UNIVERSITY OF BIRMINGHAM

Individual Differences in Adults' Mindreading: Psychometric Challenges and the Role of Social Motivation

Christina Pomareda Gillessen

This thesis is submitted to the University of Birmingham for the degree of Doctor of Philosophy.

School of Psychology

College of Life and Environmental Sciences

University of Birmingham

November 2023

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as specified in the text. No substantial part of this thesis has already been submitted or is being concurrently submitted for any degree at the University of Birmingham or any other University or similar institution. This thesis does not exceed the word limit prescribed by the relevant Degree Committee.

Summary

The last four decades have witnessed extensive scientific investigation of mindreading (or 'theory of mind'), that is, the socio-cognitive capacity to attribute mental states such as beliefs, desires and intentions to oneself and others. The present research was motivated by the observations that despite most adults having a conceptual understanding of mental states, adults exhibit substantial variation in how and to what extent they engage in mindreading and, secondly, that comprehending the nature of such variation requires the consideration of motivational factors. We investigated individual differences in mindreading in community adult populations both in terms of the accuracy of mindreading, or 'contextually justified mindreading' (see chapter 2), as well as in terms of the processes involved (see chapter 3). To achieve the latter, a unique approach was developed to capture individual differences in the structure of adults' mindreading, irrespective of response accuracy. Instead of relying on existing measures with inadequate psychometric properties, we sought to refine existing measures (see chapters 2) and to develop a novel measure of mindreading validated across two independent samples (see chapter 3). This exposed that there are reliably measurable, content-irrelevant processes and strategies that individuals employ when engaging in mindreading. Finally, this research assessed the role of individual differences in mindreading alongside social motivational factors (see chapters 2 and 4), revealing that there were positive associations between mindreading and social motivation, while both were still distinguishable constructs. The results also showed unique associations between individual differences in social motivation (but not mindreading) and social competence and mental health (see chapter 4). The present findings challenge past epistemological assumptions about how mindreading can be operationalised and measured and provide support for the notion that researchers interested in mindreading should consider social-motivational factors alongside mindreading.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisors, Professor Ian Apperly and Dr Rory Devine, whose generous supervision, continuous support and expertise have made this PhD thesis possible. I feel not only privileged to have learned so much from two brilliant researchers over the past years, but also to have had the fortune to work with thoughtful, kind and caring people. They were simply the best supervisors I could have imagined for a PhD programme, and I will always remember my time as a PhD student under their supervision fondly.

I would also like to thank all the members of my lab group, which is characterised by a spirit of collaboration, support and kindness. I thank Elaine for her help with double coding the data reported in chapter 3 and 4 and Daniel for his assistance with double coding the data reported in chapter 2. I am also particularly grateful to Sanne and Daniel for their advice and the many chats and laughs we had during the past years. I wish you only the best in your careers and lives. I am also thankful to my friends, many of whom have kept in touch despite living far away for many years.

This work is dedicated with immense love to my family. To my parents Andrea and Luis, who always support me and have given me all the love, trust and freedom needed to venture to a different country. To my sisters Isabel, Carmen, and Lucia who are my biggest inspiration. I am also immensely thankful to my partner Nadeem, for his unwavering support, love and encouragement over the course of my PhD and beyond. I feel very lucky to have these people in my life.

Lastly, I would like to thank the Economic Social Research Council (ESRC) for the studentship which allowed me to conduct my PhD research and to present some of the work reported here at an international conference in Brussels this year.

Table of Contents

1 **Introduction 9**

1.1 Key Observations Underpinning the Present Research 9

1.2 Methodological and Conceptual Challenges and Gaps in the Literature 12

1.2.1 *Past Methodological Challenges in the Assessment of Mindreading 12*

1.3 The Importance of Motivational Factors 16

1.3.1 The Correlates of Mindreading and Social Motivation 18

1.3.2 Differentiating Study Constructs 20

1.4 Grounding the Present Research in Theoretical Perspectives and Debates 22

1.5 Overview and Scope of the Present Studies 23

2 **Mindreading Quality Versus Quantity: A Theoretically and Empirically Motivated Two-factor Structure for Individual Differences in Adults' Mindreading 26**

2.1 Introduction 26

2.1.1 Measuring Individual Differences in Adults' Mindreading 27

2.1.2 The Role of Motivation in Individual Differences in Adults' Mindreading 29

2.1.3 Criterion Validity of MST and Response Appropriateness 30

2.1.4 The Present Study 30

2.2 Methods 31

2.2.1 Participants 31

2.2.2 Design and Procedure 32

2.2.3 Materials 32

2.3 Results 38

2.3.1 Analysis Plan 38

2.3.2 Missing Data 39

2.3.3 Preliminary and Descriptive Statistics 39

2.3.4 Main Analysis 40

2.4 Discussion 48

2.4.1 Evidence for a Two-Factor Structure of Mindreading 48

2.4.2 Mindreading Motivation 50

2.4.3 Social Outcome Variables 50

2.4.4 Limitations and Future Directions 51

2.4.5 Conclusion 52

3 Convergence between the MASC and a Novel Task Assessing the Structure of Mindreading 54

3.1 Introduction 54

3.1.1 The Present Studies 58

3.2 Study 1 59

3.2.1 Methods 59

3.2.1.1 Participants 59

3.2.1.2 Measures and Procedure 60

3.2.2 Results 62
3.2.2.1 Analysis Plan 62
3.2.2.2 Descriptive Statistics 63
3.2.2.3 Structural Equation Modelling 68
3.3 Study 2 69
3.3.1 Methods 69
3.3.1.1 Participants 70
3.3.1.2 Measures and Procedure 71
3.3.2 Results 72
3.3.2.1 Analysis Plan 72
3.3.2.2 Descriptive Statistics and Correlations 72
3.3.2.3 Establishing a Factor Structure for Mindreading 75
3.3.2.4 Establishing Convergent Validity 78
3.3.2.5 Structural Equation Models 80
3.3.3 Discussion 81
3.3.3.1 Sensitivity of the New Measure 81
3.3.3.2 Factor Structure of the New Measure 81
3.3.3.3 Evidence of Convergence Between the New Measure and MASC 82
3.3.3.4 Methodological and Theoretical Implications 82
3.3.3.5 Conclusion 85
4 Social Motivation, not Mindreading, is associated with Social and Mental Health Variables 88
4.1 Introduction 88
4.1.1 The Present Study 91
4.2 Method 92
4.2.1 Social Motivation Measures 93
4.2.2 Mental Health Measures 94
4.2.3 Social Outcome Measures 95
4.3 Results 95
4.3.1 Correlational Analysis 95
4.3.2 Mindreading Latent Factor 98
4.3.3 Social Motivation Latent Factor 100
4.3.4 Structural Equation Models 101
4.3.5 Moderation Analysis 103
4.4 Discussion 107
4.4.1 Mindreading, Social Competence and Mental Health 108
4.4.2 Mindreading and Social Motivation 110
4.4.3 Conclusion 111
5 General Discussion 113
5.1 Summary and Synthesis 114
5.1.1 Overview of Aims 114
5.1.2 Summary of Findings 115
5.1.3 Implications of the Present Findings and Recommendations for Future Research 118
5.1.3.1 Mindreading can be Assessed Irrespective of Response Accuracy 119
5.1.3.2 Mindreading is Underpinned by Multiple Factors 120

5.1.3.3 Mindreading Research and Data-Driven Explorations 121
5.1.3.4 Mindreading Should be Studied Alongside Social Motivation 122
5.2 Limitations 123
5.3 Conclusions 125
6 Appendix 129
7 References 154

Chapter 1

Introduction

1 Introduction

1.1 Key Observations Underpinning the Present Research

Humans possess a remarkable capacity to decipher the observed behaviours of others. This broad competence is termed 'social cognition' and encompasses a multi-dimensional and complex array of cognitive mechanisms that underlie diverse facets of human social behaviour. These include social decision-making, emotion recognition, social perception, empathy, and mindreading (Saxe, 2006). Mindreading, or the ability to infer one's own and others' mental states such as thoughts and feelings, is a pervasive everyday phenomenon that is thought to be significant to people's lives. For example, empirical evidence suggests a positive link between adept mindreading and effective social functioning (e.g., Imuta, Henry, Slaughter, Selcuk, & Ruffman, 2016; Slaughter, Imuta, Peterson, & Henry, 2015; Fink, Begeer, Peterson, Slaughter, & de Rosnay, 2015; Hughes & Devine, 2015a; Devine & Apperly, 2022). Conversely, deficits in mindreading are associated with various psychiatric and neurodevelopmental conditions, such as autism (e.g., Baron-Cohen, Leslie, & Frith, 1985; Bora, Bartholomeusz, & Pantelis, 2016). Accordingly, research into mindreading has been a focal point of extensive investigation within the field of social cognition that has motivated substantial theorising and empirical inquiry by scholars of psychology, philosophy, and cognitive science for over four decades (e.g., Quesque & Rossetti, 2020).

The present dissertation was based on two primary observations. Firstly, psychologists have often noted *variability* in terms of how individuals process social information from their environment. Even when exposed to identical stimuli, individuals tend to engage with and interpret information in distinct ways, shaped by their unique cognitive tendencies. Indeed, pronounced individual differences can be observed regarding the degree to which individuals mindread accurately (e.g., Apperly, 2012; Conway, Catmur, & Bird, 2019; Devine, 2021; Dziobek et al., 2006; Hughes & Devine, 2015; Devine, 2021). Such variability in mindreading persists in adults while most individuals demonstrating an understanding of mental states as early as the age of four or five years of age (Weimer, Warnell, Ettekal, Cartwright, Guajardo, & Liew, 2021; for a meta-analysis of 178 studies using false belief tasks, see Wellman, Cross, & Watson, 2001). Therefore, despite the claim that "*(…) we are all mind readers*" (Ames, 2004, p. 340), and the foundational role of a conceptual understanding of mental states for mindreading to occur, there remains significant heterogeneity in this socio-cognitive

phenomenon, the origin of which remains to be *understood* (Apperly, 2012). An over-arching aim of this dissertation was therefore to gain a better understanding of why individual differences in mindreading persist among adults, who have already attained the capacity to conceptualise mental states (e.g., Apperly, 2012; Conway et al., 2019; Dziobek et al., 2006).

Another crucial observation that guided the present research is that a person's motivation has a profound impact on their behaviour (e.g., Hess, 2014). Motivation encompasses complex psychological processes that drive individuals toward their goals and can be thought of as mental representations of desired outcomes (e.g., Linnenbrink-Garcia, & Patall, 2015; Gardner & Tremblay, 1994). Consequently, the extent to which individuals engage in mentalising, and the way in which they mentalise, are likely inherently linked to motivational factors (Carpenter, Green, & Vacharkulkesemsuk, 2016; Contreras-Huerta, Pisauro, & Apps, 2020). Although human behaviour is not solely determined by motivations, these motivational factors, in conjunction with an individual's unique cognitive dispositions (i.e., enduring patterns in how stimuli are processed and responded to; Mischel & Shoda, 1995), likely contribute the observed variation in mentalising. Therefore, individual differences in mindreading are unlikely to be a direct reflection of an individual's ability to accurately mentalise. Surprisingly, there has been very limited empirical or theoretical integration between motivational accounts and socio-cognitive investigations in general, as well as specifically regarding mindreading and motivation (for recent exceptions to this, see Carpenter et al., 2016; Contreras-Huerta et al., 2020; Devine & Apperly, 2022; Lockwood, Ang, Husain, & Crockett, 2017). This lack of integration may be attributed to the fact that, traditionally, motivational accounts of human behaviour have been and continue to be a primary focus of social psychology, while cognitive psychology is more concerned with human abilities. A further important goal of the present research was therefore to examine individual differences in adults' mindreading both separately and in conjunction with motivational factors.

This dissertation therefore built on the observations that (1) individuals vary in the degree to which they mentalise, and (2) that understanding the nature of such variation in mindreading requires the consideration of motivational factors. More specifically, in the present research, we sought to investigate *individual differences* in mindreading in community adult populations both in terms of the *accuracy* of mindreading, or 'contextually justified mindreading' (see chapter 2), as well as in terms of the *processes* involved in mentalising (see chapter 3). To achieve the latter, we developed a unique approach to capturing individual

differences in the structure of adults' mindreading, independently of individuals' ability to accurately infer mental states (see chapter 3). Rather than relying on existing measures of mindreading – which are oftentimes characterised by inadequate psychometric properties – we aimed to refine existing measures (see chapters 2) and to develop a novel measure of mindreading (see chapter 3). Finally, instead of investigating mindreading and its correlates in isolation, we focussed on assessing the role of individual differences in mindreading *alongside* social motivational factors (see chapters 2 and 4), with the overarching aim of investigating the *links* between social motivation and mindreading, and their respective relationships with social abilities and mental health variables (see chapter 4).

What is mindreading?

In essence, mindreading can be defined as the ability to attribute mental states such as beliefs, desires and intentions to oneself and others, and the capacity to recognise the potential disparities between one's own mental states and those of others (e.g., Wellman, 1990). However, the proliferation of research across a multitude of areas, ranging from typical human development (e.g., Devine & Hughes, 2016), to clinical populations (e.g., Eddy & Cavanna, 2015; Montag et al., 2010), and extending to nonhuman animals (e.g., Krupenye, & Call, 2019), has given rise to a wide array of terms and operationalisations of the construct. Consequently, the study of mindreading exhibits significant fragmentation. Premack and Woodruff (1978) first used the term "theory of mind" in their seminal article *"Does the chimpanzee have a theory of mind?"* in reference to chimpanzee's ability to infer human goals. Subsequently, psychologists adopted terms, including but not restricted to: 'mindreading' (e.g., Apperly, 2010), 'perspective-taking' (e.g., Hamilton, Brindley, & Frith, 2009), 'social understanding' (e.g., Carpendale & Lewis, 2006), and 'mentalising' (Frith & Frith, 2003). These terms are often used interchangeably. However, it is plausible that the different terms reflect distinct ontological assumptions about the nature of the construct they refer to, influencing perceptions of how a construct should be operationalised**.** For example, 'perspective-taking' emphasises the cognitive process of adopting another individual's viewpoint to comprehend their mental states, whereas 'social understanding' highlights a focus on the understanding of social cues (e.g., Samson, Apperly, Braithwaite, Andrews, & Bodley Scott, 2010). To address these issues of terminology and operationalisations in the context of the scope of the present dissertation, throughout the following pages, ontological and epistemological assumptions about relevant constructs are delineated. This approach seeks to facilitate comparability with the wider

academic literature and enable a more comprehensive understanding of the phenomenon under investigation. For terminological consistency, the term "mindreading", occasionally used synonymously with "mentalising" or "theory of mind", is adopted throughout this dissertation.

To ground the present research in previous findings and theoretical frameworks, the following sections will (1) offer an overview and evaluation of the current methodological and conceptual challenges to mindreading (section 1.2.1), outline the importance of considering motivational factors in the study of mindreading (section 1.3), discuss correlates of mindreading (section 1.3.2), differentiate study constructs (section 1.3.3), and ground the present research in theoretical perspectives and debates (section 1.4), before discussing the scope of the present research and providing an overview of the present studies (section 1.5).

1.2 Methodological and Conceptual Challenges and Gaps in the Literature

1.2.1 Past Methodological Challenges in the Assessment of Mindreading

Sensitivity to detect individual differences

Prior to the expansion of research into individual differences in both children's (e.g., Devine & Hughes, 2013) and adults' (e.g., Diziobek et al., 2006) mindreading, the capacity to infer one's own and others' mental states such as thoughts and feelings was thought of as an 'all-or-nothing-affair', which typically developing children would acquire by middle childhood and autistic individuals 'lack' (e.g., Povinelli & Povinelli, 1996; Senju, Southgate, White, & Frith, 2009; Baron-Cohen, 2001). However, theoretical and empirical insights have revealed that there is (a) "*developmental life after false belief*" (Carpendale & Lewis, 2006, p. 185), (b) that the relationship between mindreading and autism is more nuanced than originally asserted (e.g., Tager-Flusberg, 2007), and that there are (c) meaningful individual differences in this capacity in children (e.g., Dunn & Cutting, 1999; Dunn, Brown, Slomkowski, Tesla, & Youngblade, 1991; Dunn, & Brophy, 2005). Researchers have therefore devised a range of so-called 'advanced' mindreading tasks, that are more subtle and complex in nature to capture continuing development beyond the pre-school years in mindreading in middle childhood and adolescence (e.g., Devine & Hughes, 2013) and in adulthood (e.g., Dziobek et al., 2016; Castelli, Happé, Frith, & Frith, 2001; Conway et al., 2019; Livingston et al., 2021). However,

many of these tasks were originally designed for comparing clinical and non-clinical groups (e.g., Bradford, Hukker, Smith, & Ferguson, 2018; Preißler, Dziobek, Ritter, Heekeren, & Roepke, 2010; Livingston et al., 2021), which complicates the detection of within-group variability in broader adult populations. Nevertheless, certain tasks, such as the Movie for the Assessment of Social Cognition (MASC; Dziobek et al., 2016) have demonstrated sensitivity to detect *individual differences* in adults' mindreading (also see: Conway et al., 2019; Livingston et al., 2021; Devine & Hughes, 2019; Slaughter & Repacholi, 2004).

Convergent validity

The first task used to assess mindreading was the 'false-belief task', requiring pre-school children to either remember their own false beliefs or attribute incorrect beliefs to others (e.g., Wimmer & Perner, 1983). Subsequently, a wide array of measures was developed, including 'advanced mindreading tasks' aimed at capturing individual differences in mindreading in individuals who already demonstrated the ability to pass standard false belief tasks, such as adults (Wellman et al., 2001). Collectively, these measures appeared to assess various facets of mindreading, including the ability to infer mental states from photographs of individuals' eyes (e.g., Baron-Cohen et al., 2001), visual perspective-taking (e.g., Samson et al., 2010), recognition of deceptive intentions (e.g., Sebanz & Shiffrar, 2009), or detection of social faux pas (e.g., Baron-Cohen, O'Riordan, Stone, Jones, & Plaisted, 1999).

Despite the diverse focus of these measures, they have been employed in various studies with the common goal of assessing one specific construct – mindreading (e.g., Warnell & Redcay, 2019). However, a study by Warnell and Redcay (2019) that examined the relationship among such different measures in pre-schoolers, school-aged children and adults found that among the adult group, no single latent factor underlying mindreading could be identified (in children, weak-non-existent correlations between tasks were observed, albeit with sample size limitations preventing confirmatory factor analysis). These results provide two key insights. Firstly, current mindreading measures demonstrate limited convergent validity, meaning that performance on one measure does not necessarily correlate with performance on another measure (e.g., Warnell & Redcay, 2019; Devine 2021), and, secondly, that mindreading may not represent a single construct, a possibility that has also been raised by several other researchers in the field (e.g., Apperly, 2012; Schaafsma, Pfaff, Spunt, & Adolphs, 2015).;

Warnell & Redcay, 2019; Abu-Akel & Shamay-Tsoory, 2011; Apperly & Butterfill, 2009; Bernhardt & Singer, 2012).

Construct Validity

Notably, there is robust evidence of convergent validity for measures of mindreading in children, particularly for false-belief mindreading tasks (e.g., Cutting & Dunn, 1999; Mayes, Klin, Tercyak Cicchetti, & Cohen, 1996; Hughes, Adlam, Happé, Jackson, Taylor, & Caspi, 2000). In contrast, advanced mindreading tasks developed for adults often lack convergent validity (i.e., performance on one task does not correlate with performance on another task). One plausible explanation for this lack of convergence among many mindreading measures is that certain tasks originally designed to assess mindreading likely tap into other constructs. For instance, the Animations Task (Abell, Happe, & Frith, 2000; Castelli et al., 2001), which requires participants to explain the movements of triangles that exhibit either a simple interaction (e.g., two triangles appearing to dance), or more complex interactions (e.g., one triangle appearing to surprise another), has also been utilised to assess other constructs, such as anthropomorphism (e.g., Waytz, Morewedge, Epley, Monteleone, Gao, & Cacioppo, 2010; Tahiroglu & Taylor, 2019). Furthermore, Oakley, Brewer, Bird, and Catmur (2016) found that the Reading the Mind in the Eyes Test (RMET; Baron-Cohen et al., 2001), which involves participants selecting appropriate adjectives to match pictures of faces expressing various emotions, likely assesses emotion perception rather than mindreading.

Oakley et al.'s (2016) findings revealed a moderate significant link between RMET performance and alexithymia, a condition characterised by difficulties in understanding emotions, while autistic traits, which are usually related to mindreading, did not significantly influence RMET performance based on a sample of 42 participants, 19 of which had been diagnosed with autism spectrum disorder (ASD) and 23 made up a control group. This pattern of results was reversed for another established measure of mindreading, namely the Movie for the Assessment of Social Cognition (MASC; Dziobek et al., 2016), which requires participants to reason about the mental states of four characters shown in a realistic social interaction based on a multiple-choice question regarding the characters' mental states. Overall, these findings emphasise the challenges to establishing construct validity in the assessment of mindreading. They suggest that some widely used tasks may not exclusively tap into the intended construct of mindreading and instead capture related but distinct constructs.

Criterion Validity

Mindreading is regarded as '*an essential ability for social competence and communication*' (Navarro, 2022, p. 1). However, the presence of psychometric limitations in existing measures poses a significant obstacle to understanding the relationships between individual differences in adults' mindreading and other variables, that is, establishing criterion validity. While concerns regarding the psychometric characteristics of mindreading tasks have been raised in general (e.g., Warnell & Redcay, 2019), in comparison to the substantial body of evidence from developmental studies (e.g., Eisenberg, 2003; Devine & Apperly, 2022; Hughes & Devine, 2015a; Leslie, Friedman, & German, 2004), and clinical research (e.g., Bradford et al., 2018; Dziobek et al., 2006; Preißler et al., 2010; Livingston et al., 2021; Bora, & Berk, 2016, for a meta-analysis), relatively *fewer* studies assessing mindreading in relation to other constructs in broader adult populations have been conducted, with some notable exceptions (e.g., Devine & Hughes, 2019; Slaughter & Repacholi, 2004; Long, Cuve, Conway, Catmur & Bird, 2022; Abu-Akel, Wood, Hansen, & Apperly, 2015; Stiller & Dunbar, 2007; Richardson, Green, & Lago, 1998; Mulvey, Rizzo, & Killen, 2016).

For example, research has shown negative links between individual differences in adults' mindreading and autistic and psychotic characteristics (e.g., Abu-Akel et al., 2007), and autistic traits (e.g., Baron-Cohen et al., 2001). Such findings provide evidence for 'known-groups-validity', that is, they allow for the assessment of criterion validity based on the mindreading test performance of groups who are known to experience difficulties with social and communicative skills (e.g., individuals with schizophrenia). Specifically, consistent differences in performance on mindreading tests between such groups and control groups would be considered as providing some support for the criterion validity of a test (e.g., Hattie & Cooksey, 1984; Heimberg & Holaway, 2007). Furthermore, there are also positive associations with reduced aggression (e.g., Richardson et al., 1998) and stereotyping (e.g., Mulvey et al., 2016). However, the reliance on problematic assessment tools has likely biased some previous research findings, particularly with regard to understanding the nature of individual differences in adults' mindreading. To illustrate this concern, lower performance on the Reading the Mind in the Eyes Test (RMET; Baron-Cohen et al., 2001) has been associated with childhood adversity (e.g., Germine, Dunn, McLaughlin, & Smoller, 2015). Yet, as discussed earlier, the RMET likely assesses emotion perception rather than mindreading (e.g.,

Oakley et al., 2016). Consequently, empirical progress in the field necessitates revisiting the methods of established findings from previous research.

More specifically, uncovering potential links between individual differences in adults' mindreading and other constructs requires the use of mindreading tasks with psychometrically robust characteristics, encompassing sensitivity to detect task performance variation, convergent validity, and construct validity. Moreover, such tasks need to be administered in conjunction with assessments of other constructs that likewise demonstrate validity in the context of broader adult populations.

1.3 The importance of Motivational Factors

A fundamental premise of the present dissertation is the idea that elucidating the role of motivational factors within the context of mindreading is necessary for advancing our understanding of *why* these individual differences persist in adults who have already acquired the capacity to conceptualise mental states (relevant to chapter 2 and 3) and whether these individual differences hold meaningful associations with other constructs (relevant to chapter 2 and 4).

What is Social Motivation?

Motivation encompasses psychological processes that drive individuals toward their goals, effectively functioning as mental representations of desired outcomes (e.g., Linnenbrink-Garcia, & Patall, 2015; Gardner & Tremblay, 1994). Individual differences in motivation have been associated with various constructs, including empathy for others (e.g., Cameron, Scheffer, Hadjiandreou, & Anderson, 2022), mental health (e.g., Ratelle, Vallerand, Chantal, & Provencher, 2004), and educational success (Côté & Levine, 2000). Motivation manifests in two primary forms: as a 'trait', reflecting an individual's enduring predisposition to consistently pursue goals across different situations or over time (e.g., Wasserman & Wasserman, 2020), and as a 'state', which can fluctuate within the same individual based on the extent to which situational factors act as incentives for behaviour (e.g., Harris, Brett, Johnson, & Deary, 2016; Apperly & Wang, 2021, in Ferguson & Bradford, 2021).

Within the context of mindreading research, 'trait-like' motivations have been termed 'social motivation' (e.g., Chevallier, Kohls, Troiani, Brodkin, & Schultz, 2012) or 'mindreading motivation' (Carpenter et al., 2016). Social motivation encompasses the extent to which individuals are inclined to nurture and sustain social connections, prefer social interactions over non-social ones, and derive pleasure from social engagements (e.g., Chevallier et al., 2012), while mindreading motivation is defined as a stable inclination to engage with others' minds and mental states (Carpenter et al., 2016). Research suggests that social motivation (relevant to chapter 4) is a separable construct from general motivation (this has also been termed 'behavioural activation') and 'emotional sensitivity', that is, individuals' feelings of positive and negative affect (Ang et al., 2017). Mindreading motivation (relevant to chapter 2) is an even more specific facet of social motivation.

The relationship between Social Motivation and Mindreading

Understanding the links between social motivation and mindreading was a central focus of the current dissertation. Mindreading tasks, much like mindreading in real-life situations, implicate the engagement of individuals' memory and cognitive control processes (e.g., Apperly, Riggs, Simpson, Chiavarino, & Samson, 2006; Cane, Ferguson, & Apperly, 2017; Lin, Keysar, & Epley, 2010), effectively requiring effort and, consequently, motivation (e.g., Contreras-Huerta et al., 2020). Therefore, observed individual differences in mindreading should not be equated with an individual's capacity for accurate mentalising. Although this consideration has often been overlooked by existing studies in the field of mindreading, with a few notable exceptions (e.g., Carpenter et al. 2016; Devine & Apperly, 2022; Lockwood et al., 2017), it aligns with empirical evidence of positive associations between social motivation and performance on socio-cognitive tasks, including mindreading (e.g., Carpenter et al. 2016; Devine & Apperly, 2022) as well as cognitive empathy (e.g., Lockwood et al., 2017). Furthermore, this perspective is in line with the notion that individuals on the autism spectrum may perform less well on mindreading tasks than their peers due to diminished social motivation rather than reduced social ability (as proposed by the social motivation hypothesis**;** Chevallier et al., 2012). Therefore, a key objective that guided the present research was to establish whether adults vary independently in their mindreading/social motivation and in their mindreading performance, and to elucidate the potential connections between these two constructs.

To date, only one study in 8-to-13-year-old children (e.g., Devine & Apperly, 2022) has investigated the unique and relative impacts of social motivation and mindreading on children's social competence. This study found that children's social motivation (i.e., self-reported willingness to build and maintain social relationships) and mindreading ability made distinct contributions to predicting variation in teacher-rated social ability. However, there is currently a knowledge gap concerning (a) the relative contributions of individual differences in the domains of social motivation and mindreading to other meaningful constructs such as mental health and (b) the extent to which motivation and mindreading represent separable or overlapping constructs in *adults*. Therefore, an overarching aim of this dissertation was to investigate the *links* between social motivation and mindreading, and their respective relationships with other constructs (outlined below).

1.3.1 The Correlates of Mindreading and Social Motivation

In addition to investigating the links between social motivation and mindreading, a further aim of the present research was to investigate the unique relationships of each construct with a range of variables that may be related to individual differences in both social motivation and mindreading. As discussed earlier, it is important to acknowledge that prior findings have not always been consistent and so may have been influenced by the psychometric limitations of certain mindreading tasks (e.g., Schaafsma et al., 2015). Nevertheless, past research, especially when also drawing insights from developmental and clinical studies, still provides valuable information about potentially linked variables. The specific correlates we selected, briefly outlined below, broadly fall into broader theoretical frameworks encompassing the *Social Individual Differences Account* (e.g., Apperly, 2012; Hughes & Devine, 2015; Dunn & Cutting, 1999; Dunn et al., 1991; Dunn, & Brophy, 2005), the *Transdiagnostic Account* (e.g., Strikwerda-Brown, Ramanan, & Irish, 2019), and *Motivational Accounts* and *Cognitive Tendencies* (e.g., Scholl & Tremoulet, 2000; Villalobos, Padilla-Mora, & Trías, 2010)

Social Individual Differences

According to the social individual differences account (e.g., e.g., Apperly, 2012; Hughes & Devine, 2015; Dunn & Cutting, 1999; Dunn et al., 1991; Dunn, & Brophy, 2005), greater mindreading abilities have significant consequences for social functioning, over and above the influence of other general cognitive functions, which also affect social functioning.

Indeed, loneliness, defined as the perceived absence of a satisfactory emotional connection with others (e.g., Simcharoen, Pinyopornpanish, Haoprom, Kuntawong, Wongpakaran, & Wongpakaran, 2018), has been negatively associated with mindreading abilities in children (e.g., Devine & Hughes, 2013; Koerber & Osterhaus, 2020) and adolescents (e.g., Bosacki, Moreira, Sitnik, Andrews, & Talwar, 2020). Mindreading has, alongside memory capacity, been positively linked to the size of individuals' social networks (Stiller, & Dunbar, 2007) while it appears that no research to date has directly investigated the links between mindreading and social support.

Transdiagnostic Perspective

Transdiagnostic accounts suggest that impairments in mindreading are associated with a *variety* of mental health and neurodevelopmental conditions, (e.g., Cotter, Granger, Backx, Hobbs, Looi, & Barnett, 2018; see 'mindblindness account', Baron-Cohen, 1995). For example, worse mindreading, as indexed by performance on the MASC (Diziobek et al. 2016), has been linked with higher levels of autistic traits (e.g., Boada et al., 2020). Research also indicates that mindreading deficits are linked with depressive symptoms (e.g., Cotter et al., 2018), although there are some mixed results (e.g., Caputi & Schoenborn, 2008). Furthermore, high levels of anxiety have been associated with difficulties in interpreting others' mental states (e.g., Hezel & McNally, 2014). However, it is worth noting that some research suggests that these associations do not hold in individuals with subclinical levels of anxiety (e.g., Lenton-Brym, Moscovitch, Vidovic, Nilsen, & Friedman, 2018).

Motivational Accounts and Cognitive Tendencies

If individual differences in mindreading reflect motivational, personality or cognitive tendencies, then one might expect associations with broader worldviews such as religiosity (e.g., Scholl & Tremoulet, 2000). Indeed, research has shown that the brain structures engaged in contemplating supernatural beings and in mindreading are shared (e.g., Kapogiannis et al., 2009). Additionally, supernatural beings are commonly believed to possess human-like mental states (e.g., Scholl & Tremoulet, 2000). Authoritarian parenting styles, characterised by the promotion of an unquestioning adherence to strict rules, have been associated with reduced mindreading performance in children, such as on false-belief understanding tasks (e.g., Villalobos et al., 2010). Furthermore, previous research has indicated that anthropomorphism,

where human characteristics are attributed to non-human entities such as inanimate objects, is associated with activation in the same brain regions that are activated when individuals reason about other humans' mental states (Urquiza-Haas & Kotrschal, 2015, for a review).

It is noteworthy that the above-outlined variables, have also *independently* been linked to social motivation. For instance, low social motivation has been related to autistic traits (e.g., Chevallier et al. 2012), anxiety and depression (e.g., Hofmann & Hay, 2018). Depressed individuals, for example, have been found to show reduced interest in approaching smiling faces (e.g., Seidel, Habel, Kirschner, Gur, & Derntl, 2010). Conversely, students' social motivation has been found to correlate with their perceived social support (e.g., Tezci, Sezer, Gurgan, & Aktan, 2015). Furthermore, social motivation has been linked to the size of an individual's social network size (e.g., Qiu, Lin, Ramsay, & Yang, 2012), anthropomorphism (e.g., Epley, Waytz, Akalis, & Cacioppo, 2008; Epley, Waytz, & Cacioppo, 2007), religiosity (e.g., Van Cappellen, Fredrickson, Saroglou, & Corneille, 2017), and authoritarianism (e.g., Nettle, & Saxe, 2021).

The present dissertation aimed to integrate these relatively separate literatures by assessing the unique and relative contributions that the constructs of mindreading and social motivation regarding the above-outlined measures.

1.3.2 Differentiating Study Constructs

Motivation and Mindreading

There are notable distinctions between cognitive abilities such as mindreading and social motivation. Cognitive dispositions such as mindreading represent relatively *stable individual differences* (e.g., Hughes & Devine, 2015; Devine, 2021) which are commonly assessed via performance-based behavioural tasks (e.g., Devine & Hughes, 2013; Diziobek et al., 2016) - although exceptions to this norm exist, where self-report measures such as the Interpersonal Perception Task (IPT; Archer & Costanzo, 1993) are employed to assess mindreading (e.g., Carpenter et al., 2016). As outlined above, motivations can on the one hand be considered a relatively stable trait (e.g., Wasserman & Wasserman, 2020), but they can also vary within the same individual (e.g., Cameron & Pierce, 2002). In contrast to mindreading abilities, motivations can often be consciously accessed, and therefore *self-reported* (e.g.,

Fulmer & Frijters, 2009). In this dissertation, social motivation was therefore measured using self-report measures (relevant to Chapter 2 and Chapter 4) and mindreading was assessed using a variety of behavioural tasks (relevant to Chapter 2, 3 and 4).

Motivation and Personality

Furthermore, personality differs from motivation. Personality can be defined as an individuals' enduring behavioural patterns in both intrapersonal and interpersonal interactions (e.g., Hogan, 1991). Whereas personality traits are largely stable over time, as outlined before, motivations, which drive behaviour in pursuit of goals *can* also be made salient depending on situational factors (e.g., Harris et al., 2016; Apperly & Wang, 2021). Despite these differences, personality traits can influence the intensity of motivations, and so are expected to be related (e.g., Eysenck, 1997). In the present dissertation, it was therefore expected that personality would be related to social motivation. However, it is important to emphasise that this potential relationship does not entail redundancy between the two constructs (relevant to Chapter 4). In addition, in order to establish the validity of social motivation as a construct, it needs to be distinguished from general personality traits, such as conscientiousness or extraversion (which share similar definitions with social motivation. Personality was assessed using the Ten Item Personality Inventory (TIPI: Gosling, Rentfrow, & Swann, 2003), and where relevant, we assessed unique effects of social motivation over and above individual differences in personality (relevant to Chapter 2, 3, and 4).

