participant knows the true location of the object, but in order to answer the question correctly, they have to ignore this privileged knowledge. If the participant indicates that Sally will search in the (now empty) original container, they are said to possess ToM. In their seminal study, Wimmer and Perner (1983) showed that the majority of children at the age of six exhibit this understanding, but younger children fail. Subsequently, Baron-Cohen and colleagues (1985) showed that children diagnosed with autism failed the test at a much later age. Not only did these studies stress the importance of the emergence of ToM for typical development, but they also provided the field with a fruitful paradigm.

This paradigm was subsequently modified to show that children can exhibit a rudimentary, implicit ToM understanding even if they are not capable yet of explicitly attributing a false belief in the Sally-Anne test. Various studies showed that infants younger than two years of age look longer at the scene when the character taking on Sally's role searched for the object in its actual location (e.g. Clements & Perner, 1994; Southgate et al., 2007; Onishi & Baillargeon, 2005; Surian et al., 2007; but see Perner & Ruffman, 2005). This prolonged gaze has been taken to reflect the infant's surprise at seeing the character not behave according to her false belief. In other words, these studies showed that an implicit understanding of others' beliefs emerges before the more verbal, explicit type of understanding.

Studies like these then led Apperly and Butterfill (2009) to propose a two-systems account of false belief reasoning. According to these authors, there is an implicit system that appears relatively early in development and allows for an efficient and potentially automatic tracking of beliefs, but that is limited and inflexible. This system can then guide smooth cooperative and competitive interactions. Subsequent studies provided evidence that implicit ToM is present in both infants and adults (van der Wel et al., 2014; Bardi, Desmet, Nijhof, Wiersema, & Brass, 2016; Bardi, Six, & Brass, 2017; Buttelmann & Buttelmann, 2017; Kovács, Téglás, & Endress, 2010; Meert, Wang, & Samson, 2017; Nijhof, Brass, Bardi, & Wiersema, 2016).

In contrast, in a more recent series of studies, evidence for implicit ToM was only found for one of four utilised paradigms (Kulke, von Duhn, Schneider, & Rakozcy, 2018). Specifically, whereas the authors did find a replication of the effect reported by Low and Watts (2013), they failed to replicate the findings of Schneider and colleagues (2012a), Southgate and colleagues (2007), and Surian and Geraci (2012). In addition, Kulke and colleagues (2008) showed no correlations between the measures of implicit ToM employed in these paradigms, indicating poor convergent validity. In another study, Phillips and colleagues (2015) replicated the findings of Kovács and colleagues (2010) in a series of 13 experiments, but showed that confounding artifacts of the task design explained the ToM effects. As a result, Kulke and colleagues (2018) concluded that whereas the existence of implicit ToM has not been disproven, most currently used paradigms do not show strong validity.

The other system proposed by Apperly and Butterfill (2009) is cognitively demanding, but its flexibility allows to form an explicit understanding of someone, and of the relationship between that person and oneself. Specifically, this explicit ToM system allows to put the understanding of others into words. This is the case when a ToM task requires explicit mental reasoning about beliefs, such as in the classic Sally-Anne task and its direct successors (e.g. Wimmer & Perner, 1983; Baron-Cohen et al., 1985). In short, Apperly and Butterfill's two-systems account (2009) holds that humans have two separate systems to track others' beliefs, and which one is used depends on the situation.

Studies have broadened the application of belief tracking paradigms to investigate individual differences in ToM. The underlying notion is that, if explicit ToM is cognitively demanding, individual differences in cognitive ability should lead to differential performance on explicit ToM tasks. Such individual differences in explicit ToM performance have been shown to be explained by memory, language and intelligence (Milligan, Astington, & Dack, 2007; Brüne, 2003; Hughes, 1998; Hughes, Taffee, Happé, Taylor, Caspi, & Moffitt, 2005). Some studies have also suggested that autistic traits predict differences in ToM performance (Begeer et al., 2012; Mathersul, McDonald, & Rushby, 2013), although this finding may be (partly) explained by verbal ability (Ronald, Viding, Happé, & Plomin, 2006). In addition, Bernstein, Coolin, Fischer, Thornton, and Sommerville (2017) showed that, once explicit ToM emerges, this ability remains relatively stable throughout life, with a slight decline in later adulthood. These studies suggest that there are various factors that can contribute to individual differences in explicit ToM performance.

The above has shown that ToM can be investigated as an explicit or implicit ability and that performance can vary based on the characteristics of the person performing the ToM task. However, it remains unknown whether characteristics of the person whose beliefs are reasoned about also influence ToM performance. As was discussed in Chapter 1, some empathy-related processes can be modulated by reward. For example, it was found that the reward value associated with a face can subtly influence behavioural and neural indices of facial mimicry, although these effects are decreased for individuals high in autistic traits (e.g. Sims et al., 2012, 2014). Following the operational definitions of cognitive empathy in Chapter 1, the question arises whether reward can also influence ToM and whether this is similarly modulated by autistic traits.

Two adaptations of the original belief tracking paradigm that lend themselves well to address these questions are the Sandbox task (SB; Bernstein, Thornton, & Sommerville, 2011; Sommerville, Bernstein, & Meltzoff, 2013) and the mouse-tracking task (MT; van der Wel et al., 2014). In the SB task, the participant is presented with a Sally-Anne-type trial, but instead of making a binary response, they indicate the character's belief on a continuum (i.e. a rectangular sandbox). Similarly, the MT task provides a continuous measure of ToM, in the form of the deviation of a mouse response (e.g. away from the false belief location and towards the true belief location). Notably, the MT task has been used to show both explicit and implicit belief tracking in adults (van der Wel et al., 2014). Specifically, van der Wel and colleagues (2014) measured performance on the MT task in two groups. The first group were instructed to click on the location of an object at the end of a scenario, but they also needed to track a character's belief about this object (explicit group). The second group were simply told to click on the location of the object when prompted (implicit group). Reaction times in the explicit group slowed down when a character's belief mismatched the participant's belief, indicating explicit belief tracking. This effect was absent in the implicit group, who instead showed a mouse trajectory deviation towards the irrelevant other's belief, indicating implicit belief tracking. This trajectory deviation and the continuous measure in the SB task allow to investigate even subtle individual differences. Therefore, these two paradigms will be used to answer the following experimental questions:

1) Does reward influence explicit ToM? More specifically, does the reward value associated to a person influence the extent to which the belief of that person is tracked?
2) Relatedly, does reward influence implicit ToM?

3) In addition, do autistic traits modulate either of these effects?

2.2 Aims

The aim of the current study was to investigate the effect of reward on ToM performance, using existing paradigms. Following the majority of studies investigating ToM (e.g. Wimmer & Perner, 1983; Baron-Cohen et al., 1985; Begeer et al., 2012), belief tracking was taken as a model for ToM. In order to examine the possible effect of reward on explicit ToM, the SB task was used (Bernstein et al., 2011; Sommerville et al., 2013). In contrast to the Sally-Anne test (e.g. Wimmer & Perner, 1983; Baron-Cohen et al., 1985), where an object moves from one container to another, in the SB task an object moves from one location to another within the same container. These locations are termed the false belief location and true belief location, respectively. When the participant is asked to indicate the false belief of the scenario's protagonist, the deviation of the response towards the true belief location represents the participant's egocentric bias. Put differently, the smaller the response bias, the more the participant displays false belief understanding. Therefore, the SB task makes it possible to minutely investigate the influence of factors such as reward on ToM performance.

In order to examine the possible effect of reward on implicit ToM, the MT task (van der Wel et al., 2014) was used, as this task has been used to measure belief tracking quantified as a bias. However, in contrast to the egocentric bias used in the SB task, the implicit MT task uses what can be referred to as an altercentric bias. Participants are told to click on an object's location at the end of a trial, after that object has changed location in a similar way to other belief tracking tasks. The mouse trajectory towards that location is thought to reflect the strength of the participant's belief. During the trial scenario, an irrelevant character appears and disappears on the scene. In some cases, this character will have a different belief about the location of the object than the participant. The degree to which the participant's mouse trajectory is then attracted to the location corresponding to the irrelevant character's belief is thought to reflect implicit tracking of that character's belief.

The MT task and SB task were used in two separate sub-studies. The reason for choosing the SB task to investigate explicit ToM rather than the explicit version of the MT task was that the SB task, like the implicit MT task, represents ToM as a spatial bias. Although reaction times can reflect conflict underlying the stating of a belief (in the explicit MT task), they do not show how strongly beliefs are represented relative to each other. On the other hand, spatial bias away from a belief reflects the degree to which a belief is 'pulled' towards another belief. As such, the SB task offers more insight into the effect of reward on explicit ToM than the explicit condition of the MT task. Both tasks were preceded by an evaluative conditioning task in which the reward value associated to a character was experimentally manipulated (cf. Sims et al., 2012, 2014; Haffey et al., 2013; Panasiti et al., 2016; Trilla Gros et al., 2015). In doing so, it was possible to quantify the reward value of a character, and quantify the degree to which this reward value influenced both implicit and explicit ToM.

In addition, the Autism Spectrum Quotient (AQ; see Appendix A) was used to measure autistic traits (Baron-Cohen et al., 2001b). In doing so, it was possible to investigate the relationship between autistic traits on the one hand, and the reward-driven modulation of ToM on the other hand.

2.3 Predictions

It is predicted that the SB task and MT task will reveal an egocentric bias and altercentric bias, respectively. However, most importantly, it is predicted that reward value will influence the extent of these biases. In addition, it is expected that any effect of reward will be modulated by autistic traits. To phrase the main expectations more specifically:

1) Reward value will decrease the egocentric bias in the SB task, as measured through the one-dimensional offset from the false belief location.
2) This reward-driven modulation will be negatively related to autistic traits.
3) Reward value will increase the altercentric bias in the MT task, as measured through the deviation of the response trajectory towards the irrelevant other's belief.
4) This reward-driven modulation will be negatively related to autistic traits.

2.4 Mouse-tracking study

2.4.1 Methods

2.4.1.1 Participants

Fifty-three participants completed the MT study (*N* female = 45; mean age = 19.49, *SD* = 1.07). This sample size was based on the sample in the implicit condition of van der Wel et al.'s study (2014). The participants were all undergraduate students at the University of Reading, and they signed up through the university's School of Psychology research panel. Participants received course credit upon completion of the study. In addition, they received £2.50 as a monetary reward for their performance in the evaluative conditioning task. As described in Section 2.3.1.3 (Procedure), the performance in this task was in fact fixed. Therefore, all participants received the same monetary reward. Ethical approval for the study was received from the School of Psychology Ethics Committee and all participants provided informed consent prior to their participation.

2.4.1.2 Stimuli

2.4.1.2.1 Evaluative conditioning task

Six faces from the Fundação Educacional Inaciana (FEI) face database (de Oliveira & Thomaz, 2006; see Appendix B) were used as stimuli for the evaluative conditioning task. Because the second task, the implicit belief tracking task, involved seeing faces from their sides, this face database was used as it contains multiple fixed angles for each face. To make participants familiar with the appearance of the faces from different angles, five angles were used per face, These face images corresponded to an angle of +90°, +45°, 0°, -45°, and -90°, relative to a frontal image. All faces had a neutral expression.

In addition, images of standard playing cards were used for the decision-making aspect of each trial. Values 6 up to and including 10 were used from each of the four suits.

2.4.1.2.2 Implicit belief tracking task (mouse-tracking)

The video stimuli used for the MouseTracker (MT) task were based on the belief tracking scenarios used in the study by van der Wel and colleagues (2014). The faces that were conditioned in the evaluative conditioning task were also used for this task. See section 2.4.1.3.2 for an explanation of these scenarios.

2.4.1.3 Procedure

When the participant arrived, they were first informed about the tasks they were going to perform during the experiment: a ‘gambling task’ and an ‘attention task’. They were told that the study investigated the effect of the former on the latter, and that their performance on the gambling task would influence their earnings at the end of the experiment. The entire experiment was presented on a 17-inch monitor with a computer running Windows 7. For both tasks, instructions were read aloud as they appeared on the screen, to assure understanding on the part of the participant. The order of the experiment was fixed across participants: they first performed the evaluative conditioning task, then the belief tracking task, and finished with the questionnaire. During the experimental trials of each task, the experimenter left the testing room. Participants were debriefed at the end of the experiment.

2.4.1.3.1 Evaluative conditioning task

The evaluative conditioning task used here (see Figure 2.1A) was adapted from the task used by Sims and colleagues (2012). This particular task was chosen as it has been shown to be effective in changing the reward value associated with different faces (Neufeld & Chakrabarti, 2016; Panasiti et al., 2016).

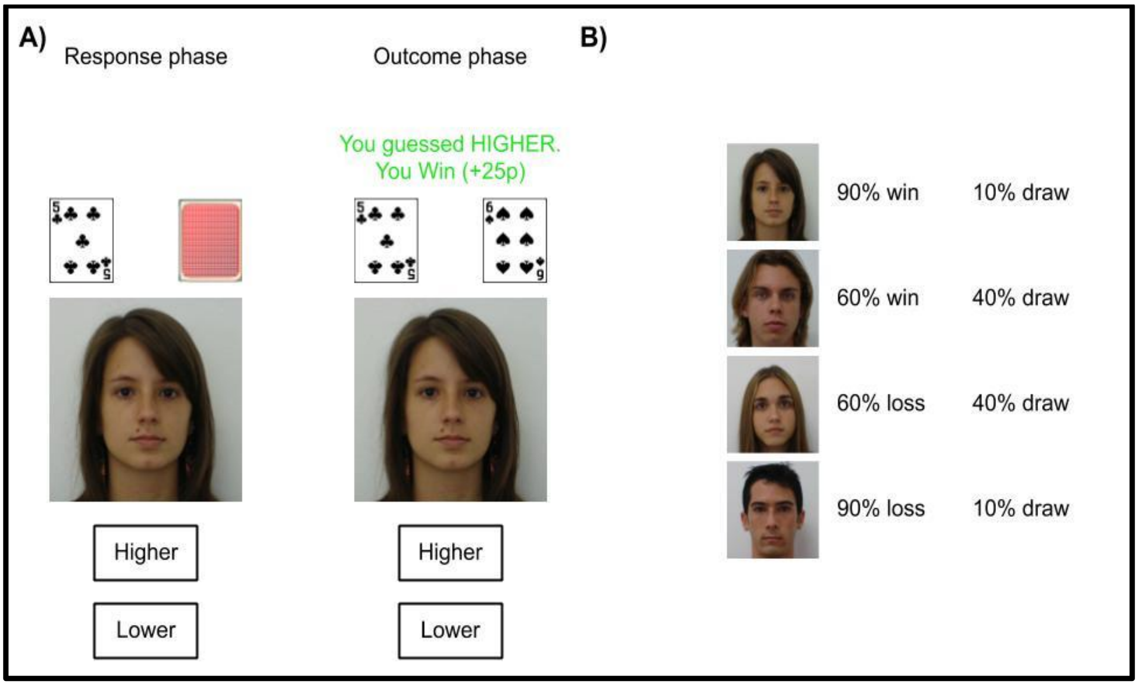


Figure 2.1 Schematic overview of the evaluative conditioning task and face-reward associations. A) During the response phase, the participant guessed whether the face-down card had a higher/lower value than the face-up card. During the outcome phase, the predetermined outcome was shown to the participant for 4,000 ms. B) One of the four counterbalanced versions used for the evaluative conditioning task. See Appendix B for all camera angles used for the face stimuli. The evaluative conditioning task and its four conditions were based on Sims and colleagues' design (2012).

The task was presented with Collector, a tool for online data collection (Garcia, Kerr, Blake, & Haffey, 2016). The screen resolution was set to 1920 x 1280 pixels for this task. During each trial, the participant was presented with two playing cards (78 x 106 pixels). The card on the left side of the screen was face-up, the card on the right face-down. Below these cards, a static picture of a face was shown (704 x 528 pixels). The participant had to guess whether the face-down card would be of a higher or lower value than the face-up card once the former was revealed. They made this guess by clicking on the corresponding response option. They were given unlimited time to make a decision. Once they had made a decision, the face-down card was revealed, together with written outcome feedback, for 4,000 ms.

During the practice phase, the value of the face-down card was always fixed, such that the decision the participant made influenced whether the outcome was a win or loss. The practice phase consisted of six trials. The two faces (one male, one female) shown during the practice phase did not return in the experimental phase.

Crucially, during the experimental phase, the outcome of each trial was predetermined, unbeknownst to the participant. This manipulation was used to condition each of four faces with a reward value (cf. Sims et al., 2012). More specifically, four faces (two male, two female) were each associated with a specific percentage of trials with a win/loss (see Figure 2.1B). In the remaining trials, the outcome was a draw. In other words, a face generally associated with a win would never be accompanied by a loss, or vice versa. When the outcome was a loss, the response feedback was “You have lost (-20p)” in red. When it was a win, the response feedback was “You have won (+25p)” in green. The reason for this numerical discrepancy is that, subjectively, losses weigh heavier than wins (Kahneman & Tversky, 1979). After a draw, the feedback was simply “Draw” in black. Reward associations were counterbalanced across four versions and face gender was never associated with win/loss.

During the 120 experimental trials, each face appeared 30 times (6 x 5 angles). Within the practice phase and experimental phase, the order of trials was fully randomised per participant.

2.4.1.3.2 Implicit belief tracking task (mouse-tracking)

The second task of the experiment, the MT task, was an adapted version of the paradigm developed by van der Wel and colleagues (2014). For this task, presentation, mouse location tracking and data saving were all done with the MouseTracker software package (Freeman & Ambady, 2010). To conform to the dimensions used by this software, the computer resolution was set to 1024x768 before running the MT task. During the task, the viewing distance was 80 cm, indicated by a marker on the desk at which the participant performed the task. This marker was aligned with, and 40 cm away from, the centre of the screen. The participant was instructed to return the computer mouse to this marker after each trial, to ensure there was sufficient desk space to respond.

Before task initiation, the participant was instructed about the trial procedure by screenshots from the task. On-screen instructions were read out from the screen by the experimenter. Then, the experimenter demonstrated the task by performing one trial from each condition. The participant was told to be fast in their responses, but that accuracy was more important.

The central instruction was to follow a ball in order to respond to its location at the end of a trial. An irrelevant object, a cube, exactly mirrored the movements of the ball. For the sake of simplicity, the following trial overview only discusses the ball movement (see Figure 2.2 for a schematic depicting both the ball and cube movements).

Each trial progressed as follows: first, two occluding blocks appeared on a background. In the left/right corner of the screen (counterbalanced), the face of a person was shown (referred to as ‘the agent’). Immediately, a ball appeared and moved behind the left/right occluder (counterbalanced). Next, the agent disappeared. Then, in the agent's absence, the ball reappeared, only to disappear again behind either its original occluder or the opposite occluder. The agent then reappeared, marking the end of the scenario video. At this point, the last frame of the video was shown as an image, together with the prompt 'NOW' in the middle of the screen. The mouse cursor appeared in the bottom middle of the screen (532 pixels below and 256 pixels to the left/right of the nearest corner of the occluders). The participant had 3,000 ms to click on the location of the ball, which was revealed after 500 ms, when the occluders vanished. The 'NOW' prompt made sure participants were encouraged to start moving the mouse before the location of the ball was revealed. If they failed to do so, a feedback window would be shown saying "Please start moving earlier on, even if you are not certain of a response yet". If the participant clicked on the wrong location, a red cross was shown as feedback for 1,500 ms. Lastly, if the participant did not respond within 3,000 ms, a feedback message ("RESPOND FASTER") was shown for 1,500 ms.

Crucially, in 50% of the trials, the location of the ball when it was revealed corresponded to where the agent had last seen it (i.e. when there had not been a switch). In the other 50%, the location did not correspond to where the agent had last seen it (i.e. when there *had* been a switch). Likewise, in 50% of the trials, the ball appeared in the location where the participant had or had not last seen it before it was revealed. As such, there were four belief conditions: Participant True - Agent True, Participant True - Agent False, Participant False - Agent False, and Participant False - Agent True.

In total, the participant performed eight practice trials (with the two faces that were used in the practice phase of the evaluative conditioning task) and 128 experimental trials. In the experimental phase, each of the four conditioned faces was presented in 32 trials (eight trials for each belief condition).

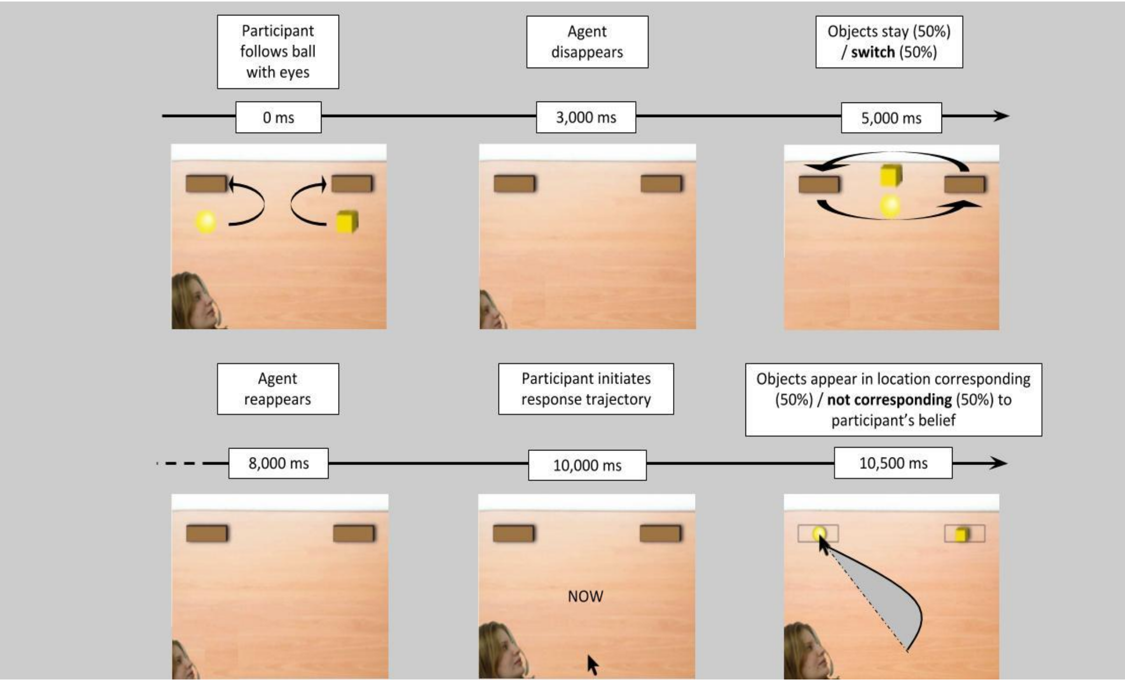


Figure 2.2 Schematic overview of a trial in the MT task. The example trial shown here represents the Participant False - Agent True condition, as, in the last panel, the target ball appears in the location where the agent has last seen it, whereas the participant has last seen it in the other location. Timestamps indicate the start of the respective trial phase. Text in bold indicates what is displayed in this example trial. The outcome variable, calculated by MouseTracker, was the area under the curve described by the mouse cursor trajectory (shown in grey in the last panel). The cube mirrored the movements of the ball, but was irrelevant to the participant. For the sake of simplicity, the above only mentions the ball.

The current belief tracking task had three notable differences compared to the task used in van der Wel et al.'s study (2014). First, to prevent the position of the face from having an influence on the mouse trajectory due to the saliency of faces, the location of the face was counterbalanced across trials. Second, the original authors used an additional stay/switch phase, but as they found that this phase had no influence on responses, this phase was omitted from this task to reduce trial duration. Third, due to software constraints, the location of the ball was revealed after 500 ms in the current version of the task, whereas in the original version this was done once the participant had moved the mouse above an imaginary horizontal line at a fixed height. The revealing of the ball in the current version was the reason why trials were not analysed if participants started moving the mouse cursor after 500 ms.

There were two experimental blocks of 64 trials, with each combination of face and belief condition occurring four times in each block. Trials within each block were fully randomised across participant. The participant started each video by clicking on a Continue button, and they were allowed to take a break after every 16 trials.

2.4.1.3.3 Trait measure

The Autism Spectrum Quotient (AQ; Baron-Cohen et al., 2001b) was used to quantify autistic traits. This questionnaire has 50 items, (e.g. "I prefer to do things with others rather than on my own."), with four response options per item (Definitely agree; Slightly agree; Slightly disagree; Definitely disagree). Cronbach's alpha coefficients of the scale's domains range between .63 and .77, indicating moderate internal consistency (Baron-Cohen et al., 2001b). For the total scale, Cronbach's alpha was found to be higher (.82; Austin, 2005). Participants filled out the AQ after completion of the evaluative conditioning task and MT task. In addition, participants provided their basic demographic information.

2.4.1.4 Data analysis

The playing card task was only used for evaluative conditioning. Due to the predetermined outcomes of decisions in the task, performance was also fixed. Therefore, data generated by that task was not further analysed.

For the MT task, one dependent variable was generated from the participant's mouse trajectory. Whereas typical task presentation software records a participant's response in terms of outcome, the MouseTracker software developed by Freeman and colleagues (Freeman & Ambady, 2010) allows to measure and quantify the mouse trajectory leading up to that response. Based on van der Wel et al.'s study (2014), the current analysis used area under the curve (AUC) as the crucial measure. AUC represents the deviation of a mouse trajectory towards an irrelevant location. This measure translates to the following in the current task: a mouse trajectory that describes a straight line towards the revealed location of the ball would have an AUC of around zero. A mouse trajectory that curves, on average, towards the location of the cube would have a positive AUC, whereas one that curves away from the location of the cube would have a negative AUC. The AUC was calculated based on time-normalised trajectories, i.e. each mouse trajectory was represented by 101 spatial coordinates that were temporally equidistant based on the total duration between initiation of the trajectory and the response click. The size of the AUC represented the surface area of MouseTracker's 2 x 1.5 coordinate system that was covered by the curve of the response trajectory.

By using AUC, it was quantified to what extent the belief of the participant and the belief of the agent influenced the mouse trajectory leading to the response. More specifically, a difference in trajectory curvature between the

Participant True versus Participant False conditions would indicate to what extent the participant had tracked the location of the ball. On the other hand, a curvature difference for the Agent True versus Agent False conditions would indicate to what extent the agent's belief had been tracked (cf. van der Wel et al., 2014). Therefore, the effect of agent's belief on AUC represents implicit belief tracking, as the participant was instructed to follow the ball (from their own perspective), whereas the agent's belief was irrelevant to the task at hand.

In addition, it was investigated whether the effect of the agent's belief on AUC was modulated by the reward value of the agent. Lastly, the degree of this modulation was represented by a single value in order to investigate the relationship with autistic traits.

All p-values reported with statistical tests below are two-tailed. For the results in this chapter and the following experimental chapters, a *p*-value lower than .05 will be regarded as statistically significant, whereas a *p*-value between .05 and .1 will be interpreted as a trend.

2.4.2 Results

2.4.2.1 Data pre-processing

The MouseTracker software package (Freeman & Ambady, 2010) automatically pre-processed the mouse trajectories, such that each trial was represented by a single AUC. The remainder of the pre-processing steps and analysis was performed with R (R Core Team, 2018).

Due to a technical issue, 979 trials were lost across participants. Specifically, the MouseTracker software would sometimes crash during a block, resulting in the loss of all data collected for that block so far. In these cases, the experimenter started the next block. In the case of three participants, this issue occurred for both experimental blocks. As these participants had no behavioural data, they also did not fill out the questionnaire. As a result, 5,805 trials remained (participant mean *N* = 109.53, *SD* = 25.02). Of the remaining trials, trials were removed for the four belief conditions individually if the response was (1) incorrect (3.8%), (2) too slowly initiated (>500 ms; 14.2%), (3) a participant-level outlier (AUC > 3 SD away from one participant's mean; 1.2%), or (4) a group-level outlier (AUC > 3 SD away from all trials' mean; 1.1%). This removal procedure was equal to the one used by van der Wel and colleagues (2014), apart from the inclusion of the second step. This step was included because the location of the ball was revealed after 500 ms.

Finally, participants were excluded from the analysis if they did not have at least one trial in each belief condition after the previous exclusion steps (three participants), or if they did not have more than two trials per condition on average (one participant). This exclusion procedure resulted in a final sample of 49 participants (42 females), whose age ranged between 18 and 25 years (*M* = 19.51, *SD* = 1.08) and whose AQ scores ranged between 2 and 26 (*M* = 13.63, *SD* = 5.07). The total number of trials was 4,518 (participant mean *N* = 92.20, *SD* = 21.65). The mean AUC per participant per condition was used for the statistical analysis.

For correlations, observations were excluded from each individual analysis separately if they were identified as bivariate outliers (Cook's distance > $4/N_{participants}$; Bollen & Jackman, 1990; see the degrees of freedom for the correlational analysis).

2.4.2.2 Implicit belief tracking (replication analysis)

It was first tested whether the implicit belief tracking effect found by van der Wel and colleagues (2014) was replicated in the current study. To this end, a 2 x 2 ANOVA was performed, with Participant's belief (True/False) and Agent's belief (True/False) as repeated-measures factors, and AUC as the dependent variable. In calculating the means used for this analysis, raw data was collapsed across the four reward conditions. There was a significant main effect of Participant's belief, $F(1,48) = 181.18$, $p < .001$, $\eta_p^2 = .791$, with the AUC being smaller when the Participant's belief was True ($M = .21$, $SE = .04$) than when it was false ($M = 1.76$, $SE = .09$). There was also a significant main effect of Agent, $F(1,48) = 7.31$, $p = .009$, $\eta_p^2 = .132$, with the AUC being smaller when the Agent's belief was True ($M = .96$, $SE = .04$) than when it was false ($M = 1.01$, $SE = .04$). There was no significant interaction effect, $F(1,48) = .19$, $p = .665$, $\eta_p^2 = .004$.

2.4.2.3 The effect of reward

To assess the effect of reward on implicit belief tracking, the four reward conditions were split up into rewarding (collapsing the 90% win and 60% win condition) and non-rewarding (collapsing the 90% loss and 60% loss condition). This step was taken to avoid too many participants having no data in a condition due to the aforementioned data loss. An implicit belief tracking bias (simply referred to as bias from now onwards) was calculated per participant by subtracting the AUC for the TT from the TF condition, subtracting the AUC for the FT from the FF condition, and taking the average of these two values. As such, the bias represented the extent to which a false Agent's belief influenced the AUC.

A paired-sample t-test was performed to investigate between the rewarding and non-rewarding conditions on bias. No significant difference was found, $t(48) = .93$, $p = .358$, $d = .17$ (see Figure 2.3).

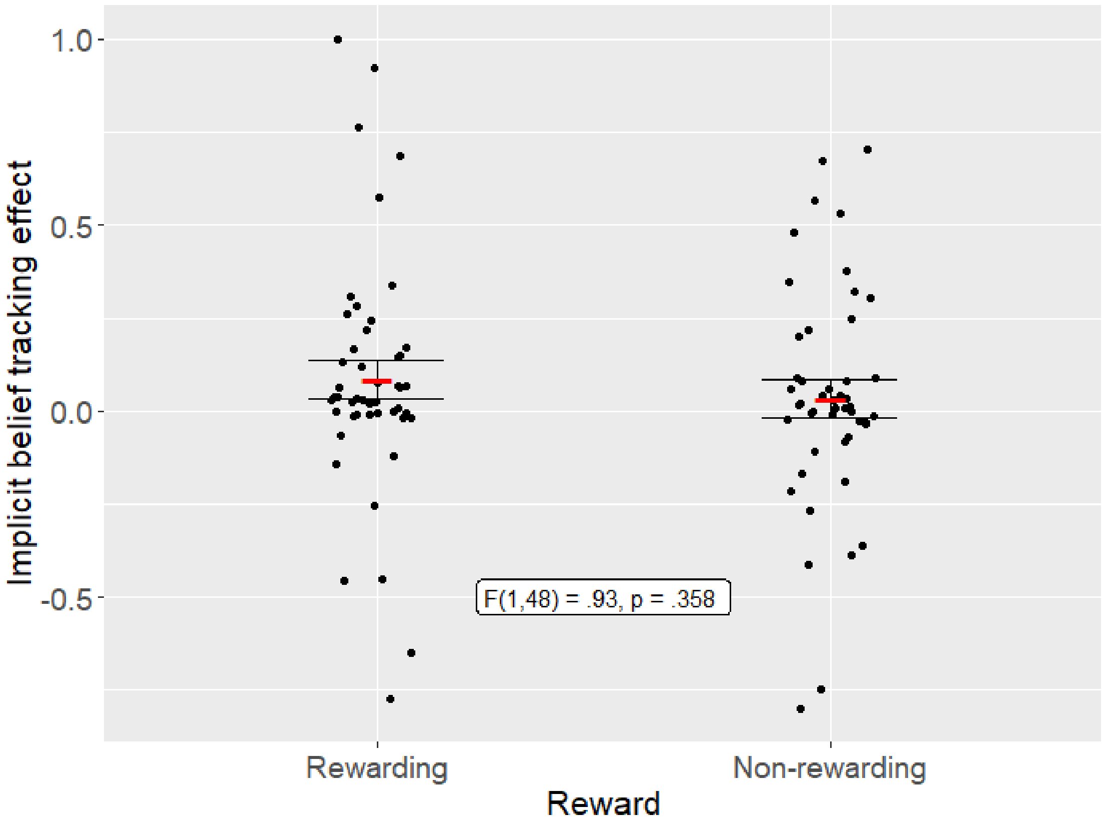


Figure 2.3 Null effect of reward on implicit belief tracking. Error bars represent 95% within-subjects confidence intervals. Dots represent individual participants.

2.4.2.4 Individual differences in the reward effect

Although there was a null effect of reward on implicit belief tracking, the *a priori* research question warranted an investigation into the relationship between the reward effect and individual differences in autistic traits. The reward effect was calculated as the difference in bias between the rewarding and non-rewarding conditions, per participant.

A Pearson's correlation was performed between AQ and this reward effect. The result was non-significant, $r(46) = -.05$, $p = .754$ (see Figure 2.4).

As mentioned in Section 2.4.2.1, 14.2% of all trials were removed due to participants being too slow to initiate movement. To investigate whether this removal resulted in the absence of effects, all analyses reported here were repeated without this specific pre-processing removal step. The results reported above did not change.

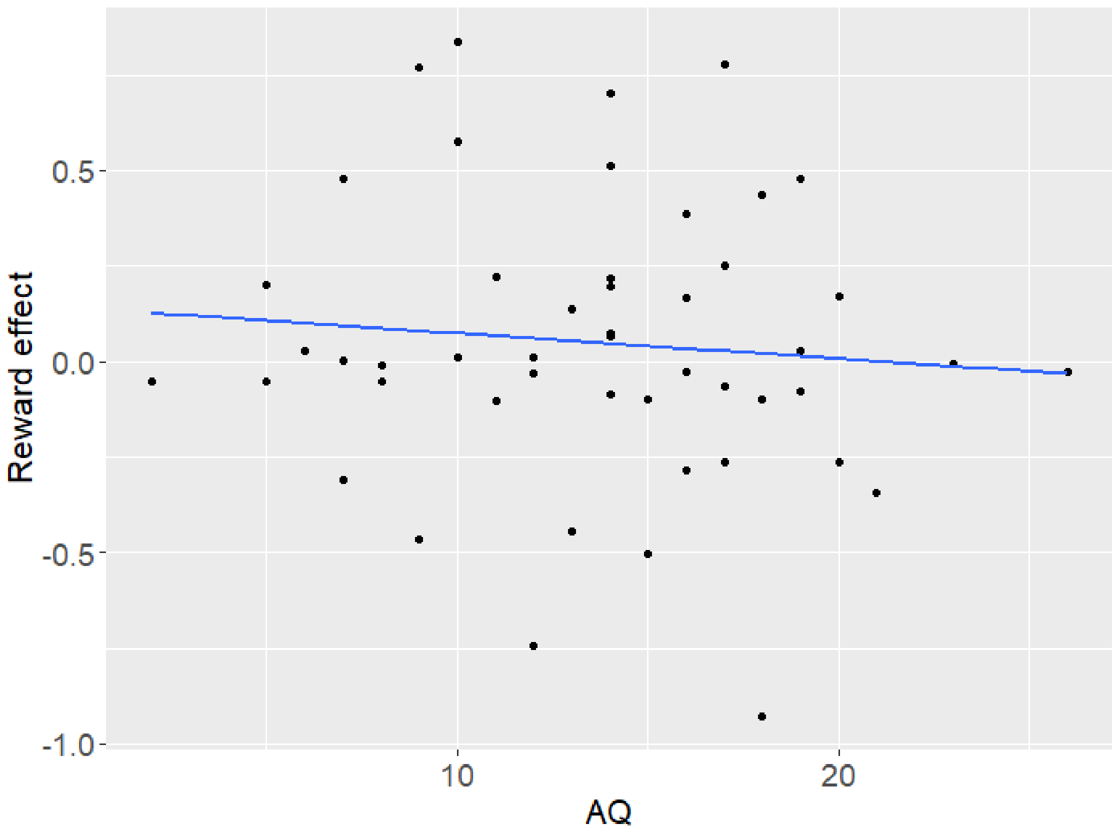


Figure 2.4 The relationship between autistic traits and the effect of reward on belief tracking bias. The reward effect was calculated by subtracting the AUC for the TT from the TF condition, subtracting the AUC for the FT from the FF condition, and taking the average of these two values.

2.4.3 Brief discussion

The results showed that the response trajectory was most influenced by the participant’s belief. This observation is to be expected, as the participant is told to track the location of the ball (from their own perspective) before they click on that location. In addition, the agent’s belief also influenced the response trajectory, albeit to a lesser extent. When the agent’s belief was false, the curvature of the trajectory was larger than when the agent’s belief was true. This observation replicates the finding in van der Wel et al.’s study (2014) and suggests that the agent’s belief is tracked implicitly, even when this belief is irrelevant to the task at hand.

However, reward value was found not to influence the effect of agent’s belief on curvature. More specifically, the reward value that had become associated with the agent in the corner of the screen through an evaluative conditioning procedure did not influence the degree to which this agent’s belief was implicitly tracked. This lack of a reward effect is further discussed in section 2.6 (Discussion).

In addition, there was no significant relationship between individual differences in autistic traits and the reward effect described above. Note that the non-significance of the reward effect indicated a general absence of the effect. However, the *a priori* research question justified performing a correlational analysis, which could have brought to light a more intricate effect of reward in the current sample. Nevertheless, based on the results, it can be concluded that the effect of reward on implicit belief tracking was absent along the full spectrum of autistic traits that was investigated in the current study.

2.5 Sandbox study

2.5.1 Methods

2.5.1.1 Participants

Sixty-five participants completed the explicit belief tracking study (*N* female = 52; mean age = 20.97, SD = 3.75). This sample size was based on the neurotypical sample in Begeer et al.'s study (2012) using the Sandbox (SB) paradigm. For the current sample, 22 participants were recruited from the University of Aarhus, Denmark, through the Cognition and Behaviour Lab research panel. Participants had a 20% chance of winning a cinema voucher. The other 43 participants were recruited from the University of Reading, and they signed up through the university's School of Psychology research panel. These participants received course credit upon completion of the study. Ethical approval for the study was received from both the School of Psychology Ethics Committee at the University of Reading and the Cognition and Behaviour Lab ethics committee in Aarhus.

2.5.1.2 Stimuli

2.5.1.2.1 Evaluative conditioning task

The evaluative conditioning task used the same playing cards as the MT study (see Section 2.4.1.2.1). In contrast, photographs taken from the chest upwards were taken from the standardised Mind Reading set (Baron-Cohen, Golan, Wheelwright, & Hill, 2004). This set shows external validity and sufficient inter-rater reliability (Golan & Baron-Cohen, 2006; Golan, Baron-Cohen, Hill, & Golan, 2006a). The selected photographs were of two males and two females, chosen for their distinguishability. Within each gender, one person was wearing a bright red shirt, the other a bright yellow shirt. All photograph depicted a neutral expression. Only photographs taken from the front of the person were used in this task.

2.5.1.2.2 Explicit belief tracking task (Sandbox)

The same faces that were shown during the evaluative conditioning task were used in the Sandbox (SB) task. The sandbox container was an open-source image of a trapezoid tray with a clickable width of 100 pixels. For the distractor task that was part of each SB trial, a 15-by-15 word search was used.

2.5.1.3 Procedure

This study used the same structure as the MT study, with an evaluative conditioning task followed by a belief tracking task, followed by a personality trait questionnaire. Before the start of the experiment, participants were told that they had a chance of winning a cinema voucher, and that their performance would determine their chance of winning it. After the participant had given informed consent, the tasks were performed in a fixed order. The entire experiment was conducted on a 13-inch Macbook screen.

2.5.1.3.1 Evaluative conditioning task

The evaluative conditioning task was largely the same as the one used in the MT study (see Section 2.4.1.3.1), with a few key differences. The participant completed eight unsupervised practice trials. In addition, each of the four faces was assigned a name, which was presented under the face during both the evaluative conditioning and SB task. Participants were instructed that, following the card game, a simple memory task involving the same faces and names would be performed. Only the 90% win and 90% loss condition were used for this task (one win and one loss condition per gender). The reason for including only these conditions was due to the SB task design (see the following section). Which two faces were rewarding/non-rewarding was counterbalanced across participants. Lastly, this task used points instead of pennies to manipulate reward value.

2.5.1.3.2 Explicit belief tracking task (Sandbox)

The explicit belief tracking task was a computerised version of the paper-and-pencil SB task (Begeer et al., 2012; Bernstein et al., 2011). The current SB task was created and conducted using Unity® (Unity Technologies, ApS, 2016). The participant was represented with 20 change-of-location scenarios (see Appendix C), adapted from scenarios used in a previous study (Sommerville et al., 2013). There were three types of scenarios: false belief (FB; eight trials), memory control (MC; eight trials), and true belief (TB; four trials). The order of these scenarios was fully randomised, but identical for different participants. All phases of each scenario were accompanied by text and stayed on-screen for 5 seconds. The participant was asked to carefully read each scenario.

On an FB trial (see Figure 2.5), the participant was first introduced to two characters. Subsequently, one character (the ‘placer’, e.g. ‘Will’) hid an object in a location (L1) along the one-dimensional continuum of a container. Then, a second character (the ‘mover’, e.g. ‘Rose’) moved the object to another location in the container (L2). This replacement was done in the absence of the placer. Next, the participant performed a word search for 20 seconds. At the end of this timespan, they indicated how many universities they had identified during the word search. This distractor task has previously been used in an SB study with an adult sample and was included to prevent the participant from using perceptual strategies to guide their response at the end of the SB trial (Coburn, Bernstein, & Begeer, 2015). After the distractor task, the placer and the container appeared again. Crucially, the participant was then asked where the placer *would look for the object*. They indicated their response by clicking inside the container. Their response was subsequently represented by an 'X'.

On an MC trial, the change-of-location scenario was different. The critical difference was in the wording of the question asked in the final phase of each trial. Specifically, the participant was asked where the placer *had put the object* before the character temporarily left the scene. Note that, for both FB and MC trials, clicking on L1 was regarded as a perfect response. However, FB trials require ignoring one's own privileged knowledge and taking the perspective of the placer, whereas MC trials only require memorising the object's location (Coburn et al., 2015).

To prevent participants from developing the rule that L1 was always the required response location, four TB trials were included. The question in the final phase of a TB trial was phrased in the same way as that in FB trials (i.e. where will the placer look for the object?). However, when the object was moved, the scenario text specified that the placer character was present during this replacement. Hence, L2 was the perfect response location.

In all trials, the placer and the mover were associated with opposite reward values, such that on half the trials the placer was associated with a 90% win value and the mover with a 90% loss value, and vice versa. As the 20 trial scenarios pitted two characters against each other, only two reward conditions were included to maintain consistency in reward value difference across trials. The 90% win and 90% loss conditions were kept to maximise the difference in reward value between reward conditions. In addition, the respective location of L1 and L2 and the distance between the locations varied across trials. This distance spanned either 25% or 50% of the container width. The direction of the object replacement was counterbalanced across trials.