Motivation and Social/Mindreading Motivation

As outlined before, social motivation (relevant to chapter 4), and mindreading motivation have been proposed to be separable from general motivation (or behavioural activation; Ang et al., 2017). Mindreading motivation, in comparison to social motivation, may be thought of as a more specific facet of social motivation that specifically pertains to engagement with others' *minds* and *mental states* (Carpenter et al., 2016), whereas social motivation encompasses inclinations to engage socially, including the exertion of social effort more generally (e.g., Chevallier et al., 2012). While individuals' general task performance, regarding any psychological experiment, is impacted by their general motivation, in the present study, we were particularly interested in *social/mindreading* motivation. To verify that we were assessing these constructs as opposed to general motivation, we accounted for individual

differences in the personality trait of conscientiousness, which has been argued to be a valid indicator of a person's motivation to perform well on a task (e.g., Preckel, Holling, & Vock, 2006).

1.4 Grounding the Present Research in Theoretical Perspectives and Debates

Concepts of Mental States

As previously noted, a conceptual understanding of mental states can be considered as a fundamental basis of mindreading. However, this understanding alone cannot account for two observations: (a) significant within-group heterogeneity in this ability (e.g., Dziobek et al., 2006; Conway et al., 2019; Apperly & Butterfill, 2009; for a focus on individual differences in mindreading in children and adolescents, see Hughes & Devine, 2015a and Devine, 2021) and (b) the continuous improvement in mindreading skills observed from middle childhood (e.g., Devine & Hughes, 2016), through adolescence (e.g., Dumontheil, Apperly, & Blakemore, 2010) to adulthood (e.g., Duval, Piolino, Bejanin, Eustache, & Desgranges, 2011). The primary reason for this is that most individuals, by the age of four or five, have acquired a basic understanding of what mental states, including beliefs, intentions and thoughts are, and have insight that other peoples' mental states can differ from their own (e.g., Wellman et al., 2001).

Social Individual Differences Account

An alternative theoretical explanation for the observed variance in mindreading abilities in adult populations beyond the preschool years is the social individual differences account (e.g., Apperly, 2012). This theory suggests that greater mindreading abilities have significant consequences for social functioning, over and above the influence of other general cognitive functions, which also affect social functioning. Notably, advanced mindreading tasks are designed to provide participants with greater challenges *specific* to the domain of mindreading. However, they are inherently more cognitively demanding compared to traditional false-belief-understanding tasks (which still pose their own non-mindreading specific cognitive demands). For instance, such tasks often require increasingly more sophisticated levels of verbal ability and executive functioning, relative to the target participants' age (e.g., Ahmed & Miller, 2011).

It is therefore plausible that performance on such tasks assesses cognitive processes *other* than mindreading (i.e., working memory capacity), and that the observed variation in 'mindreading ability' may reflect individual differences in such broader cognitive capacities. Indeed, researchers have argued, and evidence supports, that commonly used measures of mindreading likely tap into various cognitive processes (Warnell & Redcay, 2019; Hayward & Homer, 2017; Quesque & Rossetti, 2020). Thus, if meaningful links exist between individual differences in mindreading abilities in adults and other constructs, advanced mindreading tasks with robust psychometric characteristics should be able to illuminate such relationships, if other constructs (i.e., memory, executive functioning, or verbal ability) that account for the general cognitive demands of a respective task are also accounted for. In the present dissertation, we therefore aimed to account statistically for individual differences a variety of other relevant constructs (see section 1.3.3 above for greater detail on some of the variables accounted for).

1.5 Overview and Scope of the Present Studies

Considering the psychometric issues highlighted earlier (e.g., Schaafsma et al., 2015), based on theoretical frameworks, we assessed a wide array of correlates potentially related to mindreading and social motivation and adopted a multi-faceted approach to operationalising mindreading and social motivation across the studies presented here.

The aims of the study reported in **Chapter 2** are to examine whether individual differences in mindreading (conceptualised via response appropriateness *and the* quantity of mental state language) were (a) separable constructs, (b) accounted for by mindreading motivation and (c) related differentially to a set of potential correlates, specifically religiosity, authoritarianism, loneliness, social network size and anthropomorphism. The overarching aim of **Chapter 3** is to develop a unique approach to capturing individual differences in adults' mindreading. We focus on *how* mindreading differs between individuals, instead of relying on accuracy or correctness. To this end, we introduce a novel mindreading task. To validate this measure as a mindreading task, across two independent samples, we assess (a) its sensitivity to detect individual differences, (b) its factor structure and (c) the extent to which 'performance' on the task converges with an established measure of mindreading. In the studies reported in **Chapter 4**, we assess the potential links between social motivation and mindreading, and their respective relationships with social abilities and mental health, thereby bridging the gap

between the mostly distinct existing literatures on the correlates of individual differences mindreading and social motivation in adults. In **Chapter 5** all studies reported across chapters 2, 3 and 4 are synthesised into overarching conclusions, their limitations are identified and directions for future research to refine, expand and deepen the study of individual differences in adults' mindreading are outlined.

Chapter 2

Mindreading quality versus quantity: a theoretically and empirically motivated two-factor structure for individual differences in adults' mindreading

2 Mindreading quality versus quantity: a theoretically and empirically motivated two-factor structure for individual differences in adults' mindreading

2.1 Introduction

Mindreading is a concept central to social cognition. Individual differences in the ability to attribute mental states such as thoughts, beliefs, or desires to others correctly and to explain and predict behaviour taking these mental states into account, have been linked with many positive outcomes such as lowered aggression (e.g., Richardson, Green, & Lago, 1998) and reduced stereotyping (e.g., Mulvey et al., 2016). However, it is still unclear how and why mindreading varies in adults who already possess core mental state concepts of their own and others' thoughts, beliefs, and feelings (e.g., Apperly, 2012). In addition, more recent research has emphasised that motivational factors impact the manifestation of this capacity and its resulting benefits (e.g., Carpenter et al., 2016; Devine & Apperly, 2022). The present study was based on the hypothesis that adults might vary independently in their motivation for mindreading and in the degree to which their answers on mindreading tasks are appropriate. These factors have been proposed to be conceptually and empirically separate (e.g., Contreras-Huerta et al., 2020; Carpenter et al. 2016; Devine & Apperly, 2022) but are confounded in existing open-ended response measures of mindreading, which typically do not differentiate between the *quantity* of mental state terms (MST) and the *quality* of explanations given. The quantity of references to thoughts, feelings and other mental states is a potential indicator of motivation, whereas the *quality* of explanations indicates the contextual appropriateness of the inferences. Quantity and quality are clearly dissociable in principle because individuals may mentalize about someone else but do so in a way that seems at odds with a given situation. Testing whether they are distinguishable in practice can advance theory by casting light on the underlying sources of variation in adults' mindreading, as well as how they might be measured. The overall aim of the present study was to examine whether *individual differences* in adults' response quality and / or quantity of MST on two open-ended response mindreading tasks were (a) separable constructs, (b) accounted for by differences in individuals' mindreading motivation, and (c) related differentially to a variety of outcome measures that had previously been empirically or theoretically linked with mindreading; religiosity, authoritarianism, loneliness, social network size and anthropomorphism.

2.1.1 Measuring Individual Differences in Adults' Mindreading

Developmental research into mindreading has been dominated by questions about *when* and *how* children acquire mindreading concepts. Given that older children and adults already have a conceptual understanding of what thoughts, beliefs and desires are, that they may differ between people, and that they can impact behaviour, researchers have devised a wide range of tasks that are more subtle and complex in nature to assess mindreading in middle childhood and adolescence (e.g., Devine & Hughes, 2013) and in adulthood (e.g., Castelli et al., 2001; Dziobek et al., 2016). However, mindreading tasks devised to be complex enough to capture individual differences in adults' mindreading are often characterised by a lack of convergent validity; i.e., performance on one measure of mindreading does not correlate with performance on another measure (e.g., Schaafsma et al., 2015; Warnell & Redcay, 2019) or construct validity (i.e., the same task can be used to measure different constructs) (e.g., Declerck & Bogaert, 2008; Oakley et al., 2016). Some tasks used to assess mindreading such as the Animations Task (Abell et al., 2000; Castelli et al., 2001), which requires participants to explain the movements of triangles that exhibit either a simple interaction (e.g., two triangles appearing to be dancing), or a more complex interaction(e.g., one triangle appearing to surprise the other triangle), are *sometimes* also utilised to measure other constructs, such as anthropomorphism (e.g., Waytz et al., 2010; Tahiroglu & Taylor, 2019). Similarly, the Reading the Mind in the Eyes test (RMET) (Baron-Cohen, Wheelwright, Raste, & Plumb, 2001), originally developed to measure mindreading, may be more suitable as a measure of emotion perception given that the task requires participants to match suitable adjectives with pictures of faces expressing emotions (e.g., Declerck & Bogaert, 2008; Oakley et al., 2016). It is therefore unclear what "advanced" mindreading tasks really measure (e.g., Declerck & Bogaert, 2008; Oakley et al., 2016; Quesque & Rossetti, 2020). Furthermore, mindreading research in adults has predominantly focused on group-level comparisons between clinical and non-clinical groups (e.g., Bradford et al., 2018; Dziobek et al., 2006; Preißler et al., 2010) and does not provide satisfactory answers as to whether and why there are *individual differences* in adults' mindreading.

Notwithstanding these limitations, some mindreading tasks have also been successfully administered to typical adults (e.g., Devine & Hughes, 2019; Dziobek et al., 2006; Slaughter & Repacholi, 2004). The Movie for the Assessment of Social Cognition (MASC; Dziobek et

al., 2006) requires participants to reason about the mental states of four characters who are shown in a realistic social interaction based on a multiple-choice question regarding the characters' mental states. Neurotypical adults have been found to select correct responses approximately 60% of the time (Dziobek et al., 2006; Preißler et al., 2010), which indicates that the MASC is suitable for measuring individual differences in non-clinical adult samples (e.g., Luyten, Van Houdenhove, Lemma, Target, & Fonagy, 2012). It should, however, be noted that the multiple-choice format of the MASC does not allow for the differentiation of frequency of mental state terms; a quantitative marker of an individual's tendency to mindread based on the frequency with which they refer to other people's mental states, and the appropriateness of participants' mindreading responses; contextually justified mindreading.

The Silent Films Task (SFT; Devine & Hughes, 2013) has specifically been developed to detect individual differences in mindreading beyond the pre-school years. The task requires participants to answer six open-ended questions about five video clips that portray a main character interacting with a range of peripheral characters. Although this task exhibits convergent validity (e.g., scores on the SFT are strongly correlated to scores on another – strange stories – task) (Devine & Hughes, 2013), criterion validity (e.g., Devine, Kovatchev, Grumley Trayner, Smith & Lee, 2023), and test-retest reliability (Devine & Hughes, 2016), it has yet to be formally evaluated beyond early adolescence. Furthermore, the original questions used in the SFT prompt participants to reason about mental states and do not differentiate between frequency of MST and response appropriateness. Thus, despite these advances in the creation of "advanced" mindreading tasks, challenges persist in the measurement of how mindreading *varies* in adults (Apperly, 2012; Oakley et al., 2016; Quesque & Tosetti, 2020; Declerck & Bogaert, 2008; Qureshi, Monk, Samson, & Apperly, 2019; Schaafsma et al., 2015).

Finally, there are grounds to think that frequency of MST and response appropriateness may be separable constructs, although they are often confounded in open-ended response coding schemes. The frequency of spontaneously produced MST is distinct from performance on mindreading tasks in middle childhood (e.g., Meins, Fernyhough, Johnson, & Lidstone, 2006), in adults (e.g., Devine & Hughes, 2019), and in older adults (e.g., Lecce, Ceccato, & Cavallini, 2019). Furthermore, there is evidence that the use of mental state terms can vary with different interaction partners (Brown, Donelan-McCall, & Dunn, 1996), perhaps because some social interactions are more motivating than others.

2.1.2 The Role of Motivation in Individual Differences in Adults' Mindreading

A person's motivation fundamentally shapes their behaviour (e.g., socio-emotional selectivity theory; Carstensen, Isaacowitz, & Charles, 1999). Indeed, making mentalistic attributions is an effortful process that places high demands on memory and cognitive control (e.g., Apperly et al., 2006; Ferguson, Apperly, & Cane, 2017; Lin et al., 2010; Kouklari, Tsermentseli, & Auyeyng, 2018; Qureshi et al., 2020). Individual differences in motivation might therefore also influence performance on mindreading tasks (e.g., Carpenter et al., 2016; Carstensen et al., 1999; Contreras-Huerta et al., 2020; Devine & Apperly, 2022).

Mindreading motivation has been defined as a stable orientation to engage effortfully with others' minds and mental states (Carpenter et al. 2016), and it has been suggested that autistic people sometimes perform less well on mindreading tasks because they have lower social motivation rather than lower social ability (Burnside, Ruel, Azar, & Poulin-Dubois, 2018; Chevallier et al., 2012). Likewise, there is evidence that neurotypical adults' mindreading motivation is positively related to their performance on mindreading tasks (e.g., Carpenter et al., 2016). Specifically, Carpenter et al. (2016) found small positive correlations between mindreading motivation, assessed via the 13-item Mind Reading Motivation Scale and performance on the interpersonal perception test 15 (IPT-15; Archer & Costanzo, 1993), which requires participants to make judgements of videotaped real-world social situation. It has been suggested that mindreading motivation and mindreading ability are conceptually and empirically separable (e.g., Carpenter et al. 2016; Contreras-Huerta et al., 2020; Devine & Apperly, 2022), and there is evidence that social motivation and mindreading ability make distinctive contributions to explaining social outcomes (Devine & Apperly, 2022). The present study combined this observation with a revised coding scheme for mindreading that distinguishes between the quantity of mental state terms and the appropriateness of mindreading responses. We predicted that mindreading motivation might be specifically related to the quantity of mental state terms produced in participants mindreading responses.

Overall, the above-described concerns and previous research findings motivate our first objective of introducing a scoring system that distinguishes between (a) mindreading quantity, a potential indicator of motivation, operationalised as frequency of MST in responses, and (b) mindreading quality, operationalised as the "appropriateness" of participants' answers when responding to open-ended versions of the Silent Films Task (SFT; Devine & Hughes, 2013)

and the Movie of the Assessment of Social Cognition (MASC; Dziobek et al., 2016). We hypothesised that if frequencies of MST and response appropriateness are separable constructs, then response appropriateness and frequencies of MST will load onto separate latent factors (i.e., one representing response appropriateness and one frequencies of MST, a potential indicator of motivation). Alternatively, if frequencies of MST and response appropriateness test the same underlying construct (i.e., "mindreading"), we expect response appropriateness and frequencies of MST to load onto one latent factor. Finally, it is possible that individual differences in performance will be driven by characteristics of the two different tasks, rather than the psychological constructs of response appropriateness, frequencies of MST or mindreading. This effect would lead indicators (i.e., MST and appropriateness) coded from the SFT to load onto one latent factor (i.e., representing performance on the SFT) and both indicators coded from the MASC to load onto a separate factor. The possibility of shared method variance will be accounted for in our analyses.

2.1.3 Criterion validity of MST and Response Appropriateness

One way to distinguish between two closely related constructs is to investigate differential relations between the constructs and other variables (e.g., Carpenter et al., 2016). Although previous results have not always been consistent and may have been subject to some of the psychometric shortcomings of mindreading tasks (e.g., Schaafsma et al., 2015), they provide potential criteria for evaluating the different factor structures used to conceptualise mindreading in the present study. We therefore examined the distinctiveness of individual differences in mindreading quality and mindreading quantity by investigating whether these constructs might show unique associations with variables that have been linked with mindreading performance, such as measures of *social competence* (e.g., loneliness; Bosacki, Moreira, Sitnik, Andrews, & Talwar, 2020, Devine & Hughes, 2013, or social networks; Malcolm, Saxton, McCarty, Roberts, & Pollet, 2021) and *personality tendencies* such as religiosity (Atran & Norenzayan, 2004; Boyer, 2008; Vonk & Pitzen, 2017), authoritarianism (O'Reilly & Peterson, 2014), and anthropomorphism (e.g., Epley, Waytz, & Cacioppo, 2008; Urquiza-Haas & Kotrschal, 2015, for a review).

2.1.4 The Present Study

In summary, we sought to shed light on the nature of individual differences in mindreading in adults, by creating a novel scoring system that leveraged existing open-ended mindreading tasks. Our goal was to assesses whether mindreading quantity can be distinguished from mindreading quality. More specifically, we aimed to examine whether individual differences in neuro-typical adults' (*N* = 128) and frequencies of explicit references to others' mental states and response appropriateness on two open-ended response mindreading tasks were (a) separable constructs, (b) accounted for by differences in individuals' mindreading motivation, and (c) related differentially to a variety of outcome measures that had previously been empirically or theoretically linked with mindreading.

2.2 Methods

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study.

2.2.1 Participants

This study was pre-registered prior to the start of data collection. This is a link to the preregistration: https://osf.io/uerkx/). A total of 133 participants were recruited via word of mouth and via the University of Birmingham School of Psychology undergraduate database.

Power-analysis indicated that a minimum sample size of 119 participants was needed to obtain a level of acceptable statistical power of .80, and for a model that has a maximum of three latent variables and 15 indicator and observed variables. Our minimum target sample size was therefore 119 participants. All participants took part in the study in exchange for course credit and were psychology students. To be eligible participants had to be at least 18 years old, report no current or past neurological or psychiatric problems, and be English-speaking (including native and non-native English speakers). Based on these pre-registered exclusion criteria, 5 participants were excluded due to reporting a past neurological or psychiatric problem, leaving a final sample of 128 participants (113 females, 15 males, *M*age = 19.47 years, *SD* = 1.47 years, range = 18 - 28). One hundred and seven participants (83.6%) reported English as their native language, with 71.9% of participants (*N* = 92) being monolingual, 25% (*N* = 32) bilingual and 3.1% (*N* = 4) multilingual (speaking more than two languages). Most participants (95.3 %, *N* = 122) had completed secondary school or equivalent (e.g., A levels),

2.3% (*N* = 3) held a bachelor's degree, and 1.6% a master's degree (*N* = 2). Participants were invited to identify their ethnicity according to the descriptions recommended by the United Kingdom Office for National Statistics (ONS) that are comprised of five broad categories and 18 sub-categories. In order of size, the five broad categories were presented as follows: 70.4% (*N* = 90) of participants identified as White, 16.5% (*N* = 21) identified as Asian, 5.5% (*N* =7) as Black, 4.7% (*N*=6) as Mixed, and one participant (0.8%) as an ethnicity not listed. Data on religion was also recorded according to the questions proposed by the ONS: 56.3% (*N* = 72) of participants reported no religion, 21.1% (*N* = 27) identified as Christian, 2.3% (*N* = 3) of participants as Hindu, 7% (*N* = 9) as Jewish, 10.2 % (*N* = 13) as Muslim, 1.6% (*N* = 2) as Sikh, and 1.6% (*N* = 2) preferred not to answer the question.

2.2.2 Design and Procedure

As this was an individual differences study, questionnaires and measures were administered in a fixed order, following Goodhew and Edwards (2019). The rationale is that a fixed order of presentation makes it more likely that any effects of task order are equivalent for all participants, which minimizes the unexplained variance when the focus of analysis is on the pattern of individual differences between participants. A fixed order of presentation of course raises the possibility that fatigue or practice, or order effects, will result in unreliable measurement. For example, highly fatigued participants might show floor-level performance, or answer at random. All measures were therefore checked for reliability. All questions were administered via the Qualtrics survey platform. Testing was completed in one session and no feedback was provided in relation to response content. The time participants had to complete the tasks was not limited and the total duration of the study was approximately one hour. Informed written consent was obtained from all participants prior to participating and the study was approved by the Ethical Review Committee at the (*Masked*).

2.2.3 Materials

Response Appropriateness and Frequencies of Mental State Terms (MST)

Movie Assessment of Social Cognition (MASC; Dziobek et al., 2016)

Participants watched a 15-minute movie, which was stopped several times to ask participants both the open-ended (e.g., *"What do you think happened during this clip?"*) as

well as the original multiple-choice question following the open-ended question. The number of questions was reduced from the original 45 (assessed via multiple-choice questions) to 10. Participants watched the whole movie, but the movie was stopped only 10 times, with the remaining clips merged, to account for the fact that the questions were open-ended. Multiple-choice answers were scored as either correct or incorrect, with further differentiation of the incorrect answers into over-mentalizing (i.e., overly complex mental-state reasoning), no mentalising (i.e., failing to make a mental state attribution) and under-mentalising (i.e., overly simplistic mental state inferences). Scores based on the multiple-choice questions were retained for comparison purposes.

To rate response appropriateness, we developed a novel coding scheme for participants' answers to open-response questions. Two points were awarded for each one of the ten clips, if two pre-specified criteria deemed to capture the essence of the actions of the clip were met. Explicit references to mental-state terms were not required to achieve full scores, however, participants were required to either directly or indirectly attribute mental states to a character in the clips in a way that was consistent with the events shown and reflective of the question asked. Appropriate responses had to include a reference to both interacting entities. Furthermore, responses could not include speculation that could not have been logically derived from the clip. One point (reflecting a partially appropriate response) was awarded if a response met one of the two criteria. To achieve one point, participants' description had to still be related to the sequence and show rudimentary understanding of the reasons for the characters' actions without, however, reaching full understanding (i.e., responses may have been imprecise or incomplete). Zero points were awarded if none of the two criteria were met. In general, such responses were reflective of participants misunderstanding the question, were nonsensical, incoherent and / or focussed on minor aspects of the sequence. Zero-point responses would not enable the reader to reconstruct the sequences of the video clip.

Our coding approach draws from Castelli et al. (2001) who coded participants' descriptions of animated triangle cartoons in terms of coherent retelling of the scene (which they called "appropriateness") and a separate score for mindreading (which they called "intentionality"). It furthermore took inspiration from the standard scoring of the Silent Films Task (Devine & Hughes, 2013), which categorised open-ended responses according to a 3-point scale (ranging from 0 to 2). Our approach is, however, distinctive because it differentiates between the appropriateness and quantity of mindreading. An "appropriate" response required

evidence of contextually justified mindreading, whereas MST quantified the number of mental state terms regardless of their appropriateness.

Across all 10 items of the MASC participants could score from 0 (reflecting exclusively inappropriate responses) to 20 (reflecting exclusively appropriate responses). The inter-item reliability procedure was conducted by cross validating the coding of 20% of responses. After having independently coded 20% of responses, for each of the 10 items of the MASC, the inter-item correlation was calculated. At the first attempt, inter-item reliability for the appropriateness ratings of the MASC ranged from ICCs of .87 to .92, indicating that no further adjustments to the coding scheme were needed. The number of references participants made to the characters' mental states was also recorded. Following the approach taken by previous studies (e.g., Taumoepeau, Sadeghi, & Nobilo, 2019), the coding scheme for MST used in the present study was adapted from existing coding schemes (e.g., Brown et al.,1996) and extended with words that fit the general definition of what constitutes a mental state term but were not explicitly named in the existing coding schemes considered. Specifically, we did not record participants' references to their own mental states (e.g., "I think"). The total number of words within each response was also counted. All verbs, nouns or adjectives referring to cognitions such as "think", "remember", "know", desires such as "want" or "like" (not as a preposition) or emotion such as "happy", "excited", "angry" were coded as MST. We did not utilise the MST as a proportion of response length variables in the present study but used the raw scores instead. As part of the above-outlined coding procedure, inter-rater reliability was calculated for 20% of items (N = 25) and was acceptable, with ICCs ranging from .79 to .99. The coding manual including examples of how responses were coded for the MASC is available on the OSF (https://osf.io/uerkx/) and in the Appendix of this dissertation.

Silent Film Task (Devine & Hughes, 2013)

Participants watched five short black and white silent film clips showing one main character interacting with several different peripheral characters. Across the five clips, participants were asked an open-ended question (i.e., *"What do you think happened during this clip*?") after each clip. This was followed by a clip-specific question which had originally been developed for administering the task to children and adolescents (e.g., *"Why do you think the men hide?").* For the original questions, responses were coded using the standard rating scheme and retained for comparison purposes only as part of the same, within-participant, study. The

measure and standard scoring manual are available at the OSF: https://osf.io/8x73r/ and the Appendix of this dissertation contains examples of coded participant responses according to the coding scheme.

Our critical data came from an open-ended question after each clip. In line with how response appropriateness was coded for the MASC in the present study, participants' responses to this question were coded using a novel 3-point appropriateness scale indicating appropriate, partially appropriate, or inappropriate understanding. Points for the open-ended question were awarded according to the same logic as for the MASC (see description above). Specifically, for this item, to achieve two points, participants needed to (a) describe both Harold's (sits in the van) and the driver's (drives away) behaviour and (b) recognise that the driver did not know (e.g., pay attention, realise) that Harold was in the van. This could have been implicit via stressing that the driver was deaf in relation to him driving away. One point was awarded for each element, and zero points were awarded if neither element was mentioned. Across all items, minimally, participants could score 0 points and, maximally, 10 points could be achieved. Summed scores were normally distributed. Items showed acceptable inter-rater reliability, N = 25, with ICCs (calculated identically as for the MASC) ranging from .85 (for item three) to .93 (for item four). In addition, the number of mental state terms contained for each item of the SFT was counted. Inter-item reliability was acceptable, N = 25, with ICCs ranging from .74 (for item two) to .97 (for item three). The coding manual for the SFT can be found on the OSF (https://osf.io/uerkx/?view_only=a291a3440d814c3c8407fe94ad78f83b).

Mindreading Motivation (MRM) Scale (Carpenter et al., 2016)

Participants rated their level of agreement on a scale ranging from 1 (disagree completely) to 7 (agree completely) for 13 questions assessing the degree to which they are oriented towards effortfully engaging with others' perspectives and mental states. As per the standard scoring scheme, items 2, 4, 5, 7, 9, 10, 11, 12, and 13 were reversely coded. Across all items, participants could minimally score 13 points and maximally 9, α = .67.

Religiosity

Based on the United Kingdom Office for National Statistics proposed questions, participants indicated their religious affiliation, based on the question *"What is your present*

religion, if any?". No religion was coded as 0 and having a religious affiliation was coded as 1. Religious Practice was assessed via the question *"Aside from weddings and funerals, how often do you attend religious services?"*, with response options ranging from 1 (More than once a week) to 6 (Never). To assess religious upbringing, participants were asked *"As a child, were you raised in a religious home?"*, with response options of "Yes", "No" and "Don't know". Religious belief was assessed with the following question: *"Which statement comes closest to expressing what you believe about God?".* Across all items, higher scores indicated greater religiosity, $\alpha = .68$.

Social Networks (Malcolm et al., 2021)

To measure individual differences in *social network size*, we asked participants to list the initials of everyone they were in social contact within the last month, either over the telephone, text message, video communication software, email or in person. Participants were asked to exclude purely work-related contacts or casual acquaintances such as someone they have briefly encountered in the street. The total number of contacts listed was summed (M = 18.64, SD = 12.03, min = 1, max = 110). *Advice network size* (i.e., the number of people a participant discussed important matters with) was also summed, M = 8, SD = 1.77, min = 2, max = 15. Participants were able to write up to ten people's initials and were further instructed to indicate within these contacts how emotionally close they felt to their advice network contacts on a scale ranging from 1 (not close at all) to 5 (extremely close), with higher scores indicating greater closeness. The resulting scores were summed to create an overall emotional closeness score for each participant. This variable was normally distributed, M = 29.52, SD = 7.67, min = 10, max = 50.

Individual Differences in Anthropomorphism Questionnaire (IDAQ; Waytz, Cacioppo, & Epley, 2010)

Across 15 items, participants rated the extent to which natural entities, non-human animals and technological devices have "free will", "consciousness", "a mind of its own", "intentions" and can "experience emotions". Responses are rated on a scale from 0 (not at all) to 10 (very much). IDAQ ($\alpha = .85$) scores were created according to the standard coding for each participant. Scores for the IDAQ scale ranged from 18 to 154. The data was normally distributed.

Revised UCLA Loneliness Scale (RULS-6) (Wongpakaran et al. 2020)

Participants were asked to respond on a 4-point Likert scale ranging from 1 = "often" to 4 = "never" to six questions, for example, *"How often do you feel that you lack companionship?"*, assessing the degree to which participants felt lonely, a = .897. The degree of loneliness was evaluated by averaging scores across items. The items were reverse-coded, with higher scores indicating greater loneliness. Loneliness average scores ranged from 1 to 4 and the data was normally distributed, α = .67.

Ten Item Personality Measure (TIPI) (Gosling, Rentfrow, & Swann, 2003)

The 10-item measure of the Big Five (or Five-Factor Model) dimensions of personality was administered. Participants indicated the degree to which they ascribed different personality traits to themselves, using a 7-point Likert scale ranging from 1(Disagree strongly) to 7(Agree strongly). Items 2, 4, 6, 8 and 10 were reversely coded (e.g., for these items 7 was recoded with a 1, a 5 with a 3 etc.). The average of two items (standard item and reverse-scored item) made up a scale (e.g., extraverted, enthusiastic & reserved, quiet, a = .77, Critical, quarrelsome & Sympathetic, warm, a = .29, Dependable, self-disciplined & Disorganised, careless, a = .64, Anxious, easily upset & Calm, emotionally stable, a = .65, Open to new experiences, complex & Conventional, uncreative, a = .37).

Wechsler Abbreviated Scale of Intelligence (WASI – 2, Wechsler, 2001)

Current intellectual functioning (verbal ability) was assessed using the Similarities Test of the Wechsler Adult Scale of Intelligence. This vocabulary subtest has been recognized as among the most widely used measure of general mental ability. Participants completed 20 trials (with 20 items), explaining why two words were alike. For example, participants were asked: "In what way are grapes and strawberries alike?", to which the correct answer would be "both fruit". Responses were scored using the WASI manual and items were scored on a 3-point scale from 0 to 2. Scores were added up, with higher scores indicating greater expressive language ability, α = .33, *range* = 22 – 35.

Data and Supplementary Material availability Statement

The study was pre-registered on the OSF (https://osf.io/uerkx/). A latent variable modelling approach to analysis was adopted. Information of the item-level data of the SFT and MASC, descriptive statistics for each measure and correlations between the original coding of the MASC and SFT with the new coding applied in the present study can be found in the Supplementary Material.

2.3 Results

2.3.1 Analysis Plan

We began by testing measurement models for mindreading appropriateness (i.e., response appropriateness for the MASC and SFT) and frequencies of mental state terms (from responses to the MASC and SFT). We then examined whether mindreading motivation as well as a range of outcome measures were associated with individual differences in mindreading using structural equation modelling. The latent factors were regressed onto age, gender (0 = male, 1 = female), education, whether participants spoke English as their native language (0 = no, 1 = yes) and verbal ability. The data were analysed using SPSS and R, the Lavaan package (Rosseel, 2012).

A robust maximum likelihood estimator was used to account for the potential non-normal distribution of each variable where variables were continuous. For ordinal variables, diagonally weighted least squares estimation (WLSMV) was used. To evaluate whether the models fit the data, we used three recommended methods*:* the *root mean square of approximation* (RMSEA) with < 0.05 indicative of good model fit and values between 0.05 and 0.08 as adequate model fit (Kline, 2011; Brown, 2015); the *comparative fit index* (CFI); and the Tucker-Lewis index (TLI), for which values above .95 are considered as a good fit (Hu & Bentler, 1999). Where two or more models met all three criteria, we selected the simplest model that was most strongly supported by theory for further analyses. Covariates (e.g., gender, age, education, English as native language and verbal ability) were treated as separate variables. In line with recommendations for individual differences researchers, we consider correlations of .10 as a small effect, .20 a typical (or moderate) effect and > .3 as a relatively large effect (Gignac & Szodorai, 2016).

2.3.1 Missing Data

The percentage of missing values ranged from 0.8 % (N = 1) for demographic variables such as education up to 7.8% for religious upbringing (N = 10). The only reason for missing data was participants' nonresponse. A full information maximum likelihood (FIML) approach to missing data was adopted. This method has been shown to produce unbiased parameter estimates, standard errors and test statistics that are consistent and efficient when data is either missing at random (MAR) or missing completely at random (MCAR) (Enders & Bandalos, 2001). To test for patterns within the missing data, Little's Completely at Random test (MCAR; Little & Rubin, 2019) was conducted and showed that data was not systematically missing, χ^2 (1205) = 1133.893, p = .928.

2.3.3 Preliminary and Descriptive Statistics

Table A in the supplementary material (S1) shows the bivariate correlations between the item-level data of the SFT and MASC as well as the descriptive statistics for each item of the measures. For correlations between the original coding of the MASC and SFT with the new coding applied in the present study, see Table B in the supplementary material (S2). Descriptive statistics of all summed study variables can be seen in Table 2.1. The SFT mental state terms data were normally distributed, with one participant using no mental state terms at all (.8%) and two participants (1.6%) using 15 mental state terms to describe the actions in the video clips. There were no ceiling or floor effects for response appropriateness data, with no participant scoring 0 and only 4 participants achieving full scores (3.1%). For the MASC mental state terms data, 3 participants made only two mental state references (2.3%), and one participant made 75 mental state references (.8%) whilst the second highest number of MST was 51 (N = 1, .8%). The data were normally distributed. In terms of appropriateness, there were no ceiling or floor effects, with no participant scoring 0, one participant scoring 1 (.8%), and three participants scoring 2 (2.3%). No participant achieved full scores, 3 participants scored 17, (2.3%), and a further 3 participants achieved 16 across all items of the mask. Skewness and kurtosis of all study measures, except for the social network size measure, were within the acceptable ranges of -3 and +3 for skewness and -10 to +10 for kurtosis (Brown, 2006). The social network size measure was positively skewed due to an outlier (one participant indicated 110 social contacts, whilst the second highest number was 43).

Table 2.1 Descriptive Statistics for each study measure.

	M (SD)	Min	Max	Range	Skewness (SE)	Kurtosis (SE)	*N*
IDQA	68.52 (25.63)	18	154	136	.545 (.214)	.564 (.425)	128
SFT Original	8.32 (1.83)	1	12	11	-1.173 (.217)	2.47 (.430)	125
MASC Correct	7.29 (1.46)	1	10	9	-.841 (-.841)	2.67 (.425)	128
MASC Exc	1.35 (.98)	0	5	5	.699 (.214)	1.04 (.425)	128
MASC Under	.95 (.92)	0	6	6	1.56 (.214)	5.83 (.43)	128
MASC No	.414 (.682)	0	3	3	1.83 (.214)	3.51 (.43)	128
MASC APP	9.25 (3.45)	1	17	16	.162 (.217)	-.16 (.430)	125
SFT APP	5.63 (2.13)	1	10	9	-.040 (.219)	-.58 (.435)	122
MASC MST	18.54 (10.55)	2	75	73	1.94 (.217)	6.7 (.430)	125
SFT MST	7.21 (3.05)	0	15	15	.463 (.219)	-.125 (.435)	122
Loneliness	2.36 (.75)	1	4	3	-.137 (.214)	-.823 (.425)	128
Autho	2.62 (.71)	1.17	4.17	3	.178 (.215)	-.691 (.427)	127
Extraversion	3.97 (1.50)	1.50	6.50	5	-.086 (.214)	-1.111 (.425)	128
Agree	5.09 (1.03)	2.50	7	4.5	-.301 (.214)	-.612 (.425)	128
Conscien	5.39 (1.14)	2	7	5	-.889 (.214)	.437 (.425)	128
Emo Stability	3.77 (1.31)	1	7	6	.105 (.214)	-.228 (.425)	128
Openness	5.04 (1.16)	2	7	5	-.327 (.214)	-.508 (.425)	128
WASI	29.20 (2.91)	22	35	13	-.488 (.221)	-.098 (.438)	120
MRM	66.12 (7.63)	47	89	42	-.098 (.217)	.326 (.430)	125
Rel Practice	1.19 (1.28)	0	5	5	1.18 (.216)	.925 (.428)	126
Rel Belief	2.28 (1.79)	0	5	5	.370 (.216)	-1.31 (.428)	126
Rel Affiliation	.428 (.497)	0	1	1	.292 (.216)	-1.946 (.428)	126
Rel Upbringing	.423 (.496)	0	1	1	.313 (.223)	-1.935 (.442)	118
Social Net Size	18.64 (12.03)	1	110	109	3.668 (.214)	25.303 (.425)	128
Ad Net Size	8.08 (1.77)	2	15	13	-.835 (.214)	2.404 (.425)	128
Emo Ad Net Size	29.52 (7.67)	10	51	41	-.087 (.214)	-.114 (.425)	128

Note. M = mean, SD = Standard Deviation, Min = observed minimum value, Max = observed maximum value, SE = Standard Error, N = number, IDQA = Anthropomorphism; IDQA-NA = No Anthropomorphism Scale, SFT = Silent Film Task, MASC = Movie for the Assessment of Social Cognition, Exc = Excessive Mindreading (MCQ), Under = Less Mindreading (MCQ), No = No Mindreading (MCQ), APP = Appropriateness, MST = Mental State Terms, Autho = Authoritarianism; Agree = Agreeableness; Conscien = Conscientiousness; Emo Stability = Emotional Stability; Openness = Openness to Experience, WASI = Verbal Ability, MRM = Mindreading Motivation, Rel = Religious, Net = Network, Emo = Emotional, Ad = Advice.

2.3.4 Main Analysis

The latent factor structure of mindreading (i.e., whether response appropriateness and frequencies of MST are separable or overlapping) was examined. A schematic overview of the different models that were tested is shown in Figure 2.1 and the summary statistics of the different models are shown in Table 2.2. Both a one-factor (see model A.123) and two-factor

(see model C.23) model provided an acceptable fit for the data (One factor model: RMSEA = .032, CFI = .937, TLI = .927, Factor loadings = .115 - .679; Two factor model: RMSEA = .031, CFI = .943, TLI = .933, Factor loadings = .132 - .733). Additionally, the latent factors of response appropriateness and frequencies of MST were significantly positively correlated, r = .75 (see Table 2.3). Therefore, in terms of our first criterion for judging the value of coding MST and appropriateness separately versus in a single construct, although indices were marginally in favour of a two-factor solution, there is no clear evidence in favour of a one- or two-factor model. We therefore turned to our second criterion and compared the associations between the resulting factors from these models with the outcome variables.

We next compared the associations between the resulting factors from these models and the social outcome variables. Table 2.3 shows the bivariate correlations between all study variables. The latent factors for response appropriateness and frequencies of MST were strongly positively correlated, indicating that there was a large amount of shared variance between both constructs. Although there was a small significant association between response appropriateness and mindreading motivation, frequencies of MST and mindreading motivation were not significantly correlated. Anthropomorphism was weakly positively associated with response appropriateness and negatively with loneliness. Authoritarianism was weakly negatively associated with mindreading motivation.

Table 2.2 displays the summary statistics for the structural equation models that we ran to test how either response appropriateness and frequencies of MST, separately as part of a two-factor model, or jointly, as part of a one-factor model, related to other constructs of interest. Each outcome variable was separately regressed onto these two latent factors. To control for potential confounding variables, we regressed the outcome variables onto each latent factor and education, gender, age, whether they spoke English as their native language, verbal ability and personality traits. Covariates were permitted to correlate with each other in each model.

An assessment of the unique associations between mindreading motivation, frequencies of MST and response appropriateness showed that MRM was not predicted by the amount of MST people used in their responses or by the appropriateness of their responses. Likewise, a one-factor solution where frequencies of MST and response appropriateness loaded onto one factor was not a significant predictor of MRM. A one-factor solution did not predict any of the outcome variables significantly. However, for a two-factor solution there were significant,

differential associations between the latent factors of frequencies of MST and response appropriateness with religiosity. MST predicted religiosity positively while greater response appropriateness was a significant negative predictor of religiosity (see Figure 2.2 below for a simplified path diagram of the standardized estimates for significant regression paths and Table 2.4 for an overview of all tested regression paths). This finding indicates that there was valid non-shared variance between both constructs, which might be obscured by confounding or not differentiating between the degree to which participants quantitatively engage with others' mental states and the degree to which their responses are appropriate. In terms of our second criterion for judging the value of coding MST and appropriateness separately versus in a single construct, we found evidence of greater explanatory power for a two-factor solution for predicting outcomes related to mindreading.

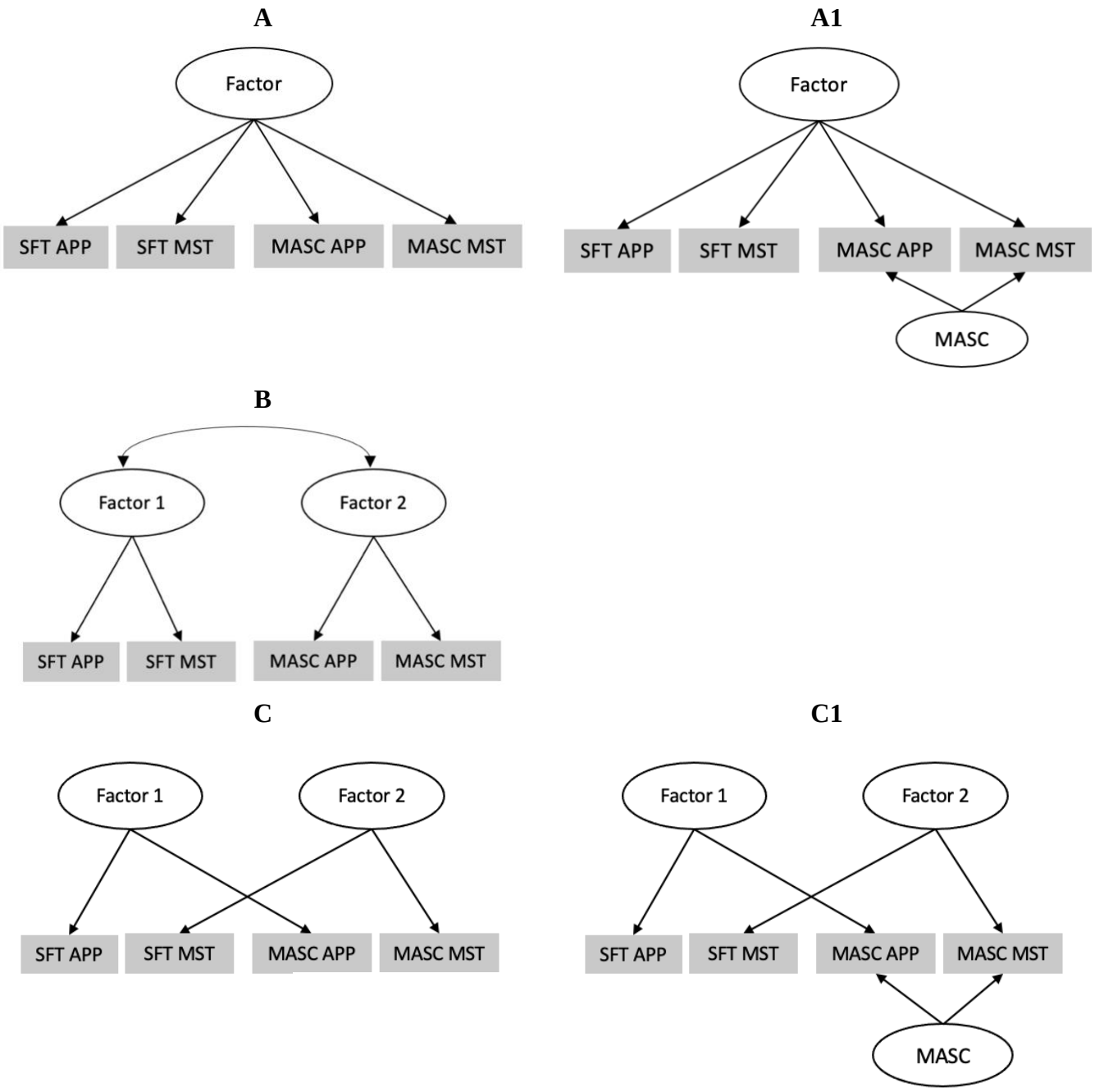

Figure 2.1. Schematic representation of the models that were tested to investigate the best fit for the item-level data of the SFT (5 items) and MASC (10 items), with each item being associated with two unique scores; a mindreading response appropriateness score and a frequencies of MST score (making 32 individual scores). **A** = One-Factor model where frequencies of MST and response appropriateness load onto one factor, regardless of task; **1** = Method factors due to the different tasks are accounted for (MASC); B = Two-Factor / Two-Task Model where both constructs (i.e., MST and response appropriateness) coded from the SFT load onto one factor (representing performance on the SFT) and both constructs coded from the MASC load onto a separate factor to test if shared task characteristics drove individual differences in response appropriateness and MST; C = Two-Factor model where response appropriateness and MST load onto separate factors.

Table 2.2. Summary statistics of different models to assess whether response appropriateness and frequencies of MST are separable or overlapping.

		RMSEA	**CFI**	**TLI**	**Factor Loadings**	**Scores from Items no longer sig. loading onto Latent Factor**			
Good Fit Indices		< .08	>.90	>.90	(range)	SFT		MASC	
Model						MST	APP	MST	APP
One-Factor	A	.061	.814	.792	.141 - .664				
Bi-Factor	A1	.055	.861	.837	.129 - .715		Q3	-	Q2, Q24
	A12	.049	.856	.835	.136 - .693	Q1, Q2	Q3	-	Q2,
	A.132	.032	.937	.927	.115 - .679	Q1, Q2	Q1, Q3	-	Q2, Q24
Two-Factor									
MASC vs SFT	B	.084	.636	.608	.153 - .743	Q2	-	-	-
Bi-factor	B2	.049	.886	.866	.157 - .663	Q2	-	-	Q1, Q2
	B24	.044	.884	.866	.171 - .690	Q2	Q1	-	-
MST vs APP	C	.058	.836	.816	.189 - .705				
Bi-factor	C1	.052	.872	.849	.129 - .659	-	-	-	Q2, Q24, Q34, Q39
	C12	.047	.867	.847	.160 - .654	Q1, Q2	-	-	Q2, Q24
	C.23	.031	.943	.933	.132 - .733	Q1, Q2, Q4,	Q1	-	Q2, Q24, Q34, Q39

Note. A = One-factor Model, B = Two-factor Model (MASC vs SFT), C = Two-factor Model (MST vs App), 1 = method factor (MASC scores), 2 = controlled for demographic variables (age, gender, education, verbal ability, personality, 3 = model fit based on most significant modification indices, 4 = method factor (MST scores), RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index. Model A.123 and C.23 provided the best fit to the data. Re-running model C.23 with all non-significant items removed showed similarly good model fit, CFI = .916, TLI = .903, RMSEA = .038. We therefore retained the original model to maximise available information.

Table 2.3. Bivariate correlations between all variables of interest.

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. LFA																				
2. LF M	.745 **																			
3. LFMA	.756**	.962**																		
4. MRM	.208*	.169	.086																	
5. Lone	-.014	-.036	-.040	-.110																
6. Anthro	.178*	.147	.137	.132	.233**															
7. Authori	.068	.064	.049	-.194*	.067	.026														
8. LF Rel	-.152	.044	.028	-.091	.070	.078	.157													
9. NetS	-.039	-.043	-.039	.112	-.139	-.078	.009	-.008												
10. AdS	.053	-.026	-.066	.112	-.139	.112	.009	.064	.156											
11. AdE	.105	.059	.033	-.121	-.353**	.092	.079	.080	.178*	.722**										
12. Edu	-.070	.160	.169	.035	.017	.037	.046	-.033	-.006	-.095	-.007									
13. Gen	.104	.058	.036	.090	.042	.197*	-.055	-.011	.054	-.039	.133	.137								
14. Nat	-.030	.112	.089	-.018	.202*	.096	.072	.134	-.150	-.032	-.017	.210*	-.083							
15. Age	-.024	.181*	.206*	.037	-.078	-.056	-.074	-.079	-.080	-.041	-.090	.551**	.030	.137						
16. VA	.206*	.097	.168	.231*	-.003	.001	-.172	.051	.188*	-.009	-.038	.058	-.009	-.093	.002					
17. Extra	.025	-.014	-.042	.147	-.294**	-.064	-.153	-.095	.231*	.004	.042	-.040	-.015	-.161	.074	-.021				
18. Agree	.194*	.112	.159	.124	-.076	.211*	-.014	.135	.113	.050	.176*	.063	.196*	-.037	-.014	.159	-.039			
19. Cons	.051	.072	.108	.131	-.081	.029	.147	.082	.098	-.038	.128	.004	.124	-.039	-.179*	-.020	-.019	.107		
20. Emo	-.090	-.049	-.020	-.251**	-.382**	-.224*	.065	.119	.024	.057	.229**	.047	-.056	.095	.110	-.021	.204*	.181*	.082	
21. Open	.037	.050	.016	.060	-.056	.027	-.155	.169	.119	-.004	-.014	.139	.181*	.129	.094	-.042	.312**	.155	-.033	.137

Note. $^{**}p < .01$. $^{*}p < .05$. LF = Latent Factor, A = Appropriateness, M = MST, MRM = Mindreading Motivation, Lone = Loneliness, Authori = Authoritarianism, Rel = Religiosity, Up = Upbringing, A = Affiliation, P = Practice, B = Belief, Net = Social Network, S = Size Ad = Advice Network, E = Emotional Closeness, Edu = Education, Gen = Gender / Sex, Nat = English Native Language, VA = Verbal Ability, Extra = Extraversion, Agree = Agreeableness, Cons = Conscientiousness, Emo = Emotional Stability, Open = Openness to Experience.

Table 2.4. Regression Summary statistics. Separate one- and two-factor models were run for each listed outcome variable individually.

Outcome Variable	Model		Regression Paths (standardized)	Standard Error	P - value
Mindreading	2-Factor	MST	.070	1.132	.627
Motivation		APP	.055	1.213	.705
	1-Factor	MST + APP	.088	.710	.349
Loneliness	2-Factor	MST	-.020	.103	.876
		APP	.013	.107	.919
	1-Factor	MST + APP	-.029	.066	.733
Religiosity	2-Factor	MST	.438	.133	.001*
		APP	-.532	.143	.001*
	1-Factor	MST+APP	-.049	-.046	.602
Authoritarianism	2-Factor	MST	.202	.107	.159
		APP	-.244	.114	.091
	1-Factor	MST + APP	.032	.067	.736
Social Network	2-Factor	MST	.039	1.827	.781
Size		APP	-.146	1.954	.307
	1-Factor	MST + APP	-.078	1.141	.396
Advice Network	2-Factor	MST	-.137	.272	.359
Size		APP	.124	.291	.359
	1-Factor	MST + APP	-.113	.169	.239
Emotional	2-Factor	MST	-.016	1.174	.910
Closeness		APP	.049	1.255	.689
	1-Factor	MST + APP	-.012	.730	.895
Anthropomorphis	2-Factor	MST	.040	3.820	.778
m		APP	.082	4.086	.570
	1-Factor	MST + APP	.082	2.384	.375

Note. *$p < .01$. MST = Mental State Terms; APP = Response appropriateness. See supplementary material S3 for details on the formation of the religiosity latent factor. Given that only two models were significant at the p=.001 level, we did not account for multiple comparisons as false positives were extremely unlikely.

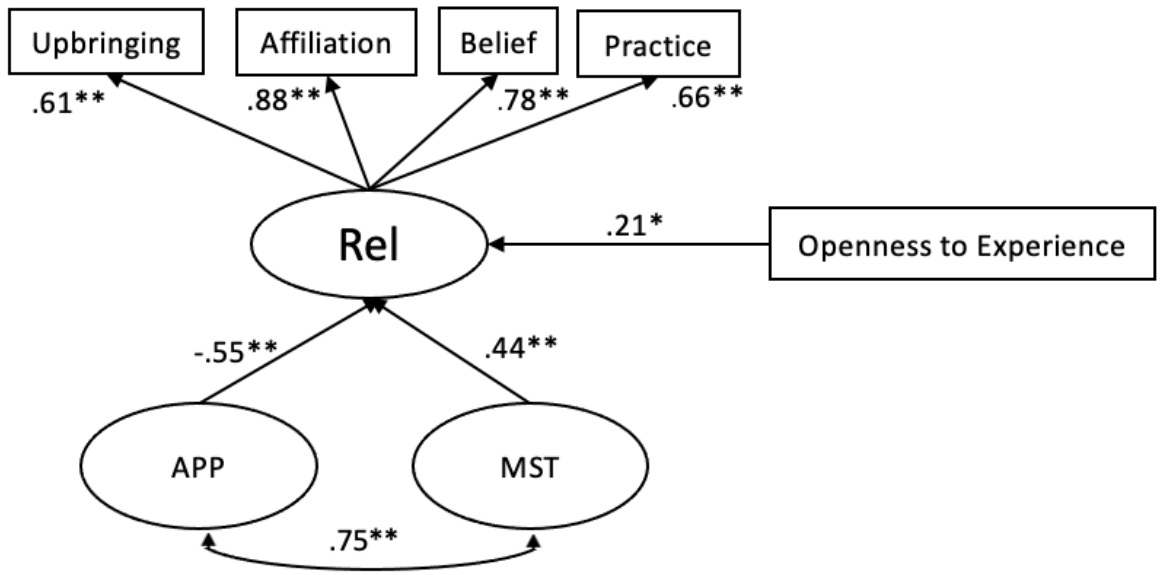

Figure 2.2 ** p<.001, *p<.05. Simplified path diagram showing standardized estimates for statistically significant regression paths for the two-factor religiosity latent factor model (see Table 4). Covariances between predictor variables are not displayed. Non-significant variables controlled for in this model are age, gender, verbal ability, English as native language, and personality (e.g., extraversion, agreeableness, conscientiousness). Rel = Religiosity; APP = Response appropriateness; MST = Mental State Terms.

2.4 Discussion

In this pre-registered study, we examined whether quantity and quality of mindreading were (a) separable constructs, (b) related differentially to a variety of outcome measures, and (c) accounted for by individual differences in mindreading motivation. To do this, we measured individual differences in adults' response "appropriateness" and their frequency of mental state terms (MST) on the Silent Film Task (SFT) and Movie for the Assessment of Social Cognition (MASC). Although response appropriateness and frequencies of MST were strongly correlated with one another, a two-factor model with frequency of MST and response appropriateness loading onto two separate but correlated factors provided greater explanatory power than a one-factor model when predicting a socially relevant outcome. Specifically, frequency of MST was positively associated with religiosity and response appropriateness was negatively associated with religiosity, whereas the one factor model did not predict any outcomes. There were no significant associations between either factor and mindreading motivation. This provides some support for the distinctiveness of mindreading quantity (measured as frequency of MST) and mindreading quality (measured as response appropriateness) in the structure of individual differences in mindreading. These results are discussed in more detail below.

2.4.1 Evidence for a Two-Factor Structure of Mindreading

A significant current challenge for the field is to advance theoretical understanding of how mindreading abilities continue to vary in adults and to devise psychometrically robust ways to measure these individual differences. Previous research has proposed that mindreading may comprise multiple component processes such as cognitive versus affective (e.g., de la Osa, Granero, Domenech, Shamay-Tsoory, & Ezpeleta, 2016; Rossetto et al., 2018), social-perceptual versus social-cognitive **(**Tager-Flusberg & Sullivan, 2000) or effortful versus automatic (Apperly & Butterfill, 2009). Whereas these proposals focus on the cognitive architecture of mindreading, we sought to characterise sources of variation in performance. This is relevant for any given component process, for "advanced" tasks like the MASC and SFT that combine multiple components, and for everyday use of mindreading outside of scientific settings.

Unlike many "laboratory" tasks (e.g., Warnell et al., 2019; Qureshi et al., 2020) the MASC and SFT are known to generate robust individual differences and to predict social

outcomes, but their coding schemes combine the quality and quantity of mindreading responses into a single scale reflecting the degree to which a response is "correct". It might be that adults' mindreading truly varies on a single dimension. However, it is not possible to test this hypothesis when mindreading quality and quantity are baked into a unidimensional coding scheme. Based upon growing evidence that motivation contributes to individual differences in mindreading (e.g., Carpenter et al. 2016; Contreras et al., 2020; Devine & Apperly, 2022), we hypothesised that some people might be disposed to mindread more than others, even if they were not especially successful in doing so. Therefore, we sought to code mindreading quality and quantity separately, and to test whether these factors were indeed reflected in the covariance structure of the tasks.

To do so, we tested (a) whether both constructs loaded onto one or two latent factors; and (b) whether a one- or two-factor model was superior at predicting outcome variables which had previously been linked with mindreading. Evidence on the first criterion was equivocal, as both a one- and two-factor solution provided an acceptable fit for the data. Whilst response appropriateness and frequencies of MST were from the same two tasks, there was a risk that either the one-factor solution, or the large amount of shared variance between the constructs in the two-factor solution, were due to shared methods rather than similarity in the underlying constructs. However, our analyses found no evidence for this because we accounted for shared method variance by running a two-task model, which did not provide a good fit for the data.

In relation to our second criterion, we found that a one-factor model (representing the shared variance between MST and response appropriateness) did not predict any of the outcomes measured in this study. In contrast, a two-factor solution revealed significant, differential associations between the MST latent factor and response appropriateness latent factor (two-factor model) with religiosity. This finding suggests that despite substantial shared variance, response quality and quantity have explanatory power as distinct constructs. It remains important for future work to determine whether the non-shared or shared variance between MST and response appropriateness predicts other criterion variables above and beyond the ones employed here.

In summary, given the support we find for both a one-factor and two-factor solution, consideration of parsimony might favour the one-factor solution. However, since the two-

factor solution has superior power at predicting outcomes, we believe there are good grounds to favour the two-factor solution.

2.4.2 Mindreading Motivation

Variation in performance on the MASC and SFT could indicate differences in motivation to reason about others (e.g., Apperly et al., 2006; Devine & Apperly, 2022; Ferguson et al., 2017; Qureshi et al., 2020; Kouklari et al., 2018; Lin et al., 2010). However, the present results indicated that self-reported motivation was not associated with either the amount of MST people used in their responses or with the appropriateness of their responses. It should be noted that minimally 13 and maximally 91 points could be achieved on the MRM scale but in the current sample, the lowest MRM score was 47, and the highest 89 (with data normally distributed between this range). This suggests that the variance of the MRM variable may have been restricted in the current study, perhaps because most of the participants were undergraduate psychology students. Likewise, a one-factor solution where frequencies of MST and response appropriateness loaded onto one factor was not associated with MRM. The present study therefore provided no evidence that frequencies of MST, and / or response appropriateness captured meaningful individual differences in self-reported mindreading motivation. However, it is important to note that the present study used a relatively brief self-report measure of mindreading motivation. It remains possible that different effects might be observed within a more diverse sample, or with more direct measures of motivation, or on tasks that assess a broader range of social motivations, including but not limited to motivation for mindreading. The origin of variation in MST remains currently unclear and future research will need to examine whether the observed lack of association between frequency of MST and mindreading motivation reflected a genuine lack of association between the two constructs.

2.4.3 Social Outcome Variables

A one-factor mindreading factor was not uniquely associated with any of the outcome variables, but there were significant, differential associations between the latent factors of frequencies of MST and mindreading response appropriateness with religiosity. Specifically, a higher degree of MST predicted religiosity (composed of the sub-components of religious belief, affiliation, practice, and upbringing) positively, whereas greater response appropriateness was a significant negative predictor of religiosity. This means that religious

participants were less likely to give appropriate (or context-sensitive) answers on two mindreading tasks and more likely to use mental state language than non-religious participants. Previous research is in line with the notion that religiosity may be negatively (or not at all) related to appropriate mentalising about other *human* minds, which differ qualitatively from supernatural agents' minds (e.g., Vonk & Pitzen, 2017; Zmigrod, Rentfrow, Zmigrod, & Robbins, 2019). The converse positive association between frequencies of MST and religiosity is likewise in line with theoretical accounts arguing that religiosity may have emerged because of the human propensity to attribute mental states (e.g., Bering & Shackelford, 2004; Norenzayan, Gervais, & Trzesniewski, 2012).

The present findings highlight the potential distinctiveness of mindreading quantity (i.e., frequency of MST) and mindreading quality (i.e., response appropriateness) in the structure of individual differences in mindreading.

2.4.4 Limitations and Future Directions

First, the measures available to investigate the constructs in the present study are limited. For instance, whilst we have used a validated self-report questionnaire to measure mindreading motivation, participants' scores on this measure were restricted in range. In addition, some of the covariate measures (i.e., personality, verbal ability) exhibited low internal reliability, which may have prevented the detection of relationships. Furthermore, participants' responses regarding their social network may have been influenced by the desire to present a socially desirable image, potentially leading to an overestimation of the number of contacts or level of emotional closeness. This may have compromised the validity of this data (e.g., Van de Mortel, 2008). Researchers often distinguish between religiosity and spirituality (McPhetres & Zuckerman, 2018), and this distinction might be relevant to the issue of mindreading. Religion is commonly perceived as more institutionalised and dogmatic than spirituality (e.g., Popp-Baier, 2010). Therefore, if spirituality is less tied to religious concepts (e.g., a personal God who is concerned with each individual's life), it may not share with religiosity its positive relation with making frequent explicit mentalistic attributions to others. Finally, it should also be noted that a large proportion of participants in the present study were female psychology students. Whilst this means we must be cautious about generalising our findings to a broader population, the possibility of reduced variance in our relatively homogeneous sample if anything reduced the likelihood that we would succeed in detecting a reliable factor structure

for our measures. Further research should use a non-university sample, as well as include a range of educational backgrounds and ages to be more inclusive and generalisable to a larger number of individuals.

2.4.5 Conclusion

In summary, our findings are consistent with the idea that response quality and quantity may be differentiable dimensions of open-ended response mindreading tasks in adults. They demonstrate valid non-shared variance between mindreading quality and quantity, which has previously been obscured by not differentiating between the degree to which participants quantitatively engage with others' mental states and the degree to which their responses are appropriate. Coding mindreading appropriateness and frequency of MST separately enabled us to detect a differential relationship between mindreading appropriateness and frequency of MST with religiosity, that would have been obscured with a one-factor solution. However, more research is needed to elucidate how exactly motivation relates to these components of mindreading. The present approach helps our understanding of why mindreading varies in adults who already understand basic mental state concepts of their own and others' thoughts, beliefs and desires and may ultimately help us to understand how an adult's propensity to effortfully engage with others' mental states (e.g., their motivation to mindread) may interact with their ability to do so appropriately.

Chapter 3

Convergence between the MASC and a Novel Task Assessing the Structure of Mindreading

3 Convergence between the MASC and a Novel Task Assessing the Structure of Mindreading

3.1 Introduction

In recent years, there has been a surge in academic research focussing on the development of advanced mindreading tasks. These tasks aim to capture the ongoing development of the ability to infer others' mental states beyond the pivotal age of four or five, a critical developmental milestone when neurotypical children attain a conceptual understanding of mental states (Weimer et al., 2021; for a meta-analysis, see Wellman et al., 2001). Indeed, 'mindreading performance' has been found to improve further throughout middle childhood (Devine & Hughes, 2016), adolescence (Dumontheil et al., 2010; Hollarek & Lee, 2022; Peterson & Wellman, 2019) and adulthood (Duval et al., 2011), with declines in later adulthood (Henry, Phillips, Ruffman, & Bailey, 2013). Furthermore, there are also normative individual differences in mindreading, even among neurotypical adults (e.g., Dziobek et al., 2006; Conway et al., 2019; Apperly & Butterfill, 2009; for a focus on individual differences in mindreading in children and adolescents, see Hughes & Devine, 2015 and Devine, 2021). However, despite the availability of many advanced mindreading measures, two key issues persist within the field. Firstly, there remains a scarcity of psychometrically robust measures suitable for assessing individual differences in mindreading among typical adults (e.g., Warnell & Redcay, 2019; Oakley et al., 2016; Bradford et al., 2018). Secondly, existing measures fall short of being able to provide insight into *why* individual differences in adults exist who have already attained the ability conceptualise mental states (e.g., Apperly, 2012; Conway et al., 2019; Dziobek et al., 2006).

Within neurotypical adults, the primary research interest lies in understanding the nature and implications of individual differences in mindreading (e.g., Dziobek et al., 2006; Conway et al., 2019; Apperly, 2012). As research in adults has predominantly employed experimental tasks that were originally developed to enable group-level comparisons between clinical and non-clinical groups (e.g., Bradford et al., 2018; Dziobek et al., 2006; Preißler et al., 2010; Livingston et al., 2021), there are relatively few psychometrically robust measures to assess individual differences in mindreading in adults (e.g., Warnell & Redcay, 2019). Furthermore, the range of suitable mindreading tasks for adults exhibit limited convergent

validity (i.e., two measures of the same construct do not correlate with each other; Warnell& Redcay, 2019), which may be because some tasks deemed to measure mindreading, such as the Reading the Mind in the Eyes Task (Baron-Cohen et al., 2001) assess other constructs, such as facial emotion recognition (e.g., Oakley et al., 2016).

Furthermore, despite significant progress in the operationalisation of mindreading, and theoretical as well as empirical evidence suggesting that mindreading is underpinned by distinct cognitive processes (e.g., Schaafsma et al., 2015; Abu-Akel & Shamay-Tsoory, 2011; Apperly & Butterfill, 2009; Gopnik & Wellman, 1992; Preston & De Waal, 2002; Poletti, Enrici, & Adenzato, 2012; Bernhardt & Singer, 2012), even existing mindreading tasks that have been successful at capturing individual differences in neurotypical adults' mindreading (i.e., the MASC; Dziobek et al., 2006) fall short of being able to explain *why* individuals may exhibit differences in their reasoning about mental states. For example, while the MASC (Dziobek et al., 2006) differentiates mentalising errors into distinct categories, such as overly complex mental state reasoning or the complete absence of mentalistic attributions, it does not illuminate potential 'structural' differences in mindreading. This is a notable limitation given that, when people attribute mental states to others, these attributions, whether deemed correct or incorrect, may stem from distinct underlying cognitive processes. To illustrate, an incorrect mental state attribution could result from having actively considered various alternative mental states that could explain a behaviour occasionally selecting the 'wrong one'. Alternatively, it may originate from difficulty or disinterest in making accurate mental state attributions, likewise leading to an inaccurate response. In both cases, individuals engage in mentalising about others in an inaccurate way but the underlying reasoning driving these mental state attributions has arguably differed.

One potential explanation for the limitations of current measures is their predominant focus on assessing the accuracy of participants' mental state attributions. What constitutes an accurate response is usually determined by the researchers who designed tasks based on criteria such as logic (e.g., White, Coniston, Rogers, & Frith, 2011; Devine & Hughes, 2013), ground-truth (i.e., true and objective data used as a reference to assess performance such as comparing inferences of others' mental states with self-reported mental states of these individuals; Long et al., 2022), or consensus scoring (e.g., Dziobek et al., 2006). However, as illustrated in the last paragraph, there is more than one way of being inaccurate. Therefore, while accuracy is a critical objective of mindreading, it may not be the most enlightening variable in terms of

understanding why there are individual differences in mindreading, perhaps especially in adults. Shifting focus from the 'content' of mentalising (i.e., the specific mental states attributed) to the structure of mentalising, could elucidate *why* adults, who already possess core mental state concepts of their own and others' mental states (e.g., Apperly, 2012), still exhibit differences in their mindreading. If such structural differences in how individuals mindread exist, and can be measured reliably, this shift in perspective can inform how mindreading is operationalised (i.e., as a unitary or multifaceted construct, as suggested by Schaafsma et al., 2015). Additionally, it offers the possibility to evaluate what such variation means (i.e., does it converge with existing measures?) as well as the possibility to re-evaluate mindreading-related 'outcomes' from a novel perspective, potentially clarifying the mixed findings concerning the implications of individual differences in neurotypical adults (e.g., Nestor, Liu, Sutherland, & Cole, 2023).

To achieve the goal of elucidating the structure of individual differences in mindreading without assessing correctness, we initially conducted a series of data-driven pilot studies without a priori hypotheses to overcome existing methodological challenges[1]. This approach was influenced by the ongoing reproducibility crisis in psychology (e.g., Van Bavel, Mende-Siedlecki, Brady, & Reinero, 2016), which has raised questions as to what extent psychological research is influenced by researchers' characteristics (i.e., researchers are, on average, unrepresentative of the general population; Haidt & Jussim, 2016) and agendas (e.g., Head, Holman, Lanfear, Kahn, & Jennions, 2015). Thus, we created a coding scheme focussing on the features of mindreading and refraining from imposing judgement about whether a specific response is correct or incorrect. Specifically, we identified four markers of structural differences in the open-ended responses of adults (or 'mindreading dimensionality'). These responses were elicited by a range of static picture stimuli of social scenarios that depicted two or three interacting individuals of differing demographic characteristics (i.e., age, gender, ethnicity). Regardless of the *specific* interpretations provided, participants' responses differed with regard to **(a)** the *quantity of mentalistic terms* used (i.e., words that describe the various

[1] This pilot work consisted of three separate online data collections (two involving student volunteers and one involving financially compensated participants) to refine stimuli and task instructions. The results from these studies informed the methodology of the two studies reported here and included analysis of frequency distributions and bivariate correlations exclusively. During these pilot studies, I maintained openness as to which other dimensions may be found in the data, however, at no point were other dimensions identified and explored. Notably, the data from these studies revealed that providing participants with examples of responses based on example picture stimuli in the introduction to the task led to greater response variation. The examples included both mindreading and non-mindreading responses. See Appendix for the complete instructions.

internal cognitive and emotional states that can occur within an individual's mind, words such as 'to know', 'to feel', 'to want', see Chapter 2, for a more detailed description), **(b)** whether or not reasons for mentalistic attributions, if they were present at all, were *explained* (i.e., written evidence of reasoning about why a certain mental state has arisen or of what behavioural consequence that mental state has), **(c)** whether or not *'uncertainty'* was expressed (i.e., expressions indicating openness as to the extent to which one can be sure of a specific interpretation, signified by words such as "may", "could", "possibly"), and **(d)** whether an *alternative interpretation* of the social scenario was provided (i.e., whether a response included at least two different interpretations or descriptions of the scenario; see Table 1 for an example of the coding scheme that was developed based on these dimensions).