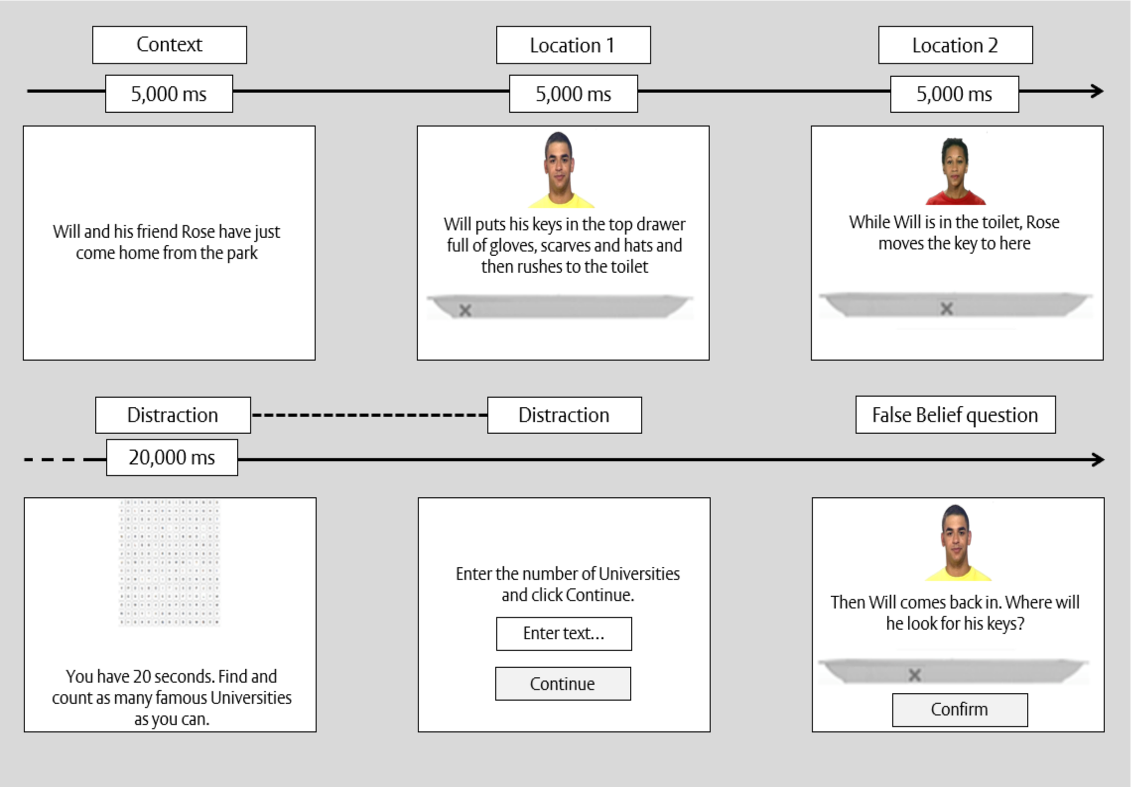


Figure 2.5 Schematic overview of a trial in the explicit belief tracking task. The trial depicted here represents the false belief condition. Timestamps indicate durations of each phase. The cross in the last panel represents a possible response.

2.5.1.3.3 Trait measure

After the evaluative conditioning and explicit belief tracking task, the participant completed the AQ (Baron-Cohen et al., 2001b). In addition, participants provided their basic demographic information.

2.5.1.3.4 Recall task

Prior to the evaluative conditioning task, the participant had been asked to remember all four name-face combinations. This instruction was included to prevent extinction of the reward manipulation during the explicit belief tracking task. To verify that the participant had remembered which name went with which face, they performed a recall task. Here, the participant was presented with all four faces, and asked to name them.

2.5.1.4 Data analysis

As the playing card task was only used for evaluative conditioning, data generated by that task was not further analysed.

For the explicit belief tracking task, one dependent variable was based on the participant's response. This calculation was done for the FB and MC trials. The dependent variable, bias, was calculated as the signed distance between L1 and the participant's actual response in N pixels, as a percentage of the length of the container. As the length of the container was 100 pixels, bias was effectively measured in pixels, representing the degree to which a response is attracted towards L2.

As was discussed in Section 2.5.1.3.2 (Sandbox task procedure), there is an important interpretative difference between the FB and MC condition. In the MC condition, the test question taps into memory, as the participant is asked where the object was originally placed. In the FB condition, however, the participant is asked where the character will look for the object. Whereas an imperfect memory trace can also contribute to a bias in the FB condition, the test question prompts the participant to make a belief attribution. Therefore, the difference in bias between the FB and MC condition represents the contribution of belief tracking over and above memory.

The TB condition, in which the perfect response was L2, was only included to prevent participants from developing the response strategy that the required response was always L1. As such, the TB trials were regarded as filler trials and excluded from the analysis.

All p-values reported with statistical tests below are two-tailed.

2.5.2 Results

2.5.2.1 Data pre-processing

Data was pre-processed with R (R Core Team, 2018). Firstly, 14 participants were excluded because they failed to link the names with the faces in the recall task. Secondly, one participant was removed from the analysis because their AQ score was 5 SD higher than the sample average. The final sample consisted of 50 participants (40 females), whose age ranged between 18 and 36 (M = 20.98, SD = 3.92) and whose AQ scores ranged between 4 and 27 (M = 13.82, SD = 4.78).

Trials were removed for the FB and MC condition separately if the response was (1) a participant-level outlier (bias in pixels > 3 SD away from one participant's mean; 0%), or (2) a group-level outlier (bias in pixels > 3 SD away from all trials' mean; cf. experiment 1; 2.1%). As such, these outlier steps were similar to the steps taken for the MT study (see section 2.4.2.1).

For correlations, observations were excluded from each individual analysis if they were identified as bivariate outliers (Cook's distance > $4/N_{participants}$; Bollen & Jackman, 1990; see degrees of freedom for the correlation analysis).

2.5.2.2 Explicit belief tracking and the reward effect

A 2 x 2 repeated-measures ANOVA was performed, with Belief (FB, MC) and Reward (Rewarding, Non-Rewarding) as within-subjects factors and bias as the dependent variable. The analysis revealed that the main effect of Belief was not significant ($F(1,49) = .02$, $p = .900$, $\eta_p^2 < .001$), although bias was numerically higher for the FB ($M = .27$ px, $SE = .49$) than for the MC condition ($M = .21$ px, $SE = .43$; see Section 2.5.3 Brief discussion, for possible explanations for this non-significant difference). Note that bias did not differ significantly from zero for either the FB condition ($t(49) = .55$, $p = .588$, $d = .078$) or MC condition ($t(49) = .632$, $d = .068$). In addition, the main effect of Reward was not significant ($F(1,49) = 1.85$, $p = .180$, $\eta_p^2 = .036$). The Reward x Belief interaction approached significance ($F(1,49) = 3.06$, $p = .087$, $\eta_p^2 = .059$).

Given its importance for the predictions in this chapter, this marginally significant interaction was further investigated. Paired t-tests showed that the difference between the FB condition ($M = -.60$ px, $SE = .55$) and the MC condition ($M = .42$ px, $SE = .53$) was not significant for rewarding targets ($t(49) = 1.47$, $p = .147$, $d = .209$). For non-rewarding targets, the difference between the FB condition ($M = 1.13$ px, $SE = .63$) and the MC condition (M = -.01, $SE = .68$) was also not significant (t(49) = 1.30, p = .199, d = .184). Alternatively, the difference between the rewarding (M = .42 px, SE = .53) and non-rewarding condition (M = -.01 px, SE = .68) was not significant for the MC condition (t(49) = .39, p = .629, d = .07), but this difference was significant for the FB condition (t(49) = 2.62, p = .012, d = .37). Specifically, bias was higher for non-rewarding targets (M = 1.13 px, SE = .63) than rewarding targets (M = -.60 px, SE = .55) (see Figure 2.6). One sample t-tests revealed that none of the biases reported here differed significantly from zero (all ps > .05), although this difference was a statistical trend for non-rewarding targets in the FB condition (t(49) = 1.80, p = .078, d = .255).

As the placer character was displayed centrally, it was possible that the effect of reward on FB bias was driven by avoidance or approach behaviour away from or towards the centre of the sandbox container. This possibility was investigated by calculating the absolute response distance away from the centre of the container for each trial. A paired-samples t-test revealed no significant difference (t(49) = .05, p = .961, d = .01) between this distance for rewarding (M = .25 px, SE = .01) versus non-rewarding (M = .25 px, SE = .01), suggesting that the reward effect on FB bias was not explained by mere avoidance or approach behaviour.

The analyses above resulted in no significant effects or trends (all ps > .1) when the 14 participants who failed the recall task were included.

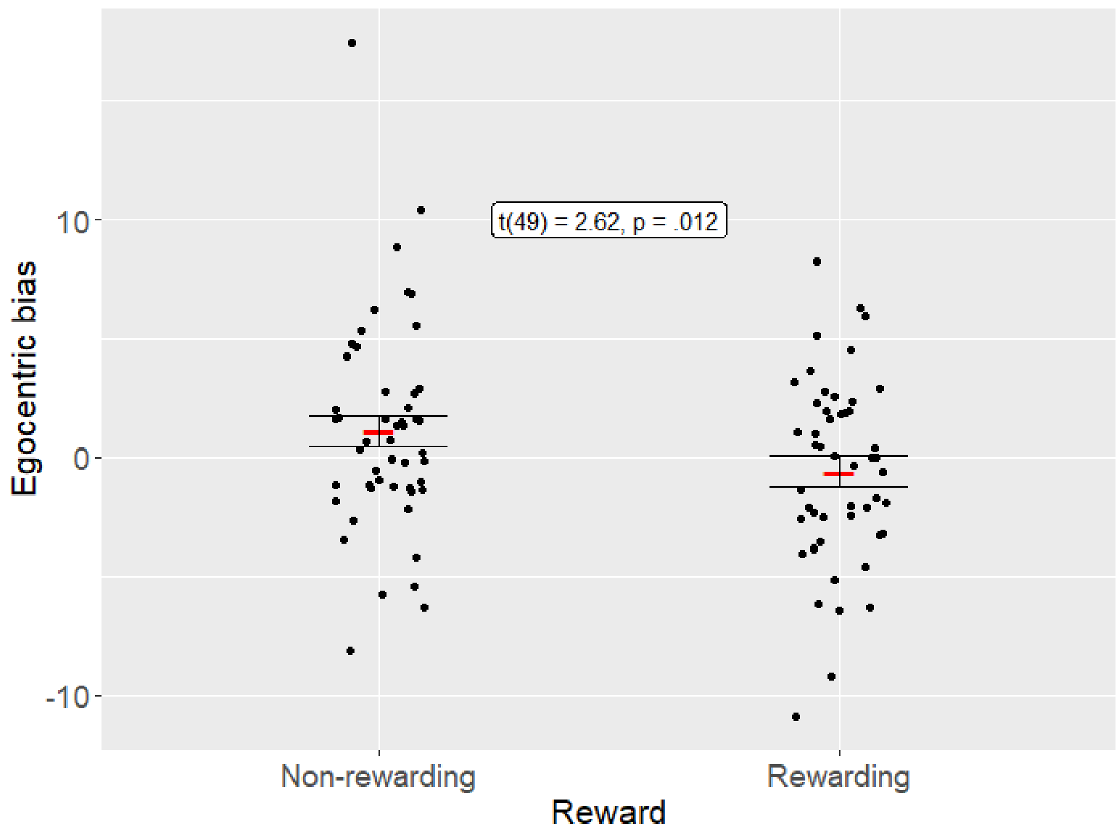


Figure 2.6 The effect of reward on explicit belief tracking in FB trials. Error bars represent 95% within-subjects confidence intervals. Dots represent individual participants.

2.5.2.3 Individual differences in the reward effect

Next, the relationship between autistic traits and the reward effect was investigated. The reward effect was calculated as the difference in FB bias between the rewarding and non-rewarding condition, per participant. A Pearson's correlation was performed between AQ and this reward effect. The result was non-significant, $r(46) = .08$, $p = .569$ (see Figure 2.7).

This result also was not significant ($p > .05$) when the 14 participants who failed the recall task were included.

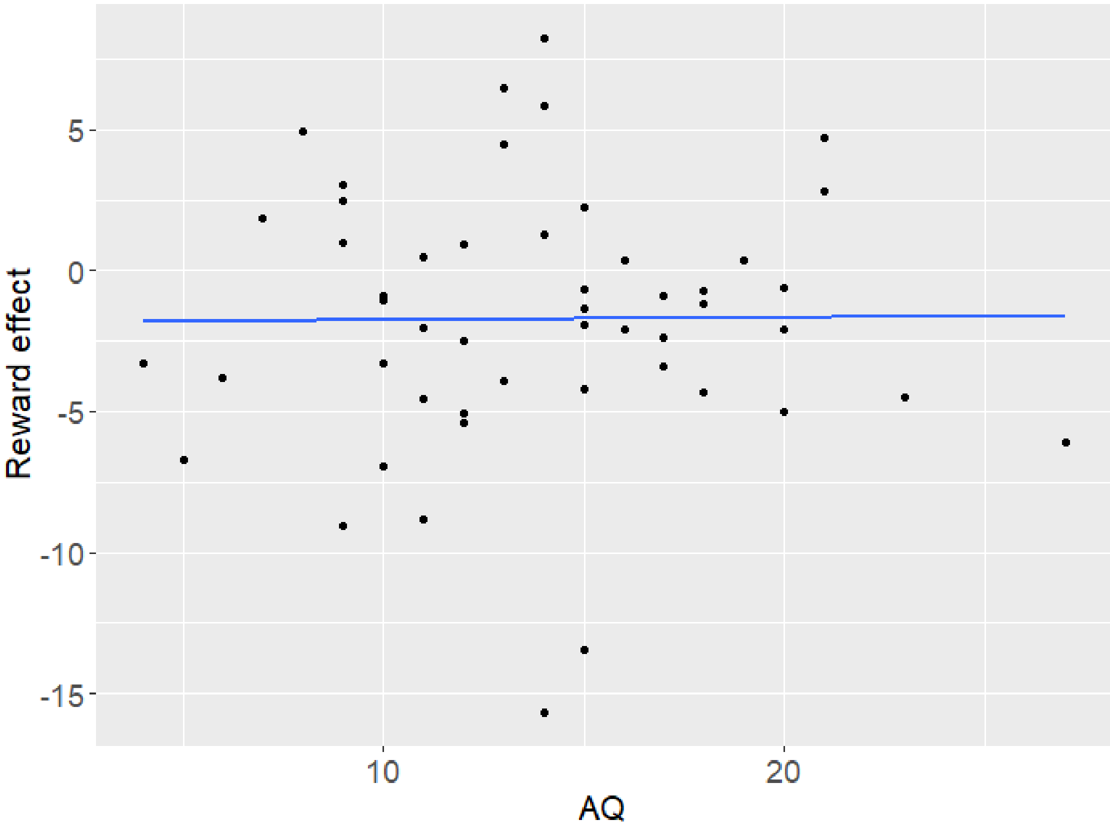


Figure 2.7 The relationship between autistic traits and the effect of reward on false belief bias. The reward effect was calculated by subtracting the participant's average bias for non-rewarding target characters from the bias for rewarding target characters on FB trials.

2.5.3 Brief discussion

The finding of a larger bias on FB trials than MC trials (e.g. Bernstein et al., 2011, Sommerville et al., 2013; Coburn et al., 2015) was not replicated in the current sample (but see Section 5.5.3 Sandbox task results, for a replication). Although this difference was numerically in the expected direction, it was less than a pixel (less than 1% of the Sandbox container width) and not statistically significant. Note also that bias did not significantly differ from zero in either the FB or MC condition. In short, there was no difference in egocentric bias when the test question tapped into false belief understanding or memory.

Although this lack of difference between the FB and MC condition was an unexpected finding, three possible explanations arise. First, in the current study, the scenarios were presented as text. In contrast, in previous studies (Begeer et al., 2012, 2016; Bernstein et al., 2011, 2017; Coburn et al., 2015; Sommerville et al., 2013), the scenarios were presented auditorily, either as recordings or narrated by the experimenter. It is possible that participants pay more attention to the exact wording of the scenarios (most importantly, the test question) when they are voiced by a human, resulting in more depth of processing due to the use of prosodic cues. Note that computerisation of the SB task likely did not result in the null finding, as a previous study replicated the FB-MC difference with a tablet-administered version of the task (Begeer et al., 2016).

Second, the current SB task consisted of 20 trials. In previous studies, the task included two (Begeer et al., 2012), four (Sommerville et al., 2013), six (Bernstein et al., 2011; Begeer et al., 2016), or nine trials (Coburn et al., 2015; Bernstein et al., 2017). The main reason for the large number in the current version was the inclusion of the reward manipulation. To test the effect of reward, it was required that every belief condition occurred for both a rewarding and non-rewarding placer. It is possible that this large number resulted in participants paying less attention to the way the test question was phrased, as they realised that the crucial information in the scenario was already given before the distraction task.

Lastly, throughout the experiment, participants had to remember which name was associated with which face/character. They were instructed to do this before the evaluative conditioning task, and that they would be tested at the end of the experiment. It is possible that this instruction effectively turned the SB task into a dual task. More specifically, whenever they were presented with a name and a face during an SB scenario, this was an occasion to reinforce their memory of the respective name-face association. When the placer returned at the end of the scenario, the participant possibly paid more attention to the face and name in favour of the phrasing of the test question. In short, all explanations given here suggest that participants distinguished between FB and MC trials less than expected.

Although the interaction between reward and belief only approached significance, the key research question warranted a further investigation of this interaction. This revealed that reward influenced bias in the FB condition, but not in the MC condition (note that the significant effect disappeared when including the participants who failed the recall task, corroborating the account above that attention to the Sandbox scenarios is important in eliciting effects). As was mentioned in Section 2.5.1.3.2 (Explicit belief tracking procedure), FB trials, in theory, require a judgment based on a belief attribution, whereas judgments in MC trials rely on a memory trace. As the main effect of belief (FB versus MC) was not significant, it is unclear whether participants in the current experiment relied on these capacities differentially for the two conditions. However, the observation that reward influenced bias in the FB condition, but not in the MC condition, suggests that the reward effect is less likely to be mediated by memory than by tracking of a belief state. Altogether, these findings provide initial evidence that reward influences explicit belief tracking.

Importantly, it was shown that the reward value of the placer character did not directly influence response location. Chen and Bargh (1999) argued that responses are predisposed towards positively valenced stimuli and away from negatively valenced stimuli. The current results suggest that such a predisposition did not occur when participants indicated their response on the sandbox container. Therefore, the mere presence of a rewarding/non-rewarding placer did not influence responses.

Lastly, there was no relationship between autistic traits and the aforementioned reward effect on FB bias. Participants higher in autistic traits were not more/less influenced by the reward value of the placer character when engaging in explicit tracking of their false belief.

2.6 Discussion

The aim of the studies discussed in this chapter was to investigate whether performance on ToM tasks is modulated by reward. For this purpose, adaptations of a commonly used belief tracking paradigm were used. Reward here was manipulated by using an evaluative conditioning paradigm which has previously been effective in associating reward value with stimuli (Hsu, Neufeld, & Chakrabarti, 2018; Neufeld & Chakrabarti, 2016; Panasiti et al., 2016; Sims et al., 2012, 2014). In addition, it was examined whether the extent of this possible reward-driven modulation was related to individual differences in autistic traits. These research questions were investigated by employing continuous belief tracking paradigms for implicit and explicit belief tracking separately.

The implicit belief tracking study employed a mouse-tracking paradigm (Freeman & Ambady, 2010). Participants responded with mouse movements when indicating their own belief, while an on-screen agent held a belief consistent or inconsistent with the participant’s belief. It was first established that participants indeed showed implicit belief tracking. Specifically, the agent’s belief had an influence on the mouse response trajectory, even when that belief was irrelevant to the task at hand (cf. van der Wel et al., 2014). However, the reward value of the agent, as conditioned through a preceding evaluative conditioning procedure, did not modulate the extent of this influence. In other words, the response trajectory did not deviate more or less towards the location corresponding to the agent's belief when that agent was rewarding or non-rewarding. Although there was no significant reward-driven modulation of implicit belief tracking, the *a priori* hypotheses allowed a further investigation of whether this modulation was related to individual differences in autistic traits. However, no such relationship was found.

The explicit belief tracking study employed a Sandbox paradigm (e.g. Bernstein et al., 2011; Sommerville et al., 2013; Coburn et al., 2015). Just as with the implicit belief tracking study, participants here had to respond to a change-of-location scenario. However, contrary to the implicit belief tracking study, the main effect found in previous studies (e.g. Bernstein et al., 2011; Sommerville et al., 2013; Coburn et al., 2015) was not replicated here. More specifically, these previous studies found that participants' responses are more biased towards their own belief in the false belief compared to the memory control condition. Coburn and colleagues (2015) argued that trials in the false belief condition rely on ignoring one's own knowledge and taking someone else's perspective, while only memorisation is needed in the memory control condition. This difference results in a more egocentric response in the false belief condition. However, the lack of difference between these conditions found here makes it unclear whether participants relied on a similar computation in both conditions or that the similarity in bias was due to other factors.

Nevertheless, the results suggested that bias in the false belief condition was larger when the placer character was rewarding, compared to when they were non-rewarding. As there was no effect of reward on bias in the memory control condition, this observation suggests that the effect of reward on performance in the Sandbox task is mediated by belief tracking, rather than memory. However, given the absence of a difference in bias between the false belief and memory control condition, this interpretation should be approached with caution. Nevertheless, the extent of the reward effect was not related to individual differences in autistic traits.

These findings pertaining to explicit and implicit belief tracking give rise to a number of interpretations. First, the studies suggest that the reward-driven modulation of belief tracking is task-dependent. Belief tracking can be influenced by the reward value associated with the person whose belief is being tracked, but only if the observer engages in explicit belief tracking. In the SB task employed here, the explicit task instructions necessitate that the participant hold a representation of the placer's belief in their working memory throughout the duration of the scenario. It is possible that the reward value of the placer can influence the depth of processing of this belief, resulting in an effect on subsequent task performance, i.e. accuracy in stating the placer's belief. The fact that the placer was represented by a face and a name in the scenario text may have reinforced this effect.

In contrast, no effect of reward was found in the MT task. This absence is likely because the belief of the agent in this task was not relevant to the task at hand. Instead, participants made their responses based on their own belief. The effect of agent's belief on the response showed that participants engaged in some degree of processing of the agent’s belief. However, the depth of processing of this belief was too shallow for the reward value associated with the agent to be of any effect. Thus, the results suggest that explicit, but not implicit, belief tracking can be influenced by reward.

In both studies discussed in this chapter there was no relationship between the reward-driven modulation of belief tracking and individual differences in autistic traits. This finding suggests that people high or low in autistic traits show similar reward-driven modulation of explicit belief tracking, and a lack of reward-driven modulation of implicit belief tracking. A previous study (Haffey et al., 2013) found that people low in autistic traits show more reward-driven modulation of mimicry of social stimuli. These authors employed a similar evaluative conditioning procedure to the one discussed in this chapter, but the task involved mimicking hands. Note that the social stimulus in Haffey et al.’s study (2013) was the only stimulus presented to the participant. In the current SB study, participants also encountered the name of the placer character in the scenario text (in addition to the object changing locations in the container). As such, the social stimulus (the image of the placer) may have prompted the participant to engage in reasoning about the placer’s belief. However, once the participant had established what the identity of the placer was, it was not necessary to pay attention to the image of the placer. Instead, the task necessitated processing the placer’s belief, the depth of which was influenced by the associated reward value. It is possible that participants higher in autistic traits attended to the image of the placer less. However, as the scenario advanced, they engaged in mental reasoning about the belief similarly to those lower in autistic traits, while being similarly influenced by the reward value of the placer.

This conjecture provides a possible explanation of why an effect of reward was found, but why this effect was not further modulated by autistic traits.

2.7 Conclusion

As was discussed in Section 2.1 (Introduction), belief tracking paradigms are often used to investigate ToM, in both children and adults. The current two studies employed paradigms in which a belief was tracked explicitly or implicitly. It was found that the reward value of the person whose belief is tracked modulates explicit, but not implicit, belief tracking. In addition, this effect was not further modulated by individual differences in autistic traits. This observation suggests that ToM, when measured as explicit belief reasoning, can be influenced by how much the target of that reasoning is considered to be rewarding. In the next chapter, it will be investigated whether a similar pattern holds for perspective taking, another process related to cognitive empathy.

3 Reward and perspective taking

Alongside Theory of Mind, perspective taking is taken as an important process underlying cognitive empathy. Perspective taking research is dominated by visual perspective taking paradigms. In these tasks, it can be quantified to which extent the self-perspective interferes with making a judgment from the other-perspective (egocentric interference). In addition, some tasks allow to quantify the inverse, namely the degree to which the other-perspective disrupts making a judgment from the self-perspective (altercentric interference). In the two studies discussed in this chapter, it was investigated whether egocentric and altercentric interference can be modulated by the reward value associated with the other person. One study employed a behavioural task in which the perspective can switch; the other employed eye-tracking to investigate solely egocentric interference. It was found that reward did not modulate either altercentric or egocentric interference. However, if the other person was non-rewarding, this reward value impeded selecting the other's perspective relative to one's own. However, this effect was not related to individual differences in autistic traits. The reward-driven modulation of perspective selection suggests that, when having to take someone else's perspective, this perspective taking happens faster when the other person is liked.

3.1 Introduction

In addition to Theory of Mind (ToM), which was discussed in the previous chapter, perspective taking is another oft-researched component of cognitive empathy. Making sense of others' actions requires that their perspective be represented. A common method of investigating perspective taking is by employing tasks that require visual perspective taking (VPT), i.e. the ability to understand what someone else can see, while often suppressing one's own viewpoint. Hence, there is a close conceptual link between ToM and VPT. As Pearson and colleagues (2013) put it, both ToM and VPT require the observer to understand that another person can have a different representation of the state of the world to theirs. For VPT, this representation is visual, whereas for ToM, it is a more abstract representation of beliefs. Indeed, some scientists have argued that VPT and ToM rely on similar cognitive mechanisms (Hamilton et al., 2009; but see Leslie, 1987).

Generally, VPT is thought to consist of two levels. VPT-1 (or a level-1 perspective) refers to a representation of *what* someone sees, whereas VPT-2 (or a level-2 perspective) is a representation of *how* someone sees an object, i.e. the 'aspectual' information (Flavell et al., 1981a, 1981b). For example, when you and someone else see the same Rubik's cube, but from opposite sides, the representation that the cube is visible to the other person relies on VPT-1, whereas the understanding that different sides with different colours are visible to the other person relies on VPT-2. As such, VPT-1 comes before VPT-2. This sequence also holds true developmentally, as VPT-1 has been shown to emerge in the second year of life (e.g. Flavell et al., 1981a, 1981b; Moll & Tomasello, 2006), whereas it takes about three years longer for VPT-2 to develop (Gzesh & Surber, 1985).

Here too lies a parallel with ToM. According to Apperly and Butterfill's two-system account (2009; see Section 2.1 Introduction), children only develop an explicit understanding that others can have a false belief around their fourth birthday (Wellman, Cross & Watson, 2001; Wimmer & Perner, 1983), whereas they already possess a rudimentary false belief understanding that is implicit and non-verbal as young as 15 months old (Onishi & Baillargeon, 2005). The system that develops first is spontaneous and effortless, but inflexible, whereas the system developing later is flexible but effortful (Apperly & Butterfill, 2009). Similarly to this dichotomy of ToM, it has been suggested that only VPT-1 can occur spontaneously, whereas VPT-2 can only be deliberate (Samson, Apperly, Kathirgamanathan, & Humphreys, 2005; Surtees, Butterfill, & Apperly, 2012). In contrast, more recent studies have found evidence that VPT-2 can also occur spontaneously (Elekes et al., 2016; Ward, Ganis, & Bach, 2019). Although the exact nature of VPT-2 relative to VPT-1 remains debated, it is clear that they are distinct processes.

As was the case with ToM, the question also arises whether reward can influence perspective taking. Do people engage in less or more perspective taking when the other person is liked? According to Pearson and colleagues (2013), two paradigms in particular lend themselves well to investigate social mechanisms in VPT, due to the social components of the respective tasks. These two paradigms are the Director task (Keysar et al., 2000) and the Dots task (Samson et al., 2010), and both focus on VPT-1.

The first VPT task discussed here is the Director task, which was developed by Keysar and colleagues (2000). Although it has since been debated whether the task taps into perspective taking or more domain-general processes (Santiesteban et al., 2015; Rubio-Fernández, 2017, for a review), this task has been used (in its original physical form, or computerised) in a variety of studies investigating VPT (e.g. Apperly et al., 2010; Dumontheil, Apperly, & Blakemore, 2010a; Dumontheil, Küster, Apperly, & Blakemore, 2010b; Epley, Morewedge, & Keysar, 2004; Santiesteban, White, Cook, Gilbert, Heyes, & Bird, 2012; Savitsky et al., 2011; Symeonidou, Dumontheil, Chow, & Breheney, 2016). In the Director task, the participant is presented with an array of shelves containing objects. A person (the 'director') is standing on the other side, giving instructions to the participant to move a certain object from one shelf to another. Importantly, some of the shelves are occluded by a board on the side of the director, meaning that some objects are only visible to the participant. In addition, the object to which the director's instructions refer belongs to a category of which there is more than one object on the shelves. Crucially, sometimes, the object that matches the instruction best *from the participant's perspective* is not the object the director can be referring to, given the occluders. Therefore, in order to respond correctly to the instructions, the participant has to suppress their own privileged information and take the director's perspective. Various studies have shown that overcoming this egocentric interference causes people to take longer and make more errors when taking the director's conflicting perspective (e.g. Apperly et al., 2010; Dumontheil, et al., 2010a, 2010b; Epley et al., 2004; Keysar et al., 2000; Santiesteban et al., 2012, 2015; Savitsky et al., 2011; Symeonidou et al., 2016). In addition, it has been shown with eye-tracking methodology, that even when participants respond correctly, their response is preceded by egocentric fixations on the occluded object (e.g Epley et al., 2004; Keysar et al., 2000; Santiesteban et al., 2012; Savitsky et al., 2011; Symeonidou et al., 2016). Therefore, the Director task provides a way of investigating egocentric interference in VPT beyond the superficial behavioural response.

The second VPT task discussed here is the Dots task (Samson et al., 2010). In this behavioural task, the participant has to respond to the number of dots shown on the walls of a virtual room. They do so by indicating whether that number corresponds to a number that was prompted at the start of each trial. On all trials, there is also an avatar shown in the middle of the room, who can only see the left or right wall. In one condition, participants have to take the other person's perspective, and respond to how many dots are visible to the avatar. Although here the own perspective is not relevant to the task at hand, participants get slowed down when the two perspectives do not match. In other words, there is egocentric interference, because it is difficult to suppress the privileged information available to the participant themselves. Notably, egocentric interference *de facto* reflects the difficulty to take another person's perspective, and is therefore commonly used to quantify perspective taking.

In another condition, participants take their own perspective. Interestingly, participants also get slowed down here when the perspectives are mismatching (i.e. a different number of dots is visible to the participant versus avatar). In other words, there is altercentric interference caused by the automatic computation of the other person's perspective. Hence, even when the participant is not instructed to take the other's perspective, it is still automatically and unintentionally computed. Therefore, the Dots task is a versatile paradigm, as it allows for the investigation of both egocentric and altercentric interference. Since its inception, the task has been used to show both egocentric and altercentric interference in other VPT studies (e.g. Marshall, Gollwitzer, & Santos, 2018; Qureshi et al., 2010; Furlanetto, Becchio, Samson, & Apperly, 2016). Importantly, a number of studies have claimed that the Dots task taps into attentional cueing mechanisms rather than perspective taking (e.g. Santiesteban, Catmur, Hopkins, Bird, & Heyes, 2014; Cole, Atkinson, Le, & Smith, 2016; Wilson, Soranzo, & Bertamini, 2017). In response, Marshall and colleagues (2018) concluded based on an experiment that the mechanism underlying interference in the Dots task depends on the nature of the stimulus central to the task. According to these authors, with a non-social stimulus (an arrow), interference relies on attentional cueing, whereas with a social stimulus (an avatar), interference does rely on perspective taking. In this thesis, the stance of Marshall and colleagues (2018) is taken, whilst acknowledging the ongoing debate.

Together, the Dots task and Director task allow to investigate several facets of VPT. The Director task only requires taking the other's perspective and can therefore only be used to examine egocentric interference. In contrast, the Dots task quantifies both egocentric and altercentric interference, as either the self- or other-perspective is taken. On the other hand, the Dots task is (currently) bound to measures of accuracy and reaction time, whereas Director task designs often also employ eye-tracking. Due to their complementary measures, the combination of these two tasks seems particularly fruitful to examine social mechanisms underlying VPT.

To this date, the studies examining whether reward value can influence VPT have been scarce. For example, such an investigation has not yet been done for the Dots task. In contrast, one study used an indirect reward manipulation in a study employing the Director task. Savitsky and colleagues (2011) let participants perform the Director task once with a friend who took on the role of the director, and once with a stranger. Surprisingly, the authors showed that, although friends and strangers are just as likely to make egocentric errors, friends took longer to overcome their egocentric interference, possibly because friends assume each other's perspectives are similar. However, as friendships can vary in their quality, it would be interesting to see whether a more experimentally controlled manipulation of reward value would have an effect on VPT too.

In addition, whereas Savitsky et al.'s study (2011) study suggested that reward value may influence VPT, it did not examine whether this effect is further modulated by autistic traits (see Section 1.1.2 for an explanation of why this is important to investigate). Whereas ToM has often been shown to be impaired in individuals with autism, the results for VPT have been mixed (see Pearson et al., 2013, for a review). An investigation of a possible relationship between autistic traits and the reward-driven modulation of VPT using more than one task could therefore shine a light on the circumstances under which a deficit appears in autism.

3.2 Aims

To investigate the possible reward-driven modulation of VPT and its relationship to autistic traits, two studies were conducted. The first used the Director task to investigate the effect of reward on egocentric interference in a context where verbal instructions are received from someone else. The second used the Dots task, allowing to investigate both egocentric and altercentric interference. By using the same evaluative conditioning procedure as the one described in Chapter 2, the current studies sought to experimentally manipulate reward value in a controlled fashion. In addition, by measuring autistic traits, it could directly be quantified to what extent autistic traits would relate to the aforementioned reward-driven modulation.

3.3 Predictions

It is predicted that egocentric interference is shown in both the Dots and Director task, whereas the Dots task will also reveal altercentric interference. Given the results in Chapter 2, it is expected that there will be no reward-driven modulation of altercentric interference. For egocentric interference, the prediction is that there will be an effect, although the hypothesis is non-directional, as the Sandbox study (Section 2.5) showed a positive effect on egocentric interference, whereas Savitsky et al.'s (2011) study showed a negative effect. In addition, as in Chapter 2, it is expected here that any reward-driven modulation will be negatively related to individual differences in autistic traits. Note that, for the remainder of this chapter, 'interference' will be used to refer to the psychological phenomenon, whereas 'bias' refers to the quantification thereof. To phrase the main predictions more specifically:

1) Reward value will influence the egocentric bias in the Dots task, as reflected by an increase in RT when the self-perspective does not match the (to be taken) other-perspective.

2) Reward value will also influence altercentric bias in the Dots task, as reflected by an increase in RT when the other-perspective does not match the (to be taken) self-perspective.
3) Reward value will influence the egocentric bias in the Director task, as reflected by the increase in looking time at the object matching the self-perspective relative to the looking time at the object matching the director's perspective.
4) All these reward-driven modulations will be negatively related to autistic traits.

3.4 Director task study

3.4.1 Methods

3.4.1.1 Participants

Forty-nine participants completed the DT study (*N* females = 40, mean age = 20.53, *SD* = 4.43). This sample size was based on the sample size used by Savitsky et al. (2011). The participants were all undergraduate students at the University of Reading. They signed up for the study through the University of Reading research panel. Students could not sign up for this study if they had already participated in the mouse-tracking study (Section 2.4) or the Dots task study (Section 3.5). Upon completion of the study, they received course credit. In addition, they received £2.50 as a fixed reward for the evaluative conditioning task (see Section 2.4.1 Methods). Ethical approval was received from the School of Psychology Ethics Committee and all participants provided informed consent prior to their participation.

3.4.1.2 Stimuli

3.4.1.2.1 Evaluative conditioning task

The stimuli that were used for the evaluative conditioning task in Chapter 2 were also used here, with one difference: photographs were cropped at the shoulders rather than at the neck (see Appendix D). In addition, five volunteers (three females, two males) provided voice recordings as auditory accompaniment for the feedback messages used in the task.

3.4.1.2.2 Director task

The two female face stimuli that were conditioned in the evaluative conditioning task, with their respective associated voices, were also used in the Director task. In addition, one other female face stimulus was used, in the practice block of the Director task. A sixth volunteer provided voice recordings for this face.

Objects for the Director Task were taken from a data set of images collected previously by a fellow lab member, Garret O'Connell. In addition, a computer-generated image of a four-by-four shelves unit was used.

3.4.1.3 Procedure

When the participant arrived, they were first informed about the tasks in the experiment. They were told the study would involve a 'gambling game' (the evaluative conditioning task), a 'communication game' (the Director task), and a questionnaire. They were told that their earnings at the end of the study would be influenced by their performance in the gambling game. In addition, it was mentioned that their eye movements would be tracked non-intrusively during the communication game. The entire experiment was presented on a 24-inch monitor with a computer running Windows 7. Instructions were read aloud to the participant before each task. During the evaluative conditioning task and the filling out of the questionnaire, the experimenter was outside of the participant's view. Participants were debriefed at the end of the experiment.

3.4.1.3.1 Evaluative conditioning task

The same evaluative conditioning task as described in Section 2.4.1.3.1 (Evaluative conditioning task) was used for this study. This design, using a 90% win, 60% win, 60% loss, and 90% loss condition, was chosen, to maintain similarity to the design used by Sims and colleagues (2012). To reduce the number of faces seen during this task, only one (female) face was used in the practice block. As explained in Section 3.4.1.3.2, instructions in the Director task were given by pre-recorded voices. To condition voices in addition to faces, voices were also used for the evaluative conditioning task, with the same face-voice pairing as in the Director task. In particular, text feedback about the outcome of the trial was accompanied by matching and appropriately valenced audio feedback ('You win', 'You lose, or 'It's a draw'). The evaluative conditioning task had two counterbalanced versions, such that each of the two faces used for the Director task was associated with a 90% win chance or 90% loss chance once.

3.4.1.3.2 Director task

The Director task was presented with MATLAB Release 2016b (The Mathworks, Inc.), using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997; Kleiner, Brainard, Pelli, Broussard, Wolf, & Niehorster, 2007). Eye-tracking was done with a Tobii X2-60 eye-tracking device (Tobii Technology) attached to the monitor. This eye-tracker recorded the two-dimensional point on the screen at which the participant looked, with a temporal resolution of 60 Hz.

The paradigm used here was an adaptation of the Director task (Keysar et al., 2000). Participants were told that they would receive instructions, from a person shown on the screen, to move a certain object from one shelf to another. They were told that shelves that had boards behind them would not be visible to the person on the screen. They were then presented with a paper example of a trial (see Figure 3.1) and asked which object they would move, and to where, given the instruction ‘Move the top cross over the paper’. They were asked to imagine that the instruction was coming from the person behind the shelves. To respond correctly on the example trial, the participant had to indicate that the object on shelf 5 (the red cross) was to be moved to shelf 11 (above the paper sheet and to the left of the black cross). If the response was correct, the experimenter tested their understanding by asking the reason for that choice. If the response was incorrect, the experimenter discussed the correct choice. Once it was clear that the participant understood the task, the experimenter explained how the participant had to use the keyboard and mouse to respond during a trial.

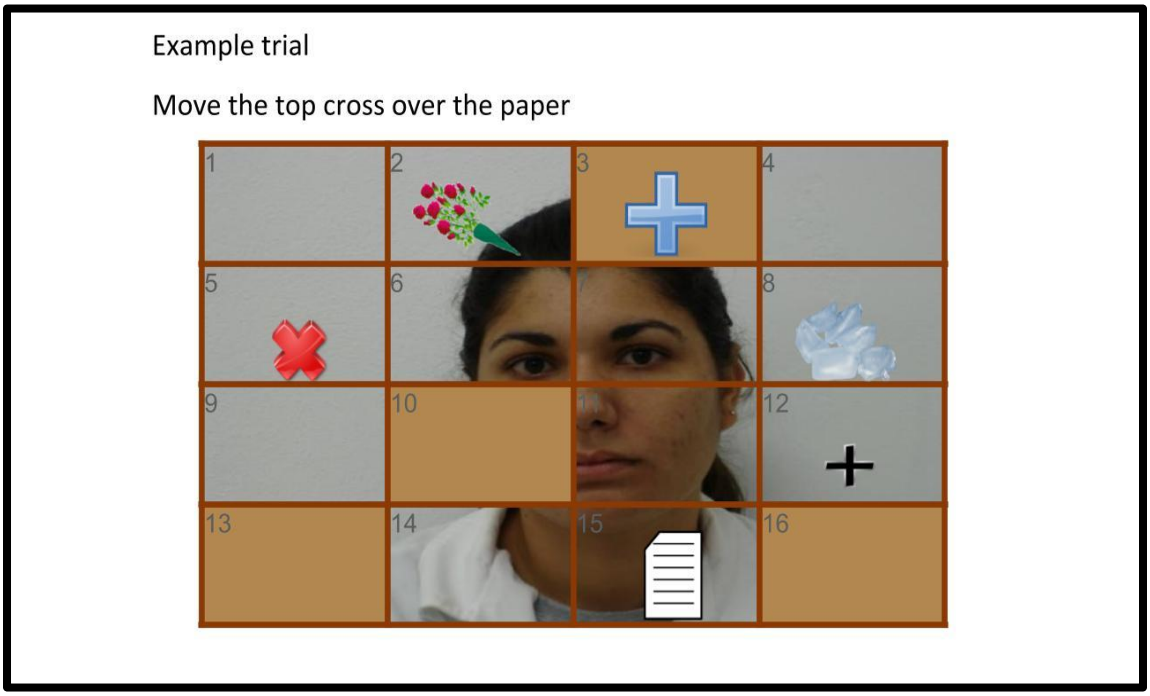


Figure 3.1 Example trial shown on paper before the Director task

Next, the eye-tracker was calibrated, with the participant sitting at approximately 60 cm from, and aligned with, the centre of the screen. The participant went through a calibration and validation step, during each of which they had to fixate on dots corresponding to the centre of the 16 shelves during the Director task. Once calibration was successful, the participant started with the practice trials.

Each trial proceeded as follows (see Figure 3.2): first, the image of a face (200 x 150 pixels) was shown in the middle of the screen, while the mouse cursor was made invisible. Concurrently, the audio instruction about which object to move and where to move it to was played. The participant had the opportunity to replay this instruction by pressing ‘R’ on the keyboard. Otherwise, they pressed the spacebar to proceed to the shelves. As soon as the shelves appeared, the eye-tracker started to record the eye movements of the participant (phase 1). The four-by-four array of shelves contained six objects visible to the participant. Between one and three of these objects were on shelves that had a wood-textured board behind them. Behind the shelves, the image of the 'director' was shown, filling the screen. The director was always one of two female heads, cropped at the shoulders. The reason for including only two of the four conditioned faces was that the characteristics of the other VPT paradigm discussed in this chapter necessitated the exclusion of faces associated with a 60% win/loss for the analysis (see Section 3.5.1.3.2). To maintain consistency across these VPT paradigms, only the female faces associated with a 90% win/loss were used here. This way, all 48 main trials could be kept for analysis. The participant was told beforehand to click on the mouse once they had made their decision about both which object to move (the target object) and where to move it to (the destination shelf). This first click resulted in the mouse cursor appearing in the middle of the shelves. Next (phase 2), the participant clicked on the shelf containing the target object, at which point the selected shelf was highlighted in yellow. Finally, they clicked on the destination shelf (phase 3). As a result, the target object moved to the highlighted destination shelf and remained there for 1,000 ms. The task then proceeded to the next trial.