Firstly, considering the frequency with which people refer to *mental state terms (MST)* as a marker of individual differences in mindreading also has empirical and theoretical support (e.g., Meins et al., 2006; Devine & Hughes, 2019; Lecce et al., 2019). While a higher quantity of MST alone does not necessarily reflect a higher quality of engagement with others' mental states, it serves as a quantitative marker of an individual's spontaneous efforts to mindread (see Chapter 2). Secondly, attribution theory (e.g., Harvey & Weary, 1984; Weiner, 1985) posits that individuals fundamentally seek to understand, and attribute causes and consequences to observed behaviour. Therefore, we expected that *explanations* would frequently accompany open-ended responses to stimuli intended to elicit mentalizing (i.e., the picture stimuli used in the present studies), elucidating more deeply *how* people mentalise. Given the highly dynamic nature of mental states in real-life contexts, it is plausible that flexibility plays a key role in successful mentalising (e.g., Apperly & Wang, 2021; Michael & Christensen, 2016). Thus, we hypothesised that the provision of alternatives and expressions of uncertainty indicate greater flexibility in mentalising. Finally, there is evidence that people frequently engage in counterfactual reasoning, that is, the cognitive process of considering alternatives that are 'counter to reality' (e.g., Harris, German, & Mills, 1996). Indeed, Guajardo and Turley-Ames (2004) reported associations between mindreading performance and counterfactual thinking, suggesting that the consideration of alternatives may be important to mindreading, too. Research also suggests that individuals differ in their cognitive control tendencies, balancing persistence with flexibility (for a review, see Hommel & Colzato, 2017). Although goal-directed actions require stable objectives, behavioural adaptability to *alternative possibilities* is essential when facing changing situational demands (Dreisbach, & Goschke, 2004).

Notably, it might be that the above-outlined markers of structural features of mindreading reflect more general reasoning processes (i.e., indicating that mindreading does not constitute a 'unique' capacity), which could account for individual differences in broader socio-cognitive phenomena, such as empathy or decision-making. Accordingly, researchers have emphasised that commonly used measures of mindreading likely tap into various cognitive processes (Warnell & Redcay, 2019; Hayward & Homer, 2017; Quesque & Rossetti, 2020). In the present studies, we therefore (a) statistically accounted for other constructs (i.e., memory, executive functioning, or verbal ability) and (b) assessed mindreading on our novel measure in relation to performance on an established measure, namely the MASC (Dziobek et al., 2016).

3.1.1 The Present Studies

In summary, for the reported research, we had three primary objectives. First, we aimed to test whether there were individual differences in 'dimensional mindreading' (i.e., MST, explanations, certainty, and alternatives) across two independent samples, as assessed in Study 1 and Study 2.

Our second goal was to examine the factor structure for 'mindreading dimensionality'. Confirmatory factor analyses (CFA) were conducted to explore whether frequencies of MST, and the presence of explanations, uncertainty, and alternatives (a) represented a single mindreading construct, (b) four distinct but correlated constructs, or (c) an overarching general mindreading factor (higher-order factor) explaining the common variance between the four first-order factors. These alternative measurement models were tested in Study 1 and cross-validated in Study 2. In Study 2, we additionally considered the possibility of a bi-factor model to verify whether method factors (i.e., due to similarity between certain picture stimuli) and more domain-specific abilities can explain variance in the indicators in addition to a more substantive factor(s) representing mindreading.

Our third aim was to test the convergent validity of the dimensional mindreading measure with an established measure based on accuracy, the Movie for the Assessment of Social Cognition (MASC; Dziobek, 2006). In Study 1, associations between both tasks were assessed using correlational analyses and in Study 2 based on Structural Equation Models (SEM). The MASC is presented as a 15-minute movie with 45 stopping points, each representing an individual test item. Each item is followed by a multiple-choice question about

characters' mental states with four alternative choices. The MASC demonstrates strong internal consistency, with Cronbach's alpha values exceeding .80 in various adult samples (e.g., Dziobek et al., 2006; Preißler et al., 2010) Test-retest reliability (r = .97, Dziobek et al., 2006) has also been reported to be high.

Given that the MASC evaluates participants' responses based on the 'correctness' of their responses, we expected small to moderate associations between the MASC and the novel approach to coding mindreading in a dimensional, as employed in in the present study.

3.2 Study 1

In Study 1, we aimed to assess the psychometric properties of the dimensional measure by (a) assessing whether there were *individual differences* across the four coded dimensions (i.e., MST, mentalistic explanations, uncertainty, and alternatives), (b) assessing convergent validity based on correlations between the observed overall scores on the dimensional mindreading task and an established measure of mindreading, the Movie Assessment of Social Cognition (MASC; Dziobek et al., 2016), and (c) to examine the factor structure of the dimensional measure by comparing, a one-factor solution, four-factor solution and higher-order factor solution (and cross-validating the best-fitting model in Study 2).

3.2.1 Method

3.2.1.1 Participants

Participants (N = 151, Mage = 19.33, SD = 2.09, *range* = 18 – 37 years, 139 females (92.7%), 11 males (7.3%) forming part of the University of Birmingham School of Psychology undergraduate research database took part in this study for university course credit. Based on power analysis, to obtain a statistical power of .80, a minimum sample size of 150 participants was needed. To be eligible, participants had to be at least 18 years old, have no current or past neurological problems, and be English-speaking (this included native and non-native English speakers). Data on ethnicity was not recorded in this study. No participant was excluded from the study. Participants provided their informed consent prior to participation and this study has been approved by the University of Birmingham STEM Research Ethics Committee.

3.2.1.2 Measures and Procedure

Movie Assessment of Social Cognition (MASC; Dziobek et al., 2016)

Participants watched a 15-minute movie, which was stopped several times to ask 45 multiple-choice questions with 4 response choices, each representing one of four distinct dimensions. Specifically, responses to these questions were scored as either correct or incorrect, with further differentiation of the incorrect answers into over-mentalising (i.e., overly complex mental-state reasoning), no mentalising (i.e., failing to make a mental state attribution) and under-mentalising (i.e., overly simplistic mental state inferences). For each item, participants received a score of 0 for the three response options not selected and a score of 1 for the response they selected, with a possible range of 0 – 45. Scores for each sub-scale were summed up to form a scale for 'correct mentalising', $\alpha = .776$, 'under mentalising', $\alpha = .538$, 'no mentalising', $\alpha = .430$, 'over-mentalising', $\alpha = .591$.

Dimensional Mindreading Task

Participants were presented with 8 picture stimuli showing different people (either two or three individuals of varied ages, genders, and ethnicities) interacting with each other. Each picture was followed by an open-ended question: *"What do you think happened in this scenario? (Please write your answer into the text box below. Most people write 20 – 200 words)."* The time participants had to describe each scenario was not restricted. Participants' written descriptions of their interpretation of the 8 scenarios were coded in terms of their word count as well as four distinct dimensions: (**a).** the number of mental state terms (MST) present within a response (i.e., the total number of MSTs present within a response), (**b).** the presence of an explanation for the mental states discussed (coded as a binary variable; 1 = explanation present, 0 = explanation not present), (**c).** the presence of 'uncertainty' within a response (coded as a binary variable; 1 = uncertainty present, 0 = uncertainty not present), and (**d).** the presence of an alternative interpretation of the situation (coded as a binary variable; 1 = alternative interpretation present, 0 = alternative interpretation not present). A full copy of the coding scheme can be found in the Appendix. Inter-rater reliability was calculated for 20% of participants and was acceptable (Krippendorf's α for MST = .88, for explanation = .73, for certainty = .85, for alternative = .79). Table 3.1 contains an example of how participants' responses were coded. To code MST, based on previous research (i.e., Taumoepeau et al.,

2019; Devine & Hughes, 2019), the coding scheme used in this study was adapted from existing coding schemes (e.g., Brown et al., 1996) and extended with words that fit the general definition of what constitutes a mental state term but were not explicitly mentioned in the existing coding schemes considered. References to participants' own mental states (e.g., *"I think")* were not coded. The approach taken to coding mental state terms was the same as the one outlined in Chapter 2.

Table 3.1. Example of how participants' responses on the dimensional mindreading task would be coded.

Participant response 	**MST**	**Explanation**	**Certainty**	**Alternative**
(…). The man *appears to be* listening intently to what the woman is saying as he is looking her in the eye. He has a slight smile on his face and appears thoughtful, so *perhaps* they are having an entertaining conversation, or she has just told him something he feels happy about.	5	1	1	1
They are on a date in a restaurant, they are discussing something over dinner with red wine. The situation is happening in the evening and both people feel awkward.	2	0	0	0
There are two people walking alongside each other. They are both looking serous because they are thinking about an unpleasant interaction they had with a client or are concerned about something more generally	2	1	0	1

Control Measures

Wechsler Abbreviated Scale of Intelligence (WASI – 2, Wechsler, 2001)

Current intellectual functioning (verbal ability) was assessed using the Similarities Test of the Wechsler Adult Scale of Intelligence (WASI). Participants completed 20 trials (with 20 items), explaining why two words were alike. For example, participants were asked: *"In what way are grapes and strawberries alike?"*, to which the correct answer would be *"both fruit".* Responses were scored using the WASI manual and items were scored on a 3-point scale from 0 to 2, with higher scores awarded for the most correct answers. Scores were added up, with higher scores indicating greater verbal ability, α = .500.

Ten Item Personality Measure (TIPI) (Gosling, Rentfrow, & Swann, 2003)

The 10-item measure of the Big Five (or Five-Factor Model) dimensions of personality was administered. Participants indicated the degree to which they ascribed different personality traits to themselves, using a 7-point Likert scale ranging from 1 (*Disagree strongly*) to 7 (*Agree strongly*). Items 2, 4, 6, 8 and 10 were reverse coded. The average of two items (standard item and reverse-scored item) made up each scale (i.e., extraverted, enthusiastic & reserved, quiet, α = .68, Critical, quarrelsome & Sympathetic, warm, α = .63, Dependable, self-disciplined & Disorganised, careless, α = .62 Anxious, easily upset & Calm, emotionally stable, α = .81, Open to new experiences, complex & Conventional, uncreative, α = .73).

3.2.2 Results

3.2.2.1 Analysis Plan

All data were analysed in SPSS (Version 29.0) and all confirmatory factor analysis models were estimated in the Lavaan software package in R (Version 5.23; Rosseel, 2012; R Development Core Team, 2016). Confirmatory factor analysis is a technique used to test and estimate relationships between observed and unobserved variables to construct a measurement model. To evaluate model fit, we used the *root mean square of approximation* (RMSEA) with values between 0.05 and 0.08 as adequate model fit (Kline, 2011; Brown, 2015); the *comparative fit index* (CFI); and the *Tucker-Lewis index* (TLI), for which values above .95 are a good fit (Hu & Bentler, 1999). An alpha level of .05 was to assess the statistical significance of the regression paths of interest. In line with recommendations for individual differences researchers, we considered correlations of .10 as a small effect, .20 a typical (or moderate) effect and > .3 as a relatively large effect (Gignac & Szodorai, 2016). Where two or more

models met all three criteria, we selected the simplest model with the lowest AIC value and considered which model was most strongly supported by theory for further analyses. Regarding missingness, a full information maximum likelihood (FIML) approach was adopted. This method has been shown to produce unbiased parameter estimates, standard errors and test statistics that are consistent and efficient when data is either missing at random (MAR) or missing completely at random (MCAR) (Enders & Bandalos, 2001). Because the data were missing at random, we used a means-and-variance adjusted weighted least squares (WLSMV) estimator to process all available observations when estimating models (e.g., Lei & Shiverdecker, 2020).

3.2.2.2 Descriptive Statistics

All variables of interest were normally distributed (Table 3.2; Figure 3.1). Notably, the observed scores of the DMT, explanation, certainty, alternative and MST were all positively inter-correlated with each other, *rs* = .212 - .440, indicating that participants who exhibited any one dimension of the DMT more, were also more likely to exhibit the remaining three sub-categories more. In addition, these positive correlations suggest that there was moderate shared variance amongst the four sub-categories of the DMT (Table 3.3). Furthermore, there were weak-moderate, positive, and significant correlations between the 'MASC correct' subscale and all sub-categories of the DMT, *rs* = .174 - .239 (Table 3.3; Figure 3.2). These correlations show that participants who had higher mindreading ability according to the MASC, were more likely to use more mentalistic terms, express more uncertainty, provide an explanation for the mental states they attribute and suggest an alternative interpretation of the social scenarios. Conversely, poorer performance on the MASC (i.e., on the over-mentalising, under-mentalising and no mentalising sub-scales) was either unrelated to performance on the DMT, or significantly negatively correlated with the sub-scales of the DMT, *rs* = -.074 - -.255 (Table 3.3; Figure 3.2). Overall, these correlations may be interpreted as reflecting good convergence between both tasks, based on the observed scores. In addition, the DMT measures do not appear to simply capture individual differences in verbal ability or personality.

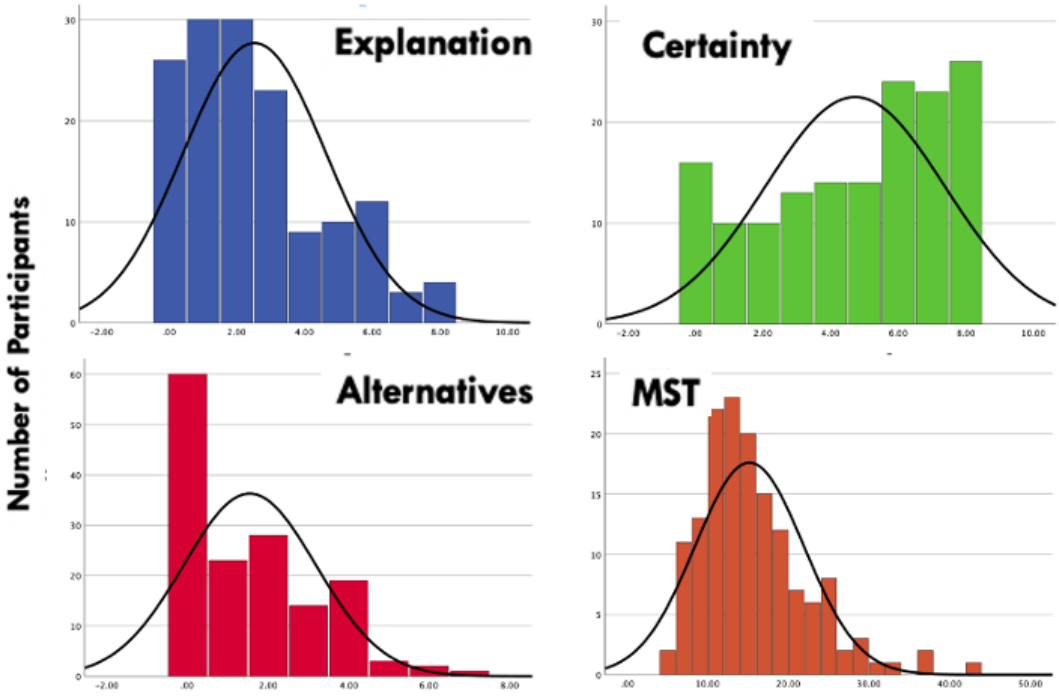

Figure 3.1. Visual representation of the distribution of scores on the DMT task based on 151 participants.

Table 3.2. Descriptive Statistics for each study measure.

Variable	*M (SD)*	*Min*	*Max*	*Range*	*Skewness (SE)*	*Kurtosis (SE)*	*N*
Explanation	3.69 (2.42)	0	8	8	.163 (.20)	-1.00 (.39)	150
Certainty	4.72 (2.66)	0	8	8	-.44 (.20)	-1.08 (.39)	150
Alternative	1.54 (1.65)	0	7	7	.87 (.20)	-.03 (.40)	150
MST	15.15 (6.76)	5	43	38	1.21 (.20)	2.08 (.40)	149
MASC Cor	32.99 (5.60)	13	42	29	-1.01 (.20)	1.09 (.39)	151
MASC Over	6.20 (3.38)	0	20	20	1.06 (.20)	1.49 (.39)	151
MASC Under	3.56 (2.49)	0	12	12	1.13 (.20)	1.40 (.39)	151
MASC No	2.07 (1.80)	0	9	9	1.29 (.20)	1.63 (.39)	151
Extraversion	3.74 (1.57)	1	7	6	.17 (.20)	-.92 (.39)	151
Agree	4.87(1.12)	1.50	7	5.5	-.28 (.20)	-.25 (.39)	151
Conscientious	5.19 (1.25)	1	7	6	-.91 (.20)	.55 (.39)	151
Emo Stability	3.69 (1.33)	1	7	6	.24 (.14)	-.48 (.39)	151
Openness	4.99 (1.17)	1.50	7	5.50	-.32 (.20)	-.17 (.39)	151
WASI	28.08 (3.43)	19	38	19	-.25 (.20)	.07 (.40)	147

Note. *M* = mean, *SD* = Standard Deviation, Min = observed minimum value, Max = observed maximum value, SE = Standard Error, *N* = number, MASC = Movie for the Assessment of Social Cognition, Over = Over-mentalising, Under = Under Mindreading, No = No Mindreading, MST = Mental State Terms; Expr. Certainty = Expressed certainty; Agree = Agreeableness; Conscientious = Conscientiousness; Emo Stability = Emotional Stability; Openness = Openness to Experience, WASI = Verbal Ability

Table 3.3. Correlations between all study variables.

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Explanation	-												
2. Expr. Cert	.378**	-											
3. Alternative	.212**	.440**	-										
4. MST	.333**	.312**	.301**	-									
5. MASC Cor	.239**	.174*	.236**	.198*	-								
6. MASC Exc	-.074	-.126	-.202*	-.082	-.757**	-							
7. MASC Less	-.248**	-.060	-.099	-.139	-.688**	.150	-						
8. MASC No	-.255**	-.215**	-.216**	-.231**	-.695**	.253**	.479**	-					
9. Extraversion	-.139	-.148	-.075	-.161*	-.028	.036	-.014	.031	-				
10. Agree	.009	.060	.109	.080	-.118	.036	.103	.087	-.140	-			
11. Conscien	.083	.056	.200*	.081	.056	-.108	.028	-.013	-.053	.113	-		
12. Emo Stab	-.141	-.147	-.019	-.024	-.098	.034	.057	.121	.084	.158	.308**	-	
13. Openness	-.042	-.148	-.048	-.153	-.039	.018	.002	.070	.313**	.043	.019	.070	-
14. WASI	.123	-.019	.043	.130	.266**	-.101	-.287**	-.212*	-.004	-.092	.103	-.072	.073

Note. MASC = Movie for the Assessment of Social Cognition, Exc = Excessive Mindreading, Less = Less Mindreading, No = No Mindreading, MST = Mental State Terms; Expr. Certainty = Expressed certainty; Agree = Agreeableness; Conscien = Conscientiousness; Emo Stab = Emotional Stability; Openness = Openness to Experience, WASI = Verbal Ability. *p<.05, **p<.01.

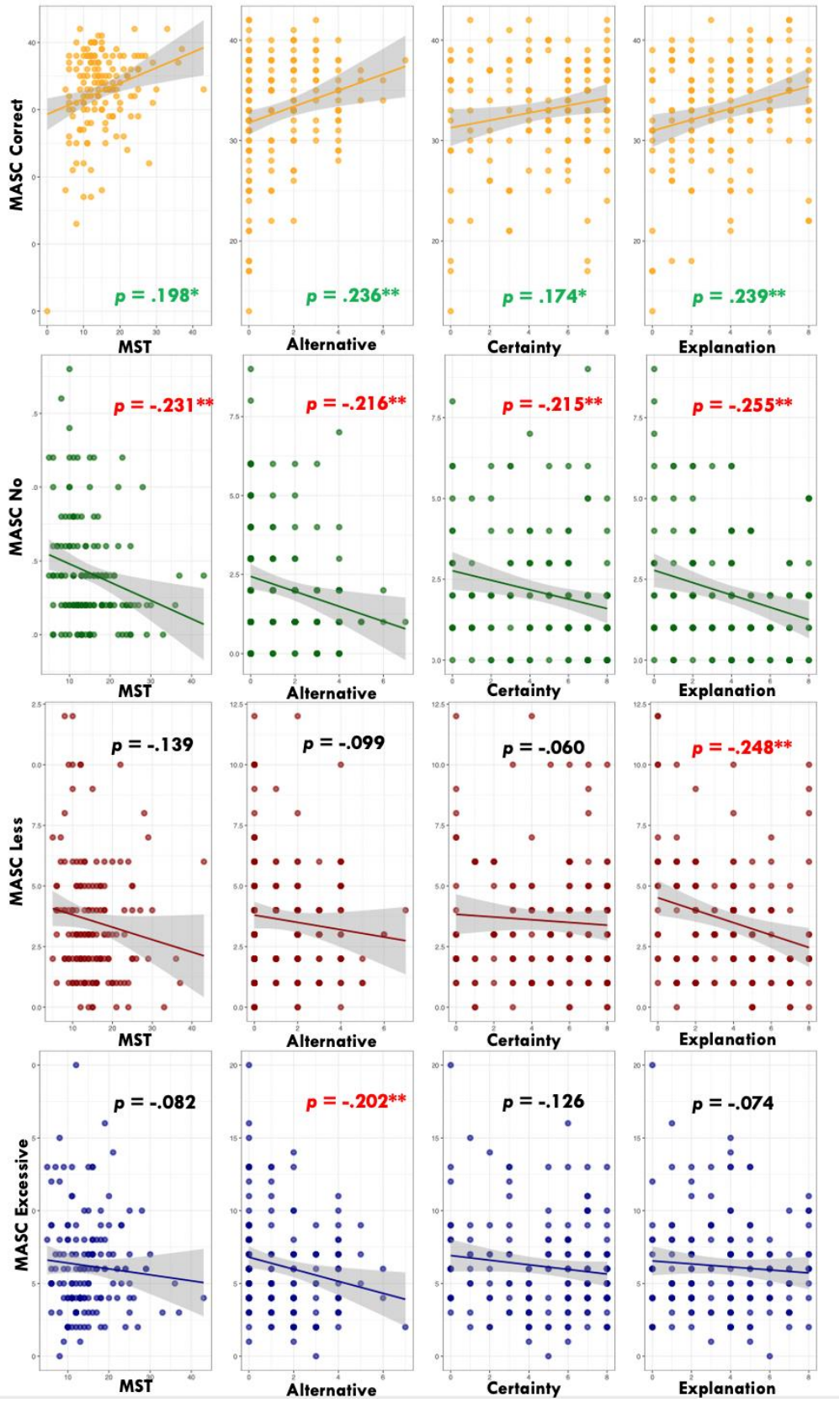

Figure 3.2. Correlation between four MASC subscales and four indicators of mindreading. Confidence intervals reflect 95% CI.

3.2.2.3 Structural Equation Modelling

Three separate CFAs were estimated to test which factor structure best fit the data. Specifically, we tested the following models (see Table 3.4 for the fit statistics for each model).

1. **A one-factor model** where all 32 indicators representing the categories of MST (8 indicators), explanation (8 indicators), certainty (8 indicators) and alternative (8 indicators) loaded onto one factor. If a one-factor model provided the best fit, this would indicate that all indicators measure the same construct (i.e., a single mindreading factor) and are statistically indistinguishable from one another.
2. **A four-factor model** were the 8 indicators representing the categories of MST, explanation, certainty and alternative respectively load onto separate but correlated factors. If this model provided the best fit, this would indicate that each domain measures a separate construct.
3. **A higher-order factor model** where in addition to the factor structure of the four-factor model, there is a higher-order factor which the first-order factors of MST, explanation, certainty, and alternative load onto. Correlations between the first-order factors are constrained to 0 and not permitted to correlate with the higher-order factor. This higher-order factor captures common variance between the four first-order constructs. If this model provided the best fit, this would indicate that there are separable domain-specific factors but that the covariance between factors can be explained by an over-arching general mindreading construct.

Table 3.4. Model fit statistics for the different measurement models tested.

	RMSEA (< .08) [CI]	CFI (>.90)	TLI (>.90)
One-factor model	.075 [.067,.084]	.635	.563
Four-factor model	.035 [.019, .047]	.924	.906
Higher-Order model	.033 [.016, .046]	.930	.914

Note. RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index. The lead indicator loading for each tested model was set to 1.

The latent factor was scaled using the lead indicator. An adjusted weighted least squares (WLSMV) estimator was used for the following models given that some of the data was categorical and the data was missing at random. Both the four-factor and higher-order model provided acceptable fit to the data, however, the AIC for the four-factor model (AIC = 8049.95) was larger than the higher order model (AIC = 8043.045). Therefore, the higher-order factor model provided the best fit for the data. To assess whether verbal ability accounted for the covariance between the indicators, we ran a further model where for the higher-order factor model, the individual indicators (all items representing MST, mentalistic explanation, MST and Alternative) of the latent factors were simultaneously regressed onto verbal ability. The model fit the data well, RMSEA = .034, CFI = .929, TLI = .912, and all factor loadings remained statistically significant .300 - .697, all ps <.05. This result indicated that the covariance between the items and between the latent factors was not explained by verbal ability.

3.3 Study 2

Study 2 was pre-registered and conducted to cross-validate and extend the results obtained in study one in a non-university sample consisting of equal numbers of males and females. Specifically, we replicated the method employed in Study 1, and cross-validated the latent factor structure of the DMT, with the addition of also testing the possibility of a bi-factor model to test whether method factors (i.e., due to similarity between certain picture stimuli) or domain-specific abilities explained the co-variance between the indicators, thereby further strengthening the notion that a higher-order factor best explained the overlap between the first-order factors. Furthermore, in Study 2, convergent validity between the MASC (Dziobek et al., 2016) and DMT was assessed using error-free latent factor scores from the DMT in a series of Structural Equation Models (SEMs). This approach extended the analysis conducted in Study 1 given that it allowed for the simultaneous consideration of multiple variables (including potentially confounding factors such as verbal ability, demographic variables, and personality traits) (Brown, 2015). Furthermore, SEM also allowed us to partial out the impact of measurement error (i.e., noise or systematic errors), thereby providing a more in-depth picture of how the DMT and MASC converge.

3.3.1 Methods

3.3.1.1 Participants

A total of 214 participants based in the United Kingdom were recruited through Prolific Academic as part of a balanced sample. Prolific is a platform for online research (Peer, Brandimarte, Samat, & Acquisti, 2017). Participants were financially compensated for their participation and provided informed consent prior to participation. Four participants were excluded due to having reported a neurological condition such as a stroke or seizures, with no participant failing to meet any of the other inclusion criteria (i.e., failing to complete at least 75% of the study, failing two attention check questions as part of the administered questionnaires, being younger than 18 years of age, very short completion times, defined as 3.5 *SDs* below the average completion time, or very long completion times, defined as 3.5 *SDs* above the mean of completion times). After exclusion, there were 210 participants remaining. Based on Monte Carlo simulations, 210 was the minimum sample size needed to obtain a level of acceptable statistical power (>.80), assuming weak-moderate loadings (.25-.30), for 32 items and moderate-strong correlations between 4 latent factors to detect significant loadings for mindreading models (outlined in the results section).

The average age for participants was 38.83 years (*SD* = 11.17 years, *range* = 18 years – 65 years), with 110 participants (52.4%) identifying as female, 99 (47.1%) as male, and one as non-binary (0.5%). In terms of education, 4 participants (1.9%) did not graduate from secondary school, 72 (34.3%) held a secondary school degree, 96 (45.7%) a bachelor's degree, 31 (14.8%) a master's degree and 7 (3.3%) a doctoral degree. 186 participants (88.6%) spoke English as their native language and 24 (11.4%) as their non-native language. 164 participants (78.1%) were monolingual, 30 (14.3%) bilingual and 16 (7.6%) spoke at least three languages fluently. In terms of religious affiliation, 136 participants (64.8%) identified with no religion, 57 (27.1%) with Christianity, and 14 participants (6.7%) with another religion. Three participants did not answer the question (1.4%). In terms of ethnicity, 166 participants identified as "White", 29 as "Mixed", 8 as "Asian", 4 as "Black" and 3 as another ethnic group. All survey items were optional or allowed the participants to indicate that they "prefer not to respond" to a particular question. Participants were invited to leave feedback at the end of the survey. Ethical approval for this study was granted by the University of Birmingham Psychology Ethics Committee (ID: ERN_09-719AP21).

3.3.1.2 Measures and Procedure

The following measures described in detail in Study 1 of this Chapter (see *3.2.1.2 Measures and Procedure*). As these measures were administered to an independent sample in Study 2, inter-item agreement is reported below.

Mindreading Tasks

Movie Assessment of Social Cognition (MASC; Dziobek et al., 2016).

Scores for each sub-scale were summed up to form a scale for 'correct mentalising', α = .740, 'under mentalising', α = .669, 'no mentalising', α = .528, and 'over-mentalising', α = .353.

Dimensional Mindreading Task

Inter-rater reliability was calculated for 20% of items and was overall acceptable (ICC for MST = .745, ICC for explanation = .823, ICC for certainty = .923, ICC for alternative = .78).

Control Measures

Wechsler Abbreviated Scale of Intelligence (WASI – 2, Wechsler, 2001)

Scores were added up, with higher scores indicating greater expressive language ability, *α* = .636, *possible range* = 0 – 40, *observed range* = 17 – 38. The data was normally distributed.

Ten Item Personality Measure (TIPI) (Gosling, Rentfrow, & Swann, 2003)

The 10-item measure of the Big Five (or Five-Factor Model) dimensions of personality was administered. Participants indicated the degree to which they ascribed different personality traits to themselves, using a 7-point Likert scale ranging from 1(*Disagree strongly*) to 7 (*Agree strongly*). Items 2, 4, 6, 8 and 10 were reverse coded. The average of two items (standard item and reverse-scored item) made up each scale (i.e., extraverted, enthusiastic & reserved, quiet, α = .77, Critical, quarrelsome & Sympathetic, warm, α = .29, Dependable, self-disciplined &

Disorganised, careless, α = .64, Anxious, easily upset & Calm, emotionally stable, α = .65, Open to new experiences, complex & Conventional, uncreative, α = .37).

Demographic Information

We recorded participants' age, gender / sex (open-ended question), educational level (multiple-choice), whether English was their first language (multiple-choice), how many languages participants spoke in total (multiple choice), and whether they had any present or past psychiatric or neurological conditions (open-ended question).

3.3.2 Results

3.3.2.1 Analysis Plan

The same approach to data analysis as outlined in Study one was taken (see section *3.2.2.1).*

3.3.2.2 Descriptive Statistics and Correlations

As indicated in Table 3.4 and in Figure 3.3, there was variation, with no floor or ceiling effects, across the four dimensions of the DMT, with all skewness and kurtosis values being within the normal range. The dimensions of MST, explanation, certainty, and alternative were also all positively inter-corelated, *rs* = .18 - .55, *ps* = .004 - .001, suggesting small to moderate shared variance between the four dimensions. Notably, performance on the WASI (assessing verbal ability) was significantly positively correlated with all dimensions of the DMT, *rs* = .16 - .25, *ps* = .02 - .001, as well as with the MASC correct subscale, *r* = .30, *p* = .001, and significantly negatively correlated with the MASC errors, *rs* = -.25 - -.14, *ps* = .037 – .001, suggesting that participants with higher levels of verbal ability performed better on the MASC, and used more mentalistic terms, gave more explanations, expressed greater uncertainty and provided more alternative explanations for the social scenarios. To ensure that any observed effects would not be driven by individual differences in verbal ability, verbal ability was included as a covariate in all further analyses.

Table 3.4. Bivariate correlations between all variables of interest.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1 MASC_Correct	-														
2 MASC_Oover	-.48**	-													
3 MASC_Under	-.81**	-.01	-												
4 MASC_No	-.72**	.02	.52**	-											
5 MST	.29**	-.05	-.30**	-.23**	-										
6 Explanation	.07	.09	-.15*	-.06	.42**	-									
7 Certainty	.20**	-.12	-.16*	-.11	.18**	.38**	-								
8 Alternative	.09	.01	-.073	-.10	.20**	.42**	.55**	-							
9 Agreeableness	.01	.01	.02	-.06	.18**	.05	.05	-.04	-						
10 Conscientiousness	-.01	.01	.02	-.02	.022	-.04	-.04	-.01	.16*	-					
11 Emotional Stability	-.03	-.01	.07	-.03	-.08	-.01	-.01	-.091	.13	.31**	-				
12 Openness to experience	.13	-.02	-.12	-.13	.17*	.09	.09	.09	.08	.11	.06	-			
13 Extraversion	.02	-.02	.01	-.03	.10	.06	.06	-.00	.11	.14*	.24**	.36**	-		
14 WASI	.30**	-.20**	-.25**	-.14*	.25**	.20**	.20**	.16*	.01	-.17*	.01	.03	-.13	-	
15 Gender	-.02	.07	-.02	-.01	-.12	.07	.07	-.06	-.10	-.14	.23**	-.04	-.15*	.03	-
16 Age	-.01	.04	-.04	.03	.05	-.02	-.02	-.02	.11	.14*	.15*	-.12	-.06	.09	.04
	Descriptive Statistics														
M (SD)	32.45 (5.39)	5.45 (2.55)	4.60 (3.25)	2.51 (2.14)	12.82 (6.14)	1.80 (1.82)	3.50 (2.49)	1.44 (1.67)	5.08 (1.14)	5.26 (1.27)	4.41 (1.45)	4.90 (1.12)	3.59 (1.62)	27.86 (4.17)	.48 (.51)
Min	3	0	0	0	0	0	0	0	2	1	1	2	1	17	0
Max	44	13	24	17	39	8	8	8	7	7	7	7	7	38	2
Range	41	13	24	17	39	8	8	8	5	6	6	5	6	21	2
Skewness	-.14	.60	2.07	2.23	1.07	1.13	.15	1.15	-.19	-.83	-.15	-.34	.12	-.21	.19
SE Skewness	.17	.17	.17	.17	.17	.17	.17	.17	.17	.17	.17	.17	.17	.17	.17
Kurtosis	4.52	.20	8.12	10.25	2.07	.78	-1.14	.81	-.69	.38	-.76	-.45	-1.05	-.10	-1.71
SE Kurtosis	.33	.32	.33	.33	.33	.33	.33	.33	.33	.34	.33	.33	.33	.33	.33
N	210	210	210	210	210	210	210	207	210	209	210	210	210	210	210

Note. WASI = Wechsler Adult Scale of Intelligence, *M* = Mean, *SD* = Standard Deviation, *SE* = Standard Error, *N* = number of participants. *p<.05, **p<.01

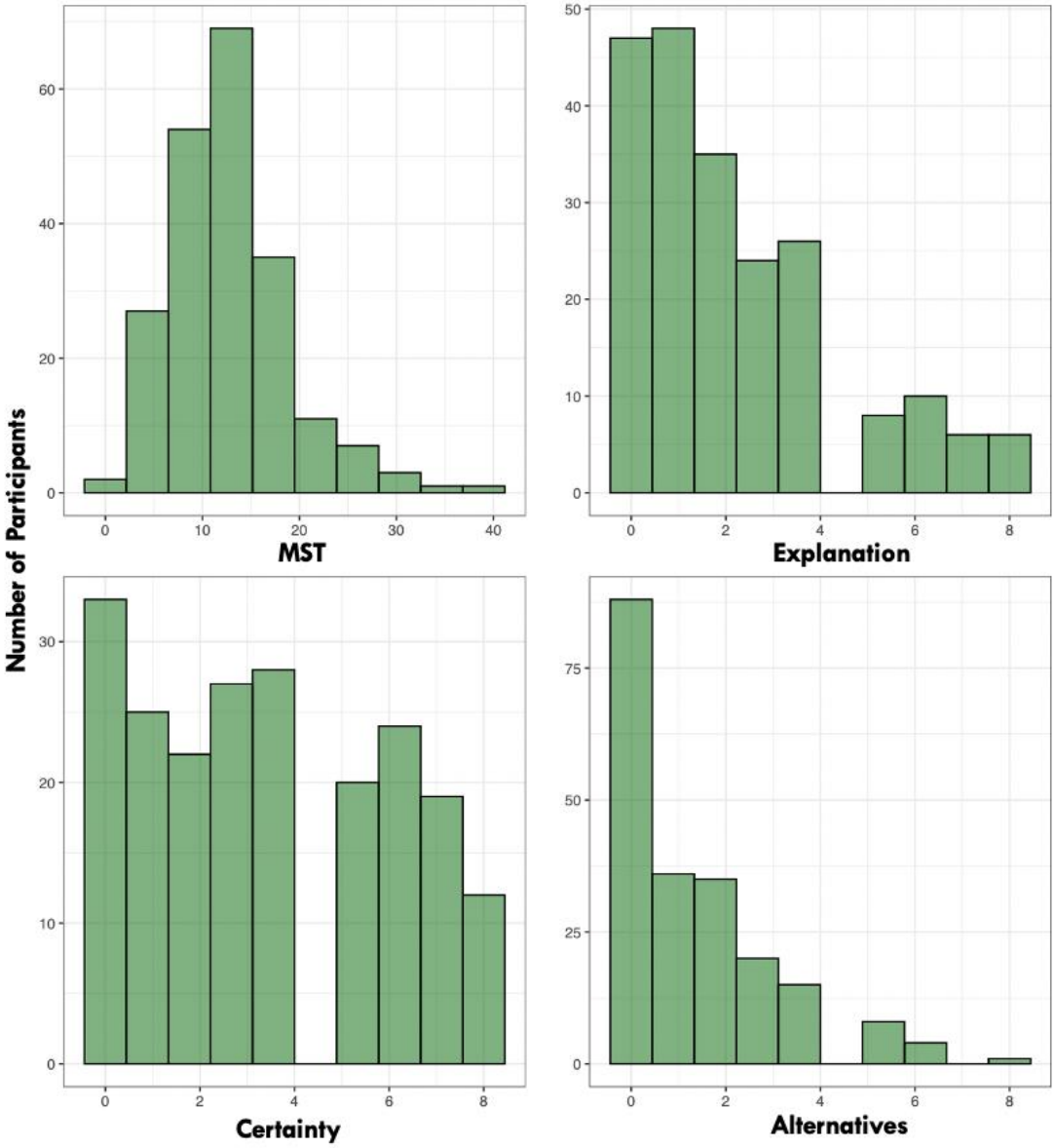

Figure 3.3. Histograms showing the distribution of scores for 210 participants for the observed overall scores of MST (8 items), explanation (8 items), certainty (8 items) and alternative (8 items).

3.3.2.3 Establishing a Factor Structure for Mindreading

An adjusted weighted least squares (WLSMV) estimator was used for the following models given that some of the data was categorical and the data was missing at random. Based on the results from Study 1, we tested alternative models to establish the factor structure of mindreading to verify whether the higher-order factor structure identified in Study 1 would replicate in a more diverse, larger non-university sample. To this end, we contrasted the four pre-registered, alternative models (Table 3.5 shows the factor structure for these models and Figure 3 shows a visual representation the best filling model).

4. **A one-factor model** where all 32 indicators representing the categories of MST (8 indicators), explanation (8 indicators), certainty (8 indicators) and alternative (8 indicators) loaded onto one factor. If a one-factor model provided the best fit, this would indicate that all indicators measure the same construct (i.e., a single mindreading factor).
5. **A four-factor model** were the 8 indicators representing the categories of MST, explanation, certainty and alternative respectively load onto separate but correlated factors. If this model provided the best fit, this would indicate that each domain measures a separate construct.
6. **A higher-order factor model** where in addition to the factor structure of the four-factor model, there is a higher-order factor which the first-order factors of MST, explanation, certainty, and alternative load onto. This higher-order factor captures common variance between the four constructs. If this model provided the best fit, this would indicate that there are separable domain-specific factors but that the covariance between factors can be explained by an over-arching general mindreading construct.
7. **A bi-factor model** where in addition to four first-order factors, there is a general factor. The correlations between the bi-factor and first-order factors were constrained to 0 (i.e., an orthogonal solution). If a bi-factor model provided the best fit, this could indicate that performance on the items of the DMT is explained by both a general mindreading factor that captures variance in all items *and* specific factors that explain the domains of the DMT.

Table 3.5. Factor Structure for mindreading (all models from pre-registration).

	RMSEA (< .08) [CI]	CFI (>.90)	TLI (>.90)
One-factor model	.074 [.065,.077]	.571	.553
Four-factor model	.042 [.033, .051]	.864	.852
Higher-Order model*	.033 [.020, .043]	.926	.912
Bi-factor model	.035 [.023, .044]	.915	.901

Note. RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index. The lead indicator loading for each tested model was set to 1. *This model is shown in Figure 3 below with additional information.

The parameter estimates were inspected for both the higher-order and the bi-factor model, given that both models provided acceptable fit to the data. Yet, no items loaded significantly onto the general factor of the bi-factor model, and the AIC for the bi-factor model (AIC = 9795.041) was larger than the higher order model (AIC = 9792.182). Thus, the higher-order factor, for which all parameter estimates were significant, was selected for further analysis, replicating previous pilot work. Figure 3.4 (shown below) provides a visual representation of this best-fitting model. Noteworthily, the MST factor exhibited the weakest loading (albeit still being significant) on the second-order factor.

To assess whether verbal ability accounted for the covariance between the indicators, we ran a further model where for the higher-order factor model, the individual indicators (all items representing MST, mentalistic explanation, MST and Alternative) of the latent factors were regressed onto verbal ability. The model fit remained acceptable, RMSEA = .033 [.022, .042], CFI = .926, TLI = .906, standardized regressions paths for the effect of verbal ability on each indicator ranged from -.001 to .066, and the standardized loadings, controlling for verbal ability, for the first-order factors remained significant, .106 - .876, all *ps* < .05, indicating that the covariance between the items and between the latent factors was not due simply to verbal ability (see Figure 3.4, illustration A). Next, in a further separate model, the covariance between gender, age, personality (i.e., extraversion, openness to experience, agreeableness, emotional stability, and conscientiousness) as well as whether participants spoke English as their first language (included to ensure that potential issues with understanding the task instructions in English could not account for variation in scores), with the higher-order latent factor was examined, CFI = .916, TLI = .902, RMSEA = .030 [.019, .038]. This model explained the links between the latent variables and each of the covariates. Here, the factor

loading of the higher-order factor remained significant, again showing that the covariance between items and between latent factor could not be for by verbal ability (see Figure 3.4, illustration B).

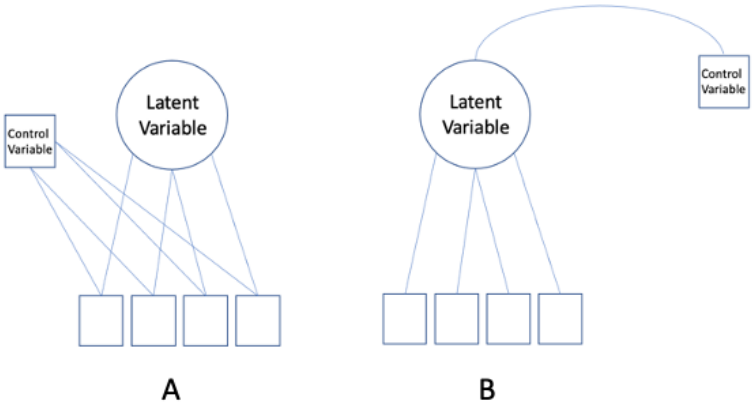

Figure 3.4. Schematic illustration showing the correlation between the latent variable and a correlate (B) and a test of whether the indicators continue to load onto the latent variable when a correlate is controlled for (B).

Given the results from these two measurement models which indicate that even when accounting for covariates, fit of the higher-order model remained acceptable, factor scores (a higher-order factor score representing shared variance between the individual indicators of MST, explanation, certainty and alternative), and four first-order factors (MST, explanation, certainty and alternative) from the higher-order model described in Table 3.5 and displayed in Figure 3.5 (where no covariates are included) were saved for further analysis.

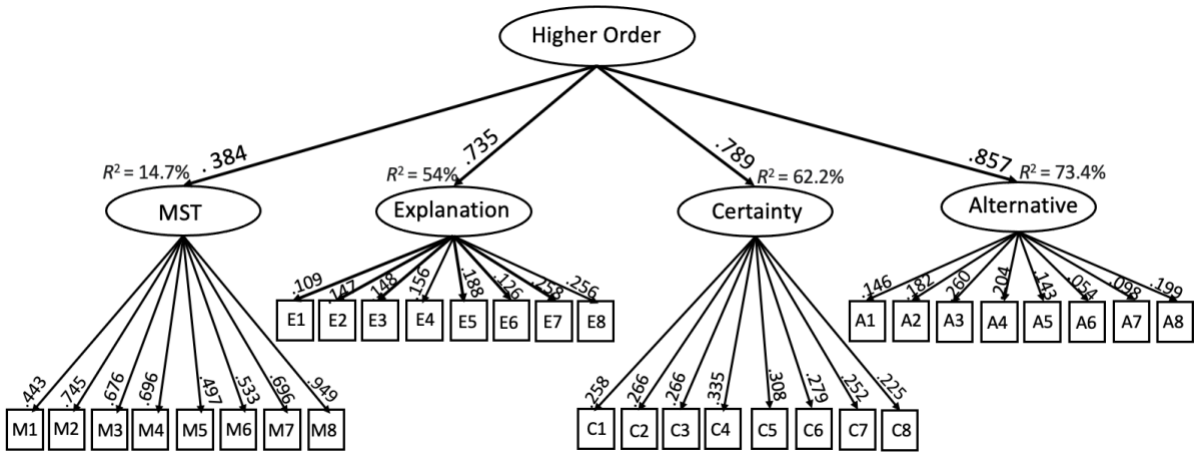

Figure 3.5 Best-fitting structural equation model for mindreading. All parameters were fully

standardized. All parameter estimates were significant at least at p <.05. Constructs coded from the same response were permitted to correlate (e.g., MST, Explanation, Certainty and Alternative scores coded from item 1), but this is not shown for simplicity. Based on modification indices, in addition, E8 and M8, M8 and C8, M5 and M6, and E5 and C4 were permitted to correlate (not shown for simplicity). This was justified given the relative similarity between picture stimuli. M = MST, E = Explanation, C = Certainty, A = Alternative. R2s for M1-M8 = 18% - 43%, R2s for E1-E8 = 7.7% - 42.6%, R2s for C1 – C8 = 26.6% - 40.4%, R2s for A1 – A8 = 5.1% - 39%.

3.3.2.4 Establishing Convergent Validity

Convergent validity was tested using the Movie of the Assessment of Social Cognition MASC; Dziobek et al., 2016), replicating the method used in Study 1. Table 3.4 includes the correlations between the latent factor scores extracted from the best-fitting mindreading model with all other study variables of interest. These associations are discussed below and visualised in Figure 3.6 below.

Bivariate Correlations between the Dimensional Mindreading Task and the MASC

As shown in Table 3.4, there were significant positive weak to moderate correlations between both (a) the higher-order mindreading latent factor as well as (b) the four first-order latent factors of the DMT with the "MASC correct" subscale, significant (except for one non-significant correlation between the "MASC correct" and "Explanation latent factor"). This suggests that individuals who scored higher on the MASC were also more likely to use more MST, provide more explanations for their mentalistic attributions, allow for more ambiguity in their interpretations, and were more likely to provide alternative interpretations of the scenarios shown. Further in line with our hypotheses, there were either no or significant negative correlations between (a) the higher-order mindreading latent factor and (b) all sub-scales of the MASC (except for the "MASC correct subscale") and all sub-categories of the dimensional mindreading task. Specifically, neither the higher-order mindreading latent factor, nor the first-order factors of the DMT correlated significantly with the "MASC Excessive" sub-scale. There were however weak to moderate significant negative correlations between the higher-order mindreading latent factor and the "MASC Under" and "MASC No" subscale. In terms of the first-order factors, the individual subscales of explanation, certainty and alternative did not significantly correlate with the "MASC No" subscale, whilst exclusively the alternative sub-

scale did not significantly correlate with the "MASC Under" subscale. All other correlations between the first-order factors and "MASC No" and "MASC Under" subscales were weak-moderate negative and significant. This suggests that participants who performed worse on the MASC were more likely to use fewer mentalistic words, provide fewer explanations for their mentalistic attributions, were more certain about their interpretations, and were less likely to provide alternative interpretations of the social scenarios.

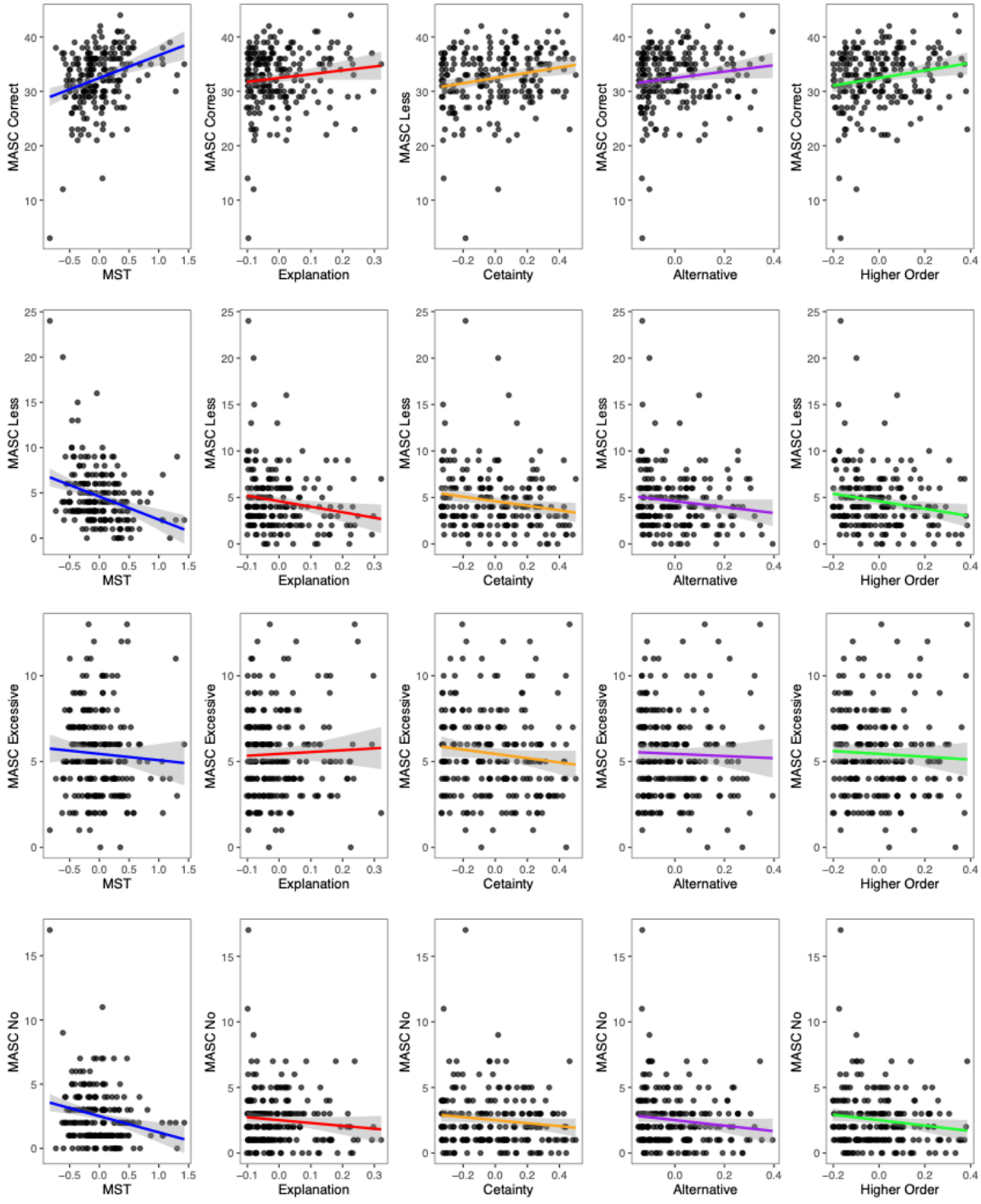

Figure 3.6. Correlations between the sub-scales of the MASC and each saved latent factor variable from the best-fitting mindreading model. The shaded areas reflect 95% confidence intervals.

3.3.2.5 Structural Equation Models

Next, convergence between the MASC, specifically, the 'MASC correct' subscale, and the DMT was assessed. The MASC correct subscale was related to the subscales of the DMT in two separate structural equation models. The MASC correct indicator was simultaneously regressed onto the higher-order mindreading latent factor scores from the best-fitting mindreading model (see Figure 3) as well as verbal ability, and, in a separate model, simultaneously regressed onto the first-order factors (MST, Explanation, Certainty, Alternative) and verbal ability. Verbal ability was included as a covariate in the models to ascertain whether any observed associations persisted above and beyond variation in verbal ability.

Whilst the higher-order factor did not remain a significant predictor of MASC correct scores once verbal ability was accounted for, in a separate model, where the higher-order factor was replaced by the first-order factors, MST and certainty were significant positive predictors of the 'MASC Correct' scale, CFI = .931, TLI = .919, RMSEA = .030 [.018,.040]. This finding suggests that a higher degree of use of mentalistic terms, and uncertainty expressed predicted greater mindreading abilities as assessed by the "MASC Correct" scale, over and above individual differences in verbal ability. With verbal ability accounted for, explanations and alternatives were no longer significant predictors of the "MASC Correct" scale.

Overall, this data suggests that participants' performance on the MASC converged with 'performance' on the dimensional mindreading task, whilst a significant degree of non-shared variance between both tasks remained. If one assumes that the MASC is a valid measure of mindreading, then the present results indicates that the DMT – or, specifically the higher-order factor which captures what is common amongst individual differences in MST, explanations, certainty and alternatives – does assess mindreading.

3.3.3 Discussion

The primary aim of our two studies was to develop a unique approach to capturing *individual differences* in the *structure* of adults' mindreading. While conventional mindreading tasks, including advanced ones, have predominantly focussed on measuring individuals' ability to accurately infer mental states (e.g., Castelli et al., 2001; Devine & Hughes, 2013; Dziobek et al., 2016), our findings suggest that it may not only be the 'accuracy' of mentalising that is relevant to understanding individual differences in adults' mindreading. In the following sections, three main findings that emerged from Study 1 and Study 2 are outlined and their theoretical and methodological implications for the study of mindreading are discussed.

3.3.3.1 Sensitivity of the New Measure

In both Study 1 and Study 2, our novel measure effectively captured individual variation across four distinct dimensions: frequency of mental state terms (MST), mentalistic explanations, expressed certainty, and alternatives. These categories were derived from participants' open-ended responses to various social picture stimuli. These findings fit with previous work suggesting that there is noteworthy variance in how adults engage in mentalising. Importantly, DMT captured variance without the need to assess the correctness of participants' responses (see Appendix the complete range of picture stimuli and coding scheme). Therefore, the present results contribute to our understanding of how the currently limited range of measurement tools available for capturing variation in adults' mindreading can be extended (e.g., Long et al., 2022).

3.3.3.2 Factor Structure of the New Measure

Secondly, structural equation modelling revealed evidence for a higher-order factor structure. The individual items representing the categories of MST, mentalistic explanation, certainty, and alternatives respectively loaded onto distinctive first-order factors, with a higher-order factor capturing the shared variance between these first-order factors (see Figure 3). This factor structure was tested in Study 1 and cross-validated in a separate sample in Study 2. Verbal ability or differences in five distinct personality traits (i.e., extraversion, conscientiousness, openness to experience, emotional stability, and agreeableness), as well as

demographic variables (i.e., age, gender, English as native language) did not to account for these observed associations between factors in either study.

3.3.3.3 Evidence of Convergence between the New Measure and MASC

In both Study 1 and Study 2, there was convergence between performance on the new DMT measure and the MASC. Correlations based on observed overall scores in Study 1 revealed weak to moderate positive associations between accuracy on the MASC and the four dimensions of the DMT (i.e., MST, explanations, certainty, and alternatives) and the higher-order mindreading factor. Conversely, there were either weakly negative or non-significant associations between the MASC error subscales and DMT latent factors. These patterns were consistent in Study 2: there were weak to moderate positive associations between the accuracy on the MASC and the latent factor scores of the DMT (except for explanations but including the higher-order latent factor). In contrast, there were either non-significant or significantly negative associations between the error subscales of the MASC and DMT latent factors (see Table 6). Structural equation modelling further revealed that the higher-order latent factor score, representing what is common across the four dimensions of MST, explanations, certainty, and alternatives, predicted better performance on the MASC (i.e., the 'MASC correct' subscale), even when accounting for verbal ability. These findings suggest that individuals who adopted a particular way of reasoning about others' mental states characterised by greater use of MST, providing explanations, expressing uncertainty, and providing alternatives, were more likely to infer others' mental states correctly. However, it is important to note that a significant degree of non-shared variance between both tasks remained, suggesting that the measures are not redundant or identical.

3.3.3.4 Methodological and Theoretical Implications

The present findings have numerous implications for the study of mindreading. The use of two diverse tasks, the MASC and the DMT, allowed for the assessment of different facets of mindreading (i.e., accuracy and structure). The MASC uses a multiple-choice format to assess individuals' ability to *select* most appropriate response to questions about movie characters' mental states. In contrast, our novel coding approach is a verbal task with an open-ended response format, requiring participants to *generate* written responses. The latter approach tapped into spontaneous or generative flexibility, referring to the ability to generate

diverse and novel ideas (e.g., Eslinger & Grattan, 1993). The current results therefore demonstrate the feasibility of assessing both the 'structure' and 'accuracy' of mindreading using methodologically distinct approaches.

However, it is worth considering that the specific four categories identified in the present study may have yielded 'meaningful' (converging) variation in adults' mindreading due to the nature of the stimuli used, that is, static, decontextualised pictures depicting unknown strangers interacting. Given the *limited* information participants were provided about the individuals depicted in the picture stimuli, higher degrees of flexibility in mindreading were logically warranted. After all, for the given stimuli, discerning the mental states of the depicted characters was significantly more challenging than in the case of many other mindreading tasks (and we had no ground truth against which to evaluate the accuracy of mentalistic inferences). Notably, the ambiguity as to what an accurate response would have been, can be seen as possessing a high degree of ecological validity, because real-life social encounters or conditions under which mentalising would spontaneously are often characterised by dynamically changing mental states that are challenging to be *known* (e.g., Cassels & Birch, 2014; Betz, Hoemann, & Barrett, 2019).

The degree to which mental states can be known differs not only across real-life contexts but also across different mindreading measures. Some tasks, like classical false-belief tasks (i.e., the Sally-Ann task; Knoll & Charman, 2000) or visual perspective taking tasks, such as the director task (e.g., Apperly & Wang, 2021), provide a relatively clear 'ground-truth' regarding the accurate mental state interpretation. For such tasks, relatively high degrees of confidence in one's interpretations is appropriate. In contrast, video-based tasks such as the MASC (Dziobek et al., 2006) or Silent Film Task (SFT; Devine & Hughes, 2013) carry more ambiguity regarding the characters' mental states, although they cue participants to mentalise and offer more context (due to the video-format) to support a *specific* mentalistic attribution compared to the static picture stimuli used in the present study. The present stimuli also did not include direct instructions for participants to mentalise. Therefore, it is plausible that the extent to which a particular 'way of mindreading', such as thinking of others' mental states *flexibly* by providing alternatives or expressing uncertainty in one's mentalistic attributions, is appropriate, critically depends on the *context* (real-life or mindreading task) within which mentalising occurs. In other words, the present study is not meant to suggest that exhibiting higher levels of the specific four categories identified in the present study is *always* most

appropriate. Instead, the present results emphasise that classifying task stimuli used to elicit mindreading in terms of their general ambiguity (i.e., how 'easily knowable' characters' mental states are) may provide a fruitful path for researchers to take. By doing so, potentially intriguing relationships between *how* people mentalise (i.e., MST, explanations, certainty, and alternatives) and how *correctly* they mentalise could be uncovered. This consideration has empirical support. Specifically, cognitive persistence and flexibility have complementary advantages. Whilst cognitive persistence enhances stability, it can lead to inflexible behaviour, however, conversely excessive flexibility could result in unproductive distractibility (e.g., Dreisbach & Goschke, 2004). Individuals could therefore benefit from using 'the right degree' of the four identified categories.

The present study identified four distinct mindreading processes underpinned by an overarching mindreading ability that explained variance on the DMT, contributing to the ongoing debate on how to operationalise mindreading (i.e., as a single or multifaceted construct; Schaafsma et al., 2015; Abu-Akel & Shamay-Tsoory, 2011; Apperly & Butterfill, 2009; Gopnik & Willman, 1994; Preston & De Waal, 2002; Poletti et al., 2012; Bernhardt & Singer, 2012). However, in addition to the factor structure identified in this study, it is also possible that there are still further components in the structure of mindreading for future research to discover, particularly if different task stimuli were used. To investigate the latter, further data-driven research may be necessary to validate or challenge existing theory-driven findings (e.g., Jack, Crivelli, & Wheatley, 2018).

The present findings also hold relevance for cross-cultural investigations in psychology. It would be intriguing to explore whether cultures exhibit differences in *how* they engage in mindreading. The methodology presented in this study could serve as a tool to make such investigations. There is already substantial developmental evidence indicating reliable cross-cultural variation in children's mindreading abilities (e.g., Naito, 2004; Wellman et al., 2001, Hughes, Devine, Ensor, Koyasu, Mizokawa, & Lecce, 2014; Fujita, Devine, & Hughes, 2022), highlighting the importance of specific experiential factors in mindreading accuracy. Therefore, accuracy on existing mindreading measures may not only be influenced by the degree of contextual information (or how 'knowable' mental states are), but also by the familiarity participants have with depicted situations and characters. Coding the *structure* of mindreading may be a powerful tool for 'fairer' assessment across difference cultures and could lead to a deeper understanding of the nature of individual differences in mindreading in

diverse populations (e.g., Norenzayan et al., 2012; Hruschka, Medin, Rogoff, & Henrich, 2018).

Furthermore, future research should investigate the relationship between performance on the DMT and measures of motivation to determine whether they represent distinct or overlapping constructs. Notably, differences in response length of participants' responses on the DMT were not accounted for given that receiving a score across the four identified dimensions required the production of more words. However, given that no incentive for a specific response as provided and differences in conscientiousness (an indicator of motivation) were accounted for, this was not considered problematic. Investigating how the MASC and DMT relate to other relevant constructs is also of interest, especially considering the non-shared variance between the two tasks (such data has been collected as part of this data collection point and is reported in Chapter 4). Indeed, in addition to understanding *why* adults vary in mindreading, it is also still unclear what such individual variation in mindreading means in terms of meaningful outcomes in individuals beyond middle childhood, and whilst advanced mindreading tasks have frequently been empirically linked with social and mental health outcomes, the evidence is not robust. To illustrate, some studies suggest mindreading deficits in depressed individuals (e.g., Nejati, Zabihzadeh, Maleki, & Tehranchi, 2012), while others report no deficits (e.g., Caputi & Schoenborn, 2008; Corcoran et al., 2011), and still others have shown improved mindreading in individuals with depressive symptoms (e.g., Harkness, Alavi, Monroe, Slavich, Gotlib, & Bagby, 2010). Lastly, investigating the relationship between traditional measures of cognitive flexibility, particularly of spontaneous flexibility (e.g., the Alternative Uses Task; Roberts, Grady, & Addis, 2020), and *how* participants mentalise would be interesting. For example, it may be that higher performance on such 'emotionally neutral' cognitive flexibility tasks may be related to a general tendency to always exhibit more 'mindreading flexibility', or it may mean that individuals are better able to appropriately adapt their 'mindreading structure' in response to the specific context they are presented with.

3.3.3.5 Conclusion

In summary, the present study provided evidence that it is possible to effectively capture structural features of mindreading in two independent samples of adult populations, without focussing on the *accuracy* of specific content. This study thereby provides initial

evidence that a diversion from current ways of assessing mindreading exclusively in terms of accuracy may be a fruitful supplementation to existing methods to take to help our understanding of *why* individuals differ in their mindreading. Our novel approach to coding open-ended responses to various social picture stimuli effectively captured individual variation across four distinct categories: the frequency of mental state terms (MST), mentalistic explanations, expressed certainty, and alternatives. Confirmatory factor analyses revealed that a higher-order factor structure reliably provided the best fit for the data (based on a test and cross-validation in a separate sample). Finally, we found evidence of convergent validity between an established mindreading measure, the MASC, and our novel task, suggesting that particular 'ways of mentalising' are associated with higher accuracy. However, our findings do however not suggest that for every context, mentalising benefits from higher degrees of the four categories identified here, and so future research will benefit from addressing the complex *interactions* between the context, structure, and accuracy of mentalising. Acknowledging the importance of linking individual differences in mindreading at the level of 'structure' with differences at the level of 'accuracy' will help to further inform our understanding of the cognitive underpinnings of individual differences in mindreading in individuals who already possess concepts of mental states.

Chapter 4

Social Motivation, not Mindreading is associated with Social and Mental Health Variables

4 Social Motivation, not Mindreading, is associated with Social Competence and Mental Health

4 Introduction

Numerous studies have demonstrated that individual differences in the capacity to infer others' mental states accurately, often referred to as theory of mind or mindreading (e.g., Wellman, 1990; Premack & Woodruff, 1978), have meaningful consequences for individuals' lives. However, much of the supporting evidence stems from developmental research (e.g., Eisenberg, 2003; Devine & Apperly, 2022; Hughes & Devine, 2015a; Sodian & Kirsten, 2010; Leslie et al., 2004) or comparisons between clinical and non-clinical adult populations (e.g., Bradford et al., 2018; Dziobek et al., 2006; Preißler et al., 2010; Livingston et al., 2021; Bora, & Berk, 2016, for a meta-analysis). Relatively less attention has been paid to the potential correlates of individual differences in mindreading among adults. Moreover, findings from studies investigating these correlates exhibit substantial heterogeneity, making it challenging to discern the strength and direction of these potential associations (e.g., Warnell & Redcay, 2019; Schaafsma et al., 2015). It is also unclear which, if any, aspects of adults' lives are influenced by individual differences in mindreading. Notably, a distinct literature emphasises the substantial impact of social motivational factors, a person's willingness to engage socially on broader social phenotypes traditionally linked with mindreading (i.e., autism; Chevallier et al., 2012). Additionally, there is a growing interest in the potential interplay between social motivation and social cognition (e.g., Contreras-Huerta et al., 2020; Carpenter et al. 2016; Devine & Apperly, 2021; Lockwood et al., 2017). However, to date, only one study in children (e.g., Devine & Apperly, 2022) has investigated the unique and relative impact of social motivation and mindreading on children's social competence, leaving a gap in our understanding the relative contribution of individual differences in the domains of social motivation and mindreading to meaningful social and mental health outcomes in adults. The overall aim of this study was to therefore to investigate the links between social motivation and mindreading, and their respective relationships with social abilities and mental health.

Ample evidence suggests that greater mindreading abilities in children are associated with positive outcomes, including prosocial behaviour (e.g., Eisenberg, 2003; Imuta et al., 2016, for a meta-analysis), improved peer relationships (e.g., Slaughter et al., 2015, for a meta-

analysis), enhanced social competence (e.g., Devine, White, Ensor & Hughes, 2016; Devine & Apperly, 2021; Devine et al., 2023), and reduced feelings of loneliness (e.g., Koerber & Osterhaus, 2020). Such evidence supports the 'social individual differences' account of mindreading, which posits that variability in mindreading in age groups beyond the age of four or five (a point in time when neurotypical children attain a conceptual understanding of mental states, e.g., Wellman et al., 2001), has meaningful consequences for individuals' social lives (e.g., Hughes & Devine, 2015a; Apperly, 2012). However, it remains relatively unclear which aspects of adults' lives are influenced by individual differences in mindreading.

To illustrate, some studies suggest mindreading deficits in individuals with depression (e.g., Nejati et al., 2012; Inoue, Yamada, & Kanba, 2006), while others report no deficits (e.g., Caputi & Schoenborn, 2008; Corcoran et al., 2008), and some indicate superior mindreading in individuals with depressive symptoms (e.g., Harkness et al., 2005). Similarly, while it has been proposed that individuals with high levels of anxiety experience deficits in social cognition affecting their ability to interpret others' mental states (e.g., Hezel & McNally, 2014), research involving participants with subclinical levels of anxiety found no evidence of impaired mindreading (Lenton-Brym, Moscovitch et al., 2018), as assessed via the Movie for the Assessment of Social Cognition (MASC; Dziobek, 2006). There is evidence of links between childhood adversity and impairments in adults' mindreading (e.g., Germine et al., 2015), however, these findings are based on the Reading the Mind in the Eyes (RMET; Baron-Cohen et al., 2001), which has been argued to assess emotion perception rather than mindreading (e.g., Oakley et al., 2016). To establish a more comprehensive picture of this conflicting evidence concerning depressive symptoms and anxiety, meta-analytic evidence should be considered (e.g., Cotter et al., 2018).

One explanation for these inconsistent findings is the possibility that individual differences in mindreading among adults, who already possess a mature conceptual understanding of mental states (e.g., Apperly, 2012), may not be the primary or unique driver of variations in meaningful social and mental health outcomes. Alternatively, if links between mindreading in adults and meaningful outcomes existed, as would be suggested by a 'social individual differences' account of mindreading (e.g., Apperly, 2012), then advanced mindreading tasks with psychometrically robust characteristics should be able to illuminate such relationships (this issue is explored in greater detail in Chapters 2 and 3). This consideration warrants the development of improved mindreading measurements (e.g.,

Warnell & Redcay, 2019; Long et al., 2022; also see Chapter 2 and 3). One important feature of such mindreading tasks is their sensitivity to detect variation in task performance. Indeed, some measures, such as the MASC (Dziobek et al., 2016), are sufficiently sensitive to detect individual differences in neurotypical adults' mindreading (also see: Conway et al., 2019; Livingston et al., 2021; Devine & Hughes, 2019; Slaughter & Repacholi, 2004). However, although worse performance on the MASC was found to predict a higher incidence of autistic traits (e.g., Boada et al., 2020), the extent to which it possesses criterion validity in terms of other social and mental health outcomes in adults is currently unclear (e.g., Warnell & Redcay, 2019). We therefore aimed to examine individual differences in adults' mindreading in relation to several outcome variables: depressive symptoms, anxiety, autistic traits, and social support (from family, friends, and significant others).