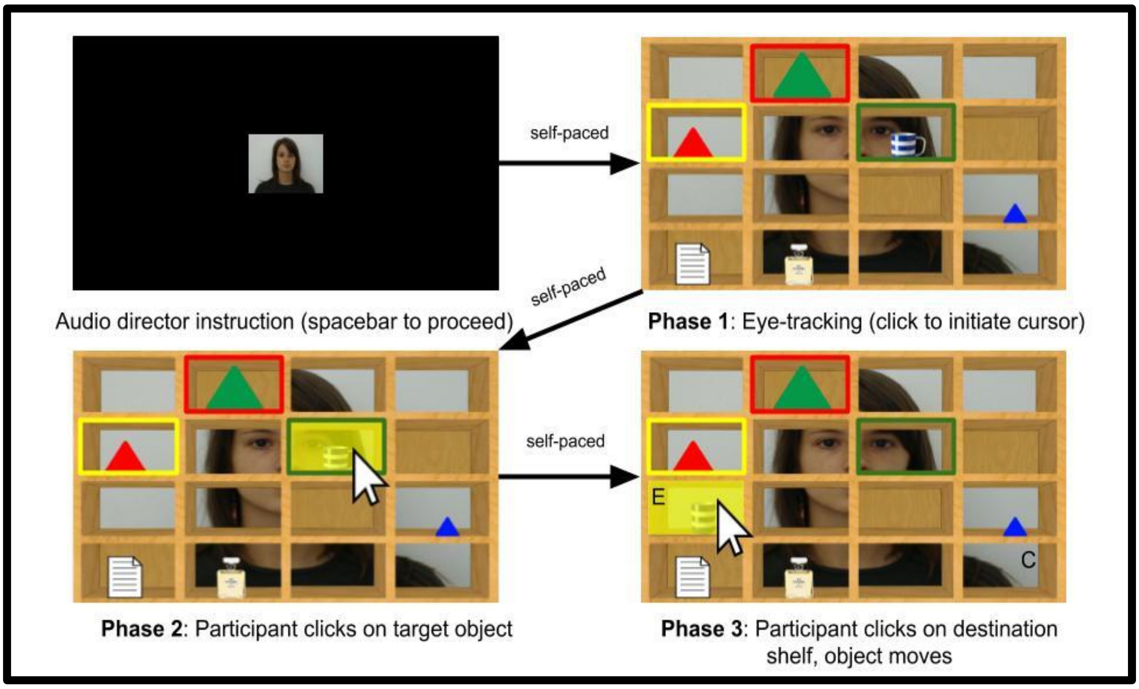


Figure 3.2 Director task trial procedure. Each trial started with an audio instruction, while the face of the 'director' was presented at fixation (e.g. 'Move the mug under the top triangle'). After the audio instruction ended and once the participant pressed the spacebar, the trial proceeded to the shelves (phase 1). Here, the participant made the decision which object to move and where to move it to, after which they clicked to initiate the cursor (enlarged here) in the middle of the screen. The trial then proceeded to phase 2, in which the participant clicked on the target shelf (a green border has been added here for illustration purposes). Their selection was highlighted, after which phase 3 started, where they clicked on the destination shelf (yellow border). The selected object then moved to this shelf, which was subsequently highlighted. The shelf configuration here was used for one Experimental trial (depicted here) and one Control trial. In the Control trial, the instruction was 'Move the mug under the *bottom* triangle'. Therefore, in phase 2, the correct response was the same, whereas in phase 3 it depended on the condition (E: destination shelf for the experimental condition; C: destination

shelf for the Control condition). Note that in both trials, the distractor was taken to be the occluded object (red border).

During the first four practice trials, the participant saw the shelves from the director's side of the shelves, with a blank background showing behind the shelves. These practice trials were included to familiarise the participant with the idea that the director could not see what was behind the boards, as their vision was blocked off. During the second four practice trials, the participant saw the shelves from the side they would be on for the rest of the experiment. After the practice trials, the participant was informed that there would be a break halfway through the task.

There were two conditions: Experimental and Control. In both conditions, there were three objects that belonged to the same semantic category (e.g. triangles), although they differed in size. These three competing objects never shared a shelf row. The remaining objects belonged to different semantic categories. One of the three competing objects (the top or the bottom one) was always behind a board. In the Control condition, the director's instruction was always unambiguous (e.g. the bottom triangle was the bottom triangle from both perspectives). In the Experimental condition, the director's instruction was always ambiguous (e.g. the triangle that matched the instruction from the director's point of view was not the same as the triangle that matched the instruction from the participant's point of view). The instruction was always to move an object below or above another object (e.g. 'move the bottom triangle over the mug'). To respond correctly, the participant had to take the director's perspective. In the Experimental condition, this meant ignoring their own privileged knowledge. In half the Experimental and Control trials, there was competition in terms of which was the destination shelf (e.g. ‘move the mug under the *top triangle*’, depicted in Figure 3.2). In the other half, there was competition regarding the target object (e.g. ‘move the *top ball* under the perfume’). The former type was a novel feature compared to previous studies, introduced to reduce repetitiveness in the task. Each Experimental trial was yoked to one Control trial in terms of the distribution of objects on the shelves.

There were 48 main trials in total, with each face appearing in 12 Experimental and 12 Control trials. Each of six object sets occurred eight times. The order of trials was pseudorandomised: there were never more than two Experimental or two Control trials in a row and each director was never shown more than twice in a row. The order of the two blocks and the order of the directors were counterbalanced across four versions of the task.

During the Director task, the experimenter sat behind the monitor opposite the participant, to monitor data quality. Two plots were produced on the experimenter’s monitor throughout the task: one showing whether the eye-tracker was picking up the participant’s eyes, at a frequency of 60 Hz; the other a superimposition of the participant’s gaze pattern onto the shelves display. The superimposed gaze pattern was colour-coded based on the phase of the trial. If either of these plots indicated faulty eye-tracking, the participant was asked to move back to their original position. If the gaze pattern plot indicated that the participant clicked to enter phase 2 before exploring the shelves, the instruction to decide before they clicked was repeated to the participant.

It should be noted that the current version of the Director task differed from previous versions in some respects (e.g. Dumontheil et al., 2010a, 2010b; Epley et al., 2004; Symeonidou et al., 2016). For instance, the use of different Director voices was a novelty, but necessary given the different conditioned faces. In addition, participants heard the instruction from the director prior to presentation of the shelves. This change was made so that participants knew what to look for on the shelves, potentially leading to less noise in their gaze behaviour leading up to their mouse response. In addition, this change allowed to implement the option to repeat the instruction, allowing participants to achieve clarity on the instruction before starting the trial proper. Lastly, some studies have included experimental and control trials in which the director was replaced by a camera (a non-social stimulus) or was absent altogether (Dumontheil et al., 2010a, 2010b; Santiesteban et al., 2015; Symeonidou et al., 2016). In these studies, it was investigated to what extent mentalising underlies perspective taking in the Director task. The current study excluded such a ‘non-social’ condition (cf. Santiesteban et al., 2012), to prevent conditioning extinction effects due to a large number of trials.

3.4.1.3.3 Trait measure

The Autism Spectrum Quotient (AQ; Baron-Cohen et al., 2001b) was used to quantify autistic traits.

3.4.1.4 Data analysis

As the playing card task was only used for evaluative conditioning, data generated by that task was not further analysed.

For each trial of the Director task the measure of interest was based on gaze behaviour before the participant clicked the mouse to activate the cursor. Note that the participant had been instructed to decide on both the target object and destination shelf before their first click. Therefore, the gaze behaviour analysis was restricted to phase 1. Previous studies have used reaction time until target object selection as the primary response measure (e.g. Santiesteban et al., 2012; Savitsky et al., 2011). However, due to restricted computer processing capacity, mouse clicks were not always registered if they rapidly succeeded previous clicks. For this reason, reaction time (of registered clicks) was not reliable, and this measure was not used.

For each trial, it was first quantified what percentage of phase 1 was spent looking at 1) the correct shelf, and 2) the distractor shelf. As mentioned before, depending on the director's instruction, there was either competition regarding which target object was to be moved (e.g. 'Move the top mug…'), or to which destination shelf the object had to be moved (e.g. '… under the top triangle'). In target competition trials, the correct shelf was the target object, whereas the distractor shelf was the occluded object from the same category. In destination competition trials, the correct shelf was the object above/under the destination shelf (which was always empty), whereas the distractor shelf was also the occluded object from the same category. For both Experimental and Control trials, the same occluded object was taken as the distractor.

It was first analysed whether eye-tracking findings of previous studies could be replicated, by looking at the degree to which the occluded object was distracting (cf. Santiesteban et al., 2012; Savitsky et al., 2011). Specifically, it was investigated whether percentage gazed at the distractor shelf was different between Experimental and Control trials.

To establish a measure of VPT, the subsequent analysis focussed on Experimental trials. The rationale behind this focus was that there is a conflict between the two perspectives on Experimental trials, so that in order to respond correctly, the participant has to discount their privileged perspective and take the distractor's perspective. Therefore, the measure for VPT was calculated by subtracting percentage gazed at the distractor shelf from percentage gazed at the correct shelf (both in phase 1). Note that the higher the value for this measure, the smaller the egocentric bias was. This difference score was used, as it reflects the dominance of the shelf corresponding to the director's perspective relative to that of the shelf corresponding to the participant's perspective. This compound measure was subsequently taken to investigate possible reward-driven modulation of VPT. Note that basing this measure only on Experimental trials allowed a comparison with the results in Chapter 5, as the Control trials of the Director task discussed in that chapter were slightly different compared to the current version (see Section 5.4.3.3). However, for completeness, the VPT measure was calculated for Control trials here too. The difference between this measure for Experimental versus Control trials was subsequently used to investigate 1) whether reward influences VPT, whilst controlling for the location of the shelves and 2) whether autistic traits correlate with individual differences in this reward effect. In addition, to qualify this second analysis, it was investigated whether autistic traits correlate with VPT itself, independent of reward. The results of these analyses are reported in Section 3.4.2.5.

In addition, another approach to investigate reward-driven modulation was taken. Savitsky and colleagues (2011) quantified VPT in the following way: first they measured the time until the first fixation (at least 100 ms) on the distractor shelf. Next, they measured the time until the last fixation on the correct shelf. Finally, the latency of this correction was taken to reflect VPT, especially on

Experimental trials. This same two-fold approach was taken here as an addition to the correct shelf versus distractor shelf approach.

For correlations, observations were excluded from each individual analysis separately if they were identified as bivariate outliers (Cook's distance > $4/N_{participants}$; Bollen & Jackman, 1990).

All p-values reported with statistical tests below are two-tailed.

3.4.2 Results

3.4.2.1 Data pre-processing

Pre-processing steps and the analysis were performed with R (R Core Team, 2018). Prior to pre-processing of trials, four participants were excluded as they had clicked on the correct target object and destination shelf in less than half the trials (see Table 3.1 for a summary of accuracy). In addition, four participants were excluded due to unusable gaze data. As a result, the final sample consisted of 41 participants (35 females), whose age ranged between 18 and 48, (*M* = 20.66, *SD* = 4.75) and whose AQ ranged between 6 and 35 (*M* = 16.20, *SD* = 6.20).

Table 3.1 Accuracy data for the Director task.

	Director condition	
Reward condition	Experimental	Control
Rewarding	11.27 (.98)	11.15 (1.33)
Non-rewarding	11.10 (1.28)	11.12 (1.38)

Note: Values represent mean (SD) of correct number of trials, out of 12, per condition.

For the remaining sample, the following steps were performed: first, it was established which of the two eyes had the most data. In order to do this, it was quantified what percentage of gaze points recorded across all trials was invalid. Gaze points were invalid if the participant blinked or looked away from the screen. Another possibility was that the eye-tracker was still initialising. When the participant continued from the director instructions to the shelves (i.e. the response phases), the eye-tracker was activated by the task script. Gaze points were invalid until the eye-tracker had located the participant's eyes. Second, trials were excluded if the participant had not clicked on the correct target object and destination shelf (7.0% of all trials). Third, all gaze points after phase 1 were removed. Fourth, the initial 333.33 ms (the first 20 gaze points) of eye gaze data were removed. As the eye-tracker had not fully initialised as soon as the shelves appeared, a trial always started with invalid data points. Next, trials were removed if for the remaining gaze points less than 50% was valid (2.8% of all trials). It was then quantified how much time was spent looking at each of the 16 shelves. If the participant had not looked at the correct shelf *and* distractor shelf (as defined in Section 3.4.1.4 Data analysis), the trial was excluded too (24.6% of the remaining trials. No further outlier analysis was performed. The final sample consisted of 1,342 trials (mean N = 32.73, SD = 7.57) across 41 participants.

Sixteen participants never repeated the audio instruction. The remaining 25 participants repeated the instruction on average on 2.37 trials (SD = 3.18).

3.4.2.2 Perspective taking

First, it was investigated whether there was a difference in percentage gazed at the distractor shelf between the Experimental and Control condition. The paired-samples t-test revealed a significant difference, $t(40) = 8.32$, $p < .001$, $d = 1.30$. Distractor gaze percentage was higher for Experimental ($M = 12.11\%$, $SE = .37$) than Control trials ($M = 8.22\%$, $SE = .39$). This finding shows that participants were more distracted by the distractor shelf when this shelf created a conflict with their own perspective than when it did not.

3.4.2.3 The effect of reward

Next, it was investigated whether there was an effect of reward on VPT in the Experimental condition. For this purpose, a paired-samples t-test was performed, comparing VPT of a rewarding versus non-rewarding director. Note that VPT was quantified as follows: per Experimental trial, the time spent looking at the distractor shelf was divided by the total duration of trial phase 1. Using the same method, a gaze percentage was calculated for the correct shelf. Finally, to establish a measure of VPT, the distractor shelf gaze percentage was subtracted from the correct shelf gaze percentage. The results showed that there was no significant difference in VPT, $t(40) = .25$, $p = .802$, $d = .04$ (see Figure 3.3), between trials with a rewarding director ($M = 8.10\%$, $SE = .80$) and a non-rewarding director ($M = 7.82\%$, $SE = .96$).

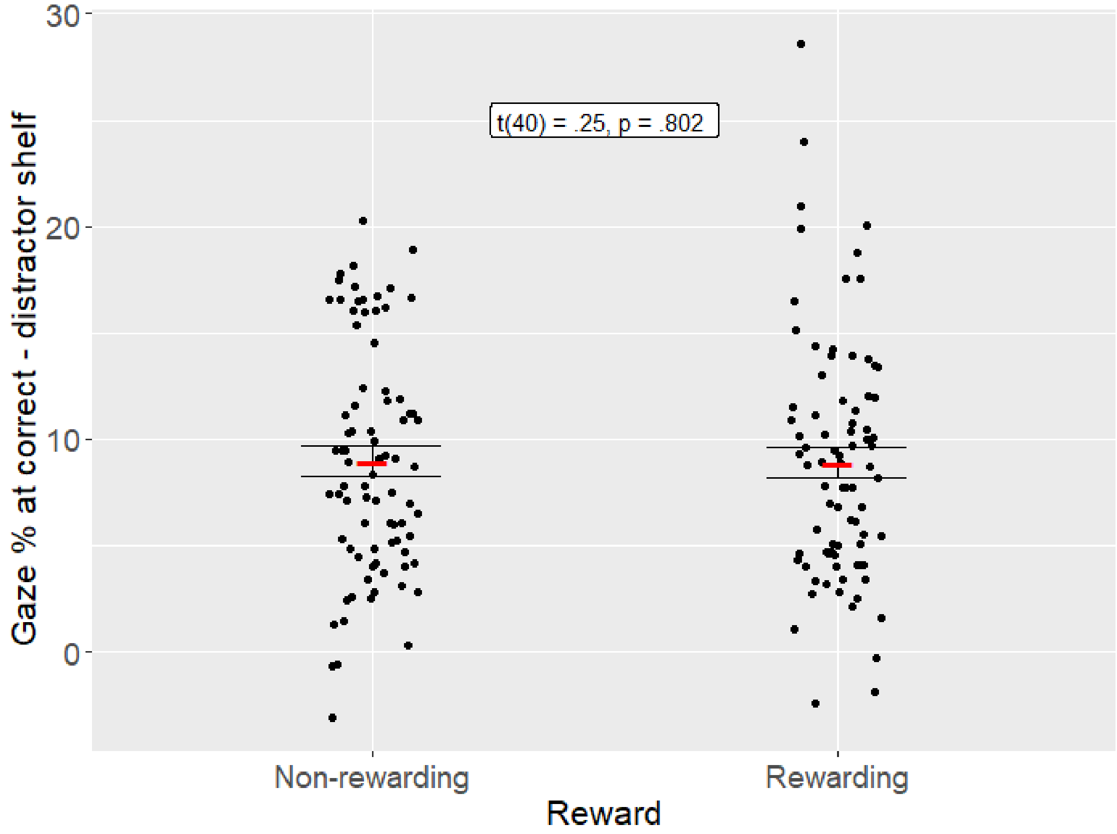


Figure 3.3 Effect of reward on visual perspective taking in the Director task. The effect was measured as the difference in percentage of a trial gazed at the correct and distractor shelf (in the Experimental condition). Error bars represent 95% within-subjects confidence intervals.

In addition, the effect of reward was investigated by examining latency to correct egocentric fixations. One participant was excluded for this step as they had no fixations on the distractor shelf when the director was rewarding. For the remaining 40 participants, it was first investigated whether egocentric fixations happened more with rewarding or non-rewarding directors. A paired-samples t-test resulted in a non-significant difference, $t(39) = .55$, $p = .584$, $d = .09$ between the rewarding ($M = 86.5\%$, $SE = 2.18$) and non-rewarding condition ($M = 87.63\%$, $SE = 2.33$). Next, a paired-samples t-test revealed no difference in the latency to correct this egocentric fixation, $t(39) = .69$, $p = .509$, $d = .11$, between the rewarding ($M = 2042.51$ ms, $SE = 140.70$) and non-rewarding condition ($M = 1964.70$ ms, $SE = 112.97$).

3.4.2.4 Individual differences in the reward effect

Third, it was investigated whether the size of the reward effect was modulated by individual differences in autistic traits. Note that the effect of reward was non-significant, but the *a priori* hypothesis warranted this analysis. The reward effect was quantified by subtracting the participant's perspective taking score in the non-rewarding condition from the rewarding condition. A Pearson correlation was performed to investigate the relationship between this composite score and individual differences in autistic traits. No significant correlation was found, $r(39) = .13$, $p = .401$. Alternatively, the reward effect was quantified as the difference between the rewarding and non-rewarding condition in the latency to correct egocentric fixations. A Pearson correlation found no significant correlation with individual differences in autistic traits, $r(38) = -.14$, $p = .393$.

3.4.2.5 Perspective taking as the difference between Experimental and Control

As an alternative approach, the effect of reward on VPT was investigated using the difference between VPT in Experimental versus Control trials, rather than in Experimental trials only. A paired-samples t-test was performed, comparing VPT of a rewarding versus non-rewarding director, taking the Experimental-Control difference score rather than the score on Experimental trials only. This test revealed that there was no significant reward effect, $t(40) = .33$, $p = .740$, $d = .05$.

As was done in Section 3.4.2.4, a Pearson correlation was performed to investigate the relationship between autistic traits and this (non-significant) reward effect. This correlation was not significant, $r(38)$, -.03, $p = .867$. Lastly, a Pearson correlation revealed that there was no significant relationship between autistic traits and individual differences in VPT when collapsing the reward conditions, $r(39)$, = -.13, $p = .388$.

3.4.3 Brief discussion

The finding that participants look more at the distractor shelf in the experimental than in the control condition was replicated (cf. Santiesteban et al., 2012; Savitsky et al., 2011). This observation indicates that, in the experimental versus control condition, participants spent relatively more time looking at the object from the crucial category that was not visible to the director. Participants still looked at this distractor shelf in the control condition, but this gaze behaviour was likely because participants looked at all three competing objects from the crucial category to establish the correct object, even during control trials. In contrast, they looked at the distractor more when it was the option corresponding to the participant's perspective. Note, however, that this measure is suboptimal in quantifying VPT, as it does not take into consideration how much the participant looked at the shelf the director *was* referring to.

Therefore, the difference between correct shelf gaze percentage and distractor shelf gaze percentage in the experimental condition was taken to measure VPT. It was found that the reward value of the director did not modulate VPT. In addition, when taking the latency to correct for egocentric fixations as a measure of VPT, there was no effect of reward either. Note that Savitsky and colleagues (2011) found indirect suggestive evidence for an effect of reward, as they reported less VPT between friends than strangers. However, the current experiment found no effect of reward using a direct experimental manipulation of reward (i.e. conditioning of faces with reward). Although in the direction of the finding reported by Savitsky and colleagues (2011), the effect of reward on the latency to correct egocentric fixations was non-significant here. This observation suggests that subtle changes in the employed paradigm can eliminate the effect of reward in the Director task.

Lastly, there was no correlation between individual differences in autistic traits and the effect of reward on VPT. Therefore, it can be concluded that effect of reward on VPT was absent along the full spectrum of autistic traits that was investigated in the current study.

3.5 Dots task study

3.5.1 Methods

3.5.1.1 Participants

Eighty-five participants completed the Dots task (*N* females = 72, *M* age = 21.25, *SD* = 5.98). The participants were all undergraduate students at the University of Reading. They signed up for the study through the University of Reading research panel or were personally recruited by the experimenters. Students could not sign up for this study if they had already participated in the mouse-tracking study (Section 2.4) or the Director task study (Section 3.4). Upon completion of the study they either received monetary compensation of £5 or course credit. In addition, they received £2.50 as a fixed reward for the evaluative conditioning task (see 2.4.1 Methods). Ethical approval was received from the School of Psychology Ethics Committee and all participants provided informed consent prior to their participation.

3.5.1.2 Stimuli

3.5.1.2.1 Evaluative conditioning task

Stimuli used for the evaluative conditioning task were similar to those used in the Director task (see Section 3.4.1.2.1), except that only female faces were used (see Section 3.5.1.3.2 (Dots task) below for an explanation). These were taken from the same database (de Oliveira & Thomaz, 2006). In addition, no voice recordings were used for feedback in this version of the evaluative conditioning task.

3.5.1.2.2 Dots task

Stimuli used for the Dots task were adapted from publicly available materials (Samson & Apperly, 2015) used for the Samson et al. (2010) study. The background for the current stimuli, an empty room, was directly taken from these materials. The avatar in the middle of the room was replaced by the photograph of a female head cropped at the shoulders, facing left or right. The same female faces as in the evaluative conditioning task were used. The images were enlarged to fill almost a third of the stimulus, to increase distinguishability (the face of the avatar in the Samson et al. (2010) study only covered a fraction of the screen). In addition, the dots on the walls of the room were enlarged relative to the original paradigm to make it clearer that they were visible to the person in the middle.

3.5.1.3 Procedure

When the participant arrived, they were first informed about the tasks in the experiment. They were told that the study would involve a ‘gambling game’ (the evaluative conditioning task) and a task involving visual scene perception, in which they had to indicate what they or someone else could see in a virtual room (the Dots task). They were told that their earnings at the end of the study would be influenced by their performance on the gambling game. The entire experiment was presented on a 17-inch monitor with a computer running Windows 7. Instructions were read aloud to the participant before each task. During the tasks, the experimenter left the experiment room. Participants were debriefed at the end of the experiment.

3.5.1.3.1 Evaluative conditioning task

The procedure for the evaluative conditioning task was the same as for the version used in the Director task study (see 3.4.1.3.1), using a 90% win, 60% win, 60% loss, and 90% loss condition. The only difference was that there was no audio feedback accompanying the text feedback shown after participant responses.

3.5.1.3.2 Dots task

The VPT task used in this study was adapted from the Dots task developed by Samson and colleagues (2010). The task in the current study was presented with E-Prime 2.0 (Psychology Software Tools). Prior to the task, participants received written instructions on-screen, which were read out by the experimenter. They were told that there would be a break halfway through the task, after which the practice phase started. The experimenter left the room during the practice phase, once they had established that the participant was showing an understanding of the task in their responses. If the participant did not show understanding of the task, the experimenter repeated the instructions, after which the main phase started.

Each trial proceeded as follows (see Figure 3.4): first, a fixation cross appeared, followed by the word “YOU” or “SHE”, followed by a digit (0, 1, 2 or 3). Each of the three stimuli was presented for 750 ms in the middle of the screen at font size 72, and was followed by a blank screen presented for 500 ms. These instructions were followed by the image of a room. In the middle of the room, a person was shown, facing left or right. A total of zero, one, two, or three dots was shown on the left and/or right wall. When the perspective instruction was “YOU” (Self), participants had to indicate whether or not they saw the number of dots that corresponded to the digit instruction (0-3). When the perspective instruction was “SHE” (Other), participants had to indicate whether or not the person in the room saw that number. They did so by pressing ‘Z’ if the number was correct (Yes), or ‘M’ when it was incorrect (No; counterbalanced). The task proceeded to the next trial upon this response. If the participant did not respond within 2,000 ms, the task also proceeded. During the practice phase (26 trials), feedback was shown for 1,500 ms after their response, or after the response window had elapsed.

The task consisted of two blocks of 56 trials in a pseudorandomised order, but fixed across participants. The rule for this pseudorandomisation was based on perspective switches. If the perspective to be taken during any given trial was different from the previous trial, that was regarded as a perspective switch. This switch happened on 50% of the trials, but never more than twice in a row. The order of the two blocks was counterbalanced. Per block there were 48 test trials, and eight filler trials. In half of the test trials, the number of dots on the walls matched the digit instruction (Yes), whereas in the other half they did not match (No). Additionally, in half of the test trials, the number of dots visible to the self- and other-perspective was the same (Consistent), whereas in the other half it was not (Inconsistent).

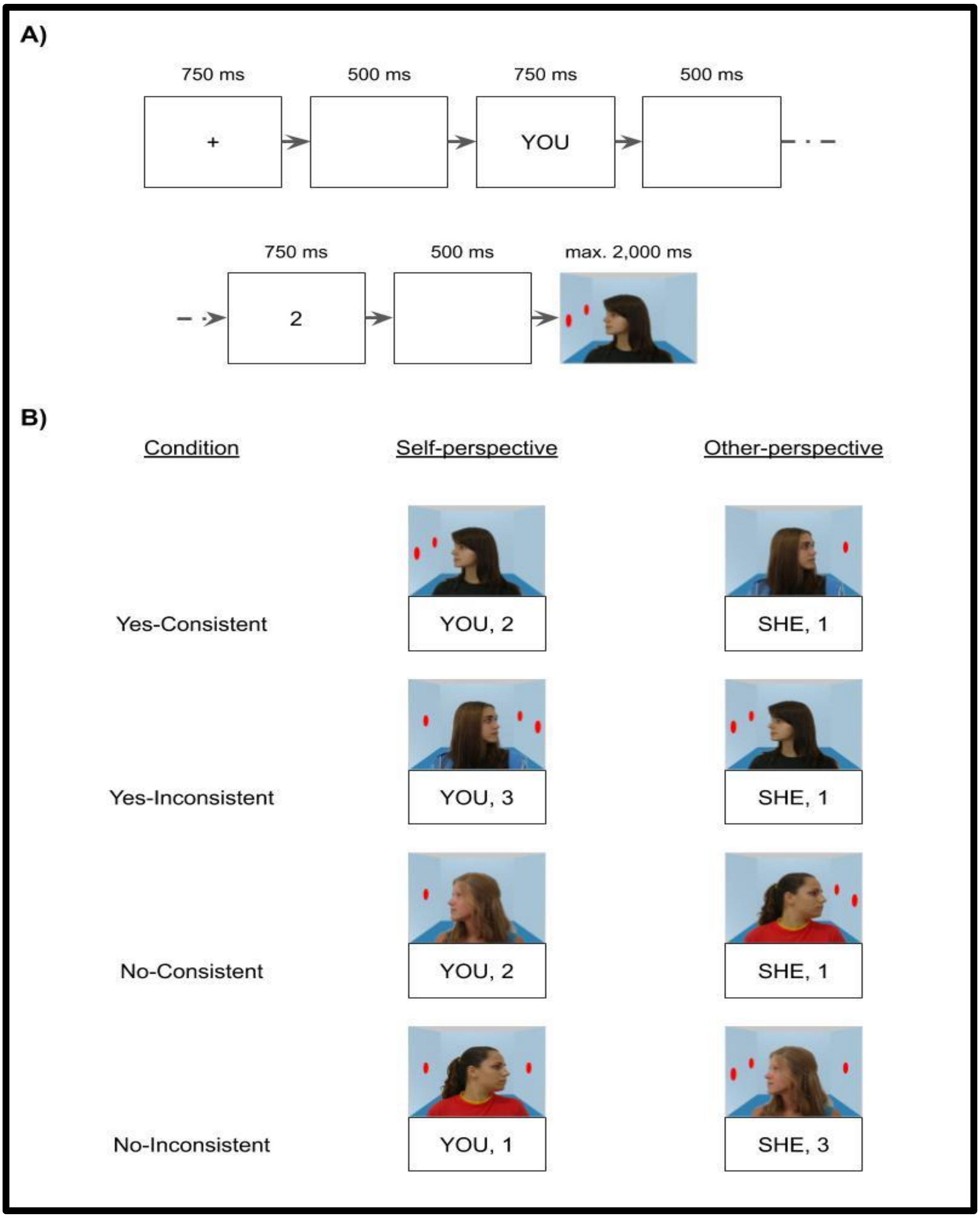


Figure 3.4 Dots perspective taking task design. A) Procedure of a trial (in the Self-Yes-Consistent condition), showing the duration per slide. B) Examples of test trials used in the Dots task conditions (based on the figure in Samson et al., 2010). The faces in the Yes conditions had been associated with a 90% win or loss reward value in the preceding evaluative conditioning task. These Yes conditions were used for the analysis.

The person shown in the middle of the screen was one of four identities. For each of the four identities, there were 24 trials. As such, for each identity there were six Self-Consistent, six Self-Inconsistent, six Other-Consistent, and six Other-Inconsistent trials. For each of these four conditions, the person faced the left/right wall three times. This variation meant that the person faced the left wall two times in one block, but one time in the other block.

For each identity there were also four filler trials. On filler trials, the digit instruction was always higher than 0, whereas no dots appeared in the room. This trial type was included so that there were also sometimes no dots visible to the self-perspective (note that on Inconsistent test trials, the self-perspective often corresponded to at least one dot, whereas none were visible to the other-perspective). Filler trials were also spread equally across identities, and consistency and perspective conditions. The summated number of dots was equal (50) for Consistent versus Inconsistent, Self versus Other, Yes versus No, and equal (25) across the four identities.

As Samson and colleagues (2010) explained, there is an imbalance in the conditions of the Dots task. Specifically, when the required response was Yes, the digit instruction always corresponded to at least one perspective. However, when the required response was No, this correspondence depended on whether the trial was in the Consistent or Inconsistent condition. In Inconsistent trials, one of the two perspectives corresponded to the instruction. On the other hand, none of the perspectives corresponded to the digit instruction in Consistent trials. This imbalance made No-Consistent trials easier to process than the other trial types. Because of this imbalance between Yes and No trials, only test trials requiring a

Yes response (half of the trials) were kept for analysis. This omission of No trials matched the approach taken by Samson and colleagues (2010).

This omission required the remaining trials to be configured in such a way that the effect of the reward manipulation could be investigated most effectively. Therefore, the two faces that were associated with a 90% win/loss chance (i.e. Rewarding/Non-Rewarding) in the evaluative conditioning task were always shown when the required response was Yes, whereas the 60% win/loss faces were always shown when the required response was No. Incidentally, no participant reported to be aware of this contingency. In addition, because of the omission it was decided that four female, rather than two male and two female, faces be used. This choice was made to prevent that gender was linked to the Yes versus No condition, or otherwise to rewarding versus non-rewarding. Note that previous studies discussed in this thesis always used the same gender for the 90% win and loss conditions.

3.5.1.3.3 Trait measure

The Autism Spectrum Quotient (AQ; Baron-Cohen et al., 2001b) was used to quantify autistic traits. After completion of this questionnaire, the participant was debriefed.

3.5.1.4 Data analysis

As the playing card task was only used for evaluative conditioning, data generated by that task was not further analysed.

For the Dots task, reaction time (RT) was taken as the primary measure of interest. As mentioned in Section 3.1 (Introduction), the Dots task can be used to quantify to what extent the irrelevant perspective is computed. This quantification can be done for both trials in which the self-perspective and the other-perspective needs to be taken. Note that, on any given trial, the information (number of dots) visible to the self- and other-perspective can be consistent or inconsistent. As Samson and colleagues (2010) showed, responses are slowed down when the information is inconsistent. This increase in RT reflects the extent of automatic computation of the irrelevant perspective. The RT difference between Inconsistent and Consistent trials for the self-perspective is the altercentric bias, i.e. interference from the information belonging to the *other's* perspective. On the other hand, this RT difference for the other-perspective is the egocentric bias, i.e. interference from the information belonging to the *own* perspective.

Here, it was investigated whether altercentric and egocentric interference were also present in the current sample. In addition, it was examined whether there was reward-driven modulation of these two types of interference. Interference was quantified by subtracting the RT for Consistent trials from the RT for Inconsistent trials. The use of this absolute difference score to quantify interference was taken from other studies employing the Dots task (Samson et al., 2010; Bukowski & Samson, 2017; Marshall et al., 2018). This calculation was performed for the Self Perspective condition to produce altercentric bias and for the Other Perspective condition to produce egocentric bias. It was then investigated whether reward influenced either of these types of interference. Lastly, it was investigated whether this possible reward-driven modulation of altercentric and/or egocentric interference was related to autistic traits. To this end, another composite score was calculated by subtracting egocentric/altercentric bias for the Non-rewarding condition from the egocentric/altercentric bias for the Rewarding condition.

For correlations, observations were excluded from each individual analysis separately if they were identified as bivariate outliers (Cook's distance > $4/N_{participants}$; Bollen & Jackman, 1990)

All p-values reported with statistical tests below are two-tailed.

3.5.2 Results

3.5.2.1 Data pre-processing

Pre-processing steps and the analysis were performed with R (R Core Team, 2018). First, practice trials and filler trials from the main phase were excluded. Second, only trials for which the required response was Yes were kept. Note that this exclusion step meant that only trials with faces conditioned with a 90% win chance or 90% loss chance ('Pos' and 'Neg' from here onwards) were kept. At this point, the full data set consisted of 4,080 trials. Third, incorrect responses were excluded (7.1%). Fourth, trials were removed if the RT was more than 3 SD away from the participant mean (1.3%) (cf. Samson et al., 2010). Note that Samson and colleagues (2010) performed their outlier removal per participant and condition. However, due to the smaller number of trials per condition in the current task because of the reward manipulation, this outlier removal was done across all conditions per participant. Lastly, data of a participant was excluded in full if they did not have at least four trials in any of the eight conditions of interest (Self - Consistent - Pos, Self - Consistent - Neg, Self - Inconsistent - Pos, Self - Inconsistent - Pos, Other - Consistent - Pos, Other - Consistent - Neg, Other -

Inconsistent - Pos, Other - Inconsistent - Neg). This procedure resulted in the exclusion of 10 participants. As a result, the final sample consisted of 75 participants (63 female), whose age ranged between 18 and 49 (M = 20.81, SD = 5.03) and whose AQ ranged between 6 and 36 (M = 16.71, SD = 6.91). The number of trials across these participants was 3,356 (mean N = 44.75, SD = 1.74).

3.5.2.2 Perspective taking and the reward effect

A 2 x 2 x 2 repeated-measures ANOVA was performed, with Perspective (Self, Other), Consistency (Consistent, Inconsistent) and Reward (Rewarding, Non-rewarding) as within-subjects factors and RT as the dependent variable. The analysis revealed a significant main effect of Perspective, $F(1,74) = 16.88$, $p < .001$, $\eta_p^2 = .186$, where RT was higher for Other (M = 736.59, SE = 18.47) than for Self (M = 702.42, SE = 18.66). In addition, there was a main effect of Consistency, $F(1,74) = 62.78$, $p < .001$, $\eta_p^2 = .459$, with RT being higher for Inconsistent (M = 747.46, SE = 18.76) than for Consistent trials (M = 691.55, SE = 18.10). However, there was no significant main effect of Reward, $F(1,74) = .03$, $p = .862$, $\eta_p^2 < .000$ (see Figure 3.5).

Two of the two-way interactions were significant. Specifically, the Perspective x Consistency interaction was significant, $F(1,74) = 11.16$, $p = .001$, $\eta_p^2 = .131$. Paired t-tests revealed that the effect of Consistency was significant when the perspective was Self, $t(74) = 3.18$, $p = .002$, $d = .368$, with an advantage of 31.18 ms for the Consistent condition. In addition, the effect of Consistency was significant when the perspective was Other, $t(74) = 7.66$, $p < .001$, $d = .885$, with a numerically larger advantage of 80.64 ms for the Consistent condition.

Alternatively, the effect of Perspective was significant in the Inconsistent condition, $t(74) = 5.12$, $p < .001$, $d = .519$, with an advantage of 58.90 ms when the perspective was Self. In contrast, the effect of Perspective was not significant in the Consistent condition, $t(74) = .81$, $p = .419$, $d = .094$, with a numerical advantage of 9.44 ms when the perspective was Self.

Additionally, the Perspective x Reward interaction was significant, $F(1,74) = 6.09$, $p = .016$, $\eta_p^2 = .076$ (see Figure 3.6). Paired t-tests revealed that, when the perspective was Self, the effect of Reward approached significance, $t(74) = 1.74$, $p = .086$, $d = .201$, with an advantage of 14.78 ms when the identity of the face was Non-rewarding. In contrast, the effect of Reward was not significant when the perspective was Other, $t(74) = 1.54$, $p = .127$, $d = .178$, with a numerical advantage of 17.15 ms when the identity of the face was Rewarding. Alternatively, in the Rewarding condition, the effect of Perspective was not significant, $t(74) = 1.50$, $p = .137$, $d = .174$, whereas it was significant in the Non-rewarding condition, $t(74) = 5.43$, $p < .001$, $d = .627$, with a numerical advantage of 50.13 ms when the Perspective was Self.

However, the Consistency x Reward interaction was not significant, $F(1,74) = .17$, $p = .681$, $\eta_p^2 = .002$.

Lastly, the three-way interaction of Perspective x Consistency x Reward was not significant, $F(1,74) = 1.02$, $p = .317$, $\eta_p^2 = .014$.

In addition, two paired-samples t-tests were performed to specifically test whether reward modulated altercentric and/or egocentric interference. To this end, the composite bias was taken for both types of interference. The analysis revealed that the effect of Reward on altercentric bias was not significant, $t(74)$

= .39, *p* = .690, *d* = .044. Likewise, the effect of Reward on egocentric bias was not significant, *t*(74) = .91, *p* = .366, *d* = .105.

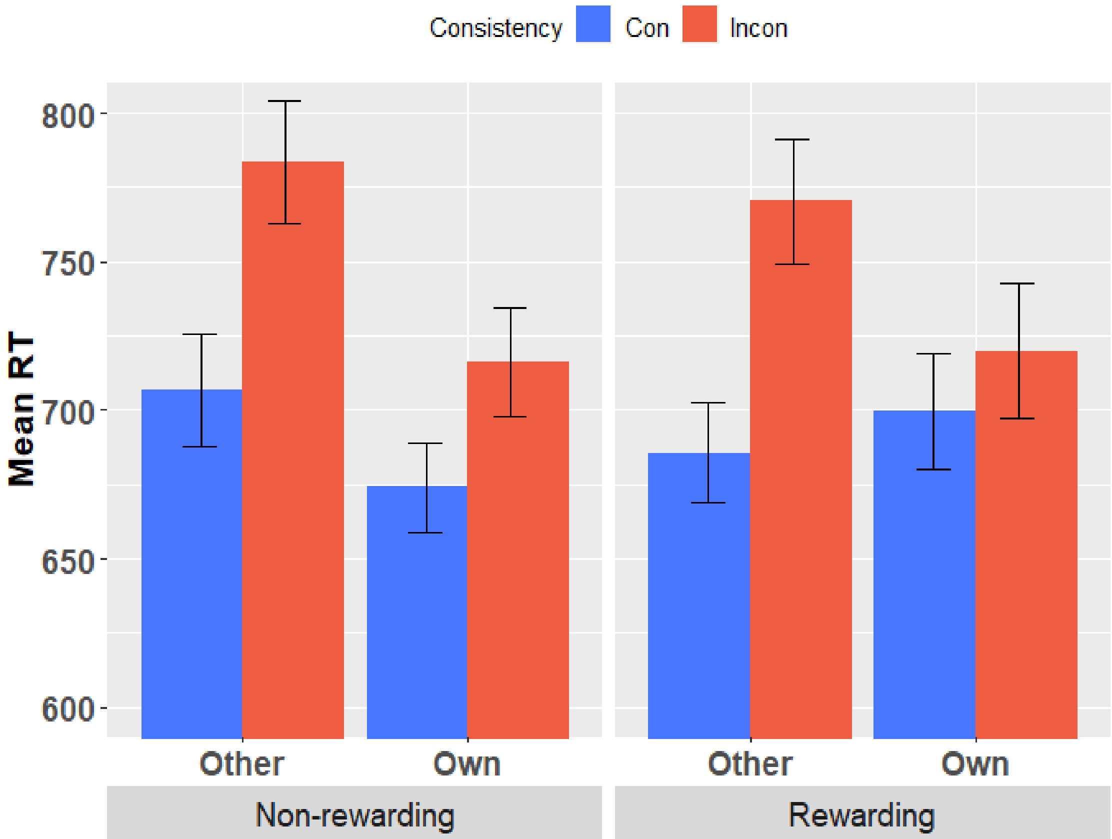


Figure 3.5 Mean reaction time in the Dots task conditions (Reward, Consistency and Perspective). Error bars indicate 95% within-subjects confidence intervals.

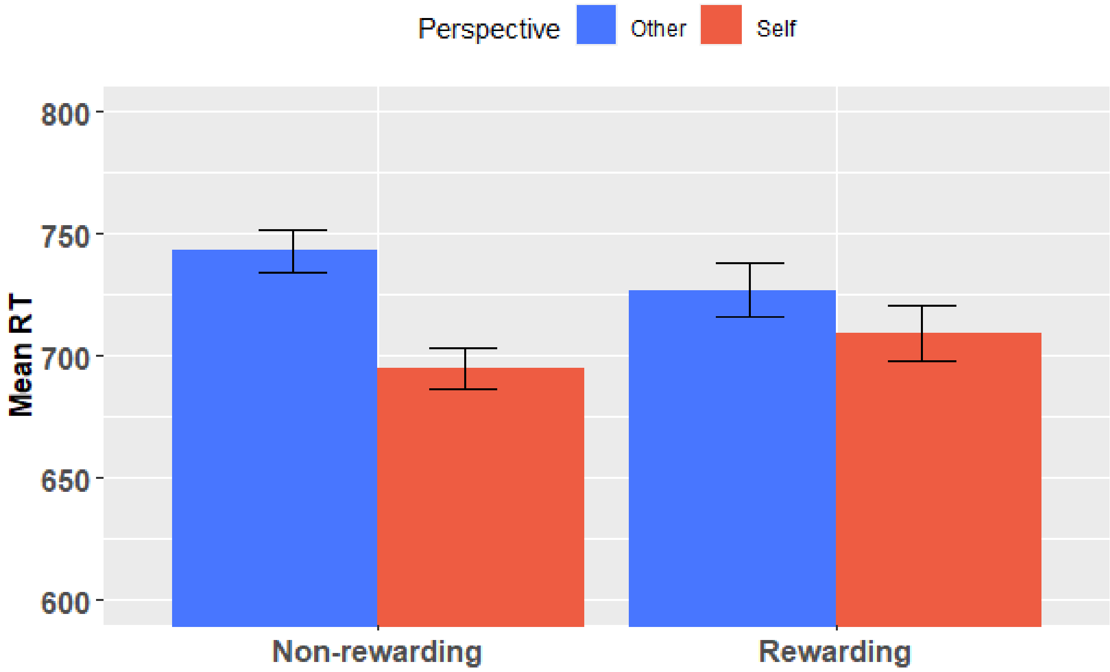


Figure 3.6 Interaction between Perspective and Reward on mean reaction time. Error bars indicate 95% within-subjects confidence intervals.

3.5.2.3 Individual differences in the reward effect

Although reward was found not to modulate altercentric or egocentric bias, an investigation into possible individual differences in the reward effect was warranted given the research questions.