Notably, several outcome variables traditionally studied in the context of mindreading research, including depression (e.g., Nestor et al., 2023, for a meta-analysis), anxiety (e.g., Hezel & McNally, 2014), autism (e.g., Baron-Cohen, 2001), and prosocial behaviour in children (e.g., Imuta et al., 2016),have also been *independently* linked to social motivation (e.g., Gandhi, Mote, & Fulford, 2022, for a meta-analysis on social motivation impairments in major depressive conditions and schizophrenia spectrum conditions; Trew, 2011; Fussner, Mancini, & Luebbe, 2018; Chevallier et al., 2012; Hofmann, & Hay, 2018; Rodkin, Ryan, Jamison, & Wilson, 2013). Social motivation encompasses a set of psychological processes relevant to individuals' propensity for social engagement, with higher social motivation generally indicating an inclination to approach desired social stimuli (e.g., Gable, 2006). Low levels of social motivation have been associated with key features of various mental health conditions such as anxiety or depressive conditions (e.g., Hofmann & Hay, 2018; Trew, 2011; Seidel, Habel, Kirschner, Gur, & Derntl, 2010). For example, individuals with depression were found to report a reduced inclination to approach smiling faces (e.g., Seidel et al., 2010; Radke et al., 2015; Fussner et al., 2018). Examining individual differences in social motivation is perhaps particularly relevant in disorders characterised by social difficulties, such as depression-anxiety spectrum disorders (e.g., Berecz, Tenyi, & Herold, 2016). However, there is also evidence of positive associations between social motivation and social outcomes in neurotypical participants. For example, school students' social motivation has been positively linked to their perceived social support from their environment (e.g., Tezci et al., 2015; Wentzel, Battle, Russell, & Looney, 2010).

Moreover, there is an emerging interest in understanding the connection between social motivation and broader aspects of social cognition (e.g., Contreras-Huerta et al., 2020; Carpenter et al. 2016; Devine & Apperly, 2022; Lockwood et al., 2017). Among adults, higher levels of social motivation have been associated with superior performance on mindreading tasks (e.g., Carpenter et al., 2016). Conversely, social apathy, characterised by low social motivation, has been found to be inversely correlated with cognitive empathy (Lockwood et al., 2017). These findings suggest that social motivation may serve as a pre-requisite for individuals to engage in accurate mentalising, and, in turn, reap the associated benefits. Alternatively, it is plausible that the mere act of mentalising, independently of accuracy, confers certain benefits. For example, a person who attempts to understand another's mental states, even if sometimes inaccurate, might still be more highly valued as a friend compared to someone who does not make such attempts at all. Finally, it may be that social motivation, perhaps distinctly from other socio-cognitive functions (e.g., Contreras-Huerta et al., 2016), significantly drives social and mental health outcomes in adults, or that both social motivation and mindreading make independent contributions to relevant outcome variables. To date, only one study has shown that social motivation and mindreading ability in 8-to-13-year-old children are separable constructs but simultaneously make distinctive contributions to predicting variation in teacher-rated social ability (Devine & Apperly, 2022). Therefore, our third aim was to assess the relative contribution of social motivation and mindreading regarding to the previously outlined social and mental health measures.

4.1.1 The Present Study

This study was pre-registered on the Open Science Framework (See Step 4 onwards: https://osf.io/3b5nq/?view_only=0380b3ea881f4ff098fcc58fd3d49e0b).

Our first aim centred on assessing the relationship between the domains of individual differences in mindreading and social motivation. Specifically, we sought to determine the extent to which these constructs were separable or overlapping. Social motivation was measured via four distinct, validated self-report measures of social motivation, the Mindreading Motivation (MRM) Scale (Carpenter et al., 2016), Social Effort and Conscientiousness Scale (SECAS; Abplanalp et al., 2022), Snaith-Hamilton Pleasure Scale (SHPS; Snaith et al., 1995), and the Social Avoidance and Distress Scale (SADS; Watson & Friend, 1969). Collectively, these scales assess various facets of the multifaceted construct of social motivation: they measure the extent to which participants are inclined to actively nurture

and sustain social connections, prefer social interactions over non-social ones, and pursue pleasure in social interactions (e.g., Chevallier et al., 2012). These scales were selected based on a systematic search of the literature and were identified to have good psychometric properties. Following this, we conducted a pilot study where correlations between the above-outlined scales in addition to a further 3 scales were inspected.

Our second aim was to examine individual differences in neurotypical adults' mindreading in relation to several outcome variables: depressive symptoms, anxiety, autistic traits, and social support (from family, friends, and significant others). To assess mindreading, we selected the Movie for the Assessment of Social Cognition (MASC; Dziobek et al., 2006) due to its sensitivity to detect individual differences in neurotypical adults (Preißler et al., 2010), strong internal consistency, as well test-retest reliability (r = .97, Dziobek et al., 2006). Additionally, we administered a novel mindreading task, the 'Dimensional Mindreading Task' (DMT). As outlined in Chapter 3, our innovative approach involved coding open-ended responses to diverse social picture stimuli, which effectively captured individual variation across four distinct domains: the frequency of mental state terms (MST), mentalistic explanations, expressed certainty, and alternatives. Across two separate data collections, performance on the DMT converged with the MASC.

Our third aim was to assess the relative predictive importance of social motivation and mindreading regarding their unique contributions to the previously outlined social and mental health outcome measures. By doing so, we sought to bridge the existing gap between separate bodies of literature on the correlates of accuracy at mindreading and social motivation in adults.

4.2 Method

In this chapter, we used the same dataset as in Chapter 3, Study 2, to investigate distinct research questions. The following measures formed part of the analysis reported in this chapter and are described in detail in Chapter 3:

1. Movie Assessment of Social Cognition (MASC; Dziobek et al., 2016)
2. Dimensional Mindreading Task (DMT)
3. Wechsler Abbreviated Scale of Intelligence (WASI – 2; Wechsler, 2001)
4. Ten Item Personality Measure (TIPI) (Gosling et al., 2003)

Measures that were not previously included in the analysis reported in Chapter 3 but are relevant to the current Chapter 4 are outlined below.

4.2.1 Social Motivation Measures

Mindreading Motivation (MRM) Scale (Carpenter et al., 2016)

Participants provided ratings on a 7-point scale (ranging from 1 to 7) for 13 questions that assessed their inclination to engage effortfully with others' perspectives and mental states (e.g., *When I am conversing with more than one person, I like to think about how one person is interpreting what another person says in the conversation*). The scoring method involved reversing the values for items 2, 4, 5, 7, 9, 10, 11, 12, and 13, following the standard procedure. The total possible score ranged from 13 to 91, with higher scores indicating a stronger motivation for mindreading. Internal consistency reliability was good, $\alpha = .818$, and items were summed together to create an overall variable, *possible range* = 13 - 91, *observed range* = 24 – 85, *M* = 61.08, SD = 10.13, *Skewness* = -.291, *SE* = .168, *Kurtosis* = -.054, *SE* = .334.

Social Effort and Conscientiousness Scale (SECAS; Abplanalp et al., 2022)

Participants rated their level of agreement on a scale ranging from 1 (*strongly disagree*) to 6 (*strongly agree*) for 16 questions assessing effort with regards to fostering social connections. Of these questions, 8 items (items 2, 6, 7, 9, 11, 12, 14, 16) assessed "social effort" (e.g., *I often arrange events with other people*), $\alpha = .863$, observed range = 10 – 45, possible range = 8 – 48, 4 items (4, 10, 13, 17) "conscientiousness" (e.g., *I usually try to help people when they are feeling down*), $\alpha = .749$, observed range = 5 – 24, possible range = 4 – 24, and 4 statements were filler items (i.e., *"I make sure to eat breakfast every morning"*). No items were reverse coded. The "social effort" and "conscientiousness" items were individually summed together to create overall subscales for each construct and additionally jointly summed together to create an overall scale, $\alpha = .872$, observed range = 16 – 68, possible range = 8 – 48. Based on the results of previous pilot work, the "conscientiousness" subscale of the SECAS was used.

Snaith-Hamilton Pleasure Scale (SHPS; Snaith et al., 1995)

Participants rated their experiences of pleasure or anhedonia on a 4-point Likert scale ranging from 0 (*strongly disagree*) to 4 (*strongly agree*) for 14 items. "*Strongly disagree*" and "*disagree*" were both scored as 1, and "*agree*" and "*strongly agree*" were both scored as 0, with a possible range of 0 – 14. Higher scores on the original scale indicate greater levels of anhedonia, however, this was reversed in the present study to enable easier interpretation and comparison with other measures. All items were summed together to create an overall variable with high scores indicating greater social pleasure, α = .832, observed range = 0 – 14.

Social Avoidance and Distress Scale (SADS; Watson & Friend, 1969)

Participants indicated whether they felt 28 statements relating to their feelings about social engagements to be either "*true*" or "*false*". For items 1, 3, 4, 6, 7, 9, 12, 15, 17, 19, 22, 25, 27 and 28, a score of 0 was awarded if participants selected "true", and a score of 1 was awarded if participants selected "false". This was reversed for the remaining items. Possible scores on the scale ranged from 0 to 28. Lower scores on the SAD are most adaptive, with higher scores indicating greater social avoidance and distress. In this study, the scale was reversed to enable easier comparison with other measures, whereby higher scores indicated greater social engagement and lower distress in social situations. Scores were summed up to create an overall variable, α = .956, observed range = 0 – 28.

4.2.2 Mental Health Measures

Anxiety: Generalized Anxiety disorder (GAD-7; Spitzer et al., 2006)

Participants were asked to indicate on a 4-point Likert-scale ranging from 0 (*not at all*) to 3 (*nearly every day*) across 7 different items how often they had been bothered by a range of different problems over the past two weeks (i.e., 'trouble relaxing'). Scores across the 7 items were added up to form an overall anxiety score, α = .92, with a possible range of 0 – 21, and higher scores indicating greater anxiety.

Patient Health Questionnaire-9 (PHQ-9; Kroenke & Spitzer, 2002)

Participants were asked to rate the frequency of depressive symptoms they experienced over the past two weeks across 9 items and via a 4-point Likert scale ranging from 0 (*not at all*) to 4 (*nearly every day*). The items are based on nine criteria for major depressive disorder

as outlines in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5). Scores from the 9 items were summed up to create an overall score, α = .899, with higher scores indicating greater prevalence of depressive symptoms.

Autism Spectrum Quotient (AQ-10; Allison, Auyeung, & Baron-Cohen, 2012)

Participants indicated their level of agreement with 10 items assessing autistic traits on a 4-point Likert scale ranging from 1 (*definitely agree*) to 4 (*definitely disagree*), i.e., *"I know how to tell if someone listening to me is getting bored"*. Per item, participants could only receive one score, with one point being awarded for either "*definitely agree*" or "*slightly agree*" on each of the items 1, 7, 8, and 10 and 1 point for "*definitely disagree*" or "*slightly disagree*" on each of the items 2, 3, 4, 5, 6 and 9. To calculate the overall score, items were added up, *α* = .644, with a possible and observed range of 0 – 10 (*M* = 3.09, *SD* = 2.13), *skewness* = 1.04 (*SE* = .17), *kurtosis* = .97 (*SE* = .33). In addition, the items specifically assessing "social functioning" were also summed to create an overall variable, *α* = .656, with a *possible and observed range* of 0 – 5, *M* = .976 (*SD* = 1.28), *skewness* = 1.31 (*SE* = .17), *kurtosis* = .95 (*SE* = .33).

4.2.3 Social Outcome Measures

Multidimensional scale of perceived social support (MSPSS; Zimet, Dahlem, Zimet, & Farley, 1988)

Participants were asked to indicate their level of agreement with 12 items assessing their perception of social support from friends, family, and significant others on a 7-point Likert scale ranging from 1 (*very strongly disagree*) to 7 (*very strongly agree*), i.e., *"There is a special person who is around when I am in need."* Each subscale was assessed by four items. The scores for these items were added up and then divided by 4 to calculate mean scores for each sub-category (significant other subscale: α = .97; family subscale: α = .93; friends subscale: α = .95).

4.3 Results

4.3.1 Correlational Analyses

Table 4.1 shows all correlations between the observed overall scores of all variables of interest in this study. The DMT scores are latent factor scores. Neither the MASC nor the DMT correlated significantly with any of the social or mental health outcome variables, except for weak significant negative correlation between the MASC correct score and autism. Conversely, the MASC errors were significantly, weak-moderately, and positively associated with autism, suggesting that individuals who performed better on the MASC possessed more autistic traits. In contrast, overall, there were significant positive, moderate-strong correlations between the social motivation self-report measures and social and mental health outcomes. It can also be noted that the MASC and DMT were both significantly correlated with verbal ability, which was statistically accounted for in any measurement models.

Table 4.1. Bivariate correlations between the extracted latent factor scores for social motivation, mindreading, and outcome measures (covariates are excluded for simplicity).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 MASC_C	---																	
2 MASC_E	-.480**	---																
3 MASC_U	-.810**	-001	---															
4 MASC_N	-.717**	.015	.519**	---														
5 MST	.291**	-.045	-.300**	-.225**	---													
6 Expla	.065	.101	-.146*	-.064	.424**	---												
7 Certainty	.197**	-.122	-.160*	-.106	.184**	.381**	---											
8Alternative	.087	-.001	-.073	-.104	.197**	.424**	.548**	---										
9 SECAS_Co	.166**	.045	-.162*	-.227**	.193**	.139*	.062	.122	---									
10 MRM	.245**	-.012	-.268**	-.195**	.367**	.172*	.169*	.104	.445**	---								
11 SHAPS	.211**	-.163**	-.122	-.152*	-.029	.039	-.092	-.019	.335**	.153*	---							
12 SADS	-.086	.058	.082	.023	.043	.011	-.009	.102	.250**	.073	.214**	---						
13 AQ-10_O	-.190**	.090	.147*	.150*	-.072	-.011	-.069	.024	-.291**	-.165*	-.267**	-.366**	---					
14 AQ-10_Sl	-.240**	.094	.184**	.211**	-.077	-.073	-.106	-.015	-.268**	-.192**	-.194**	-.270**	.834**	---				
15 SS_F	-.031	.102	-.013	-.025	-.042	.047	.062	-.033	.286**	.119	.360**	.346**	-.191**	-.142*	---			
16 SS_SO	.018	-.078	-.060	-.038	-.016	-.095	.007	-.044	.259**	.164*	.244**	.132	-.185**	-.205**	.497**	---		
17 SS_F	.088	.059	.001	-.117	.007	-.052	-.015	-.119	.453**	.154*	.363**	.341**	-.259**	-.242**	.490**	.379**	---	
18 Anx	.027	-.038	-.020	.008	.078	.014	.044	.105	-.050	.091	-.338**	-.414**	.334**	.242**	-.285**	-.144*	-.203**	---
19 Dep	.003	-.036	.022	.001	.097	.091	.098	.125	-.064	.042	-.384**	-.387**	.267**	.179**	-.400**	-.230**	-.246**	.805**

Note. MASC = Movie for the assessment of social cognition, MASC_C = MASC Correct, MASC_E = MASC Excessive, MASC_L = MASC_Under, MASC_N = MASC No, SECAS = Social Effort and Conscientiousness Scale, SECAS_Co = SECAS Conscientiousness, SECAS_SE = SECAS Social Effort, SECAS_O = SECAS Overall, MRM = Mindreading Motivation Scale, SHAPS = Snaith-Hamilton Pleasure Scale , SADS = Social Avoidance and Distress Scale , AQ_10_O = Autism Quotient Overall Score, AQ_10_S = Autism Quotient Social Score, SS_F = Social Support Family, SS_SO = Social Support Significant Other, SS_F = Social Support Friends, Anx = Anxiety, Dep = Depression.

4.3.2 Mindreading Latent Factor

A higher-order factor model (Figure 4.1) where the respective 8 indicators of MST, explanation, certainty, and alternative loaded onto individual first order-factors that was underpinned by a second/higher-order factor that captured the common variance between the four constructs provided the best fit for the data, RMSEA = .033 [CI = .020 - .043], CFI = .926, TLI = .912. These results show that there are separable domain-specific factors but that the covariance between factors can be explained by an over-arching general mindreading construct. Factor scores from the measurement model shown in Figure 4.1 were extracted for further analysis.[2]

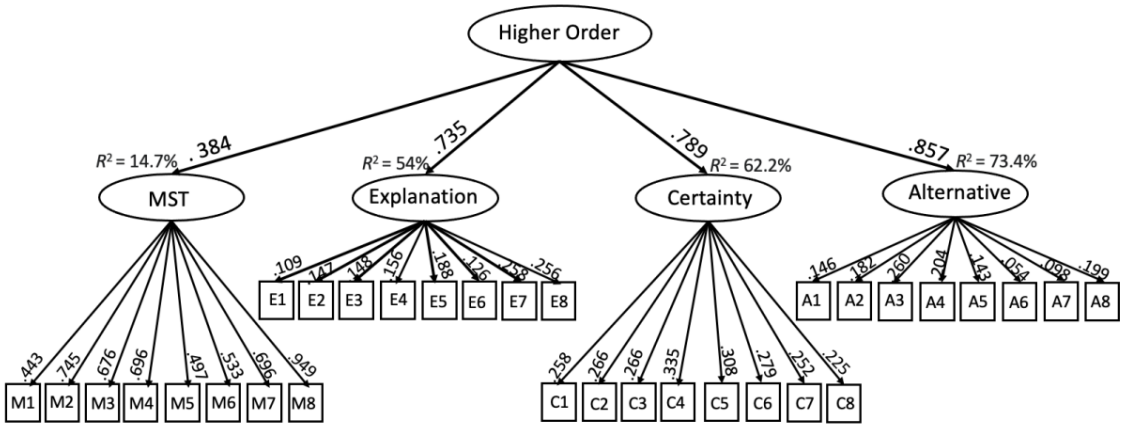

Figure 4.1. Best-fitting structural equation model for mindreading. All parameters were fully standardized. All parameter estimates were significant at least at p>.05. Constructs coded from the same response were permitted to correlate (e.g., MST, Explanation, Certainty and Alternative scores coded from item 1), but this is not shown for simplicity. Based on modification indices, in addition, E8 and M8, M8 and C8, M5 and M6, and E5 and C4 were permitted to correlate (not shown for simplicity). This was justified given the relative similarity between picture stimuli. M = MST, E = Explanation, C = Certainty, A = Alternative. R2s for M1-M8 = 18% - 43%, R2s for E1-E8 = 7.7% - 42.6%, R2s for C1 – C8 = 26.6% - 40.4%, R2s for A1 – A8 = 5.1% - 39%.

[2] Prior to extracting these factor scores, we ran separate higher-order factor models were verbal ability was accounted for by regressing the indicators representing MST, Explanations, Certainty and Alternatives onto verbal ability, and, in a further separate model, the covariance between gender, age, personality (extraversion, openness to experience, agreeableness, emotional stability, and conscientiousness) as well as whether participants spoke English as their first language with the latent factor was examined, both provided acceptable fit, CFI = .926, TLI = .906, RMSEA = .033 [.022, .042] and CFI = .916, TLI = .902, RMSEA = .030 [.019, .038], respectively (see Chapter 4 for more details).

How does the best-fitting factor structure for dimensional mindreading and MASC relate to social and mental health outcomes?

We investigated how individual variation in the performance on two mindreading tasks, the DMT (see Figure 4.1) and MASC, related to social and mental health outcomes. For the DMT, for each individual model, one social/mental health outcome variable was regressed onto the mindreading latent factors (in a series of structural equation models). This data suggests that whilst the DMS and MASC converge, variation in performance on either task was not associated with any of the social and mental health outcomes assessed in the present study, except for use of alternatives being negatively associated with social support from friends, and the performance MASC correct subscale being a significant, negative predictor of both the AQ-10 and AQ-5. This suggests that individuals with greater mindreading capacity according to the MASC, were less likely to score highly in terms of autistic traits. These associations persisted when all covariates were accounted for by regressing the social/mental health outcomes additionally onto demographic variables, personality, and verbal ability (the output from this analysis is shown in Table 4.2.

Table 4.2. Output from separate structural equation models for each listed outcome variable (DMT) and output from path analysis model for each listed outcome variable (MASC). The reported values reflect analyses once all covariates were accounted for.

	DMT							
	Regression paths					**Fit indices**		
Outcomes	**HO**	**MST**	**Expl**	**Cert**	**Alt**	**CFI**	**TLI**	**RMSEA**
SS Family	.028	.012	.111	.135	**-.285***	.916	.901	.034 [.022,.043]
SS Friends	.011	-.105	.232	.107	-.203	.919	.904	.033 [.021,.043]
SS Sig Other	-.058	-.001	-.048	.094	-.056	.914	.898	.034 [.023,.044]
Anxiety	-.081	.062	-.101	-.032	.171	.931	.918	.031 [.017,.041]
Depression	-.034	.050	-.022	.023	.112	.925	.911	.032 [.020,.042]
Autism	-.133	-.073	-.018	-.151	.142	.927	.913	.031 [.019,.041]
Autism Social	-.067	-.038	-.148	-.209	.207	.918	.903	.033 [.022,.043]

	MASC
	Regression paths
	MASC Correct
SS Family	-.037
SS Friends	.016
SS Sig Other	.539
Anxiety	.029
Depression	.007
Autism	-.195**
Autism Social	-.304**

Note. The extracted latent factor scores from the best-fitting mindreading model (Figure 1) were used for all structural equation models involving the DMT. HO = Higher-order factor, MST = Mental state terms, Expl = Explanation, Cert = Certainty, Alt = Alternative, *p = .05 **p = .01

4.3.3 Social Motivation Latent Factor

CFA with a Maximum likelihood with robust standard error as the estimator was used to test a social motivation latent variable based on four self-report measures used to assess social motivation. In line with previous pilot work and the predictions set out in our pre-registration, the model provided good fit for the data, CFI = .970, TLI = .909, RMSEA = .077 [.000, 178], standardized factor loadings ranged from .267 - .939, with the latent factor explaining between 13% and 88.1% of the variance in each indicator, all *ps* = .001, and the latent factor scores for social motivation was saved to be used in further analysis.[3]

[3] **Pilot work information:** This analysis builds upon and replicates previous pilot research conducted in March 2023, which involved 135 volunteer participants (*Mage* = 19.12 years, *SD* = 1.26 years, range = 18 – 28, consisting of 20 males, and 114 females) as part of an online study. In addition to the social motivation measures that we administered in the present study, for the pilot work, the Social Reward Questionnaire (SRQ; Foulkes, Viding, McCory, & Neumann, 2014) was considered for creating a social motivation latent factor. The four self-report measures reported here were selected for this study as they provided the best fit for constructing a social motivation latent factor.

How does Social Motivation relate to Dimensional Mindreading as assessed on the DMT? Bivariate Correlations

There were weak to moderate positive correlations between the factor scores for social motivation and mindreading (Figure 4.1) suggesting that participants who possessed higher degrees of social motivation also used more mental state terms, provided more explanations, expressed more uncertainty and provided more alternative interpretations on the dimensional mindreading task.

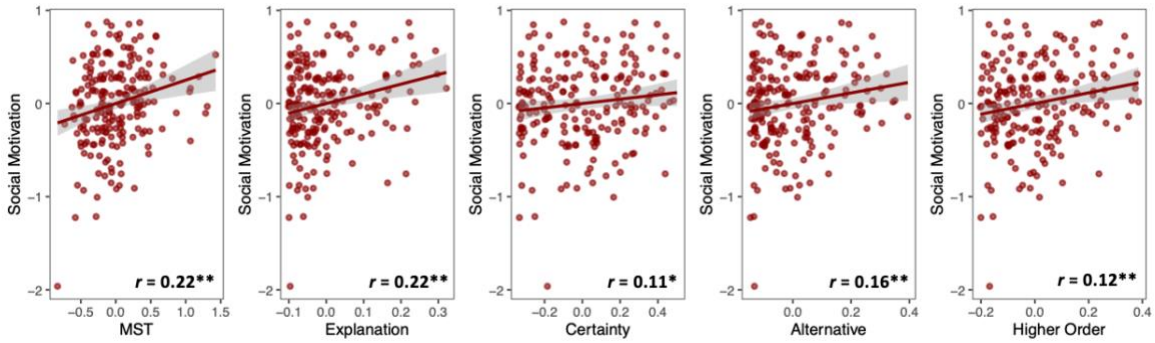

Figure 4.2. Correlations between the Social Motivation latent factor and the latent factors of the best-fitting mindreading model. The shaded areas reflect 95% confidence intervals. **p<.01, *p<.05.

4.3.4 Structural Equation Models

To assess how social motivation related to dimensional mindreading, the best-fitting higher-order mindreading model (outlined in Figure 4.1) was expanded by the social motivation latent variable (made up of the four self-report measures to assess social motivation). We then compared a model where (a) the social motivation latent factor also loaded onto the higher order factor with a second model where (b) the social motivation latent factor remained as a separate but correlated variable not loading onto the higher-order mindreading factor. If model (a) provided the best fit, then mindreading and social motivation were underpinned by a common underlying construct. In contrast, if model (b) provided the best fit, mindreading and social motivation were separable constructs. Model (a) provided an acceptable fit to the data, CFI = .918, TLI = .905, RMSEA = .032 [.021, .040], factor loadings: .272 - .898, all *ps* = .001 - .026, however, the social motivation latent variable did not significantly load onto the higher-order factor, factor-loading = .264, p = .348. Model (b) provided acceptable fit as well, CFI = .912, TLI = .900, RMSEA = .033 [.022, .041], factor

loadings: .056 - .950, all *ps* = .001-.037. Given that fit was equally acceptable, and all factor loadings were significant, it appears that social motivation and mindreading are distinguishable but correlated constructs. Social motivation and the dimensional mindreading latent factors were therefore retained as separate variables for further analysis.

How does dimensional mindreading in combination with social motivation relate to mental health and social competence ?

To investigate how dimensional mindreading in combination with social motivation related to the social and mental health outcome measures considered in this study, a series of structural equation models each with a different outcome measure and the same predictors (see Table 4.3). In each model, the respective outcome variable was regressed both onto social motivation and the mindreading higher-order latent factor. Further, the social outcome was regressed onto the demographic variables (i.e., age in years, gender, education) as well as verbal ability. All independent variables were permitted to covary in each model. Social motivation consistently significantly predicted all social outcome variables (except for anxiety), such that individuals with higher social motivation had more social support from family, friends, and significant others, as well as lower depressive symptoms and autism traits, above and beyond the impact of verbal ability and demographic variables. Conversely, the higher-order mindreading factor did not significantly predict any social outcomes, despite being consistently positively significantly related to social motivation across all structural equation models. This data suggests that peoples' motivation to mentalise and pursue social goals, rather than how they engage with mindreading tasks, drives positive social outcomes.

Table 4.3. Summary of seven individual structural equation models (each with a different social / mental health variable) and the same set of independent variables (Social Motivation, Mindreading, Age, Gender, Education and Verbal Ability).

	Social and Mental Health Outcome Variables						
	SS Family (CFI = .952, TLI = .906, RMSEA = .075 [.047,.102])	**SS Friends** (CFI = .944, TLI = .900, RMSEA = .080 [.052,.109])	**SS Sig Other** (CFI = .952, TLI = .906, RMSEA = .071 [.041,.100])	**Anxiety** (CFI = .953, TLI = .908, RMSEA = .070 [.041,.099])	**Depression** (CFI = .954, TLI = .910, RMSEA = .069 [.040,.098])	**Autism** (CFI = .952, TLI = 906, RMSEA = .071 [.042,.100])	**Autism Social** (CFI = .954, TLI = .909, RMSEA = .70 [.040,.099])
Predictors	**Regression pathways (Standard Error)**						
MR LF	-.083 (.480)	-.031 (.400)	-.078 (.451)	.102 (.475)	.130 (.470)	.006 (.514)	-.034 (.524)
SM LF	**.361 (.141) ****	**.484 (.128) ****	**.269 (.175) ****	-.117 (.150)	**-.139 (.159) ***	**-.328(.170) ****	**-.305 (.068) ****
Age	.166 (.060) **	-.089 (.061)	.047 (.070)	-.208 (.065) **	-.204 (.068) **	-.077 (.068)	-.027 (.062)
Gen	.093 (.070)	.025 (.061)	-.080 (.070)	-.097 (.069)	-.034 (.068)	-.051 (.070)	-.134 (.068)
Edu	.087 (.065)	.136 (.062)	-.012 (.067)	-.102 (.062)	-.137 (.066) *	.009 (.067)	-.074 (.064)
WASI	-.089 (.070)	-.012 (.068)	.062 (.074)	-.059 (.067)	.030 (.063)	.039 (.069)	-.035 (.066)

Note. SS = Social Support, Sig. Other = Significant Other, MR LF = Mindreading higher-order latent factor, SM LF = Social Motivation Latent Factor, Gen = Gender, Edu = Education, WASI = Verbal Ability. **p<.01, *p<.05

4.3.5 Moderation Analysis

Next, we investigated whether social motivation moderated specific relationships between mindreading and other measures (schematically shown in Figure 4.3). This analysis was motivated by the consideration that the degree to which an individual is socially motivated may impact whether 'mindreading ability' has a significant impact on any other variables. For example, mindreading ability might only matter for a given construct in the context of high levels of social motivation.

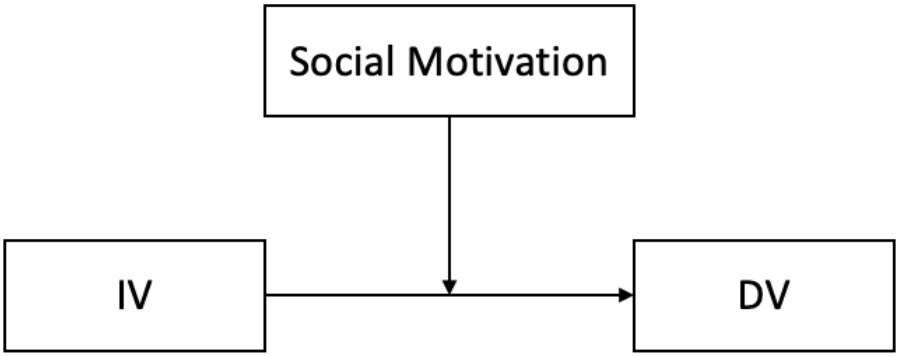

Figure 4.3. For each model, either the Higher-order mindreading latent factor or 'MASC correct' observed score was used as the IV (independent variable), social motivation was the moderator, and a social/mental health outcome variable was used as the DV (dependent variable).

Table 4.4 displays the results from the moderation analyses. These models reflect 7 different dependent variables being separately paired with six independent variables, the Higher-order, MST, Explanation, Certainty and Alternative mindreading latent factors and as well as the MASC correct observed score. These moderation analyses were carried out using the PROCESS add-on (Hayes, 2017) in R, and a percentile bootstrap approach to estimate confidence intervals was used. Specifically, 5000 bootstraps were performed to obtain robust confidence intervals for the interaction effect. Prior to the analysis, all continuous variables were mean centred (binary variables were not mean centred as this can introduce multicollinearity issues; Wissmann & Toutenburg, 2007). For all models, the following covariates were included: age, education, verbal ability, gender, and personality variables (i.e., conscientiousness, openness to experience, extraversion, agreeableness, and emotional stability). For significant moderation effects, conditional effects of the independent variable on the dependent variable at different levels of social motivation, specifically, at 1 SD below and 1 SD above the mean value, as well as at the mean value, were considered. Given the large number of analyses, we corrected for multiple tests or adopt a more stringent p-value of $p<.01$ (Chen et al., 2017).

Across all models, there were three significant interactions (see Figure 4). One of these models remained significant at the adjusted the $p<.01$ level and is therefore interpreted in further detail below. Specifically, in terms of overall autism tendencies, whilst social motivation and MST were individually not significantly related to autistic traits, *ps* = .548 and .091, respectively, there was a significant interaction between social motivation and the MST factor, $R^2 = .029$, F (1, 197) = 7.94, $p = .005$, indicating that the relationship between MST and

autistic tendencies was significantly moderated by social motivation, with an additional 2.9% of variance in autistic traits explained by the interaction term. There was no significant link between MST and autism at average and high levels of social motivation, however, at low levels of social motivation (-1SD), there was a significant negative correlation between MST and autism (p = .02). This result indicates that for individuals with lower levels of social motivation, the more MST they used, the less likely they were to possess autistic traits (Table 4.4, Model 6.1; Figure 4.4, Model 6.1).

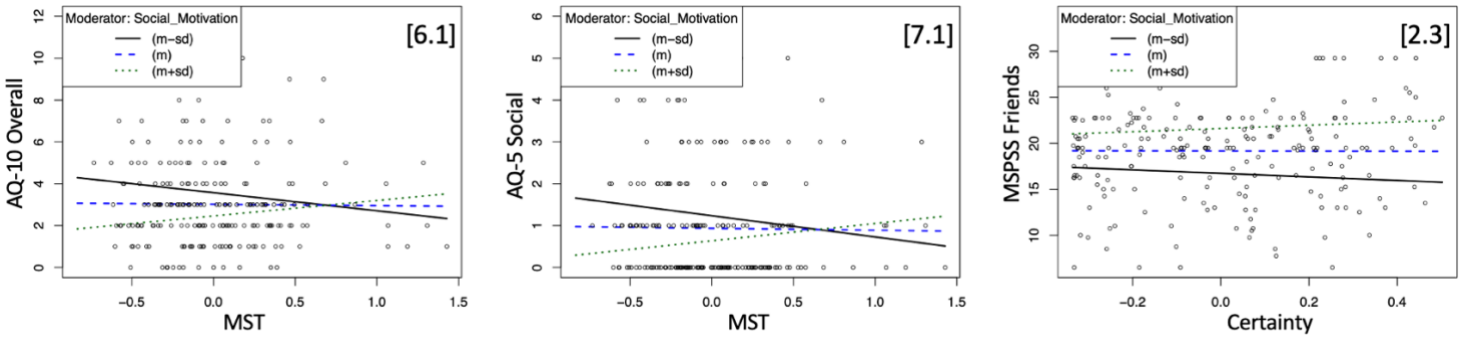

Figure 4.4. Conditional effects of the independent variable on the dependent variable at different levels of social motivation (+1SD, Mean, -1SD) for models 2.3, 6.1 and 7.1 (see Table 4.4).

Table 4.4. Output from moderation analysis.