A Pearson correlation was performed, correlating autistic traits with the composite score reflecting the reward of effect on altercentric bias. The result was non-significant, $r(70) = -.01$, $p = .941$.

Another Pearson correlation was performed, correlating autistic traits with the composite score for egocentric bias. There was a significant positive relationship, $r(72) = .24$, $p = .040$ (see Figure 3.7). Note that the bias score reflected how much more egocentric interference there was for the Rewarding versus Non-rewarding condition. Therefore, the correlation showed that this (non-significant) difference was lower for participants lower in autistic traits.

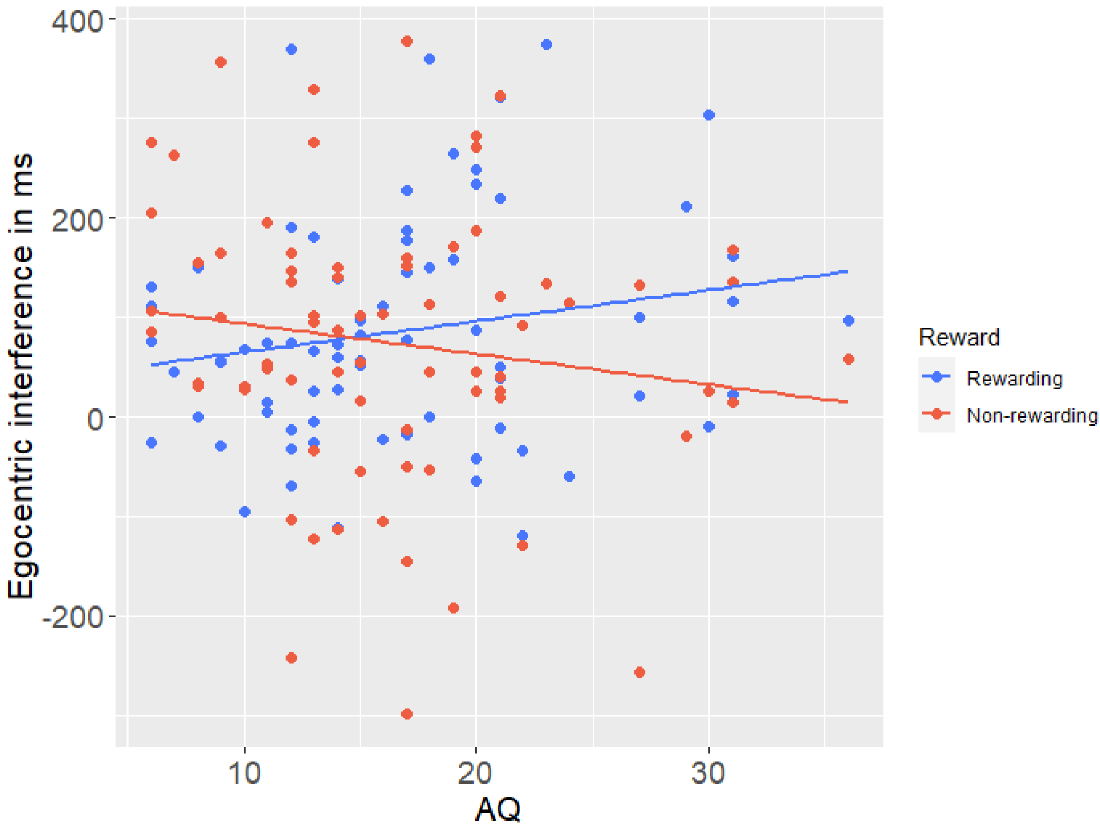


Figure 3.7 The relationship between autistic traits and the effect of reward on egocentric interference. To illustrate the effect of reward, the relationship between AQ and egocentric interference has been plotted separately for the Rewarding and Non-rewarding condition. Egocentric interference was calculated by subtracting the mean RT for Consistent trials from the mean RT for Inconsistent trials.

In addition, as it was found that reward modulated the difference between the Self and Other condition, it was also examined whether this effect of reward was related to individual differences in autistic traits. For this purpose, another two-step composite score was calculated: first, RT for the Other condition was subtracted from the RT for the Self condition. Next, the composite score was calculated by subtracting this Self-Other difference score for the Non-rewarding condition from the score for the Rewarding condition. A correlation between this composite score and autistic traits revealed a non-significant relationship between autistic traits and the effect of Reward on Self-Other, $r(70) = .03$, $p = .786$.

3.5.2.4 Individual differences in perspective taking

As an additional exploration, it was investigated whether autistic traits correlated with individual differences in perspective taking, independent of reward (cf. the analysis in Section 3.4.2.5). This investigation was done separately for egocentric interference and altercentric interference, collapsed across reward conditions. The Pearson correlations revealed that there was no significant correlation between autistic traits and egocentric interference, $r(69) = .05$, $p = .686$, nor between autistic traits and altercentric interference, $r(71) = -.05$, $p = .697$.

3.5.3 Brief discussion

The current results generally replicated the findings by Samson et al. (2010). Participants were slowed down when the self-perspective and other-perspective mismatched in terms of the visible number of dots. In addition, this disadvantage was present both when they had to take their own perspective and the other’s perspective. This observation indicates that participants exhibited both altercentric interference and egocentric interference, the latter of which was numerically larger (cf. Samson et al., 2010). The current results also showed a main effect of perspective, with participants being slower when having to take the other's perspective. The results also replicated the interaction of perspective and consistency reported by Samson et al. (2010). Specifically, the current results showed that taking the other's relative to the own perspective on consistent trials did not slow the participant down; taking the other's relative to the own perspective on inconsistent trials did slow the participant down. This pattern suggests that taking the other's perspective is particularly difficult when the information available to one's own perspective does not match the other's.

The way performance was influenced by reward was more complex. Reward did not directly influence RT, rejecting the possibility of a simple avoidance/approach mechanism. However, there was an interaction between reward and perspective. Specifically, there was only an RT difference between the self- and other-perspective when the other person had been conditioned to be non-rewarding. This finding suggests that, when you like another person, you are not slowed down when taking that person's perspective compared to your own. However, when the other person is disliked, it is more effortful to take their perspective relative to one's own.

In contrast, there was no interaction between reward and consistency. A conflict between the perspectives did not slow participants down more when the other person was rewarding or non-rewarding. In addition, neither altercentric nor egocentric interference was affected by the reward value of the other person. This finding, together with the presence of an interaction between reward and perspective, raises the following possibility: it is easier to *take* the other's perspective when that person is liked versus disliked. However, when one is taking the other's perspective, the interference of the self-perspective does not depend on whether the other is liked or not. Likewise, the interference of the other-perspective when taking one's own perspective does not depend on whether the other is liked or not.

Lastly, an examination of individual differences in the reward-driven modulation of VPT yielded mixed results. The findings showed that those low in autistic traits were *less* distracted by their own perspective (i.e. egocentric interference) when taking the perspective of a rewarding person. However, this finding should be approached with caution, as there was no significant effect of reward on egocentric (or altercentric interference). In contrast, there was no relationship between autistic traits and the (significant) reward-driven modulation of the effect of perspective.

3.6 Discussion

The aim of the two studies covered in this chapter was to investigate whether there is reward-driven modulation of perspective taking. In order to do this, two VPT paradigms, the Dots task (Samson et al., 2010) and the Director task (Keysar et al., 2000), were employed. The Dots task supplied two VPT measures: egocentric and altercentric interference. The first is a more traditional measure of VPT, as it reflects the degree to which one's own perspective interferes with taking someone else's. In contrast, the second reflects the degree to which the computation of the other's perspective interferes with judgments related to the self-perspective. In this way, the Dots task on its own can be seen as a conceptual counterpart to the Sandbox-mouse-tracking tandem discussed in Chapter 2. This is because the Dots task sometimes requires taking the other's perspective (cf. Sandbox task), and sometimes taking one's own (cf. mouse-tracking task). However, by using the Director task in addition to the Dots task, reward-driven manipulation could also be investigated in a context in which there is direct social interaction. Moreover, the Director task supplied a measure beyond reaction time and accuracy by quantifying egocentric interference in terms of eye gaze behaviour.

Across these two tasks it was shown that neither egocentric nor altercentric interference is influenced by reward. For egocentric interference this observation means the following: when you take someone else’s perspective to make a VPT judgment, it takes effort to suppress the own perspective, which may have access to privileged information. However, the effort that this suppression takes does not depend on whether the other person is liked or not. Similarly, when judging a visual scene from one’s own perspective, the visual perspective of the other person is automatically computed and may therefore interfere when making a VPT judgment from the own perspective. However, this interference from the other-perspective does not depend on whether the other person is liked or not.

For the Director task, it could be argued that the presence of the experimenter interfered with a possible effect of reward. As was explained in Section 3.4.1.3.2 (Director task procedure), the experimenter sat opposite the participant, albeit with two monitors placed between them. This setup could have resulted in the participant’s conflating the director with the experimenter, precluding conclusions about the reward value associated with the director. Nonetheless, the Dots task did not have this issue, yet also showed no effect of reward.

In contrast, it is possible that the effect of reward manifests in a different way. Specifically, it was observed for the Dots task that reward influenced the mere taking of someone else's perspective relative to one's own. In other words, it look longer to take the other-perspective relative to the self-perspective when the other person was conditioned as non-rewarding.

It has been suggested that VPT-1, as measured through the Dots task, is in fact a two-step process (Qureshi et al., 2010; Samson et al., 2010). First, there is automatic and rapid computation of both the self- and other-perspective; subsequently, one of the two perspectives is intentionally selected for the behavioural judgment (Schwarzkopf, Schilbach, Vogeley, & Timmermans, 2014; Qureshi et al., 2010). In addition, it has been suggested that reward-related neural activation can inform executive systems that guide decisions (Chevallier et al., 2012; Klein et al., 2009). It is therefore possible that a low reward value associated with an other can interfere with the intentional selection of their perspective.

Notably, there was no relationship between individual differences in autistic traits and the observed reward effect, nor between autistic traits and VPT. As Pearson and colleagues mentioned (2013), it is unclear whether there is a VPT-1 impairment in autism. The current results show that autistic traits do not influence to what extent a positive reward value facilitates perspective selection. Rather, it appears that people both high and low in autistic traits are able to learn the reward value associated with the faces shown in the Dots task, and will be more efficient in their perspective selection when this other person is liked.

3.7 Conclusion

As a parallel to the ToM-related tasks in Chapter 2, the current chapter discussed the possible reward-driven modulation of VPT. It was found that reward does not modulate the difficulty to suppress the irrelevant perspective (altercentric/egocentric interference). However, it does modulate the efficiency with which the other-perspective relative to the self-perspective can be selected. This pattern suggests that the reward-driven modulation of VPT can appear only in those tasks that tap into perspective selection. Based on the findings in Chapter 2 and 3, it can be prematurely concluded that reward influences some components of cognitive empathy, although it varies across tasks whether it is observable. In addition, the studies discussed so far have all employed tasks that operationalise cognitive empathy-related processes in a fragmentary and simplified manner. However, in real life, cognitive empathy refers to the ability to understand and interpret others' emotions, thoughts and intentions, given the context. Whereas perspective taking and ToM are important component processes of cognitive empathy, the endpoint, empathic accuracy, is more holistic. Therefore, it would be interesting to see whether and how empathic accuracy is influenced by characteristics of the target person. For this reason, the following chapter will investigate whether there is also reward-driven manipulation of empathic accuracy.

4 Reward and empathic accuracy

The previous two chapters investigated the effect of reward on task measures of Theory of Mind and perspective taking. Whereas these task measures can inform us about the component processes underlying cognitive empathy, they are but limited models when compared to the inferences about others' thoughts and feelings that are made during social interaction. For this reason, the current chapter turned to a more ecologically valid paradigm to investigate the effect of reward on the accuracy of such inferences. A previous study reported that empathic accuracy is higher towards rewarding others (friends) than non-rewarding others (strangers). In addition, empathic accuracy has been shown to be decreased in people high in autistic traits. However, it has not yet been investigated whether the effect of reward can be modulated by autistic traits. This question was the focus of the current study. A video rating task was employed in which participants watched a friend or stranger talk about emotional life events, whilst continuously rating that person's experienced valence and arousal. By comparing self-ratings to other-ratings, a measure of empathic accuracy was established. Reward value was indexed by having participants rate the friend/stranger on social closeness. It was found that reward was positively related to empathic accuracy, theoretically replicating previous findings. However, autistic traits did not modulate this relationship. It was concluded that people, regardless of their autistic traits, are better at using verbal and nonverbal cues to make inferences about the thoughts and feelings of friends versus strangers.

4.1 Introduction

Empathic accuracy (EA) refers to the ability to correctly infer others' thoughts and feelings (Ickes et al., 1990). Whereas processes such as Theory of Mind (see Chapter 2) and perspective taking (see Chapter 3) can be seen as basic processes underlying cognitive empathy (i.e. they are required in order to gain an understanding of someone else's internal state), EA applies to the eventual inferences that rely on these processes. According to Davis (1994), inferences about someone's internal state can arise through both cognitive and affective empathy processes. Affective empathy processes can aid understanding by making the perceiver resonate with the other's emotions. Yet such emotion contagion is not required for EA per se, as mental state inferences can be attained through logical deduction and the processing of available cues without feeling the emotion. Blanke and colleagues (2015) state that a variety of cues is available in social interaction, such as facial and bodily expressions, the context of the situation, prosody, and the communicated messages. EA, therefore, represents the successful integration of these cues with knowledge, reasoning and memory for making mental state inferences (Ickes, 1997; Blanke & Riediger, 2019).

In addition to verbal and nonverbal cues, another factor that can aid in the understanding of others is how much prior knowledge one possesses regarding the other person. For instance, Stinson and Ickes (1992) found that friends are better at inferring others' thoughts and feelings than strangers. In their study, pairs of friends or strangers engaged in a spontaneous interaction. Unbeknownst to them, this interaction was videotaped. After the interaction phase, participants watched the videotapes and were asked to indicate at a number of chosen time points what their own thoughts and feelings were. In addition, they made inferences about the other person's thoughts and feelings. The accuracy of these inferences was shown to be higher for friends than strangers. It was argued that this pattern was not because social interactions of friends are more cue-laden, but because they are better at reading these cues, probably because these cues evoke knowledge about their friend. In other words, friends understand each other better than strangers because they know each other more.

The method of matching self-descriptions and other-inferences employed by the Stinson and Ickes (1992) paradigm has subsequently been adapted by a plethora of studies to investigate factors influencing EA. Following the direction Stinson and Ickes (1992) took, several studies have examined the effect of relationship quality on EA. For example, Sternglanz & DePaulo (2004) showed that closer friends compared to less close friends are better at reading clear emotions, whereas they perform worse when these emotions are concealed. Similarly, there is some evidence that EA gets weaker in intimate relationships the longer they are, because there is less cognitive focus on what the other is thinking or feeling (Thomas, Fletcher, & Lange, 1997). In other words, EA is more important in relationships that are still in the consolidation phase (Ickes & Simpson, 2001; Sened et al., 2017).

Other studies have focussed more on intrapersonal rather than relationship factors. For example, it has been found that EA is stronger for younger versus older adults (Blanke et al., 2015; Blanke & Riediger, 2019); for people who understand their own emotions better (Erbas, Sels, Ceulemans, & Kuppens, 2016); for those low in attachment avoidance (Izhaki-Costi & Schul, 2011); for those whose emotion matches that of the EA target (Devlin, Zaki, Ong, & Gruber, 2014). In addition, it has been found that the expressivity of the other person increases EA (Snodgrass, Hecht, & Ploutz-Snyder, 1998; Zaki et al., 2008; but see Grant, Fetterman, Weyhaupt, Kim, & Tullett (2018) for evidence against this). A common theme in these studies is that people exhibit higher EA when they focus more on the other person and subsequently integrate the cues available to them. Indeed, it has been shown that mental state attributions are subserved by the ability to pay attention to social cues and integrate their information (Zaki, Weber, Bolger, & Ochsner, 2009).

Whereas the studies above generally utilised conventional samples, EA has also been investigated in atypical samples. Importantly, it has been shown that individuals with an autism spectrum condition, or typically developed individuals high in autistic traits, are less accurate at inferring others' thoughts and feelings (Demurie et al., 2011; Ponnet et al., 2005; Ponnet et al., 2008; aan het Rot & Hogenelst, 2014). It has been suggested that this reduced accuracy is because they find it more difficult to integrate cues (Ponnet et al., 2008). Indeed, one study showed that an administered dose of oxytocin can improve EA for those high in autistic traits, by improving attunement to social cues (Bartz et al., 2010). In short, these studies suggest that being impaired in cue-reading can result in lower EA.

Putting the above together, it can be concluded that 1) being close to someone results in an increased ability to infer their thoughts and feelings based on verbal and nonverbal cues and 2) processing of these cues is impaired in autism. Taking closeness as a proxy metric of reward value, the question arises whether autistic traits modulate this effect of closeness on EA.

4.2 Aims

The aims of the study reported in this chapter were twofold: to investigate 1) whether the effect of closeness on EA (Stinson & Ickes, 1992) could be replicated in a continuous rating paradigm and 2) whether individual differences in this effect were moderated by autistic traits. The paradigm utilised for this purpose was the video rating task developed by Zaki and colleagues (2008) and further used by Martin-Key and colleagues (2017). The use of this task was motivated by the merits of task measures with high temporal granularity (cf. the high spatial granularity of the Sandbox task and mouse-tracking task used in Chapter 2). In the video rating task, the perceiver watches a video recorded by someone else (the 'target') and provides real-time ratings on a 1-to-9 scale about how the target is feeling (see Section 4.4.3 Procedure, for more information). These ratings can then be compared over time to self-ratings that were given by the target themselves. An advantage of this approach is that an entire narrated video can be rated, as opposed to points in time chosen by the experimenters (cf. Stinson & Ickes, 1992). In addition, judgments rely on the perceiver's ability to attribute mental states, without the need to verbalise these states. In short, the video rating task has higher ecological validity compared to tasks employed in previous chapters, as it involves making continuous inferences about a real person.

Another method through which more granularity was achieved was the use of a continuous scale to rate social closeness (cf. Jones & Rachlin, 2006; O'Connell, Christakou, Haffey, & Chakrabarti, 2013). Although participants were recruited as friends or strangers (cf. Stinson & Ickes, 1992), this scale allowed to investigate the possible effect of closeness on a scale surpassing this binary distribution. Note that this combined the ecological validity of contrasting friends against strangers with a quantifiable measure of reward value.

Additionally, the investigation of EA was done for both valence and arousal. A common method of classifying emotions is the circumplex model (Russell, 1980), which uses valence and arousal as its two dimensions. The model holds that emotions range between positive and negative on the dimension of valence and between calm/passive and excited/active on the dimension of arousal. Previous studies investigating EA have generally only employed valence ratings (Zaki et al., 2008; Stinson & Ickes, 1992) or ratings about the intensity of basic emotions (Martin-Key et al., 2017). In contrast, one study compared self- and other-ratings for both valence and arousal (Erbas et al., 2016). To maximise the use of the video rating task, both valence and arousal ratings were included in the current version of the video rating task. As an exploratory analysis, it was investigated whether EA was higher for arousal or valence.

Lastly, this experiment was part of a study in which friends and strangers were recruited to participate in a real-time interaction (not covered in this thesis). For this reason, the evaluative conditioning paradigm used in the previous experimental chapters was not employed here. Instead, the use of the social closeness procedure allowed to quantify a proxy of reward value in a minute fashion, reflecting subjective real-life relationship quality.

4.3 Predictions

As the general consensus in the literature is that social closeness is related to higher EA, it was predicted that this finding would be replicated in the current study, for both valence and arousal. In addition, it was predicted that this relationship would be weaker for those higher in autistic traits, as people with an autism spectrum condition demonstrate impaired cue integration. Applying this rationale to the employed paradigm, these predictions were made:

1) Social closeness will be positively related to how accurately perceivers judge the valence and arousal experienced by the target person.
2) There will be a significant negative interaction between the effect of social closeness and autistic traits on EA for valence and arousal, as autistic traits negatively moderate this relationship.

4.4 Methods

4.4.1 Participants

Forty-two participants took part in the EA study (*N* females = 32, mean age = 20.76, *SD* = 2.61, min = 18, max = 35). This sample size was chosen based on previous studies using the video rating task (Zaki et al., 2008; Martin-Key et al., 2017). Two female participants who completed session 1 (see section 4.4.3 Procedure) did not complete session 2 and were therefore subsequently excluded. The majority of the participants were recruited through the University of Reading research panel. These participants were all undergraduate students. The remainder were recruited through word-of-mouth by the experimenters and were mostly undergraduate students. When participants were recruited, they were encouraged to ask a friend to participate. Upon completion of the study, participants received course credit, or £10. Ethical approval was received from the School of Psychology Ethics Committee and all participants provided informed consent prior to their participation.

4.4.2 Stimuli

As all stimuli were provided by the participants during the experiment, stimulus creation was intertwined with the procedure and is therefore explained in Section 4.4.3.

4.4.3 Procedure

The procedure was adapted from the Zaki et al. (2008) and Martin-Key et al. (2017) studies. Because of the pairing procedure (see Section 4.4.3.1.3 Post-session) that was employed, the study was split up into two sessions. The EA experiment reported in this chapter was part of a two-experiment study. For reasons of scope, the other experiment is not further discussed in this thesis. However, it will be mentioned below when during the procedure participants performed tasks pertaining to the other experiment. Before each task, instructions were shown on-screen and read out by the experimenter. Participants were debriefed upon completion of session 2.

4.4.3.1 Session 1

4.4.3.1.1 Video recording and self-ratings

Before the experiment started, the participant was first informed about the study. They were told that the aim of the study was to look at how people describe emotional events they have themselves experienced or that others have experienced. Additionally, they were informed which tasks they would perform during the two separate sessions.

The experimenter then commenced the video procedure. During recruitment, participants had been asked to think about two negative and two positive life experiences. The experiences could be about anything, but the participants were asked to choose ones that 1) they had not often told others about and 2) they would be comfortable sharing on camera. They were informed that their videos would only be viewed by themselves, one other participant, and the experimenters. Lastly, they were told that if they did not feel comfortable discussing an event, they could choose a different one.

In order to familiarise the participant with the rating task they would be using to rate videos, they first performed a practice trial. This task was adapted from a script kindly provided by Jamil Zaki. The task (see Figure 4.1) was run with MATLAB Release 2016b (The Mathworks, Inc.) using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007). It was shown on a 17-inch monitor, running Windows 7. The participant was first made acquainted with the 9-point valence rating scale. This scale ranged from 1 (very negative) to 9 (very positive) and started at a random, highlighted position. Upon pressing the spacebar, a 118-seconds-long example video started (532 x 315 pixels; a recording of the experimenter talking about a positive life experience). Participants wore headphones while they were watching the video. The participant was instructed to change the rating with the arrow keys on the keyboard whenever they perceived a change in emotion. The rating scale was displayed below the video. It was emphasised that they had to rate how the person felt *while describing the life experience*, rather than during the actual experience. The highlighted rating on-screen reacted in real time to the participant's key presses. To make sure that the participant adhered to the task instructions, the experimenter remained with them during the practice trial. The trial finished once the video had ended.

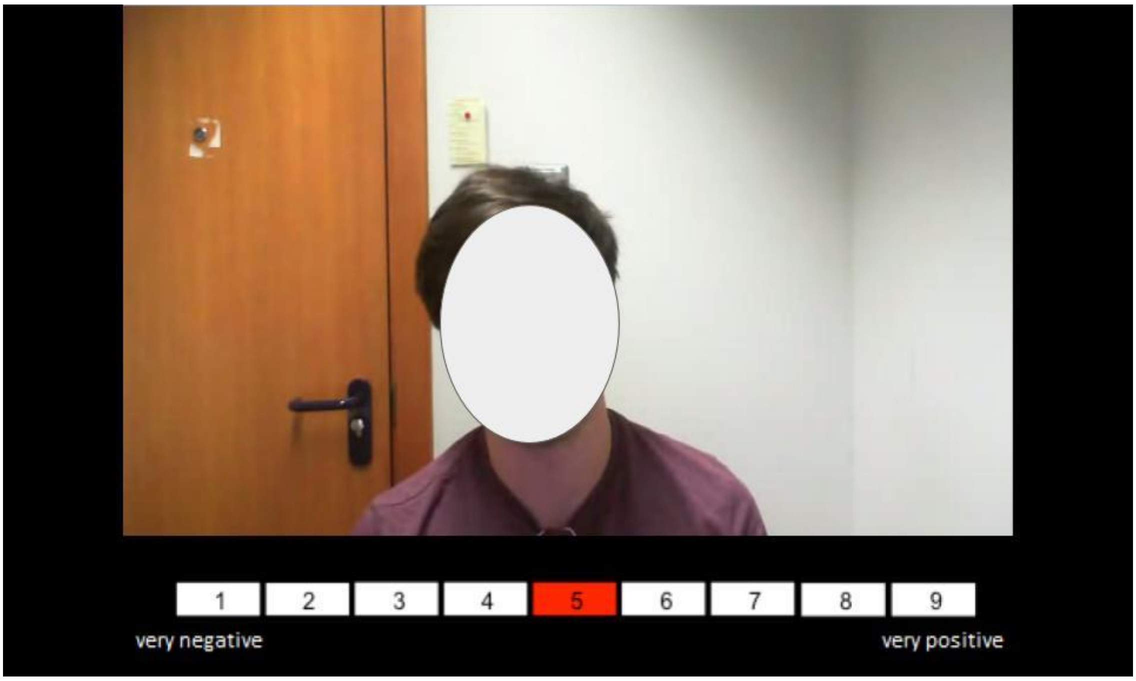


Figure 4.1 Example of a video rating task trial. The 'target' person (anonymised) is sitting in the room in which all participants recorded their videos. During the video, the rating scale was highlighted to indicate the current rating given by the participant watching the video, and changed in real time when the participant changed their rating. The trial shown here is a Valence trial. On Arousal trials, the scale ranged from 'very calm' to 'very excited'.

After the practice trial, the participant commenced the recording phase, in which they were to record videos of themselves talking about an emotional life experience. The recording phase went as follows: for each video, the participant was first asked to summarise the event on paper and to indicate on a scale of 1 to 9 how negative/positive and how calm/excited they had felt during the experience. They then discussed the experience with the experimenter. If it appeared to the experimenter that the description would not lead to a recording of at least one minute, the participant was asked to elaborate on the experience, if they were comfortable doing so. Next, they recorded a video of themselves talking about the experience with Windows LifeCam Studio, whilst the experimenter momentarily left the room. The participant decided themselves when to start and to stop recording, but were encouraged to record in between one and two minutes of footage (cf. a range of 54-158 seconds in the sample used by Martin-Key et al., 2017). After recording, they indicated on a scale of 1 to 9 how negative/positive and how calm/excited they had felt while recording the video. In addition, they indicated which primary emotion (anger, disgust, fear, happiness, sadness, surprise) was the most representative label for their overall feeling while recording. The experimenter was then notified and the recording procedure proceeded to the next video recording once it had been established that the participant was not distressed. The order of the video recording phase was positive-negative-negative-positive. This order was used so that participants would become acquainted with the camera recording while talking about a positive event, and to end the session on a positive note. To reduce variability across participants, the same 'interviewer-experimenter' conducted the video recording phase.

Next, as the participant completed a personality trait questionnaire (see Section 4.4.3.1.2 Trait measure) and another task not reported here in an adjacent room, the interviewer-experimenter prepared the videos for the rating task. After the participant had filled out the questionnaire, they started the self-rating task. The task here was the same as for the practice trial, but now the participant rated their own videos in the order that they had been recorded. They first rated all videos on valence (very negative - very positive) and subsequently on arousal (very calm - very excited). Valence was rated first because previous EA studies have generally focussed on this particular affective dimension (Zaki et al., 2008; Stinson & Ickes, 1992). Results of the current experiment regarding valence could therefore be compared to these previous studies without the potential confound that participants had already seen the video. Upon completion the participant was told that they would be contacted for the second session.

4.4.3.1.2 Trait measure

The Autism Spectrum Quotient (AQ) was used here to quantify autistic traits (Baron-Cohen et al., 2001b).

4.4.3.1.3 Post-session

The experimenters subsequently paired up participants in same-sex pairs (cf. Stinson & Ickes, 1992). If participants had signed up through a friend, it was attempted to pair them up. Otherwise, they were paired up with a random other participant. This pairing was done in such a way as to make the number of *a priori* friend pairs and stranger pairs equal.

4.4.3.2 Session 2

For this session, participants came in together. Before the tasks started, it was made sure that they were introduced to each other. The next parts were performed in adjacent rooms.

4.4.3.2.1 Social closeness questionnaire

Although participants had been assigned to the friends or strangers condition prior to participation, it was possible that ‘strangers’ were in reality acquainted. To account for this possibility and to achieve greater granularity than a binary distribution, the participants started with a short questionnaire that asked them 1) whether they knew each other, 2) to indicate the social distance between them (0: ‘Virtually the same person as me’; 100: ‘A total stranger’) and 3) a number of questions to assess the extent of their relationship. The use of a social closeness rating procedure was based on a similar method used in previous studies (O’Connell et al., 2013; Jones & Rachlin, 2006). Note that a higher score on the scale indicates more social distance. This directionality is employed in the remainder of the Methods and Results section.

4.4.3.2.2 Other-ratings

After the social distance questionnaire, the participants performed the same video rating task as described in Section 4.4.3.1.1. Here, they rated the videos that the other participant had recorded of themselves in their respective session 1. Once again, ratings were given first for valence, and second for arousal. Although the monitors used to present the video rating task were of different sizes for the two participants, the physical stimulus sizes were equal (movie: 27 x 16 cm; rating scale: 26.5 x 2.2 cm). Upon completion of the video rating task, they proceeded to perform a task together (not covered in this thesis). Afterwards, they were debriefed about the study.

4.4.4 Data analysis

In the video rating task, whenever the participant changed their rating of the video (ratings between 1 and 9), this event was logged on the level of milliseconds. Hence, the rating could be mapped minutely over time. To establish a measure of EA for each participant, the following steps were performed for the Valence and Arousal condition separately. First, the duration of the video in seconds was rounded down to the nearest even number. Any ratings beyond this cut-off point were excluded. Second, the initial 4 seconds of the ratings were excluded. This exclusion was done because the rating scale started at a random value at the beginning of each video that may not have corresponded to the participant's perception of the displayed emotion. Third, the average rating per 2-second bin (cf. Martin-Key et al., 2017) was calculated.

This procedure was performed once for the self-ratings (when participants watched their own videos) and once for the other-ratings (when participants watched the videos of their pair member). As an example: if participant A watched a video that was recorded by participant B, the self-rating was taken from participant B, and the other-rating from participant A. Subsequently, the correlation between the self- and other-ratings was calculated across the bins for a single video. Lastly, to achieve a video-level measure of EA, the correlation coefficient was subjected to a Fisher's Z transformation (cf. Martin-Key et al., 2017). This transformation results in a normal distribution of r scores, which can then be used for statistical analyses. A single value representing EA was calculated per participant by taking the mean of their four EA scores. In addition, a single self-reported social closeness score was used per participant. Finally, to achieve an estimate of how difficult it was to rate a video, self-rating variability was calculated as the SD in the self-ratings per video, for arousal and valence separately. More variability in the ratings provided by the recording participant could make it more difficult for the other person to provide accurate other-ratings. The mean of the four variability scores was taken as the difficulty level for the videos rated by the second participant. By using this measure of self-rating variability, the main analyses could be controlled for video difficulty.

For correlations, observations were excluded from each individual analysis separately if they were identified as bivariate outliers (Cook's distance > $4/N_{participants}$; Bollen & Jackman, 1990).

All p-values reported with statistical tests below are two-tailed, except for tests directly investigating the relationship between social distance and EA. As a number of studies have shown that close others are better at inferring thoughts and feelings than strangers (e.g. Stinson & Ickes, 1992; Sternglanz & DePaulo, 2004; Thomas & Fletcher, 2003), these tests were one-tailed.

4.5 Results

4.5.1 Data pre-processing

First, trials were excluded from analysis if the participant had not changed their rating throughout the video. Note that, in Section 4.4.4 (Data analysis), it was mentioned that one EA score was calculated per participant, based on their scores on videos. However, because of this first exclusion step, seven Arousal trials and six Valence trials were excluded from the total sample. Therefore, the total number of Arousal trials analysed was 149 and the total number of Valence trials was 150. As a result of this exclusion step, one female participant did not have trials in the Valence condition and was therefore excluded from analysis. The resulting sample for analysis therefore consisted of 39 participants (*N* females = 29, mean age = 20.87, *SD* = 2.67, min = 18, max = 35), with AQ scores ranging between 4 and 38 (*M* = 16.72, *SD* = 6.49).

4.5.2 Preliminary analyses

As a manipulation check, it was examined whether positive and negative videos differed in their mean valence rating. This difference was analysed with a linear mixed-effects model using the 'lme4' (Bates, Maechler, Bolker, & Walker, 2015) and 'lmerTest' (Kuznetsova, Brockhoff, & Christensen, 2017) packages in R (R Core Team, 2018). Mean self-rated valence was entered as the response variable, with the valence condition of the video (positive/negative) as the fixed effect, and random intercepts and random slopes for participants. The results showed that valence condition affected the mean valence rating, $\chi^2(1) = 83.71$, $p < .001$, with positive relative to negative videos increasing the rating by 3.84 (*SE* = .23). This observation indicates that, as expected, participants self-rated positive videos more positively than negative videos.

In addition, it was tested whether participants' EA was significantly different for positive and negative videos. This comparison was done for the Valence and Arousal condition separately (see Table 4.1 for an overview of EA across conditions). Paired t-tests revealed that EA was not significantly different between positive and negative videos in the Arousal condition, $t(38) = 1.26$, $p = .216$, $d = .20$, nor in the Valence condition, $t(38) = 1.03$, $p = .308$, $d = .17$. For this reason, EA was collapsed across video valence (positive/negative) for the remainder of the analysis.

Lastly, it was tested whether EA was significantly different when making arousal versus valence judgments. A paired t-test revealed a significant difference, $t(38) = 3.89$, $p < .001$, $d = .623$, with EA being higher for valence ($M = .47$, $SE = .06$) than for arousal ($M = .24$, $SE = .05$). Relatedly, it was explored in a follow-up analysis whether self-ratings for arousal and for valence differed in how variable they were. A paired t-test revealed no difference in self-rating variability, $t(38) = .94$, $p = .355$, $d = .150$, between the arousal ($M = 1.25$, $SE = .07$) and valence self-ratings ($M = 1.19$, $SE = .06$).

Table 4.1 Descriptives for empathic accuracy in the video rating task. Mean value (and SD) for raw correlation and Fisher's Z coefficients across all trials.

		Valence condition		
Type of rating		Positive	Negative	Total
Valence	*N*	76	74	150
	r	.35 (.40)	.43 (33)	.39 (.37)
	Z	.45 (.55)	.53 (.50)	.49 (.52)
	+	58	68	126
Arousal	*N*	73	76	149
	r	.26 (.43)	.12 (.47)	.19 (.45)
	Z	.33 (.57)	.16 (.60)	.25 (.59)
	+	58	45	103

Note: *N*: number of videos; *r*: Pearson correlation coefficient; *Z*: Fisher's *r*-to-*Z*-transformed correlation coefficient; +: number of videos with a positively signed correlation between the self- and other-rating.

4.5.3 The effect of reward

First, it was tested whether social distance (on a scale of 0-100) was different between friends and strangers. An independent-samples t-test revealed a significant difference, $t(37) = 15.75$, $p < .001$, $d = 4.91$. Social distance was on average higher for strangers ($M = 95.29$, $SE = 1.70$) than for friends ($M = 24.17$, $SE = 4.47$).

4.5.3.1 Empathic accuracy for valence

The Pearson correlation revealed a significant relationship between social distance and EA for valence, $r(34) = -.32$, $p = .029$ (see Figure 4.2).

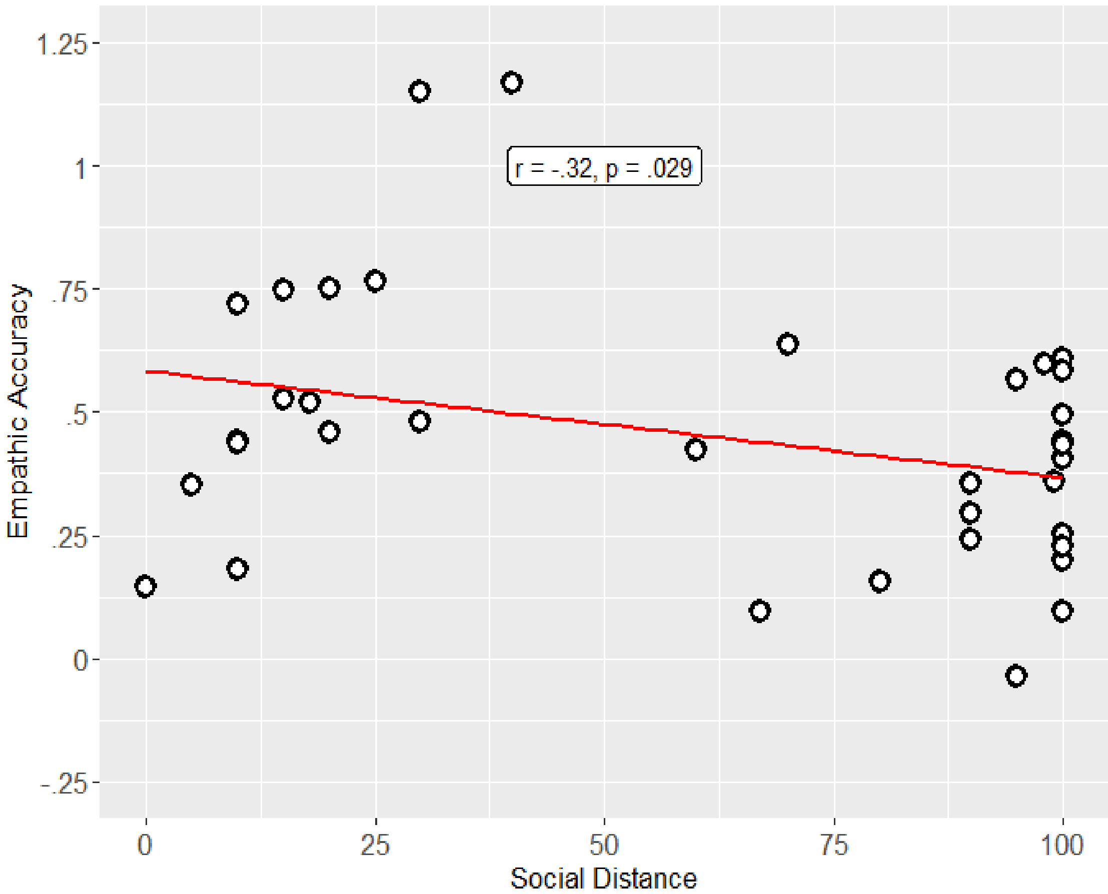


Figure 4.2 The relationship between social distance and empathic accuracy for valence. Reported *p*-value is one-tailed.

4.5.3.2 Empathic accuracy for arousal

A Pearson correlation was performed to analyse the relationship between social distance and EA for arousal. The analysis revealed a non-significant relationship, $r(35) = -.16$, $p = .177$.

4.5.4 Individual differences in the reward effect

4.5.4.1 Empathic accuracy for valence

To investigate whether AQ modulated the effect of social distance on EA for valence, a multiple linear regression was performed. The sample was taken after the Cook's procedure used prior to the analysis in Section 4.5.3.1. EA was entered in the model as the response variable, with mean-centred social distance and mean-centred AQ, and their interaction, as predictor variables. The model revealed that the predictor variables did not explain a significant amount of variance in the value of EA ($F(3,32) = 1.29$, $p = .230$, $R^2 = .11$). Mean-centred social distance significantly predicted EA ($\beta = -.33$, $p = .029$). However, mean-centred AQ did not ($\beta = .08$, $p = .652$), nor did their interaction ($\beta = -.01$, $p = .946$). Therefore, AQ did not modulate the effect of social distance on EA. The results did not change when mean-centred self-rating variability and its interaction with mean-centred AQ were added to the model as covariates.

4.5.4.2 Empathic accuracy for arousal

To investigate whether AQ modulated the effect of social distance on EA for arousal, a multiple linear regression was performed. The sample was taken after the Cook's procedure used prior to the analysis in Section 4.5.3.2. EA was entered in the model as the response variable, with mean-centred social distance and mean-centred AQ, and their interaction, as predictor variables. The model revealed that the predictor variables did not explain a significant amount of variance in the value of EA ($F(3,33) = .49$, $p = .693$, $R^2 = .04$). Mean-centred AQ did not significantly predict EA ($\beta = .11$, $p = .549$), nor did mean-centred social distance ($\beta = -.18$, $p = .299$) or their interaction ($\beta = -.07$, $p = .701$). Therefore, AQ did not modulate the effect of social distance on EA. The results did not change when the mean-centred self-rating variability and its interaction with mean-centred AQ were added to the model as covariates.

4.6 Discussion

The study reported in this chapter aimed to investigate whether social closeness influences EA, as measured with the video rating task (e.g. Zaki et al., 2008). In addition, it set out to examine whether autistic traits modulate this presumed relationship. The results showed that the finding of Stinson and Ickes (1992), i.e. that friends show higher EA than strangers, was replicated here. Specifically, the inclusion of a social closeness scale measure allowed to quantify this effect along a continuum: the more distant the other person was judged to be, the lower the perceiver's ability to infer that person's feelings.

In addition, it was found that the effect of social closeness was not further modulated by individual differences in autistic traits. This was the case neither for arousal, nor for valence. A possible explanation for this finding is that, when watching someone talk about an experience they had, there is a wealth of cues available. It is well-documented that people with autism find it difficult to recognise emotions from facial cues (e.g. Hobson, 1995; Griffiths, Jarrold, Penton-Voak, Woods, Skinner, & Munafò, 2017). Although facial expressions are an important source of emotion information, the video rating task employed here, and indeed everyday social interactions, provide a wealth of other cues, such as bodily expressions, prosody and communication content (Blanke et al., 2015). It is therefore possible that the cues that are more effectively used by those closer to each other are utilised by people higher in autistic traits just as well.

In a previous study that employed the video rating paradigm, it was found in a student sample that those higher in autistic traits showed lower EA for valence (aan het Rot & Hogenelst, 2014). This observation contradicts the lack of a direct relationship between the two as reported in the current study. As the range of AQ scores in their study and the one reported here were similar, it is unlikely that there was a selection bias in the current study. Rather, it is possible that the lack of a video selection procedure in the current study meant that it was not controlled how difficult to rate the videos were across participants. Aan het Rot and Hogenelst (2014) discarded videos that had “limited temporal variability in the continuous [self-ratings for experienced valence]”. Because of the pairing method employed in the current study, participants only saw two positive and two negative videos. Due to this small number of videos available to each participant, it was decided not to discard (less emotionally varied) videos prior to the other-ratings phase. To investigate the effect of social closeness on EA while controlling the 'difficulty' of videos, future studies could therefore first create a large video database from a social network (e.g. an academic cohort; orchestra; sports team), then make a video selection, and then have participants rate the videos and their social closeness to the persons in the videos. Although it was possible that video difficulty obscured an effect of autistic traits in the current study, the covariate analyses above showed that the null effect remained after controlling for video difficulty. Nevertheless, a video selection procedure would make it possible to control how difficult videos are *prior* to conducting the EA experiment.

Another potential limitation lies in the accuracy of self-ratings. An inherent weakness of paradigms using self-ratings is that they rely on introspection. If someone is not well aware of the valence and arousal of their emotions, it cannot be expected that other-ratings match self-ratings closely. To address this, aan het Rot and Hogenelst (2014) took the average of all other-ratings for each time point in the rated videos, and found that they were very similar to the self-ratings. Although this average only serves as a third-party proxy for what the person is feeling, it suggests that self-ratings (at least within the context of the video rating task) are in fact relatively accurate. However, a more objective real-time measure of valence and arousal (e.g. distinct neural signals reflecting the two components; see Colibazzi et al., 2010) would be preferable.

Lastly, it could be argued that watching videos twice could result in memory effects, for self-ratings, but especially for other-ratings. Specifically, if one has already watched a video, they can remember the course of that video and respond more accurately the second time. However, this memory confound is unlikely to have played a role, as the average EA for the second run (arousal ratings) was lower than for the first run (valence ratings). On the other hand, this fixed valence-arousal order could explain why a significant relationship between social closeness and EA was only found for valence. It is possible that the effect of memory suppressed a possible effect of social closeness on EA for arousal. However, it is also possible that the difference in findings between valence and arousal was meaningful. Arguably, when you are close to someone, it is more important to you that the other person feel happy than that they feel calm or excited. For this reason, it could be that close others pay relatively more attention to cues that signal valence than to cues that signal arousal. Furthermore, EA for arousal was in general lower than for valence, even though the preliminary analysis revealed that the self-ratings for arousal and valence did not differ in terms of variability/difficulty. This finding suggests that self-rated arousal is more difficult to judge by a perceiver, possibly because arousal cues tend to be internal and physiological and therefore, albeit as variable as valence cues, difficult to convey (Feldman, 1995). Nevertheless, to avoid the confounds that two successive ratings may introduce, future studies could use two-dimensional response scales (cf. Erbas et al., 2016), so participants can rate valence and arousal simultaneously. For example, adapting the task to tablet computers could allow participants to give continuous manual ratings more effortlessly.

4.7 Conclusion

Using social closeness to quantify the relationship between paired-up participants, the finding by Stinson and Ickes (1992) that friends show higher EA than strangers was conceptually replicated. This replication was done within the context of an EA paradigm that focuses on real-time ratings of emotions rather than verbal inferences. It was further found that the effect of social closeness was not further modulated by individual differences in autistic traits. As mentioned in Section 4.1 (Introduction), EA represents the integration of various information sources (e.g. verbal or non-verbal cues) to arrive at inferences about others' internal states (Blanke et al., 2015). Therefore, it is possible that people high in autistic traits can choose which cues to rely on, paying more attention to cues when the target of their inferences is liked.

The experimental chapters so far have looked at the effect of reward on cognitive empathy on the level of component processes (Chapter 2 and Chapter 3) and on a more holistic level (Chapter 4). They revealed that the effect of reward on cognitive empathy depends on the way cognitive empathy is measured. However, the question remains whether performance on these tasks that purportedly measure cognitive empathy is related. Moreover, it is possible, if reward modulates performance on some tasks (but not others), that these task measures are in some way related to each other (i.e. they reflect an overlapping construct). Therefore, it is necessary to test these task measures together in one experiment, without a reward manipulation. This way, it can 1) be investigated whether purported measures of cognitive empathy in fact relate to each other and 2) whether this relationship especially holds for tasks in which performance has been shown to be modulated by reward. In order to investigate this, Chapter 5 covers a study in which the tasks used in Chapters 2, 3 and 4 are combined into a task battery, together with other common cognitive empathy-related tasks.

5 The interrelationship between putative measures of cognitive empathy

This chapter addresses the question whether the task measures covered in previous chapters are mutually correlated. Specifically, these diverse tasks have often been used as putative indices of cognitive empathy, but their interrelationship is largely unknown. For this purpose, a study was carried out using a task battery consisting of six different task measures. Together, the selection tapped into Theory of Mind, perspective taking and empathic accuracy. The results showed that there was no significant relationship in performance between any two tasks. This lack of relationship was also the case for tasks that were similar in their design and are thought to tap into the same cognitive empathy-related process. As such, there was no evidence that putative measures of cognitive empathy, as measured in a lab-based experimental context, correlate with each other. It is proposed that task measures rely on a combination of cognitive capacities that contribute towards cognitive empathy judgments but function independently.

5.1 Introduction

As discussed in Chapter 1, cognitive empathy is an umbrella term that refers to various empathy-related processes. One effective trichotomy to operationalise cognitive empathy is that of perspective taking, Theory of Mind (ToM), and empathic accuracy (aan het Rot & Hogenelst, 2014). In the studies that followed Chapter 1, it was investigated whether reward influences performance on tasks that tap into these component processes. This research question resonates with the majority of cognitive empathy research, which focuses on the factors that impact performance on such tasks. However, research into how performance on these tasks relates to each other remains relatively sparse. Arguably, if an individual is good at exhibiting cognitive empathy, this ability should be reflected in good performance across tasks that purportedly tap into cognitive empathy. By using tasks that tap into these different processes, it can therefore be investigated whether they ultimately reflect the same construct.

The inherent challenge in answering this question lies in making a selection of tasks out of the available corpus. If one chooses tasks that are very similar in their scope, it is possible that only a certain aspect of cognitive empathy is measured. For example, there have been some studies investigating the relationship between tasks used for cognitive empathy research, but their task batteries included only two or three tasks and were limited to certain aspects of ToM (e.g. Peterson & Slaughter, 2009; Beaumont & Sofronoff, 2008; Devine & Hughes, 2013). Another common restriction in these task battery studies is that they usually only concern the non-adult population (Brent, Rios, Happé, & Charman, 2004; Carlson et al., 2002; Hughes, Adlam, Happé, Jackson, Taylor,

& Caspi, 2000). Therefore, inferences about how cognitive empathy-related processes relate to each other in the developed individual are limited. In short, a wider range of tasks and an adult participant sample are needed to investigate the relationship between tasks measuring cognitive empathy-related processes.

In addition, the previous chapters investigated the relationship between autistic traits and the influence of reward on task performance. However, because of this reward manipulation, it could not be examined whether autistic traits were directly related to task performance on the selected tasks. Some task battery studies have investigated the relationship between task measures for individuals with an autism spectrum condition versus typically developed individuals (Brent et al., 2004; Rosenblau, Kliemann, Heekeren, & Dziobek, 2015; Wilson et al., 2014). However, no study so far has looked into how autistic traits relate to differences in task performance across a wide range of cognitive empathy tasks. By measuring autistic traits, it can be examined whether and how task performance relates to quantifiable individual differences.

In short, the following questions arise:

1) How does performance on tasks reflecting separate cognitive empathy components relate to each other? The previous chapters fractionated cognitive empathy into Theory of Mind, perspective taking and empathic accuracy (cf. aan het Rot & Hogenelst, 2014), yet it is unclear whether tasks tapping into these components correlate.
2) Is performance on these tasks related to self-reported differences in autistic traits?

5.2 Aims

Based on the questions presented above, the aims of this chapter were twofold. The primary aim was to investigate whether tasks that purportedly measure processes related to cognitive empathy relate to each other. Specifically, it is possible that they map onto a single overlapping construct, or, conversely, rely on a variety of not directly related cognitive capacities. In the previous chapters, the several employed task measures tapped into different components. The Sandbox task measured explicit belief tracking, which is commonly regarded as reflecting ToM (e.g. Begeer et al., 2012). The mouse-tracking task measured implicit belief tracking, as the participant was never explicitly instructed to track the other's belief. The Director task and Dots task measured visual perspective taking: the former allowed an investigation of egocentric interference, whereas the latter additionally allowed an investigation of altercentric interference and perspective selection. Lastly, the video rating task measured empathic accuracy. Note that these tasks represent the component processes that lead to empathy-related inferences (see Figure 1.1). Therefore, all these tasks were included in a task battery in the current study, with the exception of the Dots task, which was identified as a task measure after completion of the current study. In addition, two other commonly used cognitive empathy-related tasks were selected to broaden the scope of task measures. Specifically, the revised Reading the Mind in the Eyes task (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001a) and the Hinting task (Corcoran, Mercer, & Frith, 1995) were chosen. The former has often been placed under the label 'ToM', although one could argue that labelling emotional expressions is more closely related to empathic accuracy, as it involves mental state inferences; the latter is a task that measures pragmatic language apprehension, also tapping into ToM. With this array of tasks, a representative sample of the spectrum of available tasks was used.

The second aim was to investigate whether performance on these task measures relates to self-reported autistic traits. To this end, the AQ (Baron-Cohen et al., 2001b) was used.

5.3 Predictions

Although the different task measures mentioned above are thought to reflect different processes, they have all been used in cognitive empathy research and ultimately reflect cognitive empathy ability. Therefore, it is expected that the measures used here will correlate. Combining this rationale with the findings from the previous chapters, the following predictions were made:

1) It is hypothesised that there is a positive relationship between all task measures that putatively tap into cognitive empathy.
2) It is hypothesised that autistic traits will be negatively related to performance on the selected task measures.

5.4 Methods

5.4.1 Participants

Recruitment for the current experiment took place in two separate data collection waves. The sample of the first wave consisted of 43 participants (36 female, mean age = 19.77, *SD* = 1.46, min = 18, max = 24). The sample of the second wave consisted of 40 participants (34 female, mean age = 19.83, *SD* = 2.86, min = 18, max = 36). As such, the total sample amounted to 83 participants (70 females, mean age = 19.80, *SD* = 2.23, min = 18, max = 36). The Wave 1 sample performed five tasks and were later contacted for a follow-up session to perform a sixth task. All Wave 2 participants performed all six tasks in the same session. The experimenter leading the sessions for these two data collection waves was different.

All participants were recruited through the University of Reading research panel. All participants were undergraduate students, and were granted course credit upon completion of the study. In addition, Wave 1 participants received a compensation of £5 for participating in the follow-up session. Ethical approval was received from the School of Psychology Ethics Committee and all participants provided written informed consent prior to their participation.

5.4.2 Stimuli

5.4.2.1 Director task

The shelves and the objects shown on them were taken from the same image database as for the Director task discussed in Section 3.4 (Director task study). A portrait photograph of the experimenter was shown behind the shelves to represent the director. As the Director instructions were given by the experimenter, no pre-recorded audio recordings were used. Note that the reward conditions in the version of the Director task employed in Chapter 3 required the use of a number of unique voices, as the face representing the director varied between trials. In contrast, the absence of reward conditions in the current version meant that the identity of the director could be held constant across trials. Although instructions given by the experimenter made it possible that instruction delivery varied between trials and/or participants, it was assumed that instructions coming from a real person facing the participant would be a more ecologically valid means to bringing individual differences in perspective taking to light (see also Section 5.4.3.3.).

5.4.2.2 Sandbox task

Characters in the Sandbox scenarios were represented by open-source images of cartoon characters taken from the Internet (see Appendix E), instead of the portrait photographs described in Section 2.5 (Sandbox study). The current approach was taken because, at the time of design, available stimulus sets did not provide an adequate number of images for the number of trials used in this version of the Sandbox task. In addition, the task battery was designed with possible future non-adult samples in mind. For this reason, cartoon characters were preferred to photographs of adults. Although the resulting stimulus set showed some variability in colours, styles and emotion, characters shown in a single trial were matched in terms of emotion. Scenarios were written based on the scenarios used in previous studies (e.g. Chapter 2.5 Sandbox study; Begeer et al., 2012; see Appendix F). The sandbox width was now set to 692 pixels (its border was 4 pixels on either side) to allow for more granularity. In addition, the colour of the sandbox, varied across trials. A blank canvas and a computer-generated circle were used in the filler task.

5.4.2.3 Reading the Mind in the Eyes task

All 37 monochrome stimuli of the revised version of the Reading the Mind in the Eyes task (Baron-Cohen et al., 2001a) were used in the current task (see Appendix G).

5.4.2.4 Hinting task

All 10 scenarios of the Hinting task (Corcoran et al., 1995) were used in the current task (see Appendix H).

5.4.2.5 Mouse-tracking task

The same stimuli that were used for the mouse-tracking study in Chapter 2.4 were used here, with the exception that only one of the faces used in that study was also used here.

5.4.2.6 Video rating task

The stimuli used for this study were taken from the database that was generated during the study discussed in Chapter 4. A selection of 12 videos (six positive, six negative) was made based on the following criteria: (1) the raw correlation between the self- and other-ratings for valence should be higher than .6; (2) the average valence rating for the positive videos should be higher than 6 (range: 1-9), whereas the average rating for the negative videos should be lower than 4; (3) the number of time bins for the analysis was higher than 30 (equating to over a minute of analysed video footage); (4) the life experience narrated in the video was a specific event, rather than a long-term process. In addition, the quality of the video was considered (i.e. a match between the valence of the narrator and the life experience; clear view of face; clear speech; audiovisual quality), and it was endeavoured to select videos as diverse in their subject matter as possible (see Appendix I). In half of the videos per valence, the narrator was female.

5.4.3 Procedure

As mentioned before, some tasks used for the current study were already covered in previous chapters. Although their counterparts were similar in their design, the aims of this study required alterations to be made to the design of the tasks. For each task covered below, the reader is referred to the previous chapter covering that particular task, and differences are summarised. Additionally, between the two data collection waves, minor changes were made to the study protocol. These changes are recorded and explained in Appendices J and K.

5.4.3.1 Trait measure

Upon signing up, the participant received a link to an online survey. This survey included the Autism-Spectrum Quotient (AQ; Baron-Cohen et al., 2001b), which quantifies autistic traits (see Appendix A). In addition, the participant was asked to indicate their gender and age. Upon completion of this survey, participants were invited to the lab to take part in a battery of tasks.

5.4.3.2 Pre-battery introduction

Prior to the task battery, the participant was informed about the study in general. They were told that they would perform a battery of tasks, all measuring social cognition, and that the goal was to investigate how these tasks compare to each other. They were informed about the use of a non-intrusive eye-tracker in the

Director task. In addition, it was mentioned that they would perform a task in which they would watch videos of people talking about life experiences, and would be asked whether they knew these people. Lastly, they were encouraged to take breaks between the tasks.

The order of the tasks was fixed. As the Director task required the longest setup procedure, this task was performed first. The Sandbox task and the Reading the Mind in the Eyes task were performed consecutively, as they were presented with the same software. The Hinting task was administered verbally next, to prevent participants from becoming fatigued from doing several computer tasks in a row. As a result, the mouse-tracking task was performed after the previous four tasks. As the video rating task was added at a later point, participants performed this task last.

5.4.3.3 Director task

The procedure of the Director task was generally identical to the version discussed in Section 3.4.1.3. However, some changes were made to tailor the task to the aims of the current study. In the study discussed in Chapter 3, directors were represented by images of faces that had been conditioned with a reward value prior to the Director task. Computerised directors made it possible to switch between different directors on a trial-by-trial basis. As no reward manipulation was used for the current study, only one director was needed, whose role was taken on by the experimenter. This meant that the director was both displayed behind the shelves on the screen and physically sitting behind these shelves from the perspective of the participant. The rationale behind this setup was that it would be more effective in showing individual differences in perspective taking.

The experimenter first explained the trial procedure, by leading the participant through an example trial, after which the participant was asked whether they had any questions. Next, the eye-tracker was calibrated. After calibration, there were eight practice trials, after which the 32 main trials started. For each trial, the experimenter clearly and carefully read out the director instructions and pressed the spacebar to make the shelves appear on-screen. This button press started the response phase, in which the participant made their decision about which object to move, and where to move it to. The participant then clicked to make the mouse cursor appear at fixation, after which they first clicked on the target object and then on the destination shelf.

There were three conditions: Experimental (8), Control (8), and Filler (16). On Experimental trials, the participant had to disregard their own privileged information to take the perspective of the director. There was always competition regarding the object that was to be moved, rather than the shelf where the object was to be moved to. Similar to the other Director task (Section 3.4.1.3), the Experimental condition here involved competition in a group of three objects from the same semantic category, where the best referent was visible to the participant, but not to the director (i.e. the shelf was occluded by a board). On Control trials, this semantic group only consisted of two objects, as the occluded object was replaced with an irrelevant object from a different semantic category. Therefore, the best referent for the director instruction was visible to both the director and participant. On Filler trials, the instruction referred to an object that was visible to both the director and participant, and shared its semantic category with no other object on the shelves. Each shelf configuration was used for one Experimental, one Control, and two Filler trials. The order of the trials was pseudorandomised (i.e. the same condition never occurred more than twice in a row) and fixed across participants. No feedback was given at the end of a trial.

5.4.3.4 Sandbox task

The Sandbox task employed here was similar to the task used in Section 2.5 (Sandbox task study). It was presented with Collector (Garcia et al., 2016). There were four False Belief trials, four Memory Control trials, and one True Belief trial, preceded by one practice trial from both the False Belief and Memory Control condition. Each step of a trial scenario took 6,000 ms. The character images (140 x 140 px) were presented above the sandbox.

One difference with the other Sandbox task (Section 2.5.1.3.2) was that, on Memory Control trials, the ‘mover’ character introduced a new object (e.g. a stone) to the sandbox, rather than moving the original object (e.g. a coin). This change in design was motivated by the need for a paradigm more similar to that of previous studies (e.g. Begeer et al., 2012, 2016), as the version in Chapter 2 did not replicate the main effect reported in those studies. In addition, the word search filler task described in Section 2.5.1.3.2 was replaced with a ball-tracking task. The participant used the mouse cursor to track a small purple ball (diameter: 20 pixels), moving in a sinusoidal pattern across the screen. They clicked on the ball when it arrived at its end point, prompting the start of the response phase. The ball-tracking task was included to prevent participants from using visual strategies to memorise the location of the Sandbox object. Lastly, the presence of the character images directly reflected the scenario (i.e. they were shown next to each other at the start of the scenario, and (dis)appeared when the scenario mentioned this for the respective character). For the scenarios used in this Sandbox task, see Appendix F. The order of the main trials was randomised and no feedback was given at the end of a trial.

5.4.3.5 Reading the Mind in the Eyes task

Next, the participant performed the Reading the Mind in the Eyes task (Baron-Cohen et al., 2001a), a traditional measure that taps into mentalising. The task was administered in Collector (Garcia et al., 2016) directly after the Sandbox task. First, the participant was told that they would see part of someone's face with an emotional expression. They were told to choose the best alternative out of four words to describe this emotional expression (see Appendix G for stimulus materials). There was one practice trial, after which the 36 main trials were presented in randomised order. On each trial, the stimulus was presented in the middle of the screen (500 x 330 px). The response alternatives were presented as horizontally aligned response buttons under the stimulus. Each trial had a time limit of 10 seconds. Two seconds before the time limit, the participant was warned that the time limit was imminent. The task proceeded to the next trial after the allowed time had elapsed, or once a response was given. No feedback was given.

5.4.3.6 Hinting task

The following task was the Hinting task, a task that taps into the ability to make ToM inferences (Corcoran et al., 1995). This task was administered verbally by the experimenter. In the Hinting task, the participant makes inferences about hidden meanings of what someone says. Before the task, the participant was told that the experimenter was going to read out 10 stories to them involving two people. Each story would end with one of the two people saying something, after which the experimenter would ask one or more questions about what the person said. The experimenter then read out the stories in fixed order (see Appendix H for the stories). Each story started with 2-3 sentences describing the background. This introduction was followed by Character 1 asking something or making a remark to Character 2. The hidden meaning was always that Character 1 wanted Character 2 to do something for them. Then, the experimenter asked the participant what Character 1 meant. If the participant's response matched a predefined 'ideal' response, the experimenter proceeded to the next story. Otherwise, the experimenter read out an additional prompt from Character 1 and asked another test question about the hidden meaning conveyed by Character 1. After the participant gave their response to this question (or failed to do so) the experimenter proceeded to the next story, without giving feedback on whether the response was correct or not. All the participant's responses were recorded verbatim by the experimenter.

5.4.3.7 Mouse-tracking task

The task used here was an adapted version of the mouse-tracking task covered in Section 2.4.1.3.2 (Implicit belief tracking task). In this task, the participant watched a video in which a ball moved around between occluding blocks. At the end of the video, the participant had to click on the location of the ball. During the video, there was an agent on the screen, who sometimes did not see that the ball changed location. In addition, the location of the ball, once it was revealed at the end of the video, was not always where the participant had last seen it. Therefore, there were four conditions, based on the beliefs of the participant and agent (True-True, True-False, False-True, False-False).

The task was presented with Collector (Garcia et al., 2016), in a 1024 x 768 window. The participant first watched an instruction video, in which a narrator performed an example trial from all four belief conditions. During the video, the narrator explained how the participant was to use the mouse cursor to click on the location of the ball, even when it appeared somewhere the participant was not expecting it. Next, the participant performed eight practice trials (two from each belief condition), after which their understanding of the task was tested by four statements that required a true/false answer (Click on where the ball is (true); Click on where the person on the screen last saw the ball (false); The ball will always appear in the same location after the "NOW" instruction as before it (false); The ball will sometimes appear in a different location after the "NOW" instruction than before it (true)). Once task understanding had been ascertained, the participant continued to the main trials. There were 32 trials, with eight trials for each condition. These were equally divided across, and randomised within, two blocks. After the first block, the participant was allowed a one-minute break.

Although the trial procedure was similar to the mouse-tracking task used earlier (Section 2.4.1.3.2), there were a number of differences (see Figure 5.1 for the current trial procedure). First, the agent always appeared in the bottom left corner of the screen, to maintain uniformity across trials. Second, Collector, being an online data collection platform, did not allow automatic repositioning of the mouse cursor upon the start of the response phase. For this reason, the participant was instructed to keep the mouse cursor inside a circle in the bottom middle of the screen (726 px from the top border). The diameter of this circle was set to 20 pixels, to discourage the participant from moving the mouse cursor before the response phase, whilst making it easy to return to the circle if they had inadvertently moved the cursor. The participant initiated each trial by clicking within this circle (trial initiation click), and subsequently watched the trial enfold. They then started their response trajectory by clicking again during the response phase (i.e. when the message ‘NOW’ was shown; trajectory initiation click). Therefore, the position of the mouse cursor upon the start of the response phase was similar to that in the previous mouse-tracking task (Section 2.4.1.3.2). Third, the location of the ball was revealed when the cursor’s trajectory crossed an invisible horizontal line on the screen (cf. a time-dependent reveal in the other mouse-tracking task, Section 2.4.1.3.2). This line was located 50 pixels above the top of the starting circle. This change to a ‘pixel-dependent’ reveal of the ball (cf. van der Wel et al. 2014) was made possible by the change in software used for this version of the mouse-tracking task. The rationale for this change was that it would prevent participants from waiting for the reveal to before they initiated their response. After clicking on one of the two possible locations (response click), the participant clicked on a button in the middle of the screen to proceed to the next trial (trial end click). If participants did not respond within 3,000 ms or mistakenly clicked on the location of the cube rather than the ball, they received feedback for 3,000 ms. Error feedback was included to prevent participants from making systematic mistakes by not adhering to the simple rule of clicking on the ball. As the task measure was based on the response trajectory rather than accuracy, the use of this error feedback was justified. The order of the trials was randomised across participants.

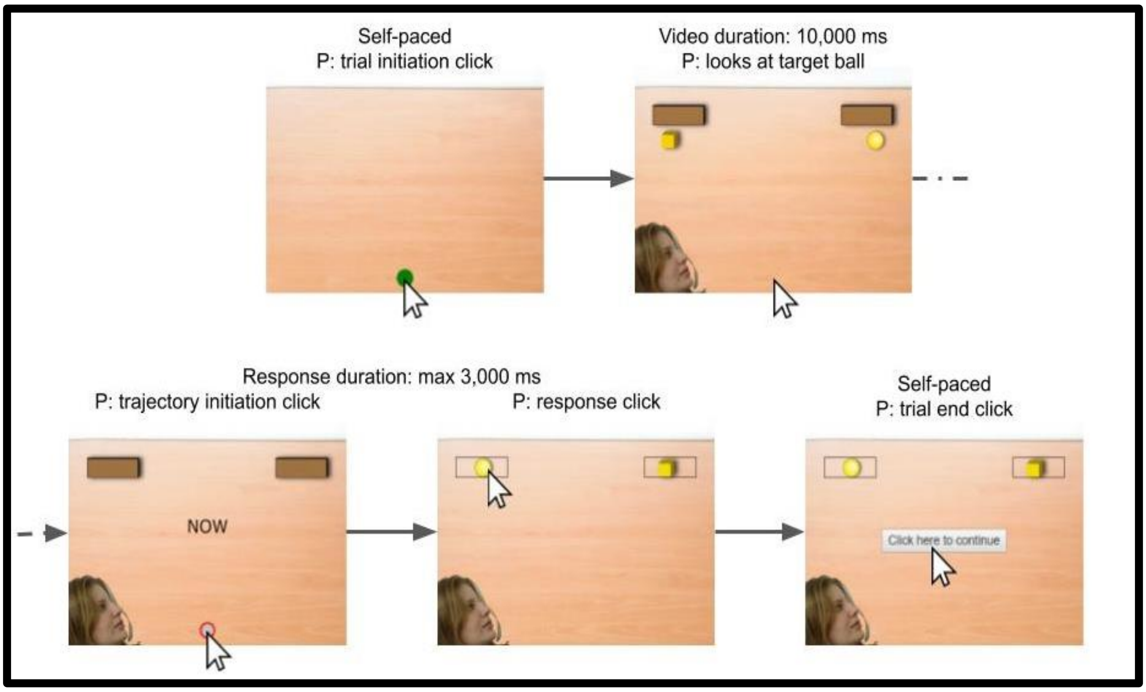


Figure 5.1 Schematic overview of a trial in the task battery MT task. Before the trial started, an empty background was shown. Once the participant ('P') clicked within the green circle, the trial was initiated and the belief scenario enfolded (see Figure 2.2 for an example of a belief scenario). Once this scenario concluded, the participant clicked within the red circle to start their response. Once the mouse cursor crossed an invisible horizontal line 50 pixels above the red circle, the ball and the cube were revealed. The participant then made their response by clicking on the location of the ball. Lastly, the participant clicked on a button to end the trial and proceed to the next. The mouse cursor and the buttons were enlarged for this overview.

There was also a difference in how the mouse trajectory data was handled in this mouse-tracking task compared to the task discussed in Section 2.4.1.4 (Mouse-tracking task data analysis). Specifically, in the other task, mouse-tracking was time-based, i.e. coordinates and their corresponding timestamps were recorded at the level of milliseconds. On the other hand, in the current task, mouse-tracking was movement-based, due to software restrictions. Put differently, mouse coordinates and timestamps were only recorded when mouse movement took place. Note that both approaches can be reconciled by using an interpolation procedure on the movement-based coordinate data to obtain coordinates on the level of milliseconds (see Section 5.4.4.5 Mouse-tracking task data analysis).

5.4.3.8 Video rating task

For Wave 1 participants, this task was not part of the main procedure. Therefore, participants were contacted for a follow-up session, to perform one more task. Twenty-two Wave 1 participants performed the video rating task as a result. All Wave 2 participants performed this task after the tasks covered above, but during the same session.

Before the video rating task, the participant indicated for each of the 10 unique narrators in the video rating task if they knew the person. In addition, the participant was asked to indicate how close they felt to them on a scale of 0-100 (0: ‘Virtually the same person as me’; 100: ‘A total stranger’).

Next, the participant was instructed about the task, performed a practice trial, and then performed the 12 main trials (see Section 4.4.3 Procedure for a trial overview). Whereas participants gave both arousal and valence ratings during separate runs in the video rating task study (Chapter 4), they only gave valence ratings in the current version. This change was implemented to maximise the number of unique videos seen by each participant while limiting the length of this final task battery task. Valence ratings were chosen over arousal ratings because the former are more commonly used in studies on empathic accuracy (Zaki et al., 2008; Stinson & Ickes, 1992) and because videos had been recorded and selected based on valence (see Section 5.4.2.6). The order of the main trials was pseudorandomised per participant, in such a way that the same valence (positive/negative) never appeared in more than two videos in a row and the same narrator never appeared in two videos in a row.

Upon completion of the video rating task, the participant was debriefed. For Wave 1 participants, this was also done after the mouse-tracking task.

5.4.4 Data analysis

For each of the tasks below, two types of task measures were calculated. First, if the task allowed it, a measure was established to test whether main effects in previous studies using the respective task could be replicated. This measure will be referred to as the 'replication measure', and is specified for each task below.

Second, individual differences in task performance were quantified. This will be referred to as the 'differences measure'. In some cases, this measure was not the same as the replication measure (albeit derived from it). The replication measure was not always specific enough to capture individual differences in task performance. For each task, it will be explained below what the differences measure was and whether it differed from the replication measure.

For the Results section (5.5.2 up to and including 5.5.8), all reported p-values are two-tailed. For the analysis investigating the relationship between task- and trait measures (Section 5.5.8), p-values were corrected for multiple comparisons using the Bonferroni method.

5.4.4.1 Director task

The measures used for the Director task analysis here were based on those explained in Section 3.4.1.4 (Director task data analysis). For the replication measure, the percentage of each Experimental trial that was gazed at the distractor shelf was calculated. Next, this percentage was calculated for Control trials. The difference between these two represented the replication measure. In the current version of the task, the distractor shelf in the Experimental condition contained the object that most closely matched the director's instruction but was only visible to the participant. This shelf was therefore the correct shelf from the participant's perspective, but not from the director's perspective. Therefore, this distractor shelf had to be disregarded to take the director's perspective. The distractor shelf in the Control condition, on the other hand, contained an irrelevant object from a different semantic category than the target object. The replication measure therefore reflected the degree to which the participant was distracted by the occluded object through egocentric interference over and above the object's position on the shelves.

For the differences measure, only the Experimental condition was taken into consideration, as it is here that taking the director's perspective is crucial to task performance. The measure was calculated by subtracting the percentage of the trial gazed at the distractor object from the percentage gazed at the target object (matching the director's perspective). Therefore, the measure reflected how much more the participant looked at the object as conveyed by the director versus the object matching the participant's perspective. As such, the replication measure and differences measure were identical to the measures covered in Section 3.4.1.4.

5.4.4.2 Sandbox task

The replication measure was based on previous papers (e.g. Begeer et al., 2012, 2016; Coburn et al., 2015). The Sandbox task requires a response on a one-dimensional continuum (see 2.5.1.3.2 Explicit belief tracking task). Bias in pixels away from the optimal response location was quantified for both the False Belief and Memory Control condition. For the replication measure, the difference between bias on these conditions was compared.

The differences measure was based on the measure explained in Section 2.5.1.4 (Sandbox task data analysis). The average bias value across False Belief trials was calculated per participant. Note that, in contrast to the other task measures employed here, a lower (bias) score reflected better performance.

5.4.4.3 Reading the Mind in the Eyes task

In contrast to the Director task and Sandbox task discussed above, the Reading the Mind in the Eyes task only provides a single total score (possible range: 0-36), i.e. the number of correct responses within the time limit. For this reason, no replication measure was established, as no comparison could be made between conditions.

The total score per participant was taken as the differences measure.

5.4.4.4 Hinting task

Similar to the Reading the Mind in the Eyes task, the Hinting task only provides a single total score. For each story, two points were added to the score if the participant gave a correct answer after the first test question, or one point if they only gave a correct answer after the second test question. As such, the possible total score ranged between 0 and 20. The task yielded no replication measure.

The total score per participant was taken as the differences measure.

5.4.4.5 Mouse-tracking task

In concordance with Section 2.4.1.4 (Mouse-tracking task data analysis), the dependent variable of interest for the mouse-tracking task was the area under the curve (AUC). This outcome represented the degree to which the participant's response trajectory was deviated from the correct location (the ball) towards the irrelevant location (the cube). Whereas in the previous study (Section 2.4.1.4), AUC was given as the output by the MouseTracker software (Freeman & Ambady, 2010), here it was calculated based on the mouse coordinates of the response trajectory for each trial. Due to an error in the script for the task in data collection wave 1, the response trajectory was tracked from the trajectory initiation click until the trial end click, whereas tracking should have been terminated at the response click (see Figure 5.1 above). To establish a measure of AUC that was equal across the two data collection waves, a truncation procedure was performed on a trial-by-trial basis for both Wave 1 and Wave 2.

This procedure to arrive at a single AUC per trial from a sequence of raw coordinates went as follows (see Figure 5.2): at the start, the response trajectory was represented by a sequence of X- and Y-coordinate pairs, with their corresponding timestamps. These coordinates corresponded to the 1024 x 768 pixels of the task window. Note that coordinates were only recorded when there was mouse cursor movement.

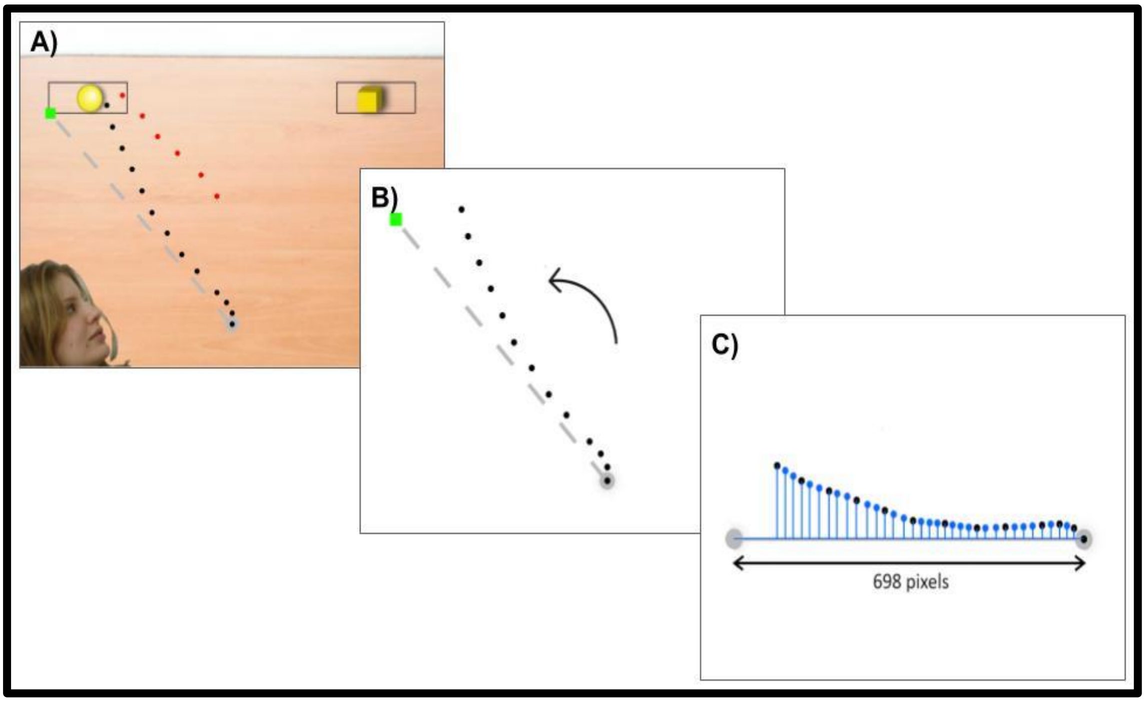


Figure 5.2 Schematic overview of the task battery mouse-tracking AUC calculation procedure. A) Example response trajectory for a trial in which the correct response location is left. Green square: outer bottom corner of the correct response location. Black dots: recorded points of the response trajectory until the truncation point (the shortest two-dimensional distance to the green square). Red dots: response trajectory excluded after truncation. Grey dashed line: optimal response trajectory. B) Enlargement of the response trajectory after truncation. C) Response trajectory after rotation. Blue dots: response trajectory coordinates interpolated between black dots. The distance to the horizontal baseline was summated across all trajectory points to arrive at the AUC, which was subsequently divided by 10,000.

First, the X- and Y-coordinate sequences were truncated at the point where the trajectory was closest to the outer bottom corner of the correct response location. This point was chosen as it reflects the greatest certainty, as expressed in the response trajectory, that this response location is the correct one. Consider the following: if a trajectory that goes up in a straight line between the response boxes can be regarded as indecision, then a trajectory going to the outer bottom corner of the correct response location has the largest angular deviation from this indecision. Therefore, for both Wave 1 and Wave 2, this truncation resulted in a trajectory end point that reflected the greatest certainty for the (correct) response that the participant made.

After this truncation, the remaining trajectory was rotated, using the hypothetical optimal trajectory (a straight line) between the middle of the start button and the outer bottom corner of the correct response location as the baseline. As such, the baseline was a horizontal line with a width of 698 pixels (X-coordinates). The actual response trajectory was rotated along with the baseline.

Third, interpolation was used to infer Y-coordinates for missing X-coordinates along the response trajectory.

Fourth, the vertical distance in pixels to the horizontal line was measured for each point of the response trajectory. Note that a trajectory could pass the same X-coordinate more than once.

Lastly, the cumulative distance to the horizontal line was taken as the trial AUC. For reasons of legibility, the total AUC was divided by 10,000.

To investigate whether the effects of participant's belief and agent's belief as reported by van der Wel and colleagues (2014) could be replicated here, the average AUC was calculated for each of the four belief conditions.

To arrive at a differences measure, the AUC for the True-True condition was subtracted from the True-False condition, the AUC for the False-True condition was subtracted from the False-False condition, and the average of the two differences was taken. As such, this average value represented the degree to which the agent's belief influenced a participant's responses.

5.4.4.6 Video rating task

The same procedure as described in Section 4.4.4 (Video rating task data analysis) was used here to arrive at a single value representing empathic accuracy per participant. As described earlier in the current chapter (Section 5.4.3.8), the current video rating task only involved valence ratings. To summarise the procedure: self- and other-ratings were averaged across time bins. The self- and other-ratings were subsequently correlated with each other. Next, the correlation coefficients across all participants' trials were subjected to a Fisher's Z transformation. Per participant, the average Z-transformed correlation coefficient was used as a reflection of empathic accuracy.

Similar to the Reading the Mind in the Eyes task and Hinting task, the video rating task does not allow a comparison between two or more conditions as a replication analysis. However, Zaki and colleagues (2008) examined the mean raw correlation coefficient and its distribution across trials, to establish whether their data was suitable to investigate individual differences between participants. Therefore, to validate the current video rating task, these descriptive statistics were compared to those reported by Zaki and colleagues (2008).

For the differences measure, each participant's empathic accuracy was represented by the average Z-transformed correlation coefficient.

5.4.5 Overview

As the current study employed a variety of task measures, Table 5.1 below shows an overview of what these tasks purportedly measure, and how these abilities are quantified in the current study (both a replication measure, and a measure used for individual differences). In addition, it shows the theorised relationship between the differences measure and cognitive empathy performance.

Table 5.1 Overview of the task measures in the task battery study

Task	Ability tapped into	Replication measure	Differences measure	Theorised relationship to cognitive empathy
Director	perspective taking	Distractor gaze % on Experimental trials vs. Distractor gaze % on Control trials	Target gaze % on Experimental trials - Distractor gaze % on Experimental trials	Positive
Sandbox	explicit belief tracking	Bias on False Belief trials vs. Bias on Memory Control trials	Bias on False Belief trials	Negative
Reading the Mind in the Eyes	ToM/ empathic accuracy	N/A	Total test score	Positive
Hinting	ToM inferences	N/A	Total test score	Positive
Mouse-tracking	implicit belief tracking	The effect of participant's and agent's belief on AUC	The effect of a False vs. True agent's belief on AUC	Positive
Video rating	empathic accuracy	Mean raw *r* correlation between narrator's and participant's valence ratings	Empathic accuracy (Z-transformed correlation between narrator's and participant's valence ratings)	Positive

Note: the Sandbox task was the only task for which the differences measure had a theorised negative relationship to cognitive empathy. The sign for correlations between this measure and other measures should be interpreted accordingly.

5.5 Results

5.5.1 Pre-processing

Pre-processing steps and the analysis were performed with R (R Core Team, 2018). For each task, data were pre-processed separately for the replication analysis. This replication analysis is described in sections 5.5.1.1 to 5.5.1.6. For each task, the sample for the replication analysis was based on all participants who had completed that particular task. For the correlation analysis, the sample was based on participants who had completed all tasks (see Section 5.6.1.7 Sample for correlations).

No procedure was used to identify trials or participants as statistical outliers. This decision was made because previous studies employing these task measures, with the exception of the mouse-tracking task (van der Wel et al., 2014), have generally not used an outlier removal procedure (e.g. Director task: Epley et al., 2004; Keysar et al., 2000; Santiesteban et al., 2012, 2015; Savitsky et al., 2011; Sandbox task: Begeer et al., 2012, 2016; Coburn et al., 2015; Reading the Mind in the Eyes task: Baron-Cohen et al., 2001a; Hinting task: Corcoran et al., 1995; Video rating task: Martin-Key et al., 2017; Zaki et al., 2008). In addition, no outlier identification was done in previous task battery studies (Wilson et al., 2014; Brent et al., 2004).

5.5.1.1 Director task

Two participants were excluded beforehand due to unusable eye gaze data. One participant was excluded because their response was not accurate (i.e. they failed to click on the correct target object and destination shelf) on all eight

Experimental trials. As a result, the final sample consisted of 80 participants (68 females), whose age ranged between 18 and 36 (M = 19.81, SD = 2.27). Only Experimental and Control trials were analysed. Due to a technical error, data for eight trials was not saved, resulting in a data set of 1,272 trials. On average, participants responded accurately on 7.71 Control trials (*SD* = .60) and on 6.96 Experimental trials (*SD* = 1.93). The subsequent pre-processing steps were based on the steps delineated in Section 3.4.2.1 (Director task data pre-processing).

First, the eye with the fewest invalid gaze data points was chosen for the gaze data analysis. Second, trials were excluded if the participant had not clicked on the correct target object and destination shelf (7.7% of all trials). Third, all gaze points after the first mouse click were removed. Fourth, the first 20 gaze data points were removed. Fifth, trials were excluded if less than 50% of the remaining gaze data points was valid (6.7% of all trials). In contrast to the pre-processing steps explained in Section 3.4.2.1, trials here were not excluded if the participant had not looked at both the target object *and* destination shelf. This difference was because Control trials in the current Director task did not have an occluded shelf that contained an object from the same category as the target object (vis-à-vis Control trials explained in Section 3.4.1.3.2 Director task procedure). Therefore, it was likely that this shelf was not looked at. For each trial, it was then calculated what percentage of the trial the participant had gazed at the correct shelf and the distractor shelf. These values were then used for further analysis. In total, 1,100 trials (range: 5-16 per participant) were included in the analysis.

5.5.1.2 Sandbox task

No data pre-processing was performed. Therefore, no individual trials were excluded from analysis; the sample remained unchanged from the sample indicated in Section 5.4.1.

5.5.1.3 Reading the Mind in the Eyes task

No data pre-processing was performed. Therefore, no individual trials were excluded from analysis; the sample remained unchanged from the sample indicated in Section 5.4.1.

5.5.1.4 Hinting task

No data pre-processing was performed. Therefore, no individual trials were excluded from analysis; the sample remained unchanged from the sample indicated in Section 5.4.1.

5.5.1.5 Mouse-tracking task

In total, 364 trials were excluded from analysis because the mouse trajectory was not fully saved. The remaining data set consisted of 2,292 trials from 83 participants. Subsequently, trials were removed if the response was not correct (1.5% of all trials; see Table 5.2 for average accuracy scores per condition). No time limit for the response initiation was used, as the reveal of the ball was dependent on the cursor crossing an invisible line rather than on time. Hence, no trials were excluded based on response initiation time. Two participants were subsequently excluded from analysis because they did not have trial data in the True-True or True-False condition, respectively. As a result, the final sample for the mouse-tracking task consisted of 81 participants (68 females), whose age ranged between 18 and 36, (M = 19.79, SD = 2.26). In total, 2,225 trials (range: 11-32 per participant) were included in the analysis for this sample.

Table 5.2 Accuracy descriptives per condition of the task battery mouse-tracking task

	Agent belief	
Participant belief	True	False
True	6.25 (2.25)	6.30 (2.13)
False	7.20 (1.22)	7.45 (.95)

Note: Values represent mean (SD) of correct number of trials, out of 8, per condition.

5.5.1.6 Video rating task

Out of the 83 participants who took part across the two data collection waves, 62 (53 female, mean age = 19.77, *SD* = 2.51, range: 18-36) completed the video rating task. For 11 videos across eight participants, the participant indicated that they knew the narrator of the video. These videos were therefore excluded from analysis. For 2.0% of the remaining videos, a correlation could not be calculated between the self- and other-ratings (e.g. the other-rating changed less than two times). Hence, these videos were also excluded from analysis. The analysis was therefore done on 718 trials across all 83 participants (range: 7-12 trials per participant).

5.5.1.7 Correlation analysis sample

After exclusions due to the pre-processing steps described above, 78 participants were included in the final sample for the correlation analysis between tasks. Since the video rating task was done only by a subset of this sample, any correlation analysis with the video rating task data was performed only on this subset (*N* = 58). Differences measures were therefore Z-scored for the smaller and larger sample separately (see Table 5.3 for descriptive statistics).