Social Support Family			**Social Support Friends**			**Social Support Significant Other**			**Anxiety**			**Depression**			**Autism**			**Autism Social Sub-scale**		
	Coefficients			**Coefficients**			**Coefficients**			**Coefficients**			**Coefficients**			**Coefficients**			**Coefficients**	
[1]	HO	-.37	**[2]**	HO	-.25	**[3]**	HO	-.40	**[4]**	HO	.44	**[5]**	HO	.54	**[6]**	HO	-.24	**[7]**	HO	-.26
	SM	.47 **		SM	.89 **		SM	.51 **		SM	.24		SM	.22		SM	-.40*		SM	-.48**
	HO x SM	-.38		HO x SM	1.62		HO x SM	-.62		HO x SM	-.86		HO x SM	.64		HO x SM	1.47		HO x SM	.41
[**1.1**]	MST	-.12	**[2.1]**	MST	-.38*	**[3.1]**	MST	-.26	**[4.1]**	MST	.17	**[5.1]**	MST	.22	**[6.1]**	MST	-.10	**[7.1]**	MST	-.01
	SM	.44 **		SM	.82**		SM	.46 *		SM	.30*		SM	.27		SM	-.30		SM	-.39*
	MST x SM	-.24		MST x SM	-.18		MST x SM	-.46		MST x SM	.15		MST x SM	.39		MST x SM	1.01*		MST x SM	.78*
[1.2]	Expl	-.36	**[2.2]**	Expl	-.07	**[3.2]**	Expl	-1.02	**[4.2]**	Expl	.50	**[5.2]**	Expl	.68	**[6.2]**	Expl	-.18	**[7.2]**	Expl	-.38
	SM	.47 **		SM	.87**		SM	.55 **		SM	.23		SM	.19		SM	-.42*		SM	-.50**
	Expl x SM	-.52		Expl x SM	1.92		Expl x SM	.03		Expl x SM	-2.19		Expl x SM	-.46		Expl x SM	1.81		Expl x SM	-.10
[1.3]	Cert	-.04	**[2.3]**	Cert	.03	**[3.3]**	Cert	-.06	**[4.3]**	Cert	.23	**[5.3]**	Cert	.29	**[6.3]**	Cert	-.24	**[7.3]**	Cert	-.24
	SM	.46 **		SM	.86**		SM	.50 **		SM	.26		SM	.23		SM	-.43**		SM	-.49**
	Cert x SM	-.27		Cert x SM	1.21*		Cert x SM	-.49		Cert x SM	-.47		Cert x SM	.68		Cert x SM	.64		Cert x SM	.09
[1.4]	Alt	-.08	**[2.4]**	Alt	-.39	**[3.4]**	Alt	-.41	**[4.4]**	Alt	.48	**[5.4]**	Alt	.57	**[6.4]**	Alt	-.10	**[7.4]**	Alt	-.11
	SM	.41		SM	.89		SM	.51**		SM	.26		SM	.23		SM	-.41*		SM	-.49**
	Alt x SM	-.09		Alt x SM	1.75		Alt x SM	-.73		Alt x SM	-.58		Alt x SM	.90		Alt x SM	1.51		Alt x SM	.55
[1.5]	MASC	-.08	**[2.5]**	MASC	-.07	**[3.5]**	MASC	.03	**[4.5]**	MASC	.06	**[5.5]**	MASC	.008	**[6.5]**	MASC	-.14*	**[7.5]**	MASC	-.17*
	SM	.41*		SM	.84		SM	.44*		SM	.36*		SM	.30		SM	-.28		SM	-.39*
	MASC x SM	-.09		MASC x SM	-.02		MASC x SM	-.07		MASC x SM	.10		MASC x SM	.08		MASC x SM	.15		MASC x SM	.07

Note. Output from moderation analysis. Social Motivation was the moderator across all models. Independent variables were either the extracted higher-order or first-order mindreading latent factor scores or the observed overall 'MASC correct sub-scale' scores. Dependent variables are shown across the 7 columns. For simplicity, the output for the covariates that were included in each displayed model - education, verbal ability, gender, personality (conscientiousness, openness to experience, extraversion, agreeableness, and emotional stability) – is not shown in the table. For degrees of freedom for the Model fit F-statistic were: df1 = 12, df2 = 197 across all models. MASC = MASC = Movie for the assessment of social cognition – Correct Sub-scale. Given the large number of analyses, we corrected for multiple tests by adopting a more stringent p-value of $p<.01$.

4.4 Discussion

The present study investigated links between social motivation and mindreading, and their relationships with social competence and mental health. In short, the present results revealed that there were weak to moderate positive associations between mindreading and social motivation, while both were still distinguishable constructs. Furthermore, the present results revealed unique associations between individual differences in social motivation (but not mindreading) and social competence and mental health (i.e., depressive symptoms, anxiety, autistic traits, and social support from family, friends, and significant others) in adults. The findings corresponding to our specific aims are summarised and discussed below.

The first aim was to elucidate the relationship between individual differences in mindreading, as assessed via the DMT, and social motivation. We aimed to discern whether these domains represented distinct or overlapping constructs. To this end, we conducted a confirmatory factor analysis (CFA) and tested a latent variable for social motivation based on data from four self-report measures, replicating prior independent pilot work. Our results revealed weak to moderately significant positive correlations between the social motivation factor scores and the various dimensions of mindreading (i.e., MST, Explanation, Certainty and Alternatives, as well as the higher-order mindreading factor) (see Figure 2). This result indicated that individuals with heightened social motivation tended to possess greater mindreading abilities. However, despite significant shared variance between both constructs, social motivation and mindreading were distinguishable. Specifically, we compared two models: one where social motivation, alongside the dimensions of MST, Explanation, Certainty and Alternatives loaded onto the higher-order factor, and another where social motivation formed a separate latent factor. Both models demonstrated acceptable fit, but in the former model, the social motivation latent factor failed to load significantly onto the higher-order factor, indicating that mindreading and social motivation were not explained by a common underlying factor. This finding aligns with existing research as it supports the notion that social motivation and mindreading are conceptually and empirically separate (e.g., Contreras-Huerta et al., 2020; Carpenter et al. 2016; Devine & Apperly, 2022, see Chapter 2).

Our second aim was to examine the potential links between individual differences in mindreading and social competence and mental health among adults. Although the Dimensional Mindreading Task (DMT) and Movie for the Assessment of Social Cognition (MASC), two methodologically distinct tasks, exhibited convergent validity, adults' performance on these tasks was not correlated with either social competence or mental health. The sole exception was a weak, albeit significant negative association between the use of alternatives on the DMT and participants' level of social support from friends, as well as a negative association between performance on the MASC and autism (see Table 2). This finding shows that individuals with greater mindreading capacity according to the MASC exhibited reduced autistic traits, corroborating prior research (e.g., Dziobek et al., 2006; White, Schry, & Maddox, 2012; Martinez et al., 2017).

The third aim was to compare the associations between social motivation and mindreading and each of the social and mental health outcomes. Across a series of structural equation models, social motivation was consistently associated with social competence and mental health variables, except for anxiety. Specifically, the results showed that individuals with higher levels of social motivation reported significantly more social support from family, friends, and significant others, as well as reduced depressive symptoms and autism-related traits. These relations remained significant after accounting for variation in age, gender, education, personality, and verbal ability. Conversely, the higher-order mindreading factor was not uniquely associated with social competence or mental health, despite its significant positive association with social motivation. Although prior research found that both social motivation *and* mindreading ability in 8-to-13-year-old children made distinctive contributions to predicting variation in teacher-rated social ability (Devine & Apperly, 2022), the present findings suggest that, in *adults*, it is the motivation to mentalise and pursue social objectives, rather than performance on specific mindreading tasks, that is significantly positively associated with social functioning and mental health variables.

4.4.1 Mindreading, Social Competence and Mental Health

Contrasting with the predictive power of social motivation, individual differences in mindreading was not uniquely associated with most of the social and mental health outcomes examined in this study. This finding aligns with prior research suggesting mixed findings between mindreading and mental health conditions (e.g., Berecz et al., 2016; Lenton-Brym et

al., 2018). For instance, a recent meta-analysis by Berecz et al. (2016) suggests that no firm conclusions can be drawn concerning the extent to which mindreading in adults is associated with depressive symptoms. However, a recent systematic review of meta-analyses across 30 clinical conditions by Cotter et al. (2018) showed consistent links between diverse measures of mindreading and mental health conditions, including depression, in adults. The apparently mixed results may, at least in part, be due to the fact some work on mental health and mindreading has used case-control designs, which select the most clinically severe cases whereas others have investigated links between mindreading and sub-clinical levels of mental health conditions, where links may be less apparent. In support of this notion, Lenton-Brym et al. (2018) found no link between mindreading and subclinical levels of anxiety. Although there is evidence of links between childhood adversity and impairments in adults' mindreading (e.g., Germine et al., 2015), these findings are based on the Reading the Mind in the Eyes (RMET; Baron-Cohen et al., 2001), which has been argued to assess emotion perception rather than mindreading (e.g., Oakley et al., 2016).

However, it is notable that our results, consistent with prior research (e.g., Livingston et al., 2023; Dziobek et al., 2006; Castelli et al., 2001; Baron-Cohen et al., 2001), reveal observable *variation* in mindreading among neurotypical adults. Thus, while there was reliable variation in performance, it is possible that this was not within the range that would have made a difference to social relationships or mental health. More specifically, if the range of performance on the DMT had been lower, then it is conceivable that it would have exhibited significant relations with potential correlates. Furthermore, a more nuanced and specific assessment of the outcome variables, focussing, for example, on the quality rather than the quantity of social support, could perhaps have revealed links between mindreading and outcomes. Indeed, this notion aligns with research, which demonstrated that mindreading in children and adolescents was uniquely associated with relatively more *complex* social interaction skills rather than basic ones (Devine & Apperly, 2022; see also Peterson, Slaughter, Moore, & Wellman, 2016). In future research, it could therefore be valuable to explore the relationship between mindreading and "advanced social skills", such as from the realm of positive psychology, which studies the conditions that contribute to optimal psychological functioning (e.g., Gable & Haidt, 2005). For example, exploring variables less likely to be influenced by social motivation, such as intellectual humility, that is, a character virtue enabling individuals to recognize their own potential fallibility when forming and revising attitudes and beliefs (e.g., Deffler, Leary, & Hoyle, 2016), could offer intriguing avenues to re-

evaluate the relative contributions of social motivation and mindreading to outcome variables. Furthermore, it might be valuable to investigate relationships between mindreading and how people form relationships with those who differ from them (i.e., cross-ethnic relationships). Indeed, there is some work suggesting that mindreading skills are inversely related to prejudice towards outgroups (e.g., Westra, 2019; Ekerim-Akbulut, Selçuk, Slaughter, Hunter, & Ruffman, 2020). Finally, given evidence of age-related improvements in mindreading (e.g., Slaughter et al., 2015) as well as evidence that as mental health conditions, such as anxiety, become less severe, so do links with individual differences in mindreading (e.g., Lenton-Brym et al., 2018), future research could also benefit from pinpointing the critical factors, such as a person's age or the degree of presence of mental health conditions at which the significance of associations typically linked to mindreading changes.

4.4.2 Mindreading and Social Motivation

We operationalised social motivation using a n extracted latent factor score based on four validated self-report measures (Carpenter et al., 2016; Abplanalp et al., 2022; Snaith et al., 1995; Watson & Friend, 1969). These measures captured various facets of social motivation, including individuals' tendencies to nurture social connections, preference for social interactions over non-social activities, and their ability to derive pleasure from social engagements (e.g., Chevallier et al., 2012). Notably, our factor structure for social motivation (i.e., a one-factor model incorporating the four self-report measures) replicated pilot work conducted on a separate sample. To ensure that 'social motivation' rather than general motivation was captured, we accounted for individual differences in 'conscientiousness' in our analyses. This emphasises the robustness and reliability of the present social motivation measure.

Firstly, there were weak, significant positive correlations between the social motivation factor score and mindreading, suggesting that individuals with higher levels of social motivation exhibited greater mindreading abilities. In the context of the DMT, this meant that participants were more likely to use mentalistic terms, express more uncertainty, provide an explanation for the mental states they attributed and suggest an alternative interpretation in the context of picture stimuli of social scenarios. Given that a higher-order model mindreading model provided the best fit to the data, this suggests that there is significant communality across

these four factors, which is captured by the higher-order factor (see chapter 3 for a more elaborate exploration).

Accordingly, existing studies have also reported similar weak-moderate positive associations between individual differences in motivation and mindreading (e.g., Carpenter et al., 2016; Devine & Apperly, 2022; Burnside et al., 2018), as well as closely related constructs such as cognitive empathy (Lookwood et al., 2017). Furthermore, this finding supports the notion that mentalistic attributions demand cognitive effort and involve memory and cognitive control processes (e.g., Contreras-Huerta et al., 2010; Apperly et al., 2006; Kouklari et al., 2018; Ferguson et al., 2017).

However, despite these associations, our results emphasise the *distinctiveness* of mindreading and social motivation, aligning with prior research (e.g., Devine & Apperly, 2022; Contreras-Huerta et al., 2020). These results suggest that the DMT does not principally assess individual differences in motivation. The distinctiveness, or meaningful non-shared variance, of social motivation and mindreading was further supported by the differential predictive relations observed between these constructs and social competence and mental health. In contrast to existing research on children and adolescents that reported distinctive contributions of both mindreading and social motivation to predicting variation in teacher-rated social ability (e.g., Devine & Apperly, 2022), we found that social motivation positively associated with social support from family, friends, and significant others, and significantly negatively predicted depressive symptoms and autistic traits. These findings are consistent with existing literature on the correlates of social motivation. For example, pupils' social motivation has been positively linked to their perceived social support from their environment (e.g., Tezci et al., 2015; Wentzel et al., 2010), and social motivation has been found to be impaired in individuals with depression (e.g., Gandhi et al., 2022, for a meta-analysis; Hofmann et al., 2011; Seidel et al., 2010), anxiety (e.g., Swain, Scarpa, White, & Laugeson, 2015), and autistic traits (e.g., Chevallier et al., 2012). Based on the present findings regarding social motivation, an aim for future research would be to investigate whether social motivation shapes one's social and mental health outcomes, or whether individuals with more ideal environmental circumstances develop higher social motivation. Longitudinal data may be the best way to address this question.

4.4.3 Conclusion

In summary, the primary objective of this study was to assess the potential links between social motivation and mindreading, and what their respective relationships and interplay with social abilities and mental health were. This investigation was aimed to bridge the gap between the mostly distinct existing literatures on the correlates of individual differences mindreading and social motivation in adult individuals. We investigated their interplay and respective influences on various social and mental health outcomes. While performance on two methodologically distinct, but converging, mindreading tasks was positively related to social motivation, confirmatory factor analysis revealed that both constructs were distinct. Our findings highlight the pivotal role of social motivation as a robust positive predictor of perceived social support, and negative predictor of depressive symptoms as well as autistic traits among a community adult sample. Conversely, individual differences in mindreading exhibited limited predictive power. Notably, social motivation, when considered as a mediator, did not significantly and interact with individual differences in mindreading to predict outcomes. However, our results, based on the measures employed in this study, do not rule out the potential significance of individual differences in adults' mindreading – it is plausible that the variance in mindreading detected by the measures employed in this study could have reached a threshold of significance if more advanced outcome measures had been tested. A deeper exploration of these intricate causal relationships and interactions could facilitate the development of tailored interventions and support systems suitable for adults.

Chapter 5

General Discussion

5 General Discussion

5.1 Summary and Synthesis

5.1.1 Overview of aims

The research presented in this dissertation was motivated by two primary observations. First, individuals exhibit substantial variation in how and to what extent they engage in mindreading, despite most adults having a conceptual understanding of mental states (e.g., Apperly, 2012). Second, comprehending the nature of such variation requires the consideration of motivational factors, given that a person's motivation fundamentally shapes their behaviour (e.g., Hess, 2014).

The first overarching aim was to understand *why* individual differences in mindreading persist among adults, despite most individuals demonstrating conceptual understanding of mental states by late childhood (e.g., Wellman et al., 2001) To achieve this, I explored the idea that individual differences in adult mindreading could be examined in terms of *both* the extent to which mindreading is accurate, or 'contextually justified' (see chapter 2) as well as in terms of the underlying *processes* involved in mentalising, independently of response accuracy (see chapter 3). Secondly, given the limited empirical or theoretical integration between motivational accounts and socio-cognitive investigations in general, as well as specifically regarding mindreading and motivation (for recent exceptions to this, see Carpenter et al., 2016; Contreras-Huerta et al., 2020; Devine & Apperly, 2022; Lockwood et al., 2017), the second overarching aim was to examine individual differences in adult mindreading both independently and in conjunction with motivational factors. This involved investigating the *links* between social motivation and mindreading, and their respective relationships with social abilities and mental health variables (see chapter 2 and 4).

The studies reported here departed from the existing psychological research in the field of mindreading in two ways. Firstly, rather than relying on existing measures of mindreading, which are oftentimes characterised by inadequate psychometric properties (e.g., Warnell & Redcay, 2019; Oakley et al., 2016; Bradford et al., 2018), we *refined* existing measures (see chapters 2) and introduced a novel approach to capture individual differences in the structure of adults' mindreading, independently of the ability to infer mental states accurately (see chapter 3). Secondly, instead of investigating mindreading and its correlates in isolation, we

focussed on assessing individual differences in mindreading *alongside* social motivational factors in adult samples (see chapters 2 and 4).

In the subsequent sections, the main aims and conclusions from each chapter are summarised (section 5.1.2), followed by a synthesis of the implications of the present findings and recommendations for future research (section 5.1.3), and overview of the limitations of the current research (section 5.3), before discussing overarching conclusions that can be drawn from the present studies (section 5.4).

5.1.2 Summary of Findings

The chapters within this dissertation share a common overarching approach and aims, but each chapter maintains a unique focus and answers distinct research questions. A summary of the main findings from each chapter is outlined below.

Chapter 2

Chapter 2 centred on the premise that some individuals might be motivated to mindread more than others, even if they were not especially successful in doing so. We therefore hypothesised that adults differ in the degree to which they exhibit (a) motivation for mindreading and (b) the extent to which their responses to mindreading tasks are appropriate (or context-sensitive). It was reasoned that existing coding schemes for tasks assessing mindreading through open-ended responses may confound these factors as they do not differentiate the frequency of mental state terms (MST), a potential indicator of motivation to mindread, from the quality of explanations provided.

Based on an innovative scoring system, we investigated whether individual differences in neurotypical adults' answer quality and / or quantity of explicit references to others' mental states on two open-ended response mindreading tasks – the Movie for the Assessment of Social Cognition (MASC; Dziobek et al., 2016) and the Silent Films Task (SFT; Devine & Hughes, 2013) were (a) separable constructs, (b) accounted for by mindreading motivation, and (c) related differentially to measures previously linked with mindreading.

Our analyses revealed that both a two-factor and a one-factor model provided an acceptable fit to the data. Contrary to our expectations, neither model showed significant associations with mindreading motivation (Carpenter et al., 2016). However, confirmatory factor analysis indicated that a two-factor model, with MST and response appropriateness loading onto separate factors, provided greater explanatory power. Specifically, MST exhibited a moderate, significant, and positive association with religiosity, while response appropriateness was negatively associated with religiosity. In contrast, the one-factor solution failed to predict any socially relevant correlates. These results provide some indication that mindreading quantity and mindreading quality may represent distinguishable constructs within the realm of individual differences in mindreading.

Chapter 3

In chapter 3, we aimed to explore a new method for measuring the *structural features* of mindreading, irrespective of specific content accuracy. This aim was motivated by two observations: (1) the scarcity of reliable and valid measures for assessing individual differences in mindreading among broader adult populations (e.g., Warnell & Redcay, 2019; Oakley et al., 2016; Bradford et al., 2018) and (2) existing measures' inability to explain the *reasons* behind the existence of individual differences in mindreading in adults who have already mastered the concept of mental states (e.g., Apperly, 2012; Conway et al., 2019; Dziobek et al., 2006).

Our novel task involved coding open-ended responses to social picture stimuli and captured individual variation across four distinct domains: the frequency of mental state terms (MST), mentalistic explanations, expressed certainty, and alternatives. Confirmatory factor analyses revealed that a higher-order factor structure, where these individual items representing the four identified domains loaded onto distinctive first-order factors, reliably fit the data best (based on a test and cross-validation in a separate sample). To the best of our knowledge, this is the first study to employ modern psychometric analysis to develop a measure of individual differences in mindreading for adults.

From a psychometric perspective, these findings demonstrate that the 'Dimensional Mindreading Task' (DMT) effectively captured performance variation across these four domains, irrespective of answer accuracy. However, questions may arise about whether DMT

performance reflects broader cognitive tendencies, such as communication style, rather than genuine variation specific to mindreading.

To address this concern, we assessed the degree of *convergent validity* between our new measure and the MASC, an established mindreading measure. We found that specific 'ways of mentalising', namely, greater use of MST, providing explanations, expressing uncertainty, and providing alternatives, were significantly positively associated with higher accuracy on the MASC. These four domains were underpinned by a higher-order factor, suggesting that there is significant common variance between the four sub-categories which may amount to an 'overarching mindreading capacity'. Furthermore, *discriminant validity* was assessed by showing that our novel measure does not simply capture individual differences in verbal ability or personality tendencies. Overall, our findings suggest that the DMT may indeed capture aspects *specific* to mindreading and that mindreading accuracy may not be the sole avenue to understanding meaningful individual differences in mindreading among adults.

Chapter 4

In chapter 4, we synthesized insights from chapters 2 and 3 with the aim to revisit the question of the potential link between social motivation and mindreading and their respective relationships with other constructs. Firstly, we aimed to determine whether the domains of mindreading and social motivation represented distinct or overlapping constructs.

Confirmatory factor analysis revealed that there were significant weak to moderate positive associations between mindreading (as assessed by the Dimensional Mindreading Task (DMT) and Movie for the Assessment of Social Cognition (MASC; Dziobek et al., 2006)) and social motivation. However, despite these associations, the results from our confirmatory factor analysis also highlighted the empirical *distinctiveness* of mindreading and social motivation.

Next, we examined how individual differences in mindreading related to various outcome variables, including depressive symptoms, anxiety, autistic traits, and social support (from family, friends, and significant others). Although the DMT and MASC, two methodologically distinct tasks, exhibited convergent validity, neither task yielded significant unique correlations with either social competence or mental health, except for weak significant and negative associations between MASC performance and autistic traits.

Our third aim was to assess the relative importance of social motivation and mindreading in relation to the above-outlined social and mental health measures in order to bridge a current gap between separate bodies of literature concerning the correlates of mindreading accuracy and social motivation (see Devine & Apperly, 2022 for an exception). Structural equation modelling showed that social motivation was associated with social competence, autism, and depressive symptoms, but not anxiety. Specifically, the results showed that individuals with higher levels of social motivation reported significantly more social support from family, friends, and significant others, as well as reduced depressive symptoms and autism-related traits. These associations remained significant after accounting for variation in age, gender, education, verbal ability, and personality (including conscientiousness and extraversion which likely overlap with general motivation). In contrast, the higher-order mindreading factor did not exhibit a unique association with social competence or mental health, despite its positive correlation with social motivation. Chapter 4 highlights the essential role of social motivation as a robust correlate of perceived social support, and negative correlate of depressive symptoms as well as autistic traits among an adult community sample.

5.1.3 Implications of the Present Findings and Recommendations for Future Research

The reported results offer four overarching insights. These insights are briefly summarised below and further elaborated on in relation to the research questions they raise for future inquiry.

Firstly, this current research challenges past epistemological assumptions about how mindreading can be operationalised and measured. The present research shows, for the first time, that there are reliably measurable, content-irrelevant processes and strategies that individuals employ when engaging in mindreading (see section 5.1.3.2 below). Secondly, the results suggest that mindreading is likely underpinned by multiple factors. We found tentative evidence for a two-factor mindreading structure (as discussed in chapter 2) and, furthermore, based on a novel mindreading task, we identified four distinct processes explained by an overarching mindreading ability (see section 5.1.3.1 below). Thirdly, the present results highlight the potential promise and feasibility of conducting data-driven or inductive research for identifying variables that reflect genuine individual differences in the processes involved in adults' mindreading (see section 5.1.3.3 below). Lastly, the present dissertation provides

support for the notion that researchers interested in mindreading should consider social-motivational factors alongside mindreading, especially if associations between mindreading and other constructs are of interest (see section 5.1.3.3 below).

5.1.3.1 Mindreading can be Assessed Irrespective of Response Accuracy

In contrast to conventional mindreading tasks, including advanced ones, which have predominantly focussed on measuring individuals' ability to infer mental states *accurately* (e.g., Castelli et al., 2001; Devine & Hughes, 2013; Dziobek et al., 2016), our findings suggest that it is possible to capture *structural features* of mindreading effectively in two independent samples from the broader adult population, without focussing on the *accuracy* of specific content. This study thereby provides initial evidence that a diversion from current ways of assessing mindreading exclusively in terms of accuracy may be a fruitful supplementation to existing methods to take to help our understanding of *why* individuals differ in their mindreading. Notably, the ambiguity of the static picture stimuli regarding what an accurate response would have been, can be seen as possessing a high degree of ecological validity, because real-life social encounters or conditions under which mentalising would spontaneously are often characterised by dynamically changing mental states that are challenging to know (e.g., Cassels & Birch, 2014; Betz et al., 2019). Furthermore, the DMT also holds potential for future research to compare the psychological correlates of individual differences in mindreading across different cultures, given that it does not depend upon understanding cultural nuances. This contrasts with, for example, the MASC (Dziobek et al., 2016) which shows a dinner party in a Western context, thereby potentially disadvantaging participants with little exposure to such contexts with regards to making accurate mentalistic attributions.

Nevertheless, future research is needed to administer the DMT to different samples for further validation. In this regard, it is important to note that different degrees of the four dimensions identified may be differentially appropriate depending on the nature of the stimuli administered. Given that the study reported in Chapter 3 utilized static, decontextualised picture stimuli of social scenarios, that is, participants were provided with very limited information and there was no ground-truth for characters' mental states, *higher* degrees of flexibility in mindreading were logically warranted. Future research would therefore benefit from classifying task stimuli used to elicit mindreading in terms of their general ambiguity (i.e., how 'easily knowable' characters' mental states are). By doing so, potentially intriguing

relationships between *how* people mentalise (i.e., MST, explanations, certainty, and alternatives) and how *correctly* they mentalise could be uncovered. For example, the DMT could be administered in relation to other mindreading tasks, or it may even be possible to code both structural features as well as accuracy from the same task, if the stimuli differed and, ideally, researchers had access to a 'ground-truth' of mental states (e.g., Long et al., 2022).

5.1.3.2 Mindreading is Underpinned by Multiple Factors

Previous research has posited that mindreading encompasses multiple component processes. These include, but are not limited to, cognitive versus affective (e.g., de la Osa et al., Rossetto et al., 2020) or effortful versus automatic mechanisms (Apperly & Butterfill, 2009). Likewise, the findings of Chapter 2 emphasise the potential distinctiveness of two dimensions within mindreading: mindreading quantity (i.e., frequency of MST) and mindreading quality (i.e., response appropriateness). Building upon this finding, Chapter 3 suggest that mindreading may be underpinned by consistent variation across four distinct domains on the Dimensional Mindreading Task (DMT): the frequency of mental state terms (MST), mentalistic explanations, expressed certainty, and alternatives.

Crucially, our findings are not mutually exclusive with existing conceptualisations of the cognitive processes involved in mindreading. Rather, they highlight the contextual nature of these processes. To illustrate, depending on the specific context (i.e., task stimuli) and a person's idiosyncratic tendencies, different levels of the four DMT dimensions could reflect either 'effortful' or 'automatic' mindreading (Apperly & Butterfill, 2009). Consider for example static, decontextualised picture stimuli where discerning characters' mental states is relatively challenging. If such stimuli are interpreted by an individual with a tendency to be certain in their mentalistic judgements, then the exhibition of higher degrees of uncertainty, more explanations for responses and the provision of alternatives might signify 'effortful mindreading', which, in this task context, would also likely be positively correlated with mindreading accuracy. Conversely, if the same picture stimuli were evaluated by a person who is inclined to provide explanations and entertain diverse possibilities in the context of mindreading, then relatively less 'effortful' or more 'automatic' processes may underpin a similar 'performance' on the DMT.

Therefore, there may be multiple 'cognitive routes' to achieving 'accuracy' or 'inaccuracy' on a mindreading task. Understanding these potentially multifaceted cognitive processes underlying mindreading is important to understand why there are individual differences in mindreading. Although being far from elucidating a full picture, the present research substantially contributes to the ongoing debate on how to operationalise mindreading.

5.1.3.3 Mindreading Research and Data-driven Explorations

Empirical research has traditionally been guided by theoretical frameworks, which are indispensable for formulating and testing hypotheses. Nevertheless, concerns persist about the extent to which such hypotheses and study designs are influenced by preconceived notions about the nature of the social world and questions have been raised about the extent to which psychological research is influenced by the personal characteristics of researchers, who often do not represent the diversity of the general population (e.g., Haidt & Jussim, 2016).

The current findings underscore the promise of future data-driven research for illuminating the origins of individual differences in adults' mindreading. For example, in addition to the factor structure of the DMT identified in this study, it is also possible that there are still further components in the structure of mindreading for future research to discover. To investigate the latter, further data-driven research may be necessary to validate or challenge existing theory-driven findings (e.g., Jack et al., 2018). Importantly, any accuracy-irrelevant domains identified as potentially relevant to mindreading should be evaluated against criteria that would indicate whether aspects *specific* to mindreading were measured instead of general forms of reasoning. This could, for example, be achieved by (a) statistically accounting for the general cognitive demands of mindreading tasks, which are *not* measuring mindreading, such as memory capacity, executive functioning, or general verbal ability, and (b) *relating* 'performance' on any accuracy-irrelevant domains to a measure that assesses accuracy (either stemming from the same or another mindreading task). Ideally, such a measure of mindreading accuracy would (a) be 'culturally appropriate', that is, it would not disadvantage certain participants due to a lack of cultural insight into the specific situation depicted, and (b) possess a 'ground truth' of what the correct mental state is, as this avoids potential researcher bias (for an example, see Long et al., 2022).

A data-driven approach also holds potential for identifying valid 'criteria' or 'correlates' of both mindreading and social motivation, particularly considering the mixed findings of previous work in this respect (e.g., Harkness et al., 2005; Nejati et al., 2012). Thus, our results do not rule out the potential significance of individual differences in neurotypical adults' mindreading. While there was reliable variation in performance it is possible that this was not within the range that would have made a difference to social relationships or mental health. Future research could adopt an approach where constructs that are not traditionally associated with mindreading are explored in terms of their potential significance. For example, it could be valuable to *explore* the relationship between mindreading and constructs from the realm of positive psychology, which studies the conditions that contribute to optimal psychological functioning (e.g., Gable & Haidt, 2005), or social-psychological constructs such as prejudice towards outgroups (e.g., Brewer, 1999).

5.1.3.4 Mindreading Should be Studied Alongside Social Motivation

Our findings highlight significant positive association between social motivation and mindreading, corroborating existing studies that have also reported similar weak-moderate positive associations between individual differences in motivation and mindreading (e.g., Carpenter et al., 2016; Devine & Apperly, 2022; Burnside et al., 2018), as well as closely related constructs such as cognitive empathy (Lockwood et al., 2017). There are also theoretical considerations supporting an association between mindreading and social motivation (e.g., see Social Individual Differences Account; Apperly, 2012; Hughes & Devine, 2015; Dunn & Cutting, 1999; Dunn et al., 1991; Dunn, & Brophy, 2005). These considerations imply that most existing mindreading research that exclusively assesses mindreading might confound variation in mindreading with variation in motivation.

Thereby, the present research adds to a growing literature emphasising the importance of studying mindreading alongside motivational factors. Furthermore, our findings stress the pivotal role of social motivation as a robust positive correlate of perceived social support, and negative correlate of depressive symptoms as well as autistic traits among neurotypical adults. Conversely, individual differences in mindreading exhibited no significant associations with measures other than social motivation, and social motivation, when considered as a mediator, did not significantly interact with individual differences in mindreading to predict outcomes. While, as outlined above, future research will need to re-consider the potential correlates of

mindreading, our findings show the motivation to mentalise and pursue social objectives appears to be independently important for adults' social functioning and mental health. However, as discussed to in Chapter 1, situationally activated motivations likely interact with an individuals' more trait-like motivation and ability in mindreading to produce variation in mindreading structure and accuracy. Future research will therefore need to consider the role of these types of motivations. Such endeavours could involve studying motivation where it is either task-relevant (i.e., social motivation is task-relevant to social stimuli) or task-irrelevant (i.e., social motivation is task-irrelevant to non-social tasks). This could elucidate the relation between social and general motivation as well as the specificity of each construct.

5.2 Limitations

Some of the results presented in this dissertation stemmed from a university sample, primarily consisting of female psychology students (chapter 2 and study 1 of chapter 3). This may impact the generalisability of the results (Hruschka, Medin, Rogoff, & Henrich, 2018). While the remaining studies involved broader adult samples with a balanced gender and age distribution, it is important to note that all studies reported in this dissertation consisted of participants who were recruited online. This presents a limitation, as online samples tend to be unrepresentative of the general population (i.e., they exclude participants who do not have internet access or a computer from participating, potentially introducing systematic bias). Nevertheless, it is worth noting that (a) most British adults should have had the technological means to access Prolific (which claims to produce representative samples) and that (b) the focus of much of our work was on psychometric evaluations of tasks in the context of individual differences (e.g., Office for National Statistics, 2023).

From a methodological perspective, most measures reported in this study had an open-ended response format. Such tasks have been claimed to be unsuitable for large-scale research (Livingston et al., 2021). However, there are now advances in the automation of scoring for open-ended mindreading tasks, that may be applied to approaches such as the present one. For instance, Devine, Kovatchev, Grumley Traynor, Smith, and Lee (2023) successfully created a freely available deep learning neural network automated scoring system to score children's and adolescents' open-ended text responses to two mindreading tasks, amongst them the Silent Films Task (Devine & Hughes, 2013). It might therefore be possible to automate the coding of

the Dimensional Mindreading Task in a similar fashion, which would allow for its validation in larger samples.

Furthermore, the study reported in chapter 2 relied on a single self-report measure to assess mindreading motivation, and participants' scores on this measure were restricted in range. This is problematic as social motivation has been argued to be comprised of several facets such as individuals' inclination actively nurture and sustain social connections, prefer social interactions over non-social ones, and pursue pleasure in social interactions (e.g., Chevallier et al., 2012). Indeed, it is a strength of chapter 4 that the assessment of social motivation included a broader range of four self-report measures, including but not limited to motivation for mindreading, that were identified from an extensive and systematic literature search. Notably, our factor structure for social motivation (a one-factor model incorporating the four self-report measures) also replicated our previous pilot work conducted on a separate sample. Furthermore, from a practical perspective, future work should investigate (a) test-retest reliability of social motivation and (b) its convergent validity with task-based measures of motivation.