Table 5.3 Descriptive statistics for the correlation analysis sample

Task/Trait measure	Mean	SD
Director		
Target gaze % on Experimental trials	17.41	7.65
Distractor gaze % on Experimental trials	7.48	5.00
Target-distractor difference*	9.93	9.66
Sandbox		
Bias in pixels on False Belief trials	16.59	44.28
Reading the Mind in the Eyes		
Total score	25.78	3.30
Hinting		
Total score	17.87	1.41
Mouse-tracking		
AUC for True agent’s belief	13.47	2.55
AUC for False agent’s belief	14.24	2.50

False-True difference**	.77	2.76
Video rating		
Mean empathic accuracy***	.67	.21
AQ score	14.88	5.24

Note: *This difference was calculated on the trial level, not as a difference of the measures above it; **This difference was calculated as the average of the Agent False-True difference for the Participant-True and Participant-False conditions. ***This reported mean and SD are for the sample of 58 participants who completed the VRT, the sample size for all other reported means and SDs was 78.

5.5.2 Director task

For the replication analysis, the percentage of a trial gazed at the distractor was compared for the Experimental versus Control condition. The paired-samples t-test revealed that there was a significant difference, $t(79) = 7.12$, $p < .001$, $d = .80$. On average, participants looked longer at the distractor on Experimental trials ($M = 7.62\%$, $SE = .56$) than on Control trials ($M = 3.15\%$, $SE = .25$).

5.5.3 Sandbox task

For the replication analysis, bias was compared for the False Belief versus Memory Control condition. The paired-samples t-test revealed that there was a significant difference, $t(82) = 2.81$, $p = .006$, $d = .31$. Bias was bigger on False Belief trials ($M = 15.90$ px, $SE = 4.77$) than on Memory Control trials ($M = -.16$

px, *SE* = 3.83). This result replicated earlier studies (e.g. Begeer et al., 2012, 2016; Coburn et al., 2015).

5.5.4 Reading the Mind in the Eyes task

No replication analysis was performed for this task. However, the distribution of scores in the full sample (*M* = 25.9, *SD* = 3.26, Min = 20, Max = 34) was similar to the distribution for a typically developed middle-aged population (*M* = 26.2, *SD* = 3.6) and student population (*M* = 28.0, *SD* = 3.5) as reported by Baron-Cohen and colleagues (2001a).

5.5.5 Hinting task

No replication analysis was performed for this task. However, the distribution of scores in the full sample (*M* = 17.89, *SD* = 1.38, Min = 13, Max = 20) was similar to the distribution for a population consisting of adults with/without anxiety or depression symptoms (*M* = 18.3, *SD* = 1.6) as reported by Corcoran and colleagues (1995).

5.5.6 Mouse-tracking task

For the replication analysis, it was investigated whether Participant's belief and Agent's belief influenced AUC. For this purpose, a 2 x 2 repeated-measures ANOVA was performed, with Participant's belief (True/False) and Agent's belief (True/False) as factors. The analysis revealed a significant main effect of Participant's belief, $F(1,80) = 203.90$, $p < .001$, $\eta_p^2 = .718$. The AUC was larger when the Participant's belief was False (*M* = 18.77, *SE* = .41) than when it was True (*M* = 9.04, *SE* = .42). There was also a significant main effect of Agent's belief, $F(1,80) = 6.74$, $p = .011$, $\eta_p^2 = .078$. The AUC was larger when the Agent's belief was False ($M = 14.32$, $SE = .28$) than when it was True ($M = 13.49$, $SE = .28$). The interaction of Participant's belief and Agent's belief was not significant, $F(1,80) = .00$, $p = .998$, $\eta_p^2 < .001$. This analysis replicated the original study (van der Wel et al., 2014).

5.5.7 Video rating task

To validate the task, it was examined whether the mean raw correlation coefficient and the distribution across trials was comparable to the results reported by Zaki and colleagues (2008). They reported a mean raw correlation coefficient of $r = .47$ ($SD = .37$), with values ranging between -.82 and .99. For the current study, the mean raw correlation coefficient was $r = .51$ ($SD = .36$), with values ranging between -.85 and .98. Therefore, there was on average a strong correlation. Importantly, the descriptive statistics were similar to those reported by Zaki and colleagues (2008), validating the current task.

5.5.8 Relationships between task- and trait measures

5.5.8.1 Analysis of individual differences

For the main analysis of interest, correlations were performed between Z-scored individual differences for the task- and trait measures. The analysis revealed that none of the correlations were significant (see Table 5.4, and Figure 5.3 for a further visualisation).

Table 5.4 Pearson correlations between task- and trait measures of the task battery

	DT	SB	RM	HT	MT	VRT	AQ
DT	-	-.04	.04	-.31	-.08	.07[a]	-.09
SB		-	-.02	.13	.15	-.21[a]	.02
RM			-	.12	.08	.35[a]	-.25
HT				-	.10	.18[a]	.02
MT					-	.14[a]	.13
VRT						-	-.07[a]
AQ							-

Note: DT = Director; SB = Sandbox; RM = Reading the Mind in the Eyes; MT = Mouse-tracking; HT = Hinting; VRT = Video rating; AQ = Autism Spectrum Quotient. No correlations statistically significant (*p* < .05, two-tailed, after Bonferroni correction). [a] correlation involved VRT and was therefore restricted to a sample of 58 participants.

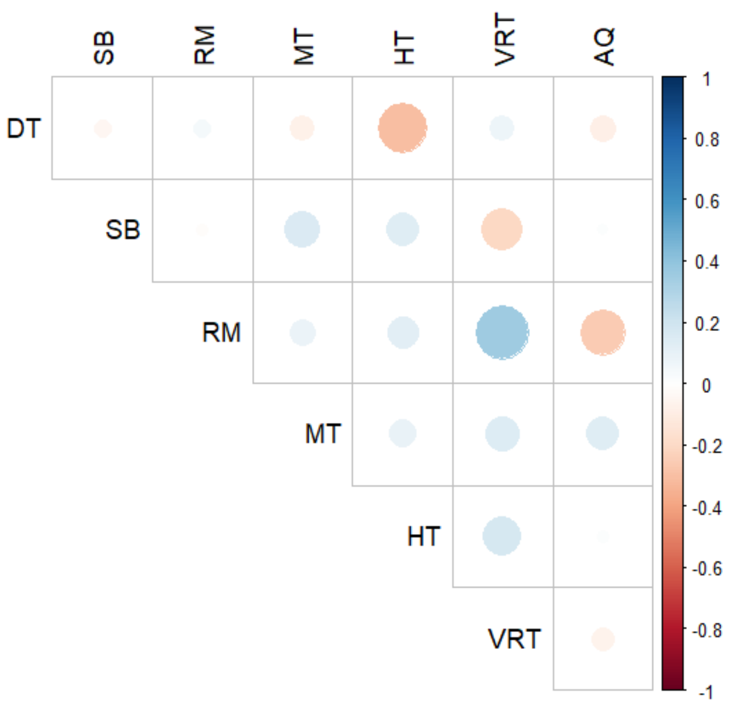


Figure 5.3 Correlogram for correlations between task- and trait measures of the task battery. Circle colour and size indicate correlation direction and magnitude.

5.5.8.2 Effect of demographic variables

The effects of age and gender on individual differences in task- and trait measures were investigated in an exploratory analysis. One participant was excluded from this analysis, as they had not provided their demographic characteristics. For each task- and trait measure, a Pearson correlation was performed with age. No correlation was significant (see Table 5.5). In addition, a Welch's independent-samples t-test was performed to investigate the effect of gender on each measure. This test was employed because of its strength when analysing unequal sample sizes. No t-test revealed a significant effect of gender (see Table 5.6). It is concluded that neither age nor gender had an effect on individual differences in task- and trait measures.

Table 5.5 Correlations between age and task battery measures

Stat.	DT	SB	RM	HT	MT	VRT	AQ
r	.04	-.04	-.04	-.10	-.02	.04	.05
p	.721	.712	.732	.409	.832	.771	.694
p (corr.)	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Note: Stat. = Statistic; DT = Director; SB = Sandbox; RM = Reading the Mind in the Eyes; MT = Mouse-tracking; HT = Hinting; VRT = Video rating; AQ = Autism Spectrum Quotient. The *p* values are corrected with the Bonferroni method. As this value was always 1.000, the uncorrected *p* values are also given as an illustration. N = 77 (57 for VRT).

Table 5.6 Effect of gender on task battery measures

Stat.	DT	SB	RM	HT	MT	VRT	AQ
t	.53	-.38	.93	.52	-.72	.55	.76
p	.600	.712	.367	.609	.478	.598	.459
p (corr.)	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Stat. = Statistic; DT = Director; SB = Sandbox; RM = Reading the Mind in the Eyes; MT = Mouse-tracking; HT = Hinting; VRT = Video rating; AQ = Autism Spectrum Quotient. The *t* values represent the female-male difference. The *p* values are corrected with the Bonferroni method. As this value was always 1.000, the uncorrected *p* values are also given as an illustration. N = 77 (57 for VRT).

5.6 Discussion

In this study it was tested whether task measures that purportedly tap into cognitive empathy processes correlate with each other. For this purpose, a selection of tasks was made that reflected the diversity of available tasks and of judgments that rely on cognitive empathy. Specifically, task measures were chosen in which either ToM, perspective taking or empathic accuracy were central. These component processes have been argued to underlie cognitive empathy (aan het Rot & Hogenelst, 2014). The task battery was complemented by the AQ, to investigate how individual differences in performance on these tasks relate to autistic traits.

In a sample of 78 participants (58 for the analyses involving the video rating task), it was shown that there was no relationship between the performance for any two task measures. In addition, task measures did not show a relationship with autistic traits. Note that the reported correlation magnitude (r = -.25) for the relationship between AQ and Reading the Mind in the Eyes task performance suggested a weak relationship. Although this magnitude lay in the range of previously reported correlations between these two measures (-.10 ≤ r ≤ -.53; Baron-Cohen et al., 2001a, 2015), the relationship did not reach significance in the current study. Previous chapters showed that reward had an effect on some cognitive empathy-related task measures (Chapter 2: Sandbox task; Chapter 3: Dots perspective taking task; Chapter 4: Video rating task), but not on others (Chapter 2: Mouse-tracking task; Chapter 3: Director task). *A priori*, it was possible that the current results would show a relationship between tasks for which an effect of reward was found earlier. This finding would have provided evidence that such task measures share the same underlying construct. However, whereas the presence of a reward effect was task-dependent, the lack of relationship between task measures was global.

A number of possible interpretations arise. First, there is the radical possibility that none of the task measures employed in this task battery reflect cognitive empathy. However, this interpretation is unlikely, as most of the tasks employed here have strong face validity.

Second, one could argue that the changes that were made to the paradigms for the current task battery resulted in the lack of correlation between tasks. However, results showed that the findings of previous studies using these tasks were replicated. The replications were as follows: in the Director task, a shelf was most distracting when it contained an object that matched the participant’s perspective (but not the director’s); in the mouse-tracking task, there was found to be implicit belief tracking of an irrelevant agent; in the Sandbox task, egocentric bias was largest on trials for which there was a conflict between one's own belief and that of someone else. The replication of the bias effect in the Sandbox task was surprising, given the non-replication of this effect in Chapter 2. However, when comparing the versions in the current study and Chapter 2, three possible contributing factors arise. First, for the task in Chapter 2, participants were asked to remember names of reward-conditioned faces during the session. In addition, the word-search distractor task required the participant's full attention. This combination could inadvertently have caused too high a cognitive load to process the Sandbox task attentively. Second, it is possible that the higher number of trials in Chapter 2 resulted in participants' construing experimental and control trials as similar, as participants learned to pay attention to only the crucial elements of scenarios. Third, it is possible that the (marginally significant) interaction between reward and belief condition in Chapter 2 suppressed the main effect of belief. In other words, the introduction of reward to the design may have resulted in the non-replication of the bias effect. In short, a replication of the bias effect may depend on the specifics of the respective Sandbox task design. Nevertheless, the replications in the current study suggest that the chosen tasks measured what they were intended to measure.

A more likely explanation of the current pattern of results is that cognitive empathy can itself be fractionated into independent components. This account was also put forward by Warnell and Redcay (2019), who published a task battery study after the study reported in this chapter had been conducted. Warnell and Redcay (2019) used a task battery to investigate coherence across a wide range of ToM measures in 4-year-old-, 6-year-old-, and school-aged children, and adults. Although the task batteries that they used differed between these sub-samples in terms of difficulty, they were conceptually similar. The study showed that, within the adult sample, there was no relationship in performance between any two tasks. The only significant relationship found across all samples was in the sample of 4-year-olds, between two types of false belief tasks. Warnell and Redcay (2019) therefore provided extensive evidence that there is minimal coherence between ToM measures.

In the current study, two additional operationalisations of cognitive empathy were used, i.e. perspective taking and empathic accuracy. In addition, both explicit and implicit measures of ToM were used (vis-à-vis only explicit measures in Warnell & Redcay, 2019). This multifaceted approach resulted in a selection of tasks that varied in their design and the demands placed on the participants. For example, the Director task relies on verbal understanding of the instructions; working memory for keeping the instruction in mind; visual attention to ignore irrelevant objects; visual perspective taking to see the shelves from the other’s point of view; and possibly other capacities. On the other hand, the Reading the Mind in the Eyes task relies on momentarily inferring an emotion from limited facial information, possibly by visual attention to emotional cues or reasoning about someone’s thoughts and feelings. As these tasks already show, various social and cognitive capacities are invoked to a varying degree depending on the judgment to be made. Therefore, it is difficult to disentangle what the relative contributions of cognitive empathy-related processes are.

Translated to an individual’s cognitive empathy ability, this observation suggests that cognitive empathy judgments can rely on a number of capacities. It is possible that individuals employ a specific mixture of cognitive capacities to arrive at inferences about others' thoughts and feelings. Based on the type of inference that is made, different capacities are employed, and to a varying degree (see Table 5.7 for the hypothesised contributions of cognitive capacities to performance on the tasks used in this study). This would be a logical explanation of why task measures do not correlate with each other. Consider the following hypothetical example (see Figure 5.4): Individual A is better at taking someone else's visual perspective, at using logical reasoning, and has a higher working memory capacity. On the other hand, individual B is better at understanding that others can have different thoughts from one's own, has a higher capacity for visual attention, and higher verbal understanding. In addition, individuals A and B are equally able to integrate different information sources to make mental state inferences, and are similar in other ancillary capacities. This pattern would explain why individual A has higher performance in one task but individual B outperforms or matches individual A in other tasks. For a third individual, this distribution of cognitive capacities would be even different, etc. This line of reasoning suggests that individuals who score higher on one cognitive empathy measure do not necessarily score higher on another.

Table 5.7 Hypothesised contributions of cognitive capacities to task performance

Task	Cognitive capacities
Director	PT, VA, WM, LR, VU
Sandbox	ToM, VA, WM, LR, VU
Reading the Mind in the Eyes	ToM, VA, LR
Hinting	ToM, WM, LR, VU
Mouse-tracking	ToM, VA, WM, LR

Video rating	ToM, II, VA, WM, LR, VU

Note: order of cognitive capacities does not reflect relative degree of contribution. ToM: Theory of Mind; PT: perspective taking; II: information integration; VA: visual attention; WM: working memory; LR: logical reasoning; VU: verbal understanding.

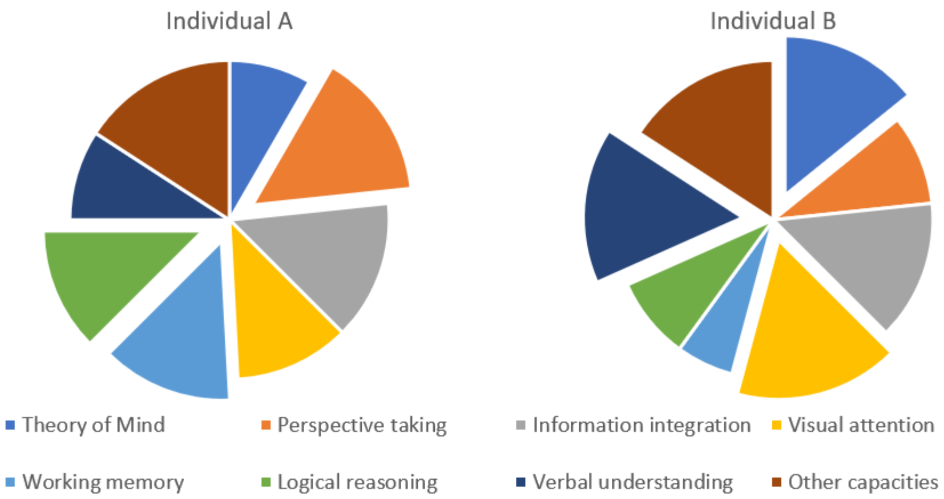


Figure 5.4 Distribution of capacities underlying cognitive empathy for two fictive individuals. Note: superior capacities are indicated by floating segments.

Relatedly, the lack of a relationship between cognitive capacities on the level of the individual could be why the current study found no relationship between task measures and autistic traits. The items of the AQ do not tap directly into the cognitive capacities mentioned above. For this reason, it would be informative to include measures of working memory, verbal understanding and other cognitive capacities in a future task battery, and see to what extent these measures explain performance on the task measures employed here. This way it could also be investigated to what extent these general cognitive capacities underlie the component processes of cognitive empathy.

Alternatively, a reason for not finding any relationship between task measures and autistic traits could be that the current study used a sample that consisted of only typically developed individuals. Rosenblau and colleagues (2015) found that there was no relationship between measures of ToM for typically developed individuals, whereas this *was* the case for participants with a diagnosis of an autism spectrum condition (ASC). It is possible that people with an ASC rely on a limited palette of their cognitive capacities when making cognitive empathy judgments, whereas typically developed individuals can switch between these capacities when switching between types of cognitive empathy inferences. Therefore, it would be informative to investigate the interrelationship between task measures within an ASC sample.

5.7 Conclusion

Taking everything together, the results showed that there was no correlation between putative measures of cognitive empathy. Interestingly, this lack of a relationship also held true for task measures for which it was shown that performance could be influenced by reward. If an interrelationship of task measures had been restricted to these tasks, this could have suggested a general mechanism through which reward influences task performance. However, such a pattern was not observed. More generally, the lack of correlation between tasks used in this battery suggests that they measure different processes related to cognitive empathy, or rely on a specific mixture of capacities that vary across individuals. For this reason, there is no clear-cut linear relationship between these task measures, or indeed, between these measures and the presumed underlying construct of cognitive empathy.

6 General discussion

This chapter discusses the findings of the preceding experimental chapters to address the hypothesis that reward can influence cognitive empathy-related processes. It first summarises the acquired evidence for reward-driven modulation of performance on tasks tapping into Theory of Mind, perspective taking and empathic accuracy. Integrating these findings, the chapter then discusses the circumstances under which an effect of reward may emerge. Additionally, possible reasons are suggested for why individual differences in the effect of reward were not found to be related to the individual's autistic traits. Subsequently, it is discussed what the reported lack of correlation between cognitive empathy-related task measures may mean for the mechanisms underlying reward-driven modulation. Lastly, the findings of this thesis are used to compare the reward-driven modulation of processes related to cognitive versus affective empathy. The chapter concludes by describing the contribution of the thesis for future research directions.

The goal of this thesis was to investigate whether reward influences cognitive empathy-related processes. As delineated in Chapter 1, empathy can be divided into an affective and a cognitive component (Baron-Cohen & Wheelwright, 2004; Decety & Jackson, 2004; de Waal, 2008; Leiberg & Anders, 2006). Whereas previous literature has established that processes related to affective empathy can be modulated by reward (Sims et al., 2012, 2014; Haffey et al., 2013; Trilla Gros et al., 2015; McIntosh, 2006) the counterpart question of whether this modulation applies to cognitive empathy-related processes remains ill-investigated. Given the complementary relationship between cognitive and affective empathy (Zurek & Scheithauer, 2017; Baron-Cohen & Wheelwright, 2004; Decety & Jackson, 2006; Dziobek et al., 2008), it was hypothesised that reward can influence cognitive empathy too. Based on existing literature, a distinction was made between task measures tapping into Theory of Mind, perspective taking, and empathic accuracy (aan het Rot & Hogenelst, 2014). This distinction led to a number of research questions. The first question was whether reward can influence (implicit and explicit) Theory of Mind. The second question was whether reward can influence perspective taking, as measured through altercentric and egocentric interference. The third question was whether reward can influence empathic accuracy. The fourth question was whether there is a relationship between autistic traits and individual differences in reward-driven modulation of performance on cognitive empathy-related tasks. Lastly, as an additional line of investigation, the interrelationship between different task measures of cognitive empathy was tested.

For a summary of the results pertaining to these research questions, see Table 6.1. These questions were addressed throughout the experimental chapters. By manipulating reward value, it was measured whether reward influences performance on tasks tapping into cognitive empathy-related processes.

In Chapter 2, the first question was tested with two adaptations of the traditional Sally-Anne test. Hence, Theory of Mind was operationalised as belief tracking. In one task, belief tracking was analysed by way of the accuracy of a belief ascribed to someone else. This task was used to examine the possible effect of reward on *explicit* belief tracking. In the other task, the task measure was based on the mouse trajectory leading up to a decision exclusively about the participant's own belief. The degree to which this decision was influenced by the irrelevant other's belief was reflected in the mouse trajectory. This task was therefore used to examine the possible effect of reward on *implicit* belief tracking.

In chapter 3, the second question was tested with two commonly used visual perspective taking tasks. In the first, only egocentric interference was measured. In this task, the participant received verbal instructions from a 'director' to move an object. Egocentric interference was then calculated using gaze duration at the object matching the director's perspective versus an object matching the participant's perspective. In the second task, both egocentric and altercentric interference were measured. Specifically, in this task, a visual judgment had to be made from the participant's own perspective or someone else's perspective. By using matching or mismatching perspectives, it could be investigated to what extent one perspective interfered with making a judgment from the other perspective. The first two research questions (Chapter 2 and Chapter 3) were tested using the same evaluative conditioning procedure for the reward manipulation.

The third question was tested in chapter 4. Here, social closeness between friends or strangers was used as a proxy measure for reward value. Empathic accuracy was measured as the correspondence between first-person affective experiences and second-person perceptions thereof, by means of a video rating task.

The fourth question was tested in Chapters 2, 3 and 4. In Chapter 2 and 3, self-reported autistic traits were correlated with reward-driven modulation of the respective measure of Theory of Mind or perspective taking. In Chapter 4, the interaction of autistic traits and social closeness on empathic accuracy was examined.

A secondary aim was to investigate the interrelationship between cognitive empathy-related task measures, for which purpose a battery of tasks was devised. The results of this study were discussed in Chapter 5.

The remainder of this chapter will elaborate on the evidence acquired to answer these research questions. First, the evidence for reward-driven modulation of cognitive empathy-related processes will be discussed (Sections 6.1.1-6.1.3). Second, these findings will be integrated to conjecture under which circumstances reward can have an effect on cognitive empathy (Section 6.1.4). Third, it will be discussed whether the effect of reward on cognitive empathy-related processes is related to autistic traits (Section 6.2). Section 6.3 then covers the results of the task battery study. Lastly, this accumulation of evidence will be used to compare the influence of reward on cognitive empathy with its influence on affective empathy (Section 6.4).

Table 6.1 Overview of findings in this thesis. The Finding column represents whether there was an effect of reward on task performance, and if so, whether reward increased or decreased task performance. It is also reported whether there was a relationship between the effect of reward and autistic traits.

Question	Task	Measure	Finding	Influence of autistic traits
Does reward influence performance on Theory of Mind tasks? (Ch. 2)	Explicit Theory of Mind: Sandbox task	Egocentric bias on false belief trials	Lower egocentric bias for positive (**increased performance**)	no
	Implicit Theory of Mind: mouse-tracking task	Influence of agent's belief on response trajectory	No difference between positive/negative	no
Does reward influence performance on perspective taking tasks? (Ch. 3)	Egocentric interference: Director task	Gaze duration ratio for target vs. distractor object on experimental trials	No gaze ratio difference between positive/negative	no
	Egocentric interference: Dots task	Reaction time for inconsistent vs. consistent *other*-perspective judgments	No RT difference between positive/negative	yes
	Altercentric interference: Dots task	Reaction time for inconsistent vs. consistent *self*-perspective judgments	No RT difference between positive/negative	no
	Perspective selection: Dots task (exploratory result)	Reaction time for self-perspective vs. other-perspective judgments	No RT difference for positive; difference for negative (**increased performance**)	no
Does reward influence performance on empathic accuracy tasks? (Ch. 4)	Emotion inferences: Video rating task	Z-transformed correlation between self- and other-ratings: valence & arousal	Positive correlation between social closeness and empathic accuracy for valence (**increased performance**)	no

Table note: in addition to the research questions listed here, it was tested whether there was an interrelationship between task measures of cognitive empathy. In a battery of six tasks (Director task, Sandbox task, Reading the Mind in the Eyes task, Hinting task, mouse-tracking task, and video rating task), the results showed no correlations between task measures, nor a correlation between these measures and autistic traits.

6.1 The effect of reward on cognitive empathy

Previous studies have consistently demonstrated positive reward-driven modulation of affective empathy-related processes (Sims et al., 2012, 2014; Trilla Gros et al., 2015; Haffey et al., 2013; McIntosh, 2006). In addition, it has been shown that mimicry, an important component process of affective empathy, results in more positive evaluations (e.g. Neufeld & Chakrabarti, 2016; Chartrand & Bargh, 1999; Lakin et al., 2003). Therefore, there is extensive evidence for a bidirectional relationship between affective empathy and positive evaluations.

Similarly, engaging in cognitive empathy-related processes towards a person (i.e. perspective taking) results in a generalised positive evaluation of that person (Erle & Topolinski, 2017; Davis et al., 1996; Vorauer & Sasaki, 2014; Todd et al., 2011; Galinsky & Moskowitz, 2000; Batson et al., 1997). Conversely, however, evidence for the reward-driven modulation of cognitive empathy-related processes has been scarce and mixed (Marsh et al., 2010; Savitsky et al., 2011; Stinson & Ickes, 1992). Therefore, it is unclear whether, or under what circumstances, a positive evaluation (i.e. reward value) results in higher performance on cognitive empathy-related measures.

6.1.1 The effect of reward on belief tracking

The findings reported in Chapter 2 show that reward value, as manipulated with an evaluative conditioning paradigm (e.g. Sims et al., 2012, 2014), influences performance on a task tapping into explicit belief tracking, but not on a different task tapping into implicit belief tracking.

For the latter, an existing mouse-tracking task was used (van der Wel et al., 2014). Implicit belief tracking was measured as the extent to which an irrelevant other's belief influenced the process leading up to the participant's judgment of their own belief. Throughout the task, the participant always had to make judgments based on their own belief, using the computer mouse to respond. The mental decision process leading up to this judgment was assumed to be reflected in the participant's mouse response trajectory. In some cases, the irrelevant other's belief corresponded to a different location than the participant's belief. Therefore, higher implicit belief tracking performance was defined as the degree to which the mouse response veered towards the *other's* belief. The results showed that this performance was not influenced by the reward value associated with the irrelevant other.

In contrast, for the explicit belief tracking task (Sandbox task; Bernstein et al., 2011; Sommerville et al., 2013), performance was reflected in the bias in a participant's decision, rather than in the process leading up to this decision. Specifically, participants tracked a character's belief about the location of an object throughout a scenario. Subsequently, the participant was asked where the character would look for the object (a false belief question), or where the object had been originally placed (a memory control question). The bias of the participant's response was then measured on a continuous scale. For the false belief condition, the bias was higher when the target character was non-rewarding. In other words, reward increased performance in the false belief condition. Previous studies (e.g. Bernstein et al., 2011; Sommerville et al., 2013; Coburn et al., 2015) reported a difference in bias between the false belief and memory control condition. They concluded that judgments in the false belief condition rely on a capacity that is not required for memory control judgments (i.e. belief tracking). However, this difference was not found in the current experiment. It is therefore unclear whether participants relied on belief tracking for their judgments in the false belief condition and what mediated the reward effect. Yet, as there was no effect of reward on performance in the memory control condition, it is more likely that the reward effect was mediated by belief tracking rather than memory. Altogether, the results provide a preliminary suggestion that reward benefits belief tracking.

6.1.2 The effect of reward on perspective taking

Chapter 3 reported findings on the effect of reward on performance in perspective taking tasks. Two visual perspective taking paradigms were taken to measure perspective taking. Reward was manipulated with an evaluative conditioning procedure (e.g. Sims et al., 2012, 2014). The results showed that reward does not influence egocentric interference or altercentric interference. However, the speed with which someone else's perspective is taken relative to one's own does get influenced by reward.

One task, the Director task (Keysar et al., 2000), involved verbal communication. Participants received instructions from a character (the 'director') to move a certain object between shelves. As the director had incomplete visual access to the visual scene, their perspective had to be taken in order to respond correctly. Specifically, on experimental trials, participants had to ignore the distractor object that ostensibly matched the verbal instruction, but was not visible to the director. The ratio of gaze duration at the target versus distractor object was taken as a measure of egocentric interference, and hence of task performance. The reward value associated with the director had no influence on task performance.

The other task, the Dots task (Samson et al., 2010), did not involve verbal communication, yet also relied on judgments in response to visual scenes. The participant had to respond to a number of dots in a visual scene, either from their own perspective, or that of a character in that scene. Crucially, this number was either consistent or inconsistent between the two perspectives. Egocentric interference was measured as the degree to which participants were slowed down in their other-perspective judgment when the perspectives were inconsistent versus consistent. Altercentric interference was measured using the same method, but for self-perspective judgments. The results showed that neither egocentric interference, nor altercentric interference, was influenced by the reward value associated with the character in the visual scene.

Conversely, reward value influenced the speed with which the other- versus the self-perspective was taken in the Dots task. Specifically, when the other person was associated with a negative reward value, the participant took longer in making judgments from the other- versus the self-perspective. This difference was absent when the other person was associated with a positive reward value. Therefore, it can be concluded that the relative difficulty with which someone else's perspective is selected is reduced with reward. This dovetails with the two-step perspective taking process described in Section 3.6: the self- and other-perspective are first automatically computed. Next, one of the two perspectives is intentionally selected for the behavioural judgment (Schwarzkopf et al., 2014; Qureshi et al., 2010). The findings suggest that it is this second step where reward can have an effect. Put differently, reward improves explicit perspective selection when taking someone else's perspective.

6.1.3 The effect of reward on empathic accuracy

In chapter 4, it was investigated whether reward value, quantified as social closeness, influenced empathic accuracy. Participants were paired up as either friends or strangers and rated each other on social closeness. Using an existing video rating task (Zaki et al., 2008; Martin-Key et al., 2017), participants continuously rated videos of themselves on valence and arousal. These same videos were then rated by their pair member. Empathic accuracy was based on the correlations between the self- and other-ratings. It was found that if participants reported themselves to be closer to the other person, their empathic accuracy for valence (but not arousal) towards this other person was higher. Therefore, it can be concluded that reward increases empathic accuracy, at least for inferences about valence. This conceptually replicated the findings reported by Stinson and Ickes (1992), who contrasted friends with strangers in a similar empathic accuracy paradigm.

6.1.4 Task-dependence of reward effects

Although the studies reported above do not suggest a universal positive effect of reward on cognitive empathy-related processes, some observations can be made. If there is an effect of reward, it is positive. In addition, the pattern of the current findings, in conjunction with previous findings, suggests that the effect of reward is task-dependent. This section discusses under what circumstances reward may have an effect, taking into account previous findings in the literature.

The first potential factor to influence whether reward can affect performance on cognitive empathy-related tasks is the importance of the other person for the task at hand (i.e. do they need to be explicitly processed in order for the participant to respond accurately?). For instance, in the mouse-tracking task (Chapter 2), the target person is irrelevant to the task at hand, i.e. the participant is instructed to respond based solely on their own belief. The results of the study employing this task showed no effect of reward. In contrast, the results for the Sandbox task study (Chapter 2) did show an effect. In this task, the participant is presented with two characters. To respond correctly to the test question in the Sandbox task, the participant should keep track of one of the two characters' belief. In order to do this, the participant has to be able to distinguish which character is the one that places the object in the sandbox container, and which one moves it. Similarly, two previous studies employing the action identification task showed an effect of reward on attributing intentions to actions (Kozak et al., 2006; Marsh et al., 2010). In this task, the participant is actively encouraged to think about the other person performing these actions. In the video rating paradigm developed by Stinson and Ickes (1992) and employed in Chapter 4, the other person is the very source of the verbal and nonverbal information required to make inferences. Lastly, in the Dots task experiment covered in Chapter 3, it is necessary that the participant take the other person's visuospatial perspective in half of the trials. Selecting the correct perspective to make a decision from is crucial to performance, and this selection is where the reward effect takes place. For all of these four tasks (Sandbox; action identification; video rating; Dots), reward-driven modulation of performance was reported. Therefore, it can be concluded that a facilitatory effect of reward is observed only in tasks in which the other person is central to the task at hand.

However, experiments employing the Director task provide evidence against this interpretation. It could be argued that the director is central to the task at hand, as they are the person giving instructions to the participant. Yet no effect of reward was found in the Director task in Chapter 3. Moreover, Savitsky and colleagues (2011) found that people were better at taking someone's perspective when that person was a stranger than when they were a friend. Although the friends versus strangers design is an indirect manipulation of reward, the finding suggests an impeding, rather than facilitatory, effect of reward on perspective taking. An interesting difference between the task used by Savitsky and colleagues (2011) and the Director task in Chapter 3 (and the majority of tasks discussed in this thesis) is that Savitsky and colleagues (2011) employed physically present directors, whereas the task in Chapter 3 was computerised. It is possible that, when the director is shown behind the shelves on the computer screen, the participant processes the director's instructions without paying attention to the director themselves, precluding an effect of the reward associated with the director. This possibility could be investigated by including a no-director condition in the task, which can be used to reveal uniquely social effects in the director condition, i.e. whether mentalising of directors takes place. If participants show a similar (lack of reward-driven modulation of) performance in the director and no-director condition, this would suggest that the computerised task in Chapter 3 is not conducive to paying attention to the director and hence mentalising. Alternatively, it can be argued that a real friend has more reward value than an evaluatively conditioned face. The Director task may not be sensitive to reward-driven modulation with this type of experimental manipulation. Regardless of the methodological differences between these task versions, the case of the Director task suggests that the need to process the other person does not categorically result in reward-driven facilitation of performance on cognitive empathy-related tasks.

Alternatively, the effect of reward may only emerge when the task measure for the cognitive empathy-related task is explicit. The Sandbox task, the video rating task, the action identification paradigm and the Dots task all use an explicit response measure to quantify performance (e.g. reaction time and accuracy derived from binary or continuous decisions). On the other hand, the Director task experiment in the current thesis only employed eye-gaze as a measure, as the design did not lend itself to an investigation of reaction times. As eye-gaze behaviour is commonly taken as an implicit response measure (e.g. Schneider et al., 2012a; Onishi & Baillargeon, 2005), it is possible that the effect of reward could not be observed here. In addition, the mouse-tracking experiment reported here, which showed no effect of reward, used an implicit measure based on the response trajectory. In short, it is possible that the effect of reward relies on the explicitness of the employed measure; the importance of the other person for the task at hand; or a combination of the two.

The emerging pattern is that the presence and direction of reward-driven modulation of performance can differ based on the way cognitive empathy-related processes are measured. This observation dovetails with Warnell and

Redcay's proposal (2019) that cognitive empathy tasks (in their example: Theory of Mind tasks) in reality tap into various cognitive capacities, such as working memory, language understanding and executive function. It is possible that there is a relationship between reward and some, but not other, cognitive capacities (see also Section 6.3).

Another possibility is that the presence and direction of a reward effect relies on the extent that verbal understanding plays a role in the task. It has been argued that being close to someone can result in a perceived self-other overlap (Aron, Aron, Tudor, & Nelson, 1991). It has subsequently been shown that this self-other overlap can impede perspective taking (Savitsky et al., 2011) and that inducing self-other differentiation can facilitate perspective taking (Sassenrath, Sassenberg, & Scholl, 2014). Notably, both these studies involved experiments in which an ambiguous verbal message was communicated to the participant. A perceived self-other overlap may result in the assumption that the meaning intended by the other person matches the participant's perception of the world. This assumption may be made especially when the communicator is a friend, rather than a computerised, conditioned avatar. In short, these studies provide evidence that a high reward value can impede perspective taking in the case of semantic ambiguity of a communicated message.

In contrast, when the task requires judgments that do not rely on interpreting ambiguous verbal messages, the effect of reward is generally facilitatory. This facilitation was shown in several studies involving mental state attribution, i.e. the attribution of intentions (Kozak et al., 2006; Marsh et al., 2010) and empathic accuracy (Stinson & Ickes, 1992). This interpretation was further corroborated by the findings reported in this thesis for the Sandbox task (Chapter

2) and the video rating task (Chapter 4). One possible mechanism for this facilitatory effect of reward is that, when someone is liked, there is increased depth of processing, possibly due to more attention allocated to the other person. This conjecture is further supported by the current findings in the Dots task (Chapter 3), where it was shown that the selection of the other person's perspective (relative to one's own) was sped up when that person was liked. It is possible that, when someone else is liked, more cognitive resources are employed to try to understand their point of view.

In short, the combination of the current data and previous findings suggests that the effect of reward on performance on cognitive empathy-related measures depends on the task at hand. Although in some cases a higher reward value can result in egocentricity, it generally leads to more processing of the other person's mental state. Returning to the model formulated in Chapter 1 (Figure 1.1), the current findings suggest that performance on tasks that tap into Theory of Mind, perspective taking, and empathic accuracy can be improved by reward. In other words, it can be concluded that the effect of reward can take place for different component processes underlying empathy-related inferences.

6.2 Individual differences in the effect of reward

In addition to examining possible reward-driven modulation of performance on cognitive empathy-related tasks, another primary aim of this thesis was to investigate whether this modulation is related to autistic traits. It has previously been reported that autistic traits can alter reward-driven modulation of affective empathy-related processes (Sims et al., 2012, 2014; Trilla Gros et al., 2015; Haffey et al., 2013). For cognitive empathy-related processes, this relationship has only been found for the attribution of intentions to actions (Marsh et al., 2010). However, this relationship has not been examined for belief tracking, a more direct measure of Theory of Mind. In addition, this relationship between autistic traits and reward-driven modulation has not been examined yet for perspective taking or empathic accuracy.

For each of the experiments reported in this thesis that used a reward manipulation, it was subsequently measured whether autistic traits altered the reward-driven modulation of cognitive empathy performance. Note that the primary effect of reward was not always found (i.e. for implicit belief tracking; altercentric interference in perspective taking; egocentric interference in perspective taking). Individual differences in reward-driven modulation of task performance were only related to autistic traits once. This was the case for the effect of reward on egocentric interference in the Dots task (Section 3.5). This finding suggests that one's own perspective interferes less with taking the other's perspective when that person is liked, but only for perspective-takers low in autistic traits. However, as the effect of reward was not significant, it cannot be concluded that this relationship between autistic traits and the reward effect magnitude is meaningful. In contrast, when reward-driven modulation of task performance was observed, individual differences in this modulation were on no occasion related to autistic traits.

The absence of this relationship suggests that people high and low in autistic traits show no observable difference in reward-driven modulation of cognitive empathy-related processes tested in this set of studies. Note that this finding is inconsistent with the generally negative relationship between autistic traits and the effect of reward on *affective* empathy-related processes. It has been shown that automatic facial mimicry, an important affective empathy-related process, already occurs rapidly (i.e. at 17-40 ms) when a participant is presented with a face (Sonnby-Bergström, 2002). It is therefore possible that the effect of reward on facial mimicry is also rapid and bypasses conscious reasoning. This rapid response, or reflex even, is not facilitated by reward in those high in autistic traits (see Sims et al., 2012, 2014). On the other hand, for those processes for which a reward effect was reported in this thesis, the time scale was longer and allowed for conscious reasoning. It is therefore possible that those low and high in autistic traits alike benefit from reward, as the reward value can be consciously represented and can influence depth of processing of others' mental states.

6.3 The interrelationship between cognitive empathy measures

The aim of the study covered in Chapter 5 was to see whether (1) task measures purportedly tapping into different components of cognitive empathy (Theory of Mind, perspective taking, and empathic accuracy) are related to each other and (2) whether the answer to this question could elucidate the found pattern of reward-driven modulation.

The correlational analysis showed that there was no relationship between any two task measures. Note that these task measures were all free of a reward manipulation, and the sample consisted of participants who had not yet participated in the experiments involving a reward manipulation. A possible interpretation of this general absence of a relationship is that different cognitive empathy-related tasks all tap into a different mixture of underlying cognitive capacities (cf. Warnell & Recay, 2019). Based on the selection of tasks in this thesis (and Chapter 5 in particular), it is clear that they are characterised by a diversity in task demands. It has been suggested that it is difficult to disentangle different components of cognitive empathy (Preston & de Waal, 2002). By the same token, it is difficult to disentangle the contributions of different cognitive capacities to performance on individual tasks tapping into cognitive empathy. Therefore, it is likely that various cognitive empathy tasks are not directly related because these cognitive capacities vary between individuals.

As no relationship was found between cognitive empathy measures, this investigation did not identify a possible mechanism through which reward can influence cognitive empathy. Hence, it is likely that reward-driven modulation of cognitive empathy occurs through its effect on underlying cognitive capacities.

6.4 A comparison between cognitive and affective empathy

Based on the above, it can be concluded that reward-driven modulation takes place for both cognitive and affective empathy components. However, whereas the direction of this effect is generally positive for affective empathy, it is task-dependent for its cognitive counterpart. It is possible that the effect of reward on affective versus cognitive empathy may rely not only on different behavioural, but also different neural mechanisms.

For affective empathy, the effect of reward may rely on a direct functional connection between brain regions involved in reward processing and mimicry (Sims et al., 2014). When an emotional expression (or a bodily expression in general) is observed, this expression is neurally and behaviourally matched by the observer (e.g. Field et al., 1982; Meltzoff & Moore, 1977; Stern, 1977; Sims et al., 2012, 2014; Bastiaansen et al., 2009). The perception of another person directly results in action (behaviour) on the part of the observer (cf. the perception-action link; Preston & de Waal, 2002).

On the other hand, for cognitive empathy, the effect of reward may rely more on whether the observer engages in conscious reasoning about the other person's mental states. During such reasoning, a representation of the other person is held in working memory. That person's reward value subsequently determines the depth of processing of that person and their mental states. Cognitive empathy arguably depends on more varied underlying capacities than affective empathy does, as was suggested by the task-dependence of the reward effect (Section 6.1.4). Therefore, it is likely that there is no such straightforward functional connection between regions involved in cognitive empathy and the reward system as there is for affective empathy.

This neural account is also a possible explanation of why autistic traits influence the reward-driven modulation of affective empathy, but not of cognitive empathy. More precisely, it has been shown that the functional connectivity between frontal regions involved in affective empathy-related processes and striatal regions involved in reward processing is reduced for people high in autistic traits (Sims et al., 2014). For cognitive empathy-related processes, autistic traits may not play a modulatory role if there is no such direct functional connectivity with reward processing. Rather, the extent of the reward effect may depend more on individual differences in general cognitive functioning.

6.5 Potential limitations and future directions

First, for all experiments covered in this thesis, it was attempted to replicate previous findings before examining a possible reward effect. Generally, results reported by previous studies using the same paradigms were replicated. An exception, however, was the Sandbox task. In Chapter 2, there was no significant difference between bias for false belief versus memory control trials. For this reason, the reward effect was subsequently investigated for false belief trials only, as, theoretically, this is where Theory of Mind ability is most apparent. However, note that the false belief versus memory control difference has previously been taken as a measure of belief tracking performance, as this difference measure controls for the contribution of processes shared by both conditions (e.g. Coburn et al., 2015). Therefore, the reward effect observed in Chapter 2 should be interpreted with caution, as it cannot be excluded that the effect was due to some other process than belief tracking.

On the other hand, the false belief versus memory control difference was replicated in Chapter 5. As discussed in Chapter 5, it is possible that the version used in Chapter 2 put too many demands on the participant, making a direct investigation of belief tracking difficult. For this reason, when adapting the Sandbox task for future studies, distracting features of the design should be kept to a minimum.

In the majority of experiments reported here, the manipulation of reward was central. For the tasks using the evaluative conditioning procedure, there was always one person associated with positive reward, and one with negative reward. Due to this binary manipulation, it is unclear whether any reward effect is due to facilitation by positive reward, or interference from negative reward. In order to give a more definitive answer about the effect of reward, studies should include a neutral condition too.

Relatedly, in the experiments using the conditioning procedure, it was not validated whether the manipulation was successful. Critics could suggest that the absence of a reward effect can be explained by unsuccessful reward conditioning. Supporting this suggestion, Dawson, Rissling, Schell, and Wilcox (2007) showed that people need to be aware of the evaluative conditioning in order for it to have an effect. On the other hand, several papers have already shown that the manipulation used in this thesis resulted in reward-driven modulation of behaviour (e.g. Sims et al., 2012, 2014). Panasiti and colleagues (2016) also showed that this manipulation influenced performance on the Implicit Association Test, which is a measure of the strength of conditioning (Greenwald, McGhee, & Schwartz, 1998). In addition, Neufeld and Chakrabarti (2016) showed that (not) noticing the manipulation in the conditioning procedure had no effect on behavioural proxy measures of the success of conditioning. Therefore, it can be surmised that the evaluative conditioning procedure was also successful in the current experiments, even when no reward effect was observed. Nevertheless, future studies could use more explicit measures of reward value (e.g. vignettes; see Kozak et al., 2006; Marsh et al., 2010) or directly test the success of conditioning with self-reported liking scales (cf. social closeness scale).

A last point related to reward is that the question remains to what extent reward as manipulated through (1) an evaluative conditioning procedure and (2) social closeness are reflective of the same underlying construct. As was mentioned in Chapter 4, being close to someone implies a shared, and often long, history. On the other hand, reward value as manipulated through the evaluative conditioning procedure is a novel and rapidly acquired association with the other person. Stinson and Ickes (1992) argued that empathic accuracy is higher for friends than for strangers because of their shared knowledge, or shared history. However, a more recent study found that shared knowledge does not predict the correct interpretation of ambiguous messages communicated by friends versus strangers (Pollmann & Krahmer, 2018). Instead, the authors postulated that friends are more motivated than strangers to engage in the reasoning about conveyed meanings. This explanation suggests that the effect of both social closeness and conditioned reward value on cognitive empathy performance is mediated by motivation. However, to be able to compare between various positive evaluations, including for example similarity and familiarity, more research should be done on how they directly relate to each other.

Lastly, the paradigms employed in Chapter 3 were level-1 visual perspective taking tasks, tapping into the judgment of whether someone else could see an object, as opposed to how this object looked to this other person. In their review, Pearson and colleagues (2013) showed that level-1 perspective taking is mostly intact in autism, whereas level-2 visual perspective taking is impaired. Mirroring the Dots task, recent paradigms have been developed that can be used to quantify both altercentric and egocentric interference in level-2 visual perspective taking (Elekes et al., 2016; Ward et al., 2019). Using such paradigms, it could be examined whether there is also reward-driven modulation of level-2 visual perspective taking, and whether individual differences in this modulation are related to autistic traits. As the majority of the participants in the current studies did not reach the clinical cut-off score for the AQ, it is recommended that such a study include a larger range of autistic traits than is reported here. Alternatively, such a study could be conducted with a sample of people with a diagnosis of high- or low-functioning ASC.

6.6 Conclusions

In the current thesis, it was investigated whether reward influences cognitive empathy-related processes. This question was addressed using a variety of task measures in a laboratory setting, tapping into Theory of Mind, perspective taking, and empathic accuracy. The results generally demonstrated subtle egocentric biases when adult participants had to reason about someone else's mental states. More importantly, results show that reward-driven modulation was task-dependent, i.e. reward had a positive effect on task performance, but not for all employed tasks. In addition, it was found that individual differences in this reward-driven modulation were in no case related to autistic traits. These observations suggest that social motivation plays a role in some cognitive empathy-related processes, and not differentially so for people high versus low in autistic traits. However, as no clear pattern of reward-driven modulation was found across the task measures, and as performance on these measures was shown not to be interrelated, the findings do not elucidate possible mechanisms underlying the effect of reward on cognitive empathy. For this reason, more research is needed to clarify under which experimental or quotidian circumstances reward can influence cognitive empathy. Nevertheless, whereas previous studies showed a relatively generalised positive effect of reward on affective empathy-related processes, the current results suggest that reward can also, in some cases, have a positive influence on cognitive empathy.

References

Adolphs, R. (2002). Recognizing emotion from facial expressions: Psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews, 1*(1), 21-62. DOI:10.1177/1534582302001001003

Altschuler, E.L., Vankov, A., Hubbard, E.M., Roberts, E., Ramachandran, V.S., Pineda, J.A. (2000, November). *Mu wave blocking by observer of movement and its possible use as a tool to study theory of other minds*. Poster session presented at the 30th Annual Meeting of the Society for Neuroscience, New Orleans, LA.

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

Apperly, I. A. (2011). *Mindreaders: The cognitive basis of "theory of mind"*. Hove, UK: Psychology Press.

Apperly, I. A. (2012). What is "theory of mind"? Concepts, cognitive processes and individual differences. *Quarterly Journal of Experimental Psychology, 65*(5), 825-839. DOI:10.1080/17470218.2012.676055

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review, 116*(4), 953-970. DOI:10.1037/a0016923

Apperly, I. A., Carroll, D. J., Samson, D., Humphreys, G. W., Qureshi, A., & Moffitt, G. (2010). Why are there limits on theory of mind use? Evidence from adults' ability to follow instructions from an ignorant speaker. *The Quarterly Journal of Experimental Psychology, 63*(6).

DOI:10.1080/17470210903281582

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. *Journal of Neuroscience, 31*(40), 14243-14249. DOI:10.1523/jneurosci.0963-11.2011

Aron, A., Aron, E. N., Tudor, M., & Nelson, G. (1991). Close relationship as including other in the self. *Journal of Personality and Social Psychology, 60*(2), 241-253. DOI:10.1037/0022-3514.60.2.241

Ashwood, K. L., Gillan, N., Horder, J., Hayward, H., Woodhouse, E., McEwen, F. S., . . . Murphy, D. G. (2016). Predicting the diagnosis of autism in adults using the Autism-Spectrum Quotient (AQ) questionnaire. *Psychological Medicine, 46*(12), 2595-2604. DOI:10.1017/s0033291716001082

Astington, J. W., & Baird, J. A. (2005). *Why language matters for theory of mind*. Oxford, UK: Oxford University Press. DOI: 10.1093/acprof:oso/9780195159912.001.0001

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. DOI:10.1016/j.paid.2004.04.022

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.

Bardi, L., Desmet, C., Nijhof, A., Wiersema, J. R., & Brass, M. (2016). Brain activation for spontaneous and explicit false belief tasks overlaps: New fMRI evidence on belief processing and violation of expectation. *Social Cognitive and Affective Neuroscience*. DOI:10.1093/scan/nsw143

Bardi, L., Six, P., & Brass, M. (2017). Repetitive TMS of the temporo-parietal junction disrupts participant’s expectations in a spontaneous Theory of Mind task. *Social Cognitive and Affective Neuroscience, 12*(11), 1775-1782. DOI:10.1093/scan/nsx109

Baron-Cohen, S., Bowen, D. C., Holt, R. J., Allison, C., Auyeung, B., Lombardo, M. V., . . . Lai, M.-C. (2015). The “Reading the Mind in the Eyes” Test: Complete absence of typical sex difference in ~400 men and women with autism. *PLOS ONE, 10*(8), e0136521. DOI:10.1371/journal.pone.0136521

Baron-Cohen, S., Golan, O., Wheelwright, S., & Hill, J. J. (2004). *Mind reading: The interactive guide to emotions*. London, UK: Jessica Kingsley Limited.

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

Baron-Cohen, S., & Wheelwright, S. (2004). The Empathy Quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders, 34*(2), 163-175. DOI:10.1023/b:jadd.0000022607.19833.00

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001a). The “Reading the Mind in the Eyes” Test Revised Version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, 42*(2), 241-251. DOI:10.1111/1469-7610.00715

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E.

(2001b). The Autism-Spectrum Quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*(1), 5-17. DOI:10.1023/a:1005653411471

Bartz, J. A., Zaki, J., Bolger, N., Hollander, E., Ludwig, N. N., Kolevzon, A., & Ochsner, K. N. (2010). Oxytocin selectively improves empathic accuracy. *Psychological Science, 21*(10), 1426-1428. DOI:10.1177/0956797610383439

Bastiaansen, J. A., Thioux, M., & Keysers, C. (2009). Evidence for mirror systems in emotions. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1528), 2391-2404. DOI:10.1098/rstb.2009.0058

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48. DOI:10.18637/jss.v067.i01

Batson, C. D. (2009). These things called empathy: Eight related but distinct phenomena. In J. Decety & W. Ickes (Eds.), *The Social Neuroscience of Empathy*. Cambridge, MA: MIT Press. DOI:10.7551/mitpress/9780262012973.003.0002

Batson, C. D., Polycarpou, M. P., Harmon-Jones, E., Imhoff, H. J., Mitchener, E. C., Bednar, L. L., . . . Highberger, L. (1997). Empathy and attitudes: Can feeling for a member of a stigmatized group improve feelings toward the group? *Journal of Personality and Social Psychology, 72*(1), 105-118. DOI:10.1037/0022-3514.72.1.105

Bavelas, J. B., Black, A., Lemery, C. R., & Mullett, J. (1986). "I show how you feel": Motor mimicry as a communicative act. *Journal of Personality and*

Social Psychology, 50(2), 322-329. DOI:10.1037//0022-3514.50.2.322

Beaumont, R. B., & Sofronoff, K. (2008). A new computerised advanced Theory of Mind measure for children with Asperger syndrome: The ATOMIC. *Journal of Autism and Developmental Disorders, 38*(2), 249-260. DOI:10.1007/s10803-007-0384-2

Begeer, S., Bernstein, D. M., Aβfalg, Azdad, H., Glasbergen, T., Wierda, M., & Koot, H. M. (2016). Equal egocentric bias in school-aged children with and without autism spectrum disorders. *Journal of Experimental Child Psychology, 144*, 15-26. DOI:10.1016/j.jecp.2015.10.018

Begeer, S., Bernstein, D. M., van Wijhe, J., Scheeren, A. M., & Koot, H. M. (2012). A continuous false belief task reveals egocentric biases in children and adolescents with Autism Spectrum Disorders. *Autism, 16*(4), 357-366. DOI:10.1177/1362361311434545

Begeer, S., Malle, B. F., Nieuwland, M. S., & Keysar, B. (2010). Using Theory of Mind to represent and take part in social interactions: Comparing individuals with high-functioning autism and typically developing controls. *European Journal of Developmental Psychology, 7*(1), 104-122. DOI:10.1080/17405620903024263

Behrens, T. E., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. (2008). Associative learning of social value. *Nature, 456*(7219), 245-249. DOI:10.1038/nature07538

Bernhardt, B. C., & Singer, T. (2012). The neural basis of empathy. *Annual Review of Neuroscience, 35*(1), 1-23. DOI:10.1146/annurev-neuro-062111-150536

Bernstein, D. M., Coolin, A., Fischer, A. L., Thornton, W. L., & Sommerville, J.

A. (2017). False-belief reasoning from 3 to 92 years of age. *PLOS ONE, 12*(9). DOI:10.1371/journal.pone.0185345

Bernstein, D. M., Thornton, W. L., & Sommerville, J. A. (2011). Theory of Mind through the ages: Older and middle-aged adults exhibit more errors than do younger adults on a continuous false belief task. *Experimental Aging Research, 37*(5), 481-502. DOI: 10.1080/0361073X.2011.619466

Blair, R. (2005). Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations. *Consciousness and Cognition, 14*(4), 698-718. DOI:10.1016/j.concog.2005.06.004

Blair, R. (2008). The amygdala and ventromedial prefrontal cortex: Functional contributions and dysfunction in psychopathy. *Philosophical Transactions of the Royal Society B: Biological Sciences, 363*(1503), 2557-2565. DOI:10.1098/rstb.2008.0027

Blanke, E. S., Rauers, A., & Riediger, M. (2015). Nice to meet you—adult age differences in empathic accuracy for strangers. *Psychology and Aging, 30*(1), 149-159. DOI:10.1037/a0038459

Blanke, E. S., & Riediger, M. (2019). Reading thoughts and feelings in other people: Empathic accuracy across adulthood. *Progress in Brain Research, 247*, 305-327. DOI: 10.1016/bs.pbr.2019.02.002

Boele, S., van der Graaff, J., de Wied, M., van der Valk, I. E., Crocetti, E., & Branje, S. (2019). Linking parent–child and peer relationship quality to empathy in adolescence: A multilevel meta-analysis. *Journal of Youth and Adolescence, 48*(6), 1033-1055. DOI:10.1007/s10964-019-00993-5

Bollen, K. A. & Jackman, R. W. (1990) Regression diagnostics: An expository treatment of outliers and influential cases, in J. Fox and J. S. Long (eds.); *Modern Methods of Data Analysis*, 257-91. Newbury Park, California: Sage Publishing.

Braadbaart, L., Williams, J. H., & Waiter, G. D. (2013). Do mirror neuron areas mediate mu rhythm suppression during imitation and action observation? *International Journal of Psychophysiology, 89*(1), 99-105. DOI:10.1016/j.ijpsycho.2013.05.019

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433-436. DOI:10.1163/156856897X00357

Brent, E., Rios, P., Happé, F., & Charman, T. (2004). Performance of children with Autism Spectrum Disorder on advanced Theory of Mind tasks. *Autism, 8*(3), 283-299. DOI:10.1177/1362361304045217

Brüne, M. (2003). Theory of mind and the role of IQ in chronic disorganized schizophrenia. *Schizophrenia Research, 60*(1), 57-64. DOI:10.1016/s0920-9964(02)00162-7

Bukowski, H., & Samson, D. (2017). New insights into the inter-individual variability in perspective taking. *Vision, 1*(1), 8. DOI:10.3390/vision1010008

Burnside, K., Wright, K., & Poulin-Dubois, D. (2017). Social motivation and implicit theory of mind in children with autism spectrum disorder. *Autism Research, 10*(11), 1834-1844. DOI:10.1002/aur.1836

Buttelmann, F., & Buttelmann, D. (2017). The influence of a bystander agent's beliefs on children's and adults' decision-making process. *Journal of Experimental Child Psychology, 153*, 126-139. DOI:10.1016/j.jecp.2016.09.006

Bzdok, D., Langner, R., Schilbach, L., Engemann, D. A., Laird, A. R., Fox, P. T., & Eickhoff, S. B. (2013). Segregation of the human medial prefrontal cortex in social cognition. *Frontiers in Human Neuroscience, 7*. DOI:10.3389/fnhum.2013.00232

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

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, 364*(1528), 2369-2380. DOI:10.1098/rstb.2009.0048

Charman, T., & Baird, G. (2002). Practitioner review: Diagnosis of autism spectrum disorder in 2- and 3-year-old children. *Journal of Child Psychology and Psychiatry, 43*(3), 289-305. DOI:10.1111/1469-7610.00022

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. DOI:10.1037//0022-3514.76.6.893

Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin, 25*(2), 215-224. DOI:10.1177/0146167299025002007

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. DOI:10.1016/j.tics.2012.02.007

Clements, C. C., Zoltowski, A. R., Yankowitz, L. D., Yerys, B. E., Schultz, R. T., & Herrington, J. D. (2018). Evaluation of the social motivation hypothesis of autism: A systematic review and meta-analysis. *JAMA Psychiatry, 75*(8), 797. DOI:10.1001/jamapsychiatry.2018.1100

Coburn, P. I., Bernstein, D. M., & Begeer, S. (2015). A new paper and pencil task reveals adult false belief reasoning bias. *Psychological Research, 79*(5), 739-749. DOI:10.1007/s00426-014-0606-0

Cole, G. G., Atkinson, M., Le, A. T. D., & Smith, D. T. (2016). Do humans spontaneously take the perspective of others? *Acta Psychologica, 164*, 165-168. DOI:10.1016/j.actpsy.2016.01.007

Colibazzi, T., Posner, J., Wang, Z., Gorman, D., Gerber, A., Yu, S., . . . Peterson, B. S. (2010). Neural systems subserving valence and arousal during the experience of induced emotions. *Emotion, 10*(3), 377-389. DOI:10.1037/a0018484

Colombi, C., Liebal, K., Tomasello, M., Young, G., Warneken, F., & Rogers, S. J. (2009). Examining correlates of cooperation in autism. *Autism, 13*(2), 143-163. DOI:10.1177/1362361308098514

Corcoran, R., Mercer, G., & Frith, C. D. (1995). Schizophrenia, symptomatology and social inference: Investigating “theory of mind” in people with schizophrenia. *Schizophrenia Research, 17*(1), 5-13. DOI:10.1016/0920-9964(95)00024-g

Dapretto, M., Davies, M. S., Pfeifer, J. H., Scott, A. A., Sigman, M., Bookheimer, S. Y., & Iacoboni, M. (2006). Understanding emotions in others: Mirror neuron dysfunction in children with autism spectrum disorders. *Nature Neuroscience, 9*(1), 28-30. DOI:10.1038/nn1611

Davis, M. H. (1994). *Empathy: A social psychological approach*. Boulder, CO: Westview Press.

Davis, M. H., Conklin, L., Smith, A., & Luce, C. (1996). Effect of perspective taking on the cognitive representation of persons: A merging of self and other. *Journal of Personality and Social Psychology, 70*(4), 713-726. DOI:10.1037/0022-3514.70.4.713

Davis, H. L., & Pratt, C. (1995). The development of children's theory of mind: The working memory explanation. *Australian Journal of Psychology, 47*(1), 25-31. DOI:10.1080/00049539508258765

Dawson, M. E., Rissling, A. J., Schell, A. M., & Wilcox, R. (2007). Under what conditions can human affective conditioning occur without contingency awareness? Test of the evaluative conditioning paradigm. *Emotion, 7*(4), 755-766. DOI:10.1037/1528-3542.7.4.755

Dawson, G., Webb, S. J., & Mcpartland, J. (2005). Understanding the nature of face processing impairment in autism: Insights from behavioral and electrophysiological studies. *Developmental Neuropsychology, 27*(3), 403-424. DOI:10.1207/s15326942dn2703_6

Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews, 3*(2), 71-100. DOI:10.1177/1534582304267187

Decety, J., & Jackson, P. L. (2006). A social-neuroscience perspective on empathy. *Current Directions in Psychological Science, 15*(2), 54-58. DOI:10.1111/j.0963-7214.2006.00406.x

Demurie, E., De Corel, M., & Roeyers, H. (2011). Empathic accuracy in adolescents with autism spectrum disorders and adolescents with attention-deficit/hyperactivity disorder. *Research in Autism Spectrum Disorders, 5*(1), 126-134. DOI:10.1016/j.rasd.2010.03.002

Deruelle, C., Rondan, C., Gepner, B., & Tardif, C. (2004). Spatial frequency and face processing in children with autism and Asperger syndrome. *Journal of Autism and Developmental Disorders, 34*(2), 199-210. DOI:10.1023/b:jadd.0000022610.09668.4c

Devine, R. T., & Hughes, C. (2013). Silent films and strange stories: Theory of Mind, gender, and social experiences in middle childhood. *Child Development, 84*(3), 989-1003. DOI:10.1111/cdev.12017

Devlin, H. C., Zaki, J., Ong, D. C., & Gruber, J. (2014). Not as good as you think? Trait positive emotion is associated with increased self-reported empathy but decreased empathic performance. *PLOS ONE, 9*(10), e110470. DOI: 10.1371/journal.pone.011047

Dichter, G. S., Richey, J. A., Rittenberg, A. M., Sabatino, A., & Bodfish, J. W. (2012). Reward circuitry function in autism during face anticipation and outcomes. *Journal of Autism and Developmental Disorders, 42*(2), 147-160. DOI:10.1007/s10803-011-1221-1

Dimberg, U., Andréasson, P., & Thunberg, M. (2011). Emotional empathy and facial reactions to facial expressions. *Journal of Psychophysiology, 25*(1), 26-31. DOI:10.1027/0269-8803/a000029

Dinstein, I., Thomas, C., Humphreys, K., Minshew, N., Behrmann, M., & Heeger, D. J. (2010). Normal movement selectivity in autism. *Neuron, 66*(3), 461-469. DOI:10.1016/j.neuron.2010.03.034

Dumontheil, I., Apperly, I. A., & Blakemore, S.-J. (2010a). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science, 13*(2), 331-338. DOI:10.1111/j.1467-7687.2009.00888.x

Dumontheil, I., Küster, O., Apperly, I. A., & Blakemore, S.-J. (2010b). Taking perspective into account in a communicative task. *NeuroImage, 52*(4), 1574-1583. DOI:10.1016/j.neuroimage.2010.05.056

Dziobek, I., Rogers, K., Fleck, S., Bahnemann, M., Heekeren, H. R., Wolf, O. T., & Convit, A. (2008). Dissociation of cognitive and emotional empathy in adults with Asperger syndrome using the Multifaceted Empathy Test (MET). *Journal of Autism and Developmental Disorders, 38*(3), 464-473. DOI:10.1007/s10803-007-0486-x

Edwards, K., & Low, J. (2017). Reaction time profiles of adults' action prediction reveal two mindreading systems. *Cognition, 160*, 1-16. DOI:10.1016/j.cognition.2016.12.004

Elekes, F., Varga, M., & Király, I. (2016). Evidence for spontaneous level-2 perspective taking in adults. *Consciousness and Cognition, 41*, 93-103. DOI:10.1016/j.concog.2016.02.010

Elekes, F., Varga, M., & Király, I. (2017). Level-2 perspectives computed quickly and spontaneously: Evidence from eight- to 9.5-year-old children. *British Journal of Developmental Psychology, 35*(4), 609-622. DOI:10.1111/bjdp.12201

Elsabbagh, M., Divan, G., Koh, Y.-J., Kim, Y. S., Kauchali, S., Marcín, C., . . . Fombonne, E. (2012a). Global prevalence of autism and other pervasive developmental disorders. *Autism Research, 5*(3), 160-179. DOI:10.1002/aur.239

Elsabbagh, M., Mercure, E., Hudry, K., Chandler, S., Pasco, G., Charman, T., . . . BASIS Team (2012b). Infant neural sensitivity to dynamic eye gaze is associated with later emerging autism. *Current Biology, 22*(4), 338-342. DOI:10.1016/j.cub.2011.12.056

Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology, 40*(6), 760-768. DOI:10.1016/j.jesp.2004.02.002

Erbas, Y., Sels, L., Ceulemans, E., & Kuppens, P. (2016). Feeling me, feeling you: The relation between emotion differentiation and empathic accuracy. *Social, Psychological and Personality Science, 7*(3), 240-247. DOI: 10.1177/1948550616633504

Erle, T. M., & Topolinski, S. (2017). The grounded nature of psychological perspective-taking. *Journal of Personality and Social Psychology, 112*(5), 683-695. DOI:10.1037/pspa0000081

Eslinger, P. J. (1998). Neurological and neuropsychological bases of empathy. *European Neurology, 39*(4), 193-199. DOI:10.1159/000007933

Fareri, D. S., Niznikiewicz, M. A., Lee, V. K., & Delgado, M. R. (2012). Social network modulation of reward-related signals. *Journal of Neuroscience, 32*(26), 9045-9052. DOI:10.1523/jneurosci.0610-12.2012

Fawcett, C., & Liszkowski, U. (2012). Mimicry and play initiation in 18-month-old infants. *Infant Behavior and Development, 35*(4), 689-696. DOI:10.1016/j.infbeh.2012.07.014

Feldman, L. A. (1995). Valence focus and arousal focus: Individual differences in the structure of affective experience. *Journal of Personality and Social Psychology, 69*(1), 153-166. DOI:10.1037//0022-3514.69.1.153

Feldman, R. (2007). Mother-infant synchrony and the development of moral orientation in childhood and adolescence: Direct and indirect mechanisms of developmental continuity. *American Journal of Orthopsychiatry, 77*(4), 582-597. DOI:10.1037/0002-9432.77.4.582

Field, T. M., Woodson, R., Greenberg, R., & Cohen, D. (1982). Discrimination and imitation of facial expression by neonates. *Science, 218*(4568), 179-181. DOI:10.1126/science.7123230

Flavell, J. H. (1977). The development of knowledge about visual perception. *Nebraska Symposium on Motivation, 25*, 43-76. DOI: N/A

Flavell, J. H., Abrahams Everett, B., Croft, K., & Flavell, E. R. (1981a). Young children's knowledge about visual perception: Further evidence for the Level 1-Level 2 distinction. *Developmental Psychology, 17*(1), 99-103. DOI:10.1037/0012-1649.17.1.99

Flavell, J. H., Flavell, E. R., Green, F. L., & Wilcox, S. A. (1981b). The development of three spatial perspective-taking rules. *Child Development, 52*(1), 356. DOI:10.2307/1129250

Focquaert, F., & Vanneste, S. (2015). Autism spectrum traits in normal individuals: A preliminary VBM analysis. *Frontiers in Human Neuroscience, 9*. DOI:10.3389/fnhum.2015.00264

Freeman, J. B., & Ambady, N. (2010). MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behavior Research Methods, 42*(1), 226-241. DOI:10.3758/brm.42.1.226

Frith, C. D., & Frith, U. (2006). The neural basis of mentalizing. *Neuron, 50*(4), 531-534. DOI:10.1016/j.neuron.2006.05.001

Furlanetto, T., Becchio, C., Samson, D., & Apperly, I. (2016). Altercentric interference in level 1 visual perspective taking reflects the ascription of mental states, not submentalizing. *Journal of Experimental Psychology: Human Perception and Performance, 42*(2), 158-163. DOI:10.1037/xhp0000138

Galef, B. G. (1988). Imitation in animals: History, definition and interpretation of data from the psychological laboratory. In T. Zentall & B. G. Galef (Eds.), *Comparative Social Learning* (pp. 3-28). Hillsdale, NJ: Erlbaum.

Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology, 78*(4), 708-724. DOI:10.1037/0022-3514.78.4.708

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119*(2), 593-609. DOI:10.1093/brain/119.2.593

Gallese, V. (2007). Before and below 'theory of mind': Embodied simulation and the neural correlates of social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences, 362*(1480), 659-669. DOI:10.1098/rstb.2006.2002

Gallese, V. (2003). The roots of empathy: The shared manifold hypothesis and the neural basis of intersubjectivity. *Psychopathology, 36*(4), 171-180. DOI:10.1159/000072786

Garcia, M., Kerr, T., Blake, A. & Haffey, A. (2016). Collector: platform for online data collection. Retrieved from https://github.com/gikeymarcia/Collector

Gillespie, S. M., McCleery, J. P., & Oberman, L. M. (2014). Spontaneous versus deliberate vicarious representations: Different routes to empathy in psychopathy and autism. *Brain, 137*(4). DOI:10.1093/brain/awt364

Golan, O., Baron-Cohen, S., Hill, J. J., & Golan, Y. (2006a). The “Reading the Mind in Films” Task: Complex emotion recognition in adults with and without autism spectrum conditions. *Social Neuroscience, 1*(2), 111-123. DOI:10.1080/17470910600980986

Golan, O., Baron-Cohen, S., & Hill, J. (2006b). The Cambridge Mindreading (CAM) face-voice battery: Testing complex emotion recognition in adults with and without Asperger syndrome. *Journal of Autism and Developmental Disorders, 36*(2), 169-183. DOI:10.1007/s10803-005-0057-y

Golan, O., & Baron-Cohen, S. (2006). Systemizing empathy: Teaching adults with Asperger syndrome or high-functioning autism to recognize complex emotions using interactive multimedia. *Development and Psychopathology, 18*(2). DOI:10.1017/s0954579406060305

Gordon, A. C., & Olson, D. R. (1998). The relation between acquisition of a Theory of Mind and the capacity to hold in mind. *Journal of Experimental Child Psychology, 68*(1), 70-83. DOI:10.1006/jecp.1997.2423

Grant, B. J., Fetterman, Z., Weyhaupt, M. B., Kim, M., & Tullett, A. M. (2018). It takes two: A replication. *Journal of Research in Personality, 72*, 58-63. DOI:10.1016/j.jrp.2016.06.023

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*(6), 1464-1480. DOI:10.1037/0022-3514.74.6.1464

Griffiths, S., Jarrold, C., Penton-Voak, I. S., Woods, A. T., Skinner, A. L., &

Munafò, M. R. (2017). Impaired recognition of basic emotions from facial expressions in young people with Autism Spectrum Disorder: Assessing the importance of expression intensity. *Journal of Autism and Developmental Disorders, 49*(7), 2768-2778. DOI:10.1007/s10803-017-3091-7

Gzesh, S. M., & Surber, C. F. (1985). Visual perspective-taking skills in children. *Child Development, 56*(5), 1204-1213. DOI:10.2307/1130235

Haffey, A., Press, C., O'Connell, G., & Chakrabarti, B. (2013). Autistic traits modulate mimicry of social but not nonsocial rewards. *Autism Research, 6*(6), 614-620. DOI:10.1002/aur.1323

Hamilton, A. F. de C., Brindley, R. M., & Frith, U. (2007). Imitation and action understanding in autistic spectrum disorders: How valid is the hypothesis of a deficit in the mirror neuron system? *Neuropsychologia, 45*(8), 1859-1868. DOI:10.1016/j.neuropsychologia.2006.11.022

Hamilton, A. F. de C., Brindley, R. M., & Frith, U. (2009). Visual perspective taking impairment in children with autistic spectrum disorder. *Cognition, 113*(1), 37-44. DOI:10.1016/j.cognition.2009.07.007

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1994). *Emotional Cognition*. Cambridge, UK: Cambridge University Press.

Hermans, E. J., Putman, P., & van Honk, J. (2006). Testosterone administration reduces empathetic behavior: A facial mimicry study. *Psychoneuroendocrinology, 31*(7), 859-866. DOI:10.1016/j.psyneuen.2006.04.002

Heyes, C. (1993). Imitation, culture and cognition. *Animal Behaviour, 46*(5), 999-1010. DOI:10.1006/anbe.1993.1281

Hobson, R. P. (1995). *Autism and the development of mind*. Hove, UK: Psychology Press.

Hoffman, M. (1977). Empathy, its development and prosocial implications. *Nebraska Symposium on Motivation, 25*, 169-217. DOI: N/A

de Houwer, J., Thomas, S., & Baeyens, F. (2001). Association learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin, 127*(6), 853-869. DOI:10.1037/0033-2909.127.6.853

Hsu, C.-T., Neufeld, J., & Chakrabarti, B. (2018). Reduced reward-related neural response to mimicry in individuals with autism. *European Journal of Neuroscience, 47*(6), 610-618. DOI:10.1111/ejn.13620

Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology, 16*(2), 233-253. DOI:10.1111/j.2044-835x.1998.tb00921.x

Hughes, C., Adlam, A., Happé, F., Jackson, J., Taylor, A., & Caspi, A. (2000). Good test-retest reliability for standard and advanced false-belief tasks across a wide range of abilities. *Journal of Child Psychology and Psychiatry, 41*(4), 483-490. DOI:10.1017/s0021963099005533

Hughes, C., Jaffee, S. R., Happé, F., Taylor, A., Caspi, A., & Moffitt, T. E. (2005). Origins of individual differences in Theory of Mind: From nature to nurture? *Child Development, 76*(2), 356-370. DOI:10.1111/j.1467-8624.2005.00850_a.x

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., & Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biology, 3*(3).

DOI:10.1371/journal.pbio.0030079

Iacoboni, M., & Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nature Reviews Neuroscience, 7*(12), 942-951. DOI:10.1038/nrn2024

Ickes, W. J. (1997). *Empathic Accuracy*. New York, NY: The Guilford Press.

Ickes, W. J., & Simpson, J. (2001). Motivational aspects of empathic accuracy, in G.J.O. Fletcher and M.S. Clark (eds.); *Interpersonal Processes: Blackwell Handbook in Social Psychology*, 229-249. Oxford, UK: Blackwell.

Ickes, W. J., Stinson, L., Bissonnette, V., & Garcia, S. (1990). Naturalistic social cognition: Empathic accuracy in mixed-sex dyads. *Journal of Personality and Social Psychology, 59*(4), 730-742. DOI:10.1037/0022-3514.59.4.730

Izhaki-Costi, O., & Schul, Y. (2011). I do not know you and I am keeping it that way: Attachment avoidance and empathic accuracy in the perception of strangers. *Personal Relationships, 18*(3), 321-340. DOI:10.1111/j.1475-6811.2010.01292.x

Jolliffe, D., & Farrington, D. P. (2006). Development and validation of the Basic Empathy Scale. *Journal of Adolescence, 29*(4), 589-611. DOI:10.1016/j.adolescence.2005.08.010

Jones, A. P., Happé, F. G E.., Gilbert, F., Burnett, S., & Viding, E. (2010). Feeling, caring, knowing: Different types of empathy deficit in boys with psychopathic tendencies and autism spectrum disorder. *Journal of Child Psychology and Psychiatry, 51*(11), 1188-1197. DOI:10.1111/j.1469-7610.2010.02280.x

Jones, B., & Rachlin, H. (2006). Social Discounting. *Psychological Science, 17*(4), 283-286. DOI:10.1111/j.1467-9280.2006.01699.x

Jones, C. R., Simonoff, E., Baird, G., Pickles, A., Marsden, A. J. S., Tregay, J., . . . Charman, T. (2017). The association between theory of mind, executive function, and the symptoms of autism spectrum disorder. *Autism Research, 11*(1), 95-109. DOI:10.1002/aur.1873

Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263-291. DOI:10.2307/1914185.

Kanske, P., Böckler, A., Trautwein, F., Lesemann, F. H., & Singer, T. (2016). Are strong empathizers better mentalizers? Evidence for independence and interaction between the routes of social cognition. *Social Cognitive and Affective Neuroscience, 11*(9), 1383-1392. DOI:10.1093/scan/nsw052

Kanske, P., Böckler, A., Trautwein, F.-M., & Singer, T. (2015). Dissecting the social brain: Introducing the EmpaToM to reveal distinct neural networks and brain–behavior relations for empathy and Theory of Mind. *NeuroImage, 122*, 6-19. DOI:10.1016/j.neuroimage.2015.07.082

Kawamichi, H., Sugawara, S. K., Hamano, Y. H., Makita, K., Kochiyama, T., & Sadato, N. (2016). Increased frequency of social interaction is associated with enjoyment enhancement and reward system activation. *Scientific Reports, 6*(1). DOI:10.1038/srep24561

Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: the role of mutual knowledge in comprehension. *Psychological Science, 11*(1), 32-38. DOI:10.1111/1467-9280.00211

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults.

Cognition, 89(1), 25-41. DOI:10.1016/s0010-0277(03)00064-7

Keysers, C., & Gazzola, V. (2014). Dissociating the ability and propensity for empathy. *Trends in Cognitive Sciences, 18*(4), 163-166. DOI:10.1016/j.tics.2013.12.011

Klein, J. T., Shepherd, S. V., & Platt, M. L. (2009). Social attention and the brain. *Current Biology, 19*(20). DOI:10.1016/j.cub.2009.08.010

Kleiner, M., Brainard, D., Pelli, D., Broussard, C., Wolf, T., & Niehorster, D. (2007). What's new in Psychtoolbox-3? *Perception, 36*(14), 1-16. DOI:10.1068/v070821

Knafo, A., Zahn-Waxler, C., Davidov, M., Hulle, C. V., Robinson, J. L., & Rhee, S. H. (2009). Empathy in early childhood: Genetic, environmental and affective contributions. *Annals of the New York Academy of Sciences, 1167*(1), 103-114. DOI:10.1111/j.1749-6632.2009.04540.x

Kohls, G., Peltzer, J., Schulte-Rüther, M., Kamp-Becker, I., Remschmidt, H., Herpertz-Dahlmann, B., & Konrad, K. (2011). Atypical brain responses to reward cues in autism as revealed by event-related potentials. *Journal of Autism and Developmental Disorders, 41*(11), 1523-1533. DOI:10.1007/s10803-011-1177-1

Kohls, G., Schulte-Rüther, M., Nehrkorn, B., Müller, K., Fink, G. R., Kamp-Becker, I., . . . Konrad, K. (2013). Reward system dysfunction in autism spectrum disorders. *Social Cognitive and Affective Neuroscience, 8*(5), 565-572. DOI:10.1093/scan/nss033

Kohls, G., Chevallier, C., Troiani, V., & Schultz, R. T. (2012). Social ‘wanting’ dysfunction in autism: Neurobiological underpinnings and treatment implications. *Journal of Neurodevelopmental Disorders, 4*(1).

DOI:10.1186/1866-1955-4-10

Koolschijn, P. C. M. P., Geurts, H. M., van der Leij, A. R., & Scholte, H. S. (2015). Are autistic traits in the general population related to global and regional brain differences? *Journal of Autism and Developmental Disorders, 45*(9), 2779-2791. DOI:10.1007/s10803-015-2441-6

Kozak, M. N., Marsh, A. A., & Wegner, D. M. (2006). What do I think you're doing? Action identification and mind attribution. *Journal of Personality and Social Psychology, 90*(4), 543-555. DOI:10.1037/0022-3514.90.4.543

Kovács, A. M., Téglás, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science, 330*(6012), 1830-1834. DOI:10.1126/science.1190792

Kühn, S., Müller, B. C., van Baaren, R. B., Wietzker, A., Dijksterhuis, A., & Brass, M. (2010). Why do I like you when you behave like me? Neural mechanisms mediating positive consequences of observing someone being imitated. *Social Neuroscience, 5*(4), 384-392. DOI:10.1080/17470911003633750

Kulke, L., von Duhn, B., Schneider, D., & Rakoczy, H. (2018). Is implicit theory of mind a real and robust phenomenon? Results from a systematic replication study. *Psychological Science, 29*(6), 888-900. DOI:10.1177/0956797617747090

Kuznetsova, A., Brockhoff P. B., & Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26. DOI: 10.18637/jss.v082.i13.

Lakin, J. L., Jefferis, V. E., Cheng, C. M., & Chartrand, T. L. (2003). The chameleon effect as social glue: evidence for the evolutionary significance of nonconscious mimicry. *Journal of Nonverbal Behavior, 27*(3), 145-162. DOI:10.1023/a:1025389814290

Laurita, A. C., Hazan, C., & Spreng, R. N. (2019). An attachment theoretical perspective for the neural representation of close others. *Social Cognitive and Affective Neuroscience, 14*(3), 237-251. DOI:10.1093/scan/nsz010

Leekam, S. R., López, B., & Moore, C. (2000). Attention and joint attention in preschool children with autism. *Developmental Psychology, 36*(2), 261-273. DOI:10.1037//0012-1649.36.2.261

Leekam, S. R., & Ramsden, C. A. H. (2006). Dyadic orienting and joint attention in preschool children with autism. *Journal of Autism and Developmental Disorders, 36*(2), 185-197. DOI:10.1007/s10803-005-0054-1

Leiberg, S., & Anders, S. (2006). The multiple facets of empathy: a survey of theory and evidence. In S. Anders, G. Ende, M. Junghöfer, J. Kissler, & D. Wildgruber (Eds.), *Progress in Brain Research, Vol. 156* (pp. 419-440). Amsterdam, the Netherlands: Elsevier. DOI: 10.1016/S0079-6123(06)56023-6

Leslie, A. M. (1987). Pretense and representation: The origins of "theory of mind." *Psychological Review, 94*(4), 412-426. DOI:10.1037/0033-295x.94.4.412

Lin, A., Rangel, A., & Adolphs, R. (2012). Impaired learning of social compared to monetary rewards in autism. *Frontiers in Neuroscience, 6*. DOI:10.3389/fnins.2012.00143

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

Lundqvist, L., & Dimberg, U. (1995). Facial expressions are contagious. *Journal of Psychophysiology, 9*(3), 203-311. DOI: N/A

Majdandžić, J., Amashaufer, S., Hummer, A., Windischberger, C., & Lamm, C. (2016). The selfless mind: How prefrontal involvement in mentalizing with similar and dissimilar others shapes empathy and prosocial behavior. *Cognition, 157*, 24-38. DOI:10.1016/j.cognition.2016.08.003

Marsh, A. A., Kozak, M. N., Wegner, D. M., Reid, M. E., Yu, H. H., & Blair, R. J. R. (2010). The neural substrates of action identification. *Social Cognitive and Affective Neuroscience, 5*(4), 392-403. DOI:10.1093/scan/nsq004

Marsh, L. E., & Hamilton, A. F. de C. (2011). Dissociation of mirroring and mentalising systems in autism. *NeuroImage, 56*(3), 1511-1519. DOI:10.1016/j.neuroimage.2011.02.003

Marshall, J., Gollwitzer, A., & Santos, L. R. (2018). Does altercentric interference rely on mentalizing?: Results from two level-1 perspective-taking tasks. *PLOS ONE, 13*(3), e0194101. DOI:10.1371/journal.pone.0194101

Martin-Key, N., Brown, T., & Fairchild, G. (2017). Empathic accuracy in male adolescents with Conduct Disorder and higher versus lower levels of callous-unemotional traits. *Journal of Abnormal Child Psychology, 45*(7), 1385-1397. DOI:10.1007/s10802-016-0243-8

Mathersul, D., Mcdonald, S., & Rushby, J. A. (2013). Understanding advanced theory of mind and empathy in high-functioning adults with autism spectrum disorder. *Journal of Clinical and Experimental Neuropsychology, 35*(6), 655-668. DOI:10.1080/13803395.2013.809700

The Mathworks, Inc. (2016). MATLAB Release 2016b [Computer software]. Natick, MA, United States.

McIntosh, D. M. (2006). Spontaneous facial mimicry, liking and emotional contagion. *Polish Psychological Bulletin, 37*(1), 31-42. DOI: N/A

McIntosh, D. N., Reichmann-Decker, A., Winkielman, P., & Wilbarger, J. L. (2006). When the social mirror breaks: Deficits in automatic, but not voluntary, mimicry of emotional facial expressions in autism. *Developmental Science, 9*(3), 295-302. DOI:10.1111/j.1467-7687.2006.00492.x

Meert, G., Wang, J., & Samson, D. (2017). Efficient belief tracking in adults: The role of task instruction, low-level associative processes and dispositional social functioning. *Cognition, 168*, 91-98. DOI:10.1016/j.cognition.2017.06.012

Meltzoff, A. (2002). Elements of a developing theory of autism. In A.N. Meltzoff & W. Prinz (Eds.), *The Imitative Mind: Development, Evolution, and Brain Bases* (pp. 19-41). New York, NY: Cambridge University Press. DOI: 10.1017/CBO9780511489969

Meltzoff, A., & Moore, M. (1977). Imitation of facial and manual gestures by human neonates. *Science, 198*(4312), 74-78. DOI:10.1126/science.897687

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and Theory of Mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development, 78*(2), 622-646.

DOI:10.1111/j.1467-8624.2007.01018.x

Moll, H., & Tomasello, M. (2006). Level 1 perspective-taking at 24 months of age. *British Journal of Developmental Psychology, 24*(3), 603-613. DOI:10.1348/026151005x55370

Murray, A. L., Booth, T., Mckenzie, K., Kuenssberg, R., & O'Donnell, M. (2014). Are autistic traits measured equivalently in individuals with and without an Autism Spectrum Disorder? An invariance analysis of the Autism Spectrum Quotient Short Form. *Journal of Autism and Developmental Disorders, 44*(1), 55-64. DOI:10.1007/s10803-013-1851-6

Murray, R. J., Debbané, M., Fox, P. T., Bzdok, D., & Eickhoff, S. B. (2015). Functional connectivity mapping of regions associated with self- and other-processing. *Human Brain Mapping, 36*(4), 1304-1324. DOI:10.1002/hbm.22703

Murray, R. J., Schaer, M., & Debbané, M. (2012). Degrees of separation: A quantitative neuroimaging meta-analysis investigating self-specificity and shared neural activation between self- and other-reflection. *Neuroscience & Biobehavioral Reviews, 36*(3), 1043-1059. DOI:10.1016/j.neubiorev.2011.12.013

Mutter, B., Alcorn, M. B., & Welsh, M. (2006). Theory of Mind and executive function: Working-memory capacity and inhibitory control as predictors of false-belief task performance. *Perceptual and Motor Skills, 102*(3), 819-835. DOI:10.2466/pms.102.3.819-835

Naber, F. B. A., Bakermans-Kranenburg, M. J., van IJzendoorn, M. H., Dietz, C., van Daalen, E., Swinkels, S. H., . . . van Engeland, H. (2007). Joint attention development in toddlers with autism. *European Child &*

Adolescent Psychiatry, 17(3), 143-152. DOI:10.1007/s00787-007-0648-6

Neufeld, J., & Chakrabarti, B. (2016). Empathy modulates the rewarding effect of mimicry. *Scientific Reports, 6*(1). DOI:10.1038/srep27751

Neuhaus, E., Beauchaine, T. P., & Bernier, R. (2010). Neurobiological correlates of social functioning in autism. *Clinical Psychology Review, 30*(6), 733-748. DOI:10.1016/j.cpr.2010.05.007

Nijhof, A. D., Brass, M., Bardi, L., & Wiersema, J. R. (2016). Measuring mentalizing ability: A within-subject comparison between an explicit and implicit version of a ball detection task. *PLOS ONE, 11*(10). DOI:10.1371/journal.pone.0164373

Oberman, L. M., Winkielman, P., & Ramachandran, V. S. (2009). Slow echo: Facial EMG evidence for the delay of spontaneous, but not voluntary, emotional mimicry in children with autism spectrum disorders. *Developmental Science, 12*(4), 510-520. DOI:10.1111/j.1467-7687.2008.00796.x

O'Connell, G., Christakou, A., Haffey, A. T., & Chakrabarti, B. (2013). The role of empathy in choosing rewards from another's perspective. *Frontiers in Human Neuroscience, 7*. DOI:10.3389/fnhum.2013.00174

de Oliveira, L. L., & Thomaz, C. E. (2006). *Captura e Alinhamento de Imagens: Um Banco de Faces Brasileiro*. (Undergraduate dissertation). Retrieved from: https://fei.edu.br/~cet/iniciacaocientifica_LeoOliveira_2006.pdf.

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

Oostenbroek, J., Suddendorf, T., Nielsen, M., Redshaw, J., Kennedy-

Costantini, S., Davis, J., … & Slaughter, V. (2016). Comprehensive longitudinal study challenges the existence of neonatal imitation in humans. *Current Biology, 26*(10), 1334-1338. DOI:10.1016/j.cub.2016.03.047

Osterling, J. A., & Dawson, G. (1994). Early recognition of children with autism: A study of first birthday home videotapes. *Journal of Autism and Developmental Disorders, 24*(3), 247-257. DOI:10.1007/bf02172225

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. DOI:10.1017/s0954579402002031

Panasiti, M. S., Puzzo, I., & Chakrabarti, B. (2016). Autistic traits moderate the impact of reward learning on social behaviour. *Autism Research, 9*(4), 471-479. DOI:10.1002/aur.1523

Pearson, A., Ropar, D., & Hamilton, A. F. de C. (2013). A review of visual perspective taking in autism spectrum disorder. *Frontiers in Human Neuroscience, 7*:652. DOI:10.3389/fnhum.2013.00652

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research, 91*(1), 176-180. DOI:10.1007/bf00230027

Pelli, D. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4) 437-442. DOI:10.1163/156856897X00366

Pellicano, E. (2007). Links between theory of mind and executive function in young children with autism: Clues to developmental primacy.

Developmental Psychology, 43(4), 974-990. DOI:10.1037/0012-1649.43.4.974

Perner, J., & Ruffman, T. (2005). Infants’ insight into the mind: How deep? *Science, 308*(5179), 214-216. DOI:10.1126/science.1111656

Peterson, C. C., & Slaughter, V. (2009). Theory of mind (ToM) in children with autism or typical development: Links between eye-reading and false belief understanding. *Research in Autism Spectrum Disorders, 3*(2), 462-473. DOI:10.1016/j.rasd.2008.09.007

Phillips, J., Ong, D. C., Surtees, A. D. R., Xin, Y., Williams, S., Saxe, R., & Frank, M. C. (2015). A second look at automatic Theory of Mind: Reconsidering Kovács, Téglás, and Endress (2010). *Psychological Science, 26*(9), 1353-1367. DOI:10.1177/0956797614558717

Piaget, J. (1946). *Symbol formation in the child.* Paris, FR: Delachaux & Nestlé.

Pineda, J. A. (2005). The functional significance of mu rhythms: Translating “seeing” and “hearing” into “doing”. *Brain Research Reviews, 50*(1), 57-68. DOI:10.1016/j.brainresrev.2005.04.005

Pollmann, M. M. H., & Krahmer, E. J. (2018). How do friends and strangers play the game Taboo? A study of accuracy, efficiency, motivation, and the use of shared knowledge. *Journal of Language and Social Psychology, 37*(4), 497-517. DOI:10.1177/0261927X17736084

Ponnet, K., Buysse, A., Roeyers, H., & De Corte, K. (2005). Empathic accuracy in adults with a Pervasive Developmental Disorder during an unstructured conversation with a typically developing stranger. *Journal of Autism and Developmental Disorders, 35*(5), 585-600. DOI:10.1007/s10803-005-0003-z

Ponnet, K., Buysse, A., Roeyers, H., & De Clercq, A. (2007). Mind-reading in young adults with ASD: Does structure matter? *Journal of Autism and Developmental Disorders, 38*(5), 905-918. DOI:10.1007/s10803-007-0462-5

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*(4), 515-526. DOI:10.1017/s0140525x00076512

Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences, 25*(1), 1-20. DOI:10.1017/s0140525x02000018

Psychology Software Tools, Inc. (2012). E-Prime 2.0 [Computer software]. Sharpsburg, PA, United States.

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

R Core Team (2018). R: A language and environment for statistical computing [Computer software]. Vienna, Austria: R Foundation for Statistical Computing.

Richey, J. A., Rittenberg, A., Hughes, L., Damiano, C. R., Sabatino, A., Miller, S., . . . Dichter, G. S. (2014). Common and distinct neural features of social and non-social reward processing in autism and social anxiety disorder. *Social Cognitive and Affective Neuroscience, 9*(3), 367-377. DOI:10.1093/scan/nss146

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual*

Review of Neuroscience, 27(1), 169-192. DOI:10.1146/annurev.neuro.27.070203.144230

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research, 3*(2), 131-141. DOI:10.1016/0926-6410(95)00038-0

Ronald, A., Viding, E., Happé, F., & Plomin, R. (2006). Individual differences in theory of mind ability in middle childhood and links with verbal ability and autistic traits: A twin study. *Social Neuroscience, 1*(3-4), 412-425. DOI:10.1080/17470910601068088

Rogers, S. J., Hepburn, S. L., Stackhouse, T., & Wehner, E. (2003). Imitation performance in toddlers with autism and those with other developmental disorders. *Journal of Child Psychology and Psychiatry, 44*(5), 763-781. DOI:10.1111/1469-7610.00162

Rolls, E. T. (2015). Limbic systems for emotion and for memory, but no single limbic system. *Cortex, 62*, 119-157. DOI:10.1016/j.cortex.2013.12.005

Rosenblau, G., Kliemann, D., Heekeren, H. R., & Dziobek, I. (2015). Approximating implicit and explicit mentalizing with two naturalistic video-based tasks in typical development and Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders, 45*(4), 953-965. DOI:10.1007/s10803-014-2249-9

Rosset, D. B., Rondan, C., Da Fonseca, D., Santos, A., Assouline, B., & Deruelle, C. (2008). Typical emotion processing for cartoon but not for real faces in children with Autistic Spectrum Disorders. *Journal of Autism and Developmental Disorders, 38*(5), 919-925. DOI:10.1007/s10803-007-0465-2 aan het Rot, M., & Hogenelst, K. (2014). The influence of affective empathy and autism spectrum traits on empathic accuracy. *PLOS ONE, 9*(6). DOI:10.1371/journal.pone.0098436

Rubio-Fernández, P. (2017). The director task: A test of Theory-of-Mind use or selective attention? *Psychonomic Bulletin and Review, 24*(4), 1121-1128. DOI:10.3758/s13423-016-1190-7

Ruff, C. C., & Fehr, E. (2014). The neurobiology of rewards and values in social decision making. *Nature Reviews Neuroscience, 15*(8), 549-562. DOI:10.1038/nrn3776

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161-1178. DOI:10.1037/h0077714

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

Sabbagh, M. A., Moses, L. J., & Shiverick, S. (2006). Executive functioning and preschoolers' understanding of false beliefs, false photographs, and false signs. *Child Development, 77*(4), 1034-1049. DOI:10.1111/j.1467-8624.2006.00917.x

Sagi, A., & Hoffman, M. L. (1976). Empathic distress in the newborn. *Developmental Psychology, 12*(2), 175-176. DOI:10.1037//0012-1649.12.2.175

Samson, D., & Apperly, I. A. (2015). *Level 1 Visual Perspective-Taking Task*. Retrieved from https://figshare.com/articles/Level_1_Visual_Perspective

_Taking_Task/1455943

Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. (2010). Seeing it their way: Evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance, 36*(5), 1255-1266. DOI:10.1037/a0018729

Samson, D., Apperly, I. A., Kathirgamanathan, U., & Humphreys, G. W. (2005). Seeing it my way: A case of a selective deficit in inhibiting self-perspective. *Brain, 128*(5), 1102-1111. DOI:10.1093/brain/awh464

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

Santiesteban, I., Shah, P., White, S., Bird, G., & Heyes, C. (2015). Mentalizing or submentalizing in a communication task? Evidence from autism and a camera control. *Psychonomic Bulletin & Review, 22*(3), 844-849. DOI:10.3758/s13423-014-0716-0

Santiesteban, I., White, S., Cook, J., Gilbert, S. J., Heyes, C., & Bird, G. (2012). Training social cognition: from imitation to Theory of Mind. *Cognition, 122*(2), 228-235. DOI:0.1016/j.cognition.2011.11.004

Sassenrath, C., Sassenberg, K., & Scholl, A. (2014). From a distance…: The impact of approach and avoidance motivational orientation on perspective taking. *Social Psychological and Personality Science, 5*(1), 18-26. DOI:10.1177/1948550613486672

Savitsky, K., Keysar, B., Epley, N., Carter, T., & Swanson, A. (2011). The closeness-communication bias: Increased egocentrism among friends versus strangers. *Journal of Experimental Social Psychology, 47*(1), 269-273. DOI:10.1016/j.jesp.2010.09.005

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: The role of the temporo-parietal junction in “theory of mind”. *NeuroImage, 19*(4), 1835-1842. DOI:10.1016/s1053-8119(03)00230-1

Saxe, R., & Wexler, A. (2005). Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia, 43*(10), 1391-1399. DOI:10.1016/j.neuropsychologia.2005.02.013

Schmitz, T. W., & Johnson, S. C. (2007). Relevance to self: A brief review and framework of neural systems underlying appraisal. *Neuroscience & Biobehavioral Reviews, 31*(4), 585-596. DOI:10.1016/j.neubiorev.2006.12.003

Schneider, D., Bayliss, A. P., Becker, S. I., & Dux, P. E. (2012a). Eye movements reveal sustained implicit processing of others' mental states. *Journal of Experimental Psychology: General, 141*(3), 433-438. DOI:10.1037/a0025458

Schneider, D., Lam, R., Bayliss, A. P., & Dux, P. E. (2012b). Cognitive load disrupts implicit Theory-of-Mind processing. *Psychological Science, 23*(8), 842-847. DOI:10.1177/0956797612439070

Schultz, R. T. (2005). Developmental deficits in social perception in autism: The role of the amygdala and fusiform face area. *International Journal of Developmental Neuroscience, 23*(2-3), 125-141. DOI:10.1016/j.ijdevneu.2004.12.012

Schuwerk, T., Vuori, M., & Sodian, B. (2015). Implicit and explicit Theory of

Mind reasoning in autism spectrum disorders: The impact of experience. *Autism, 19*(4), 459-468. DOI:10.1177/1362361314526004

Schwarzkopf, S., Schilbach, L., Vogeley, K., & Timmermans, B. (2014). “Making it explicit” makes a difference: Evidence for a dissociation of spontaneous and intentional level 1 perspective taking in high-functioning autism. *Cognition, 131*(3), 345-354. DOI:10.1016/j.cognition.2014.02.003

Sened, H., Lavidor, M., Lazarus, G., Bar-Kalifa, E., Rafaeli, E., & Ickes, W. (2017). Empathic accuracy and relationship satisfaction: A meta-analytic review. *Journal of Family Psychology, 31*(6), 742-752. DOI:10.1037/fam0000320

Senju, A., Southgate, V., White, S., & Frith, U. (2009). Mindblind eyes: An absence of spontaneous Theory of Mind in Asperger syndrome. *Science, 325*(5942), 883-885. DOI:10.1126/science.1176170

Shamay-Tsoory, S. G., Tomer, R., Berger, B. D., & Aharon-Peretz, J. (2003). Characterization of empathy deficits following prefrontal brain damage: The role of the right ventromedial prefrontal cortex. *Journal of Cognitive Neuroscience, 15*(3), 324-337. DOI:10.1162/089892903321593063

Shamay-Tsoory, S. G., Aharon-Peretz, J., & Perry, D. (2009). Two systems for empathy: A double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain, 132*(3), 617-627. DOI:10.1093/brain/awn279

Simner, M. L. (1971). Newborn's response to the cry of another infant. *Developmental Psychology, 5*(1), 136-150. DOI:10.1037/h0031066

Sims, T. B., Neufeld, J., Johnstone, T., & Chakrabarti, B. (2014). Autistic traits modulate frontostriatal connectivity during processing of rewarding faces. *Social Cognitive and Affective Neuroscience, 9*(12), 2010-2016. DOI:10.1093/scan/nsu010

Sims, T. B., van Reekum, C. M., Johnstone, T., & Chakrabarti, B. (2012). How reward modulates mimicry: EMG evidence of greater facial mimicry of more rewarding happy faces. *Psychophysiology, 49*(7), 998-1004. DOI:10.1111/j.1469-8986.2012.01377.x

Snodgrass, S. E., Hecht, M. A., & Ploutz-Snyder, R. (1998). Interpersonal sensitivity: Expressivity or perceptivity? *Journal of Personality and Social Psychology, 74*(1), 238-249. DOI:10.1037//0022-3514.74.1.238

Sodian, B., Schuwerk, T., & Kristen, S. (2015). Implicit and spontaneous Theory of Mind reasoning in Autism Spectrum Disorders. In Michael Fitzgerald (Ed.), *Autism Spectrum Disorder - Recent Advances*. Rijeka, Croatia: IntechOpen. DOI:10.5772/59393.

Sommerville, J. A., Bernstein, D. M., & Meltzoff, A. N. (2013). Measuring beliefs in centimeters: Private knowledge biases preschoolers' and adults' representation of others' beliefs. *Child Development, 84*(6), 1846-1854. DOI:10.1111/cdev.12110

Sonnby-Borgström, M. (2002). Automatic mimicry reactions as related to differences in emotional empathy. *Scandinavian Journal of Psychology, 43*(5), 433-443. DOI:10.1111/1467-9450.00312

Soto, J. A., & Levenson, R. W. (2009). Emotion recognition across cultures: The influence of ethnicity on empathic accuracy and physiological linkage. *Emotion, 9*(6), 874-884. DOI:10.1037/a0017399

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

Stel, M., van den Heuvel, C., & Smeets, R. C. (2008). Facial feedback mechanisms in Autistic Spectrum Disorders. *Journal of Autism and Developmental Disorders, 38*(7), 1250-1258. DOI:10.1007/s10803-007-0505-y

Stel, M., & Vonk, R. (2010). Mimicry in social interaction: Benefits for mimickers, mimickees, and their interaction. *British Journal of Psychology, 101*(2), 311-323. DOI:10.1348/000712609x465424

Stern, D. N. (1977). *The first relationship: Infant and mother*. Cambridge, MA: Harvard University Press.

Sternglanz, R. W., & DePaulo, B. M. (2004). Reading nonverbal cues to emotions: the advantages and liabilities of relationship closeness. *Journal of Nonverbal Behavior, 28*(4), 245-266. DOI:10.1007/s10919-004-4158-7

Stinson, L., & Ickes, W. J. (1992). Empathic accuracy in the interactions of male friends versus male strangers. *Journal of Personality and Social Psychology, 62*(5), 787-797. DOI:10.1037//0022-3514.62.5.787

Sullivan, M., Finelli, J., Marvin, A., Garrett-Mayer, E., Bauman, M., & Landa, R. (2007). Response to joint attention in toddlers at risk for Autism Spectrum Disorder: a prospective study. *Journal of Autism and Developmental Disorders, 37*(1), 37-48. DOI:10.1007/s10803-006-0335-3

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science, 18*(7), 580-586. DOI:10.1111/j.1467-

9280.2007.01943.x

Surian, L., & Geraci, A. (2012). Where will the triangle look for it? Attributing false beliefs to a geometric shape at 17 months. *Developmental Psychology, 30*(1). DOI:10.1111/j.2044-835X.2011.02046.x

Surtees, A. D. R., & Apperly, I. A. (2012). Egocentrism and automatic perspective taking in children and adults. *Child Development, 83*(2), 452-460. DOI:10.1111/j.1467-8624.2011.01730.x

Surtees, A. D. R., Butterfill, S. A., & Apperly, I. A. (2012). Direct and indirect measures of level-2 perspective-taking in children and adults. *British Journal of Developmental Psychology, 30*(1), 75-86. DOI:10.1111/j.2044-835x.2011.02063.x

Swettenham, J., Baron-Cohen, S., Charman, T., Cox, A., Baird, G., Drew, A., . . . Wheelwright, S. (1998). The frequency and distribution of spontaneous attention shifts between social and nonsocial stimuli in autistic, typically developing, and nonautistic developmentally delayed infants. *Journal of Child Psychology and Psychiatry, 39*(5), 747-753. DOI:10.1111/1469-7610.00373

Symeonidou, I., Dumontheil, I., Chow, W.-Y., & Breheny, R. (2016). Development of online use of theory of mind during adolescence: An eye-tracking study. *Journal of Experimental Child Psychology, 149*, 81-97. DOI:10.1016/j.jecp.2015.11.007

Thomas, G., & Fletcher, G. J. O. (2003). Mind-reading accuracy in intimate relationships: assessing the roles of the relationship, the target, and the judge. *Journal of Personality and Social Psychology, 85*(6), 1079-1094. DOI:10.1037/0022-3514.85.6.1079

Thomas, G., Fletcher, G. J. O., & Lange, C. (1997). On-line empathic accuracy in marital interaction. *Journal of Personality and Social Psychology, 72*(4), 839-850. DOI:10.1037//0022-3514.72.4.839

Thompson, N. M., Uusman, A., Gross, J. J., & Chakrabarti, B. (2019). Empathy and emotion regulation: An integrative account. In N. Srinivasan (Ed.), *Emotion and Cognition. Progress in Brain Research, 247* (pp. 273-304). Amsterdam, the Netherlands: Elsevier. DOI:10.1016/bs.pbr.2019.03.024

Todd, A. R., Bodenhausen, G. V., Richeson, J. A., & Galinsky, A. D. (2011). Perspective taking combats automatic expressions of racial bias. *Journal of Personality and Social Psychology, 100*(6), 1027-1042. DOI:10.1037/a0022308

Tomasello, M., Savage-Rumbaugh, S., & Kruger, A. C. (1993). Imitative learning of actions on objects by children, chimpanzees, and enculturated chimpanzees. *Child Development, 64*(6), 1688-1705. DOI:10.1111/j.1467-8624.1993.tb04207.x

Trilla Gros, I., Panasiti, M. S., & Chakrabarti, B. (2015). The plasticity of the mirror system: How reward learning modulates cortical motor simulation of others. *Neuropsychologia, 70*, 255-262. DOI:10.1016/j.neuropsychologia.2015.02.033

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review, 117*(2), 440-463. DOI:10.1037/a0018963

Ungerer, J. A., Dolby, R., Waters, B., Barnett, B., Kelk, N., & Lewin, V. (1990). The early development of empathy: Self-regulation and individual differences in the first year. *Motivation and Emotion, 14*(2), 93-106.

DOI:10.1007/bf00991638

Unity Technologies ApS (2016). Unity [Computer software]. Copenhagen, Denmark.

Verhofstadt, L., Devoldre, I., Buysse, A., Stevens, M., Hinnekens, C., Ickes, W. J., & Davis, M. (2016). The role of cognitive and affective empathy in spouses' support interactions: An observational study. *PLOS ONE, 11*(2). DOI:10.1371/journal.pone.0149944

Vorauer, J. D., & Sasaki, S. J. (2014). Distinct effects of imagine-other versus imagine-self perspective-taking on prejudice reduction. *Social Cognition, 32*(2), 130-147. DOI:10.1521/soco.2014.32.2.130

de Waal, F. B. M. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annual Review of Psychology, 59*(1), 279-300. DOI:10.1146/annurev.psych.59.103006.093625

de Waal, F. B. M., & Preston, S. D. (2017). Mammalian empathy: behavioural manifestations and neural basis. *Nature Reviews Neuroscience, 18*, 498-509. DOI:10.1038/nrn.2017.72

Ward, E., Ganis, G., & Bach, P. (2019). Spontaneous vicarious perception of the content of another’s visual perspective. *Current Biology, 29*(5), 874-880. DOI:10.1016/j.cub.2019.01.046

Warnell, K. R., & Redcay, E. (2019). Minimal coherence among varied theory of mind measures in childhood and adulthood. *Cognition, 191* (pre-print). DOI:10.31234/osf.io/dmtke

van der Wel, R. P. R. D., Sebanz, N., & Knoblich, G. (2014). Do people automatically track others’ beliefs? Evidence from a continuous measure. *Cognition, 130*(1), 128-133. DOI:10.1016/j.cognition.2013.10.004

Wellman, H., Cross, D., & Watson, J. (2001). Meta-analysis of theory of mind development: The truth about false belief. *Child Development, 72*(3), 655-684. DOI:10.1111/1467-8624.00304

Westra, E. (2016). Spontaneous mindreading: A problem for the two-systems account. *Synthese, 194*(11), 4559-4581. DOI:10.1007/s11229-016-1159-0

Williams, J. H. G., Whiten, A., Suddendorf, T., & Perrett, D. I. (2001). Imitation, mirror neurons and autism. *Neuroscience & Biobehavioral Reviews, 25*(4), 287-295. DOI:10.1016/s0149-7634(01)00014-8

Wilson, C. E., Happé, F., Wheelwright, S. J., Ecker, C., Lombardo, M. V., Johnston, P., … & Murphy, D. G. M. (2014). The neuropsychology of male adults with high-functioning autism or Asperger syndrome. *Autism Research, 7*(5), 568-581. DOI:10.1002/aur.1394

Wilson, C. J., Soranzo, A., & Bertamini, M. (2017). Attentional interference is modulated by salience not sentience. *Acta Psychologica, 178*, 56-65. DOI:10.1016/j.actpsy.2017.05.010

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

Zahn-Waxler, C., & Radke-Yarrow, M. (1982). The development of altruism: Alternative research strategies. In N. Eisenberg (Ed.), *The Development of Prosocial Behavior*, (pp. 109-137). New York, NY: Academic Press.

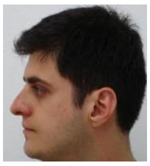
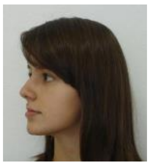
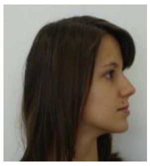
Zaki, J., Bolger, N., & Ochsner, K. (2008). It takes two: The interpersonal nature of empathic accuracy. *Psychological Science, 19*(4), 399-404.

DOI:10.1111/j.1467-9280.2008.02099.x

Zaki, J., Weber, J., Bolger, N., & Ochsner, K. (2009). The neural bases of empathic accuracy. *Proceedings of the National Academy of Sciences, 106*(27), 11382-11387. DOI:10.1073/pnas.0902666106

Zurek, P. P., & Scheithauer, H. (2017). Towards a more precise conceptualization of empathy: an integrative review of literature on definitions, associated functions, and developmental trajectories. *International Journal of Developmental Science, 11*(3-4), 57-68. DOI:10.3233/dev-16224

Appendix B - Faces shown in implicit belief tracking study (Chapter 2.4)

	-90°	-45°	0°	+45°	+90°
Prac F					
Prac M					
Main F1					
Main F2					
Main M1					
Main M2					

Prac: used for practice trials in the evaluative conditioning task; Main: used for main trials in the evaluative conditioning task. The -90° and +90° Main images were used for the agents in the implicit belief tracking task (see Chapter 2.4.1.3). For the implicit belief tracking task, the faces (without background) were rotated at a pitch of 45°.

Appendix C - Change-of-location scenarios used in explicit belief tracking study (Chapter 2.5)

Numbers in each scenario represent sequentially presented sections. The filler task was presented between section 3 and 4.

False belief scenarios

2) 1. Will and his friend Rose have just come home from the park. 2. Will puts his keys in the top drawer full of gloves, scarves and hats and then rushes to the toilet. 3. While Will is in the toilet, Rose moves the keys to here. 4. Then Will comes back in. Where will he look for his keys?

3) 1. Mark and his friend Will are planting flowers in their garden plot to surprise their other housemate. 2. Mark buries a flower here and then goes to the shed to get a shovel. 3. While Mark is gone, Will decides to move the flower and buries it here. 4. Then Mark comes back. Where will he look for the flower?

6) 1. Maria and her friend Mark are moving house and they are packing their mugs into a box full of tissue paper. 2. Maria buries a mug here in the box and then goes to the kitchen to get a drink of water. 3. While Maria is in the kitchen, Mark moves the mug to here. 4. Then Maria comes back in. Where will she look for the mug?

10) 1. Will and his friend Mark are taking decorations off the Christmas tree and putting them back in the box. 2. Will takes the star off the tree and puts it in the box here. He then leaves to get more decorations. 3. While Will is gone, Mark moves the star to here where it is covered by tinsel. 4. Then Will comes back in. Where will he look for the star?

11) 1. Mark and his friend Maria have just been clothes shopping and bought some socks. 2. Mark folds up their socks and puts them in the drawer here. He then goes outside to get some more shopping from the car. 3. While Mark is outside, Maria moves the socks to here. 4. Mark then comes back in. Where will he look for the socks?

14) 1. Rose and her friend Maria are running a bubble bath for their housemate who has had a stressful day. 2. Rose picks up the soap from the side and accidentally drops it here. She then leaves to go and get a towel. 3. While Rose is gone, Maria reaches into the bath and knocks it over to here. 4. When Rose comes back in, where will she look for the soap?

17) 1. Maria and her friend Rose are making a tray bake cake. 2. Maria buries an almond here and then leaves to get a drink of water. 3. While Maria is gone, Rose moves the almond and buries it here. 4. Then Maria comes back. Where will she look for the almond?

19) 1. Rose and her friend Will have just come home from the shops and are putting the groceries away. 2. Rose puts the frozen peas in the freezer here and

then leaves to go to the toilet. 3. While Rose is in the toilet, Will moves the peas to here. 4. Then Rose comes back in. Where will she look for the peas?

Memory control scenarios

4) 1. Will and his friend Mark are taking down their clothes from the washing line. 2. Will puts his t-shirt here in the basket and then goes inside to get a drink. 3. While Will is inside, Mark moves the t-shirt to here. 4. Then Will comes back outside. Where did he put the t-shirt before going to get a drink?

5) 1. Mark and his friend Maria have just come home from the supermarket and are putting the food away. 2. Mark puts the cheese away in the fridge here and then leaves to go to the toilet. 3. While Mark is in the toilet, Maria moves the cheese to here. 4. Then Mark comes back in. Where did he put the cheese before going to the toilet?

9) 1. Maria and her friend Mark are out camping and are halfway through setting up the tent. 2. Maria puts the peg bag here at the front of the tent. Then, she leaves to go and peg round the back. 3. While Maria is pegging, Mark opens the bag to get a peg and then places the bag down here. 4. Then Maria comes back round the front. Where did she put the peg bag before leaving?

8) 1. Mark and his friend Will have just come back from IKEA and are putting their new plates away in the cupboard. 2. Mark puts the plates here and then goes back to the car to get the rest of the items. 3. While Mark is outside, Will moves

the plates to here in the cupboard. 4. Then Mark comes back in. Where did he put the plates before going outside?

13) 1. Rose and Maria have just finished watching Toy Story on DVD. 2. Rose puts the DVD in the box here and then goes to the kitchen to put the kettle on. 3. While Rose is in the kitchen, Maria moves the DVD to here. 4. Then Rose comes back in. Where did she put the DVD before going to put the kettle on?

15) 1. Rose and her friend Will are packing a suitcase before their flight. 2. Rose puts her wash bag here in the suitcase and then goes downstairs to get her shoes. 3. While Rose is downstairs, Will moves her wash bag to here. 4. Then Rose comes back in. Where did she put her wash bag before going downstairs?

16) 1. Will and his friend Rose are preparing for their friend’s birthday party by wrapping up their joint present together. 2. After wrapping, Will puts the present in the drawer here to keep it safe. He then leaves to go and get ready. 3. While Will is getting ready, Rose opens the drawer to check the present and moves it to here. 4. Then Will comes down to check on the present. Where did he put the present before leaving to go and get ready?

18) 1. Maria and her friend Rose are clearing out old clutter that they don’t need anymore. 2. Maria puts her old teddy bear in the bin here and then goes inside to get more clutter. 3. While Maria is inside, Rose moves the bear to here in the bin. Then Maria realises she wants to keep the bear. 4. She goes back to the bin. Where did she put the bear before going inside?

True belief scenarios

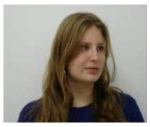
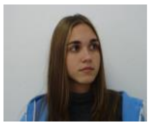
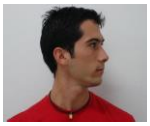
1) 1. Mark and his friend Will have just come home from the shops and are putting their groceries away. 2. Mark puts the ice cream in the freezer here. 3. In view of Mark, Will moves the ice cream to here to make space for more items. Mark then leaves to go to the toilet. 4. Then Mark comes back down. Where will he look for the ice cream?

7) 1. Will and his friend Mark are planting flowers in the garden plot. 2. Will plants the flower here. 3. While Will is watching, Mark decides to move the flower to here. Will then leaves to go and blow his nose. 4. When Will comes back out, he wants to look at the flower. Where will he look for it?

12) 1. Rose and her friend Maria are putting the plates away in the cupboard. 2. Rose puts the plates here. 3. While Rose is watching, Maria moves the plates to here. Rose then leaves to get her phone from the other room. 4. Then Rose comes back and wants to get a plate. Where will she look for them?

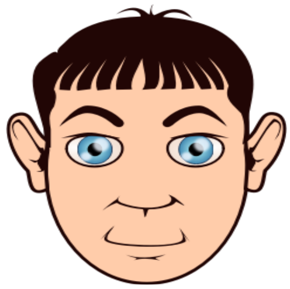
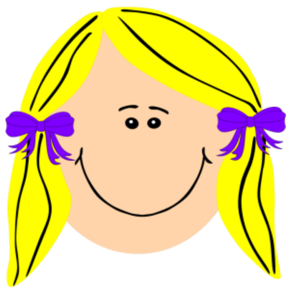
20) 1. Maria and her friend Rose are taking the Christmas decorations down. 2. Maria takes a decoration and places it here in the box. 3. While Maria is watching, Rose moves the decoration to here. Maria then leaves to get a snack. 4. Then Maria comes back. Where will she look for the decoration?

Appendix D - Faces shown in the Director task study (Chapter 3.4)

	-90°	-45°	0°	+45°	+90°
Prac F					
Main F1					
Main F2					
Main M1					
Main M2					
Dir prac F	X	X		X	X

Prac: used for practice trials in the evaluative conditioning task; Main: used for main trials in the evaluative conditioning task. The highlighted photographs of 'Main F1', 'Main F2', and 'Dir F3' were used for the Director task.

Appendix E - Cartoon images used for characters in the Sandbox task (Chapter 5.4.2.2)

Scenario	**‘Placer’ character**	**‘Mover’ character**
Practice 1		
Practice 2		
1		
2		

3		
4		
5		
6		
7		
8		

9		

Appendix F - Change-of-location scenarios used in the Sandbox task (Chapter 5.4.2.2)

Numbers in each scenario below represent sequentially presented sections. The filler task was presented between section 5 and 6.

False Belief scenarios

Practice 1) 1. Paula and her friend Laura are burying flower bulbs. 2. Paula buries a flower bulb in the sandbox, as shown below. 3. After that, Paula goes to the shed to get a watering can and Laura is on her own. 4. In her absence, Laura decides to move the flowerbulb and bury it in a new location. 5. Laura leaves the flower bulb here and goes inside to get a drink. 6. When Paula returns with the watering can, where will she look for the flower bulb ?

Main 3) 1. Steven and Kevin are tidying the living room. 2. Steven puts his book on the bookshelf, as shown below. 3. After that, Steven goes to get the vacuum cleaner and Kevin is left on his own. 4. While Steven is gone, Kevin moves the book to another location on the shelf. 5. Kevin leaves it there and leaves the room to find a duster. 6. When Steven returns with the vacuum cleaner, where will he look for his book?

Main 4) 1. Philip and Tom are unpacking after a holiday. 2. Philip puts his suitcase under his bed, as shown below. 3. After that, Philip leaves the room to do his laundry and Tom is left on his own. 4. While Philip is gone, Tom moves the suitcase to another location under Philip's bed. 5. Tom leaves it there and leaves the room to bring his laundry downstairs. 6. When Philip returns after starting his laundry, where will he look for the suitcase ?

Main 6) 1. Sandra and her sister Amelia are standing next to their mum's freezer. 2. Sandra places a tub of ice cream in the freezer. 3. After that, Sandra goes to ask Mum about dinner time and Amelia is on her own. 4. In her absence, Amelia decides to store the tub of ice cream in a different location. 5. Amelia leaves the ice cream here in the freezer and goes to help Mum set the table. 6. When Sandra returns to take the ice cream out again for dessert, where will she look for the item?

Main 7) 1. Dean and Shane are putting things in the attic. 2. Dean stores a box of old toys in the attic, as shown below. 3. After that, Dean goes downstairs to find more items and Shane is left on his own. 4. While Dean is gone, Shane moves the box of toys to another part of the attic. 5. Shane leaves it there and goes to help Dean with the other items. 6. When Dean returns to the attic, where will he look for the box of toys ?

Memory Control scenarios

Practice 2) 1. Sam and Jack are organising their old toys. 2. Sam puts a toy in the box, as shown below. 3. After that, Sam leaves to find another box and Jack is on his own. 4. In his absence, Jack adds a ball to the box, in a different location.

5. Jack leaves it there and leaves the room to look for more items. 6. Sam then returns to the room. Where did he put the toy before he left?

Main 1) 1. John and Michael are playing in the sandbox. 2. John buries a coin in the sandbox, as shown below. 3. After that, John goes to the shed to get a spade and Michael is on his own. 4. In his absence, Michael decides to bury a stone in a different location. 5. Michael then goes inside to get a snack. 6. John then returns with the spade. Where did he bury the coin before he left?

Main 2) 1. Jane and Mary have come home from grocery shopping. 2. Jane puts her coat in the hallway closet, as shown below. 3. After that, Jane goes into the kitchen to put away the groceries and Mary is left in her own. 4. While Jane is gone, Mary opens the closet and puts her umbrella in another location within the closet. 5. Mary leaves it there and goes to help Jane with the groceries. 6. Jane then returns after putting away the groceries. Where did she leave her coat before she left?

Main 8) 1. Louise and Kate are tidying their room. 2. Louise puts her shoes in the large storage chest, as shown below. 3. After that, Louise leaves to get a snack and Kate is left on her own. 4. While Louise is gone, Kate puts her hoodie in the storage chest. 5. Kate leaves it there and goes to join Louise. 6. Louise then returns to her room. Where did she leave her shoes before she left?

Main 9) 1. Ben and Jake are putting away the dishes. 2. Ben puts the frying pan in the cupboard, as shown below. 3. After that, Ben leaves to answer the phone and Jake is on his own. 4. While Ben is gone, Jake puts a bowl in another location inside the cupboard. 5. Jake leaves it there and goes to watch television. 6. Ben then returns from answering the phone. Where did he leave the frying pan before he left?

True belief scenario

Main 5) 1. Nancy and Claire accidentally break their sister's necklace. 2. Nancy hides the necklace in a drawer, as shown below. 3. After that, Nancy wants to leave the room to see where their sister is. 4. However, before Nancy leaves, she sees Claire moves the necklace to another location inside the drawer. 5. Claire leaves it there and follows Nancy to see where their sister is. 6. When Nancy returns, where will she look for the necklace?

Appendix H - Scenarios used in the Hinting task (Section 5.4.2.4) with possible correct answers

Scenario 1

George arrives in Angela's office after a long and hot journey down the motorway. Angela immediately begins to talk about some business ideas. George interrupts Angela saying: "My, my! It was a long, hot journey down that motorway!"

Question: What does George really mean when he says this?

Answer: George means either "Can I have a drink" and/or "Can I have a few minutes to settle down after my journey before we start talking business".

Add: George goes on to say: "I'm parched!"

Question: What does George want Angela to do?

Answer: George wants Angela to get him or offer to get him a drink.

Scenario 2

Melissa goes to the bathroom for a shower. Anne has just had a bath. Melissa notices the bath is dirty so she calls upstairs to Anne: "Couldn't you find the Ajax, Anne?"

Question: What does Melissa really mean when she says this?

Answer: Melissa means “Why didn’t you clean out the bath” or “Go and clean out the bath now”.

Add: Melissa goes on to say: "You're very lazy sometimes, Anne!"

Question: What does Melissa want Anne to do?

Answer: Melissa wants Anne to clean out the bath.

<u>Scenario 3</u>

Gordon goes to the supermarket with his mum. They arrive at the sweetie aisle. Gordon says: "Cor! Those treacle toffees look delicious."

Question: What does Gordon really mean when he says this?

Answer: Gordon means “Please buy me some sweets, mum”

Add: Gordon goes on to say: “I'm hungry, mum."

Question: What does Gordon want his mum to do?

Answer: Buy him some sweets.

<u>Scenario 4</u>

Paul has to go to an interview and he's running late. While he is cleaning his shoes, he says to his wife, Jane: "I want to wear that blue shirt but it's very creased."

Question: What does Paul really mean when he says this?

Answer : Paul means “Will you iron my shirt for me please?”

Add: Paul goes on to say: "It's in the ironing basket."

Question: What does Paul want Jane to do?

Answer: Iron his shirt.

<u>Scenario 5</u>

Lucy is broke but she wants to go out in the evening. She knows that David has just been paid. She says to him: "I'm flat broke! Things are so expensive these days."

Question: What does Lucy really mean when she says this?

Answer: Lucy means “Will you lend me some money David ?” or “Will you take me out tonight and pay?”

Add: Lucy goes on to say: "Oh well, I suppose I'll have to miss my night out."

Question: What does Lucy want David to do?

Answer: She wants David to lend her money or offer to take her out and pay.

<u>Scenario 6</u>

Donald wants to run a project at work but Richard, his boss, has asked someone else to run it. Donald says: "What a pity. I'm not too busy at the moment."

Question: What does Donald really mean when he says this?

Answer: Donald means “ Please change your mind Richard and give the project to me”

Add: Donald goes on to say: "That project is right up my street."

Question: What does Donald want Richard to do?

Answer: Change his mind and give the project to him to run.

<u>Scenario 7</u>

Rebecca's birthday is approaching. She says to her Dad: "I love animals, especially dogs."

Question: What does Rebecca really mean when she says this?

Answer: “Will you buy me a dog for my birthday Dad?”

Add: Rebecca goes on to say: "Will the pet shop be open on my birthday, Dad?"

Question: What does Rebecca want her dad to do?

Answer: To say he’ll buy her a dog for her birthday / buy her a dog for her birthday

<u>Scenario 8</u>

Betty and Michael moved into their new house a week ago. Betty has been unpacking some ornaments. She says to Michael: "Have you unpacked those shelves we bought, Michael?"

Question: What does Betty really mean when she says this?

Answer: Betty means “Will you put those shelves up now please?”

Add: Betty goes on to say: "If you want something doing you have to do it yourself!"

Question: What does Betty want Michael to do?

Answer: Put the shelves up.

Scenario 9

Jessica and Max are playing with a train set. Jessica has the blue train and Max has the red one. Jessica says to Max: "I don't like this train."

Question: What does Jessica really mean when she says this?

Answer: Jessica means “I want your train and you can have mine.”

Add: Jessica goes on to say: "Red is my favourite colour."

Question: What does Jessica want Max to do?

Answer: Swap trains.

Scenario 10

Patsy is just getting off the train with three heavy cases. John is standing behind her. Patsy says to John: "Gosh! These cases are a nuisance."

Question: What did Patsy really mean when she said this?

Answer: Patsy means “ Would you help me with my luggage please”

Add: Patsy goes on to say: "I don't know if I can manage all three."

Question: What does Patsy want John to do?

Answer: Help her with her cases.

Appendix I - Description of the subject matter of the videos in the Video rating task (Chapter 5.4.2.6)

Video number	**Narrator ID***	**Valence**	**Narrator gender**	**Subject matter**	**Length (m:s)**
1	4	Positive	Female	Mother was walking again after long period of rehabilitation	1:13
2	12	Positive	Female	Getting accepted to a University	1:36
3	18	Positive	Male	Enjoying Reading festival and making friends	1:18
4	20	Positive	Female	Got a puppy	1:41
5	25	Positive	Male	Bonding with a colleague and a warm encounter later on	1:36
6	25	Positive	Male	Making music with a friend and finding out a friend listened to the mixtape on repeat	1:20
7	12	Negative	Female	Parents moving away from her	1:33
8	24	Negative	Female	Crashing car into another at night	1:16
9	30	Negative	Male	Threatened by a gun while on a trip to a waterfall	1:28
10	31	Negative	Male	Fellow student killed by car	1:42
11	32	Negative	Male	His car was hit by someone else	1:08

12	35	Negative	Female	Grandfather had a stroke on New Year's Eve; difficulty in caring for him afterwards	1:52

* Note: Same ID indicates that the videos had the same narrator

Appendix J - Differences between Wave 1 and Wave 2 for the task battery Director task

Aspect	Wave 1	Wave 2
Calibration	5 points (corners and centre) Calibration only	16 points (centre of all shelves) Calibration and validation (test of calibration quality)
Restart after crash	From the start	From a selected trial
Background behind shelves and director	White	Blue
Response instruction	N/A	Decide on object and destination before mouse activation click
Data monitoring	N/A	Plots visible to experimenter, showing eye gaze pattern and proportion of invalid gaze points

Reasons for differences: the program running the task occasionally crashed. When this happened in Wave 1, the task could only be restarted from the beginning. To avoid attrition and repeated trials, the version for Wave 2 included the option to skip the trials that the participants had already completed. The calibration procedure, response instruction and data monitoring were changed or added to safeguard data quality, in terms of precision of eye-tracker data and participant's adherence to the task instructions. The background was changed to blue to match the physical background behind the experimenter during the task.

Appendix K - Differences between Wave 1 and Wave 2 for the task battery mouse-tracking task

Aspect	Wave 1	Wave 2
Mouse location tracking	Between initialisation of the response trajectory and the button click to proceed to the next trial	Between initialisation of the response trajectory and the response click
Redo practice trials with slow/incorrect response?	Yes	No
In case of incorrect response to the four task understanding questions	Redo two practice trials	Window appears that reminds the participant that they should click on the location of the ball, and that it can sometimes appear in a different place to where they last saw it
Initiation RT (click to start trajectory)	Not saved	Saved

Reasons for differences: mouse location tracking and logging of the initiation reaction time should have functioned in Wave 1 as in Wave 2, but an overlooked error in the script caused suboptimal data recording. This error was rectified for Wave 2. The other two aspects were changed because it was observed in Wave 1 that participants generally showed more task understanding if they were simply reminded after the practice trials to always click on the ball in the task.