Finally, it is important to acknowledge the inherent limitation of the present research to understand the direction of associations between variables. For instance, chapter 4 highlights significant associations between social motivation and a range of other variables. However, it remains unclear whether social motivation influences one's social functioning and mental health or if individuals with favourable social support systems are likely to develop higher social motivation. To address such questions, longitudinal research tracking individuals over time is necessary. More specifically, in future research, it is crucial to consider that social motivation is likely a dynamic process influenced by various factors, and, in turn, influences those same factors. Researchers will therefore need to develop measures to quantify and characterise the relative contributions of these processes over time to establish a comprehensive causal model.

5.3 Conclusions

In this doctoral thesis, the first overarching aim was to understand why individual differences in mindreading persist amongst adults, despite most individuals demonstrating a conceptual understanding by late childhood. We explored the *both* the accuracy of mindreading, or 'contextually justified' as well as in terms of the underlying *processes* involved in mentalising, independently of response accuracy. Furthermore, given the limited empirical or theoretical integration between motivational accounts and socio-cognitive investigations in general, as well as specifically regarding mindreading and motivation, the second overarching aim was to examine individual differences in adult mindreading both independently and in conjunction with motivational factors. This involved investigating the *links* between social motivation and mindreading, and their respective relationships with social abilities and mental health variables.

The findings reported in Chapter 2 support the notion that response quality and quantity may be differentiable dimensions of open-ended response mindreading tasks in adults. They demonstrate valid non-shared variance between mindreading quality and quantity, which has previously remained hidden by not differentiating between the degree to which participants quantitatively engage with others' mental states and the degree to which their responses are appropriate. This research aids our understanding of why mindreading varies in adults who already understand basic mental state concepts of their own and others' thoughts, beliefs and desires and may ultimately help us to understand how an adult's propensity to effortfully engage with others' mental states (e.g., their motivation to mindread) may interact with their ability to do so appropriately. Chapter 3 built on these findings by introducing a novel approach to capture individual differences in the *structure* of adults' mindreading, independently of the ability to infer mental states accurately. The results from this chapter show, for the first time, that it is possible to effectively capture structural features of mindreading in two independent samples of adult populations, without focussing on the *accuracy* of specific content, thereby providing initial evidence that a diversion from current ways of assessing mindreading exclusively in terms of accuracy may be a fruitful supplementation to existing methods to take to help our understanding of *why* individuals differ in their mindreading. Our novel approach to coding open-ended responses to various social picture stimuli effectively captured individual variation across four distinct categories: the frequency of mental state terms (MST), mentalistic

explanations, expressed certainty, and alternatives. Confirmatory factor analyses revealed that a higher-order factor structure reliably provided the best fit for the data (based on a test and cross-validation in a separate sample). Finally, we found evidence of convergent validity between an established mindreading measure, the MASC, and our novel task, suggesting that particular 'ways of mentalising' are associated with higher accuracy. Importantly, these findings do however not suggest that for every context, mentalising benefits from higher degrees of the four categories identified here, and so future research will benefit from addressing the complex *interactions* between the context, structure, and accuracy of mentalising. Acknowledging the importance of linking individual differences in mindreading at the level of 'structure' with differences at the level of 'accuracy' will help to further inform our understanding of the cognitive underpinnings of individual differences in mindreading in individuals who already possess concepts of mental states.

Finally, in Chapter 3, we assessed the potential links between social motivation and mindreading, and what their respective relationships and interplay with social abilities and mental health were. This investigation was aimed to bridge the gap between the mostly distinct existing literatures on the correlates of individual differences mindreading and social motivation in neurotypical adults. We investigated their interplay and respective influences on various social and mental health outcomes. While performance on two methodologically distinct, but converging, mindreading tasks was positively related to social motivation, confirmatory factor analysis revealed that both constructs were distinct. Our findings highlight the pivotal role of social motivation as a robust positive predictor of perceived social support, and negative predictor of depressive symptoms as well as autistic traits among neurotypical adults. Conversely, individual differences in mindreading exhibited limited predictive power. Notably, social motivation, when considered as a mediator, did not significantly and interact with individual differences in mindreading to predict outcomes. However, our results, based on the measures employed in this study, do not rule out the potential significance of individual differences in neurotypical adults' mindreading – it is plausible that the variance in mindreading detected by the measures employed in this study could have reached a threshold of significance if more advanced outcome measures had been tested. A deeper exploration of these intricate causal relationships and interactions could facilitate the development of tailored interventions and support systems suitable for adults.

Appendix

Appendix

Chapter 2

Table A. *Bivariate correlations and descriptive statistics for the item-level data of the SFT and MASC.*

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1MASC_1_A																														
2MASC_2_A	.071																													
3MASC_13_A	.016	.122																												
4MASC_18_A	.072	.195*	.246**																											
5MASC_24_A	.078	.216*	.258**	.254**																										
6MASC_28_A	.224*	.119	.186*	.365**	.197*																									
7MASC_30_A	.155	.067	.042	.146	.064	.320**																								
8MASC_34_A	.133	.038	.178*	.125	.162	.190*	.155																							
9MASC_39_A	.103	.172	.226*	.314**	.317**	.132	.091	.293**																						
10MASC_45_A	.076	.150	.122	.331**	.220*	.207*	.262**	.117	.306**																					
11SFT_1_A	.129	.090	.186*	.094	.142	.036	.063	-.046	.122	.142																				
12SFT_2_A	.051	.082	.137	.097	.037	.103	.095	.080	.046	.200*	.163																			
13SFT_3_A	.154	.131	.075	.094	.195*	.096	.182*	.012	.001	.182*	.103	.158																		
14SFT_4_A	.192*	.018	.081	.070	-.066	.216*	.164	.027	.080	.143	.231**	.154	.232**																	
15SFT_5_A	.182*	.152	.308**	.102	.150	.234**	.080	.203*	.225*	.096	.058	.225*	.217*	.383**																
16MASC_1_M	.334**	-.045	.050	.244**	-.041	.063	.103	.235**	.201*	.130	.038	.037	.026	.240*	.174															
17MASC_2_M	.036	.287**	.255**	.188**	.134	.212*	.237**	.047	.181*	.135	.105	.221*	.145	.246**	.276**	.375**														
18MASC_13_M	.074	.217*	.243**	.149	.065	.196*	.304**	.098	.246**	.152	.257**	.189*	.118	.251**	.309**	.313**	.467**													
19MASC_18_M	.018	.210*	.322**	.721**	.361**	.288**	.310**	.242**	.358**	.416**	.111	.208*	.007	.064	.171	.275**	.371**	.334**												
20MASC_24_M	-.003	.255**	.303**	.216*	.471**	.181*	.227*	.182*	.250**	.293**	.072	.169	.083	.078	.157	.002	.373**	.293**	.387**											
21MASC_28_M	.120	.004	.085	.181*	.076	.289**	.358**	.136	.162	.278**	.032	.159	-.020	.212*	.218*	.377**	.450**	.472**	.381**	.394**										

22MASC_30_M	.162	.010	-.005	.058	.161	.340**	.436**	.218*	.144	.214*	.060	.133	.049	.172	.248**	.103	.260**	.314**	.334**	.237**	.525**									
23MASC_34_M	.178*	.047	.273**	.168	.088	.189**	.349**	.469**	.339**	.236**	.096	.135	-.003	.218*	.360**	.337**	.325**	.442**	.390**	.330**	.464**	.385**								
24MASC_39_M	.130	.153	.209*	.392**	.257**	.304**	.213*	.137	.562**	.322**	.200*	.231*	.014	.237**	.241**	.239**	.402**	.370**	.436**	.412**	.292**	.314**	.381**							
25MASC_45_M	.079	.053	.128	.205*	.131	.360**	.280**	.231**	.185*	.535**	.164	.083	.145	.228*	.155	.165	.216*	.358**	.291**	.325**	.429**	.392**	.398**	.333**						
26SFT_1_M	.125	-.002	.078	.181*	-.041	.112	.076	.026	-.021	.151	.065	.232**	-.159	.161	.190*	.031	.043	.136	.118	.083	.190*	.010	.084	.150	.047					
27SFT_2_M	.035	-.024	-.025	.187*	-.077	.076	-.018	-.022	-.034	.060	.045	.229**	-.014	.093	.056	.261**	.141	.109	.124	.043	.147	.005	.127	.098	.157	.112				
28SFT_3_M	.178*	.073	.049	.155	.055	.054	.221*	.065	.088	.170	.078	.169	.353**	.229**	.266**	.007	.293**	.196*	.109	.301**	.263**	.277**	.244**	.153	.217*	.118	.018			
29SFT_4_M	.239**	.080	.072	.084	.001	.166	.104	.170	.143	.220*	.150	.152	.217*	.687**	.479**	.224*	.193*	.143	.156	.132	.176	.185*	.317**	.169	.260**	.148	.074	.331**		
30SFT_5_M	.010	.171	.056	-.021	.088	.070	.081	.129	.197*	.016	.050	.157	.068	.052	.431**	.045	.108	.295**	.054	.207*	.196*	.225*	.247**	.158	.261**	.053	.234**	.190*	.140	
Descriptive Statistics																														
N	128	127	128	127	127	126	126	126	126	126	126	126	127	125	125	128	128	128	127	127	126	126	126	126	126	126	127	128	125	125
M (SD)	.96 (.62)	1.10 (.80)	.76 (.66)	.95 (.66)	.79 (.65)	1.13 (.78)	.98 (.65)	.75 (.65)	1.18 (.61)	.63 (.70)	1.06 (.82)	1.42 (.64)	1.08 (.70)	.85 (.72)	1.18 (.69)	1.09 (1.35)	2.20 (1.79)	2.27 (2.13)	2.02 (1.47)	2.11 (1.64)	1.61 (1.61)	1.55 (1.40)	2.14 (1.92)	2.37 (1.57)	1.22 (1.48)	.94 (1.11)	.98 (.70)	1.88 (1.10)	1.40 (1.22)	1.98 (1.22)
Skew	.02	-.19	.29	.05	.22	-.243	.023	.299	-.120	.645	-.104	-.641	-.108	.235	-.253	2.14	3.37	1.75	.65	1.11	1.92	1.58	1.14	1.17	1.31	1.24	.16	.63	.82	1.04
Kurtosis	-.35	-1.39	-.73	-.71	-.66	-1.32	-.59	-.71	-.44	-.75	-1.52	-.55	-.92	-1.03	-.91	8.54	20.21	4.48	.44	1.81	5.56	4.11	1.24	2.21	1.06	.93	-.51	.26	.24	1.80
Range	0 - 2	0 – 2	0 – 2	0 – 2	0 – 2	0 – 2	0 – 2	0 – 2	0 – 2	0 – 2	0 – 2	0 – 2	0 – 2	0 – 2	0 – 2	0 – 9	0 - 15	0 - 12	0 - 7	0 - 8	0 - 10	0 - 8	0 - 9	0 - 9	0 – 6	0 - 4	0 - 3	0 - 5	0 - 5	0 - 7

Note. *Note.* $^{**}p < .01$. $^{*}p < .05$. SFT = Silent Film Task, MASC = Movie for the Assesment of Social Cognition, A = Appropriateness, M = Mental State Terms, N = Number, M = Mean, SD = Standard Deviation, SE = Standard Error

Table B. Comparison original and new coding MASC and SFT.

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1SFT_App_ Total_Score (summed)															
2MASC_App_ Total_Score (summed)	.361**														
3SFT_MST_ Total_Score (summed)	.549**	.275**													
4MASC_MST_ Total_Score (summed)	.384**	.633**	.429**												
5 SFT_Original_ Total_Score (summed)	.210*	.285**	.297**	.349**											
6 MASC_Original_Total_ Correct ToM	-.017	.213*	-.077	.077	.223*										
7 MASC_Original Total_Excessive ToM	-.086	-.144	.123	-.089	-.064	-.641**									
8 MASC_Original_Total_ Less ToM	-.012	-.133	-.087	-.032	-.275**	-.566**	-.065								
9 MASC_Original_Total_ No ToM	.180*	-.061	.111	.012	.007	-.446**	.016	-.051							
10 MST Latent Factor (MASC & SFT combined)	.494**	.400**	.595**	.915**	.332**	-.030	-.022	.004	.092						
11 App_Latent Factor (MASC & SFT combined)	.784**	.574**	.591**	.656**	.324**	-.028	-.059	.005	.139	.745**					
12 One-factor LF (MASC & SFT combined)	.470**	.635**	.533**	.980**	.363**	.057	-.085	-.048	.065	.945**	.734**				
13 MST Latent Factor (based on SFT only)	.749**	.369**	.805**	.428**	.065	-.070	.044	-.050	.154	.566**	.753**	.525**			
14 APP Latent Factor (based on SFT only)	.840**	.379**	.728**	.423**	.113	-.068	.013	-.029	.166	.560**	.841**	.516**	.966**		
15 MST Latent Factor (based on MASC only)	.364**	.595**	.431**	.984**	.095	.063	-.093	-.033	.046	.920**	.648**	.983**	.407**	.399**	
16 APP Latent Factor (based on MASC only)	.330**	.902**	.279**	.777**	.099	.203*	-.188*	-.108	-.015	.543**	.528**	.775**	.356**	.351**	.758**

Note. **$p < .01$. *$p < .05$

Religiosity Latent Factor

The latent factor structure of the four religiosity indicators was investigated. There were moderate to strong positive correlations between each of the four indicators (see Table 1). The fit of a one-factor solution in which each of the four religiosity indicators was loaded onto one single latent factor was examined. This model provided excellent fit to the data, χ^2 (11) = 8.413, p = .676, CFI = 1.0, TLI = 1.027, RMSEA = 0 (CI = 0 - .078). All items loaded onto the single religiosity factor with standardized loadings ranging from .65 (religious upbringing) to .87 (religious affiliation), all ps < .001. When age and gender, education, verbal ability and whether participants had English as their native language was controlled for, model fit was still excellent: χ^2 (11) = 8.766, p = .643, CFI = 1.0, TLI = 1.021, RMSEA = .0 (CI = 0 = .077).

CODING SCHEME MENTAL STATE TERMS

General information: Answers should be coded in terms of their quantity (i.e., word count and number of articulated mental states) and their quality (i.e., appropriateness given the task requirements. Each participant should receive two distinct scores – a quantity and a quality score per item.

QUANTITY

Category	Rule
Total Word Count	• Record the total number of words contained within a response.
Total MST Words	• Record the total number of explicit mental state words (i.e., desire, cognition and emotion terms) contained within a response. Simply "trying to" does not count as a mental state term. References to participant's own mental states (e.g., I don't know, I think) should be recorded under "MST-own". See MST list* below for more detail.

Sample List of Mental-State Terms

Desire	**Emotion**	**Cognitive**	**Fixed Expressions**
Want	Happy	Think	"To look for"
Like (not as a	Angry	Believe	(desire, in the
preposition)	Sad	Know	context of
Love	Upset	Wonder	wanting
Dream	Excited	Expect	something)
Hope	Frightened	Pretend	
Wish	Worried	Forget	"Fed up"
Keen on	Fed Up	Remember	(emotion)
Prefer	Pleased	Imagine	
Care	Proud		

Desire Favourite Need Invite Propose Suggest Offer Rather Fancy	Scared Cross Fun Grumpy Guilty Irritated, Hassled, Nervous, Proud (un)comfortable Feel Uneasy Aversive Insecure Concerned Apprehensive	Suppose (not 'should') Guess Lie Cheat Bet Assume Idea Secret Clever Mistake Plan Realise Notice Understand Opinion Persuade Agree Convince Doubtful Indecisive	"To come up with" (cognitive) "To figure out" (cognitive) "To act out"

Appropriateness

QUALITY

General information: Decide how appropriate an answer is (regardless of the quantity of words or MST). Each answer should be scored into one of three possible ordinal categories (e.g., 0,1,2) depending on how accurately it reflects the sequence. Answers are not scored down for spelling/ grammar. If participants give several alternative responses the most appropriate alternative should be scored.

Appropriateness Scale (for the SFT and MASC):

General Guidance for Silent Film Task and Triangles Task

Category	Score	Rule
Uninterpretable	-99	• An answer that cannot be classified in terms of its appropriateness as it is not understandable (i.e., "I don't know, missing, unintelligible, containing 3 or less words)
Appropriate	2	• The answer is appropriately related to the question asked. Sufficiently detailed description that enables reconstruction of important* elements of the sequence based on the response. Includes reference to *both* interacting entities. Does not include speculation which cannot be derived from the clip.
Partially appropriate	1	• Description related to sequence, but imprecise or incomplete. Some rudimentary understanding of the reasons for characters' actions without reaching full understanding. Might only focus on one interaction partners' behaviour and include speculation which cannot be derived from the clip.

Inappropriate	0	• Misunderstanding of the question, nonsensical descriptions, wrong descriptions, focus on minor aspect of sequence, incoherent (reader unable to reconstruct clip).

Category	Score	Rule
Uninterpretable	-99	• An answer that cannot be classified in terms of its appropriateness as it is not understandable (i.e., "I don't know, missing, unintelligible, containing 3 or less words)
Appropriate	2	• The answer is appropriately related to the question asked. Sufficiently detailed description that enables reconstruction of important* elements of the sequence based on the response. Includes reference to *both* interacting entities. Does not include speculation which cannot be derived from the clip.
Partially appropriate	1	• Description related to sequence, but imprecise or incomplete. Some rudimentary understanding of the reasons for characters' actions without reaching full understanding. Might only focus on one interaction partners' behaviour and include speculation which cannot be derived from the clip.
Inappropriate	0	• Misunderstanding of the question, nonsensical descriptions, wrong descriptions, focus on minor aspect of sequence, incoherent (reader unable to reconstruct clip).

MASC_Q1 (1)

Category	Score	Examples
Appropriate **- *both (a) and(b) need to be fulfilled; (c) cannot not be fulfilled**	2	**(a) Michael visits unexpectedly and / OR compliments Sandra** **(b) Sandra's reaction to compliment (i.e., feels uncomfortable, surprised). Can be indirect via awareness that the compliment appeared inappropriate in the context of the interaction (i.e., Michael was uninvited, Sandra just asked a question but was complimented, Michael is creepy)** **(c) Does *not* mention that Sandra is pleased, thankful, or equally returns Michael's compliment** • Michael went to Sandra's house unannounced. He then started to compliment her even though Sandra had asked what he was doing there • The man has come to see the woman after not seeing each other for a while, and he comments on her changed appearance. Possibly uninvited and unwanted as she seemed uncomfortable. • Michael came into Sandra's house and started to complement her. Sandra look flustered and a bit perturbed but tried to be nice as Michael was a bit pushy with his talking. • Michael went to Sandra's house unannounced. He then started to compliment her even though Sandra had asked what he was doing there
Partially appropriate **-Either (a) or (b) need to be fulfilled;**	1	**(a) Michael visits unexpectedly and / OR compliments Sandra** **(b) Sandra's reaction to compliment (i.e., feels uncomfortable, surprised). Can be indirect via awareness that the compliment appeared inappropriate in the context of the interaction (i.e., Michael was uninvited, Sandra just asked a question but was complimented, Michael is creepy)**

(c) cannot be fulfilled		**(c) Does *not* mention that Sandra is pleased, thankful, or equally returns Michael's compliment** • The women unexpectedly answers the door to Michael who compliments her hair. • Michael entered the house where Sandra lives. They greet each other rather uncomfortably and exchange pleasantries. Michael asks Sandra if she got a new haircut which she confirms. • Michael is complimenting Sandra. • he arrives at her house unexpectedly and compliments her new hairstyle
Inappropriate **-Neither (a) not (b) are fulfilled; (c) can (but does not have to) be fulfilled**	0	**(a) Michael visits unexpectedly and / OR compliments Sandra** **(b) Sandra's reaction to compliment (i.e., feels uncomfortable, surprised). Can be indirect via awareness that the compliment appeared inappropriate in the context of the interaction (i.e., Michael was uninvited, Sandra just asked a question but was complimented, Michael is creepy)** **(c) Does *not* mention that Sandra is pleased, thankful, or equally returns Michael's compliment** • Sandras old friend has come to her house after a long time and they have reunited. • it is mike's first visit to sandra's house • Susan perhaps surprises Michael. he then notices a change in her hairstyle and compliments her for which she thanks him. • a man showed up at an old friends house • Sandra is seeing Michael after a long while at her place • The male is visiting his friend. • One of the girls opens the door and there is a man stood there. They ask each other how they are and he compliments her hair. • He comes to meet her at her house • Micheal turned up to Sandras house I'm not sure if she knew that he was coming and then they begin to complement each other as if they haven't seen each other in a while

MASC_Q2 (2)

Category	Score	Examples
Appropriate **- *both (a) and(b) need to be fulfilled**	2	• **Sandra wants to meet Michal together with other people (i.e., feels uncomfortable; unhappy about Michael cancelling group-meeting)** • **Michael has contrary wishes, he wants to see her alone (e.g., persuading her). Participant needs to be explicit that he wants to meet Sandra alone – it is not enough to say he wants to play tennis instead of having dinner.** • The man is trying to persuade Sandra to go and play tennis with him but she wants to have dinner with other people instead.

		• Micheal and Sandra talked about a past meeting at which Sandra met Cliff. She wanted to hang out with him again however, Micheal suggested that they should go out alone. • the woman thought they were meeting up as a group not just the two of them and she feels uncomfortable
Partially appropriate **-Either (a) or (b) need to be fulfilled**	1	**(a) Sandra wants to meet Michal together with other people (i.e., feels uncomfortable; unhappy about Michael cancelling group-meeting)** **(b) Michael has contrary wishes, he wants to see her alone (e.g., persuading her). Participant needs to be explicit that he wants to meet Sandra alone – it is not enough to say he wants to play tennis instead of having dinner.** • They are reminiscing about a past time and Sandra mentions she met his friend Cliff there. She suggested they were meant to have dinner together and looks a bit annoyed when he said they'll do that another time. • they are talking about a time they went out together, but she mentions his friend Chris who she may like more • Sandra and Michael talk about a trip they took together where she met his friend cliff. She proposes they all make dinner at her house but Michael wants to go for a game of tennis instead.
Inappropriate **-Neither (a) not (b) are fulfilled**	0	**(a) Sandra wants to meet Michal together with other people (i.e., feels uncomfortable; unhappy about Michael cancelling group-meeting)** **(b) Michael has contrary wishes, he wants to see her alone (e.g., persuading her). Participant needs to be explicit that he wants to meet Sandra alone – it is not enough to say he wants to play tennis instead of having dinner.** • Sandra tells Michael that she remembers Cliff • They continued chatting, now about Cliff. • The man and woman are remembering another night they spent together with company fondly. The woman remembers an agreement to have dinner at her house, the man inquires about a game of tennis he wants to play with her instead of dinner • she is rejecting him, because she likes cliff. • They are discussing the plans they had for that night. they are also talking about his friend Clive

MASC_Q13 (3)

Category	Score	Examples
Appropriate **- *both (a)**	2	**(a) Cliff reveals that reason for him being sad is break-up from long-term relationship and / or Cliff's current reaction (upset, sad)**

and(b) need to be fulfilled		**(b) Sandra's intention (e.g., flirt with Cliff, fancies him) AND / OR Sandra's reaction to Cliff's break-up (i.e., disappointed, sad)** • Sandra is trying to flirt with Cliff but Cliff is unaware of it as he is talking about how he is still connected to hi sex girlfriend • After cracking a joke, Sandra and Cliff seem to be more comfortable around each other. She tells him he's in a good mood unlike the time they first met at the pub. Cliff opens up and tells Sandra that he had just broken up with his girlfriend of 4 years at the time and sadly mentions that even though they were separated, he still feels so connected.
Partially appropriate **-Either (a) or (b) need to be fulfilled**	1	**(a) Cliff reveals that reason for him being sad is break-up from long-term relationship and / or Cliff's current reaction (upset, sad)** **(b) Sandra's intention (e.g., flirt with Cliff, fancies him) AND / OR Sandra's reaction to Cliff's break-up (i.e., disappointed, sad)** • Sandra is trying to get to know Cliff because she fancies him. • cliff was talking to sandra about his previous relationship and how he still feels connected to his ex • sandra arranged with michael and betty that they would have dinner and invite cliff so she can flirt with him
Inappropriate **-Neither (a) not (b) are fulfilled**	0	**(a) Cliff reveals that reason for him being sad is break-up from long-term relationship and / or Cliff's current reaction (upset, sad)** **(b) Sandra's intention (e.g., flirt with Cliff, fancies him) AND / OR Sandra's reaction to Cliff's break-up (i.e., disappointed, sad)** • The lady and Cliff are talking about his romantic relationships. • They talked about themselves • Michael fancies the girl but she doesn't like him back, she like Cliff. Cliff is still thinking about his ex girlfriend and the other is being dragged into this. They're all about to have dinner with each other. • Cliff tells Sandra about his break up • Sandra invites Michael, Cliff and Betty to her house for a dinner on Saturday night. At first, Cliff and Betty are not too sure whether they want to go but they are persuaded by their friends; Sandra and Michael. Cliff is then the first one to come to Sandra's house and then spends time talking with Sandra and getting to know her. • The man is talking about his ex girlfriend • Mike wanted to invite his friend cliff and Sandra Betty. Sandra doesn't like Mike, finds him overly flirty but likes Cliff so wanted to meet up with Mike so he'd invite Cliff. Mike really likes Sandra and invites cliff because she's inviting Betty, he believes it'll be a double date where Cliff can get to know Betty.

MASC_Q18 (4)

Category	Score	Examples
Appropriate **- *both (a) and(b) need to be fulfilled**	2	**(a) Michal's intention of trying to impress Sandra with story about helping lady** **(b) Sandra's reaction (i.e., she thinks he's bragging, doesn't like it)** • Sandra seems uncomfortably by Michael's presence but Michael seems unaware. He is recalling a story from his day and seems to be

		bragging/trying to impress Sandra with how he was trying to be a good person and help a stranger • Sandra and Cliff were talking about his holiday and dogs and cats. They got interrupted by Michael arriving and giving her flowers, he then told a story very loudly about him helping a lady that got hit by a bike Michael is trying to flirt with Sandra. Sandra is not interested but is trying to be polite. Michael is showing off about being a hero to impress her. • Cliff is talking to Sandra about his holiday in Sweden and how his dog ran away. Sandra feels sympathetic towards him. Michael arrives and gifts flowers to a disinterested Sandra. He tells her about a story where he helped a woman in order to impress her.
Partially appropriate **-Either (a) or (b) need to be fulfilled**	1	**(a) Michal's intention of trying to impress Sandra with story about helping lady** **(b) Sandra's reaction (i.e., she thinks he's bragging, doesn't like it)** • he is trying to impress sandra • Michael is talking about what happened during the day to impress Sandra- he makes out as if he was a hero when the car hit the woman. • The man is bragging about helping a lady who was in an accident • Cliff and Sandra talk about her ex boyfriend, Sweden and pets. Then comes Micheal and brings Sandra flowers. He touches her and tells her a story about something that happened to him that day. However, Sandra did not seem pleased. • Michael slaps Sandra on the knee and starts talking about how he witnessed a motorcyclist hit a woman and that he helped her and yelled at the driver. Cliff looks uncomfortable and Sandra looks disinterested.
Inappropriate **-Neither (a) not (b) are fulfilled**	0	**(a) Michal's intention of trying to impress Sandra with story about helping lady** **(b) Sandra's reaction (i.e., she thinks he's bragging, doesn't like it)** • Michael tells everyone that he saw a cyclist hit a woman. • Mike arrives and brings Sandra flowers. He talks about helping a lady who had been knocked over. • Mike liked telling a story about an event that happened to him before he arrived at Sandra's.

MASC_Q24 (5)

Category	Score	Examples
Appropriate **- *both (a) and(b) need to be fulfilled**	2	**(a) Michal's intention of trying to impress Sandra / show off by opening champagne bottle** **(b) Sandra's AND / OR Betty's reaction to Michael's attempt to impress them (e.g., Sandra thinks he's showing off, Betty doesn't like his behaviour)** • All 4 guests sit at the dinner table, Sandra seems uncomfortable with Michael's behaviour and attitude. Michael seems very braggy and likes to be the center of attention, taking control of the situation in an attempt to impress Sandra. Betty also seems as though she does not enjoy Michael's presence or behaviour • they are all sitting at the table having the champagne and sandra is being put off by micheal being overly confident

		• Sandra is laughing at Michael as he's trying to impress her and Betty is like 'lol this guy'.
Partially appropriate **-Either (a) or (b) need to be fulfilled**	1	**(a)Michal's intention of trying to impress Sandra / show off by opening champagne bottle** **(b)Sandra's AND / OR Betty's reaction to Michael's attempt to impress them (e.g., Sandra thinks he's showing off, Betty doesn't like his behaviour)** • Michael wants to impress Sarah again • Micheal is trying to impress sandra again by attempting to open the bottle of champaigne • Betty is acknowledging how annoying Michael is. • the men followed them into the kitchen, Michael tries to show off again
Inappropriate **-Neither (a) not (b) are fulfilled**	0	**(a)Michal's intention of trying to impress Sandra / show off by opening champagne bottle** **(b) Sandra's AND / OR Betty's reaction to Michael's attempt to impress them (e.g., Sandra thinks he's showing off, Betty doesn't like his behaviour)** • They were all sitting down to drink wine, and Michael offered to open an pour it. • Michael wanted to open the champagne for them. • They all moved to the kitchen

MASC_Q28 (6)

Category	Score	Examples
Appropriate **- *both (a) and(b) need to be fulfilled**	2	**(a) Michael being sexist by suggesting cooking is a woman's job** **(b) Bettys reaction to Michael's remark (e.g., she gets annoyed or does not like it). This can be implicit by simply stressing that Betty suggests Michael cuts the onions, an uncomfortable task.** • michael made a rude comment so betty made one back The women go to make the dinner and when Cliff suggests Michael should go to help, he makes a sexist remark that it's a woman's job. Betty, who took offense suggested he chops onions as an insult. • Michael makes a misogynistic comment and it annoys betty • mike makes a sexist remark about cooking being for women and so betty tells him he should cut the onions • Michael expresses his dislike for cooking. He believes it is 'ladies work'. Betty offers that he should cut onions. She is being sarcastic in her approach towards him. Michael is rude. • Michael is bitter and misogynistic. Betty is not impressed and bites back.
Partially appropriate **-Either (a) or (b) need to be fulfilled**	1	**(a) Michael being sexist by suggesting cooking is a woman's job** **(b) Bettys reaction to Michael's remark (e.g., she gets annoyed or does not like it). This can be implicit by simply stressing that Betty suggests Michael cuts the onions, an uncomfortable task.** • Sandra and Betty start making dinner, with no sardines, Micheal is misogynistic when cliff suggests he helps with the cooking • They share champagne and start making dinner. Michael makes a obviously sexist comment and says that cooking is meant for women, he is reluctant to help

		• They are making dinner by Michael doesn't want to help he thinks it is women's work • Michael made a sexist comment
Inappropriate **-Neither (a) not (b) are fulfilled**	0	**(a) Michael being sexist by suggesting cooking is a woman's job** **(b) Bettys reaction to Michael's remark (e.g., she gets annoyed or does not like it). This can be implicit by simply stressing that Betty suggests Michael cuts the onions, an uncomfortable task.** • Sandra has an obvious attraction for Cliff and Micheal is starting to get jealous and offended • They go back to dinner • Sandra and Betty begin the cooking and Cliff volunteers himself and Michael to help. • They start to make dinner (with no sardines • Michael tries too hard to impress Sandra but then he realises that Sandra likes Cliff.

MASC_Q30 (7)

Category	Score	Examples
Appropriate **- *both (a) and(b) need to be fulfilled**	2	**(a) Michael makes a nasty joke aimed at Betty (it is important to mention that the joke is aimed at Betty, not enough to say that Michael is generally annoying)** **(b)Betty's reaction (e.g., annoyed)** • They continue to prepare dinner and Cliff is willing to help in any way. Instead of helping, Michael makes rude comments towards Betty and her eating habits. Betty is visibly offended/shocked by his statement • cliff is being helpful in the kitchen with the wine and the cutting but micheal continues to be rude and betty is getting more annoyed
Partially appropriate **-Either (a) or (b) need to be fulfilled**	1	**(a) Michael makes a nasty joke aimed at Betty (it is important to mention that the joke is aimed at Betty, not enough to say that Michael is generally annoying)** **(b) Betty's reaction (e.g., annoyed)** • cliff wants to help more with the cooking- michael asks what else goes in the sauce- when betty says 2 cups of cream he says that she would want 5 cups in it to insult her • Michael is trying to be funny by being rude about Bettys weight • Michael passive aggressively called Betty fat • Michael gets ignored and proceeds to make a nasty comment about betty
Inappropriate **-Neither (a) not (b) are fulfilled**	0	**(a) Michael makes a nasty joke aimed at Betty (it is important to mention that the joke is aimed at Betty, not enough to say that Michael is generally annoying)** **(b) Betty's reaction (e.g., annoyed)** • They finish cutting the vegetables, Cliff asks about the wine, Michael asks about the sauce

		• Micheal made a comment about hoe betty would use more cream than needed whilst cliff opens a bottle of wine • The girls continue to make food and Micheal continues to be annoying • Micheal is starting to be rude

MASC_Q35 (8)

Category	Score	Examples
Appropriate **- *both (a) and(b) need to be fulfilled**	2	**(a) Michael justifying his nasty jokes / comments (e.g., jokes about it)** **(b) Betty's reaction (either that she's questioning, not happy or starting to like him – all are logical possibilities)** • Sandra and Cliff are getting along and having pleasant conversation. As a good friend, Betty doesn't want to interrupt the pair and so makes polite conversation with Michael although she doesn't appear to be a big fan. Michael in response jokes about his rude nature, but Betty doesn't seem amused • Michael is picking up on Betty's dislike of him and he is testing her to see what she actually thinks of him and he justifies and qualifies his behaviour in a moral self-righteous way by claiming to be honest, which he actually isn't • micheal starts to feel left out from the conversation. betty notices this and tries to make a conversation about his and cliff's friendship. wherein micheal justifies his impolite behaviour at which she giggles
Partially appropriate **-Either (a) or (b) need to be fulfilled**	1	**(a) Michael justifying his nasty jokes / comments (e.g., jokes about it)** **(b) Betty's reaction (either that she's questioning, not happy or starting to like him – all are logical possibilities)** • Michael explains why he acts the way he does. • Betty and Michael are actually getting along. • Betty has decided to be more open to Michael and give him a chance
Inappropriate **-Neither (a) not (b) are fulfilled**	0	**(a) Michael justifying his nasty jokes / comments (e.g., jokes about it)** **(b) Betty's reaction (either that she's questioning, not happy or starting to like him – all are logical possibilities)** • Betty asks Michael how long he's know Cliff • Betty and Micheal are chatting • Betty starts a conversation with Michael whilst Sandra talks to Cliff • betty and billy are talking and she is trying to get to know more about him. • Cliff and Sandra are talking. Betty asks Micheal about his friendship with cliff. • sandra and cliff are involved in conversation so betty tries to start a conversation with Michael • They are now having dinner and talking

MASC_Q39 (9)

Category	Score	Examples
Appropriate **- *both (a), (b) and (c) need to be fulfilled**	2	**(a) Sandra annoyed at Betty** **(b) Michael is bragging** **(c) Sandra is annoyed by Michael** • sandra is annoyed at betty because she keeps forgetting the rules and she is annoyed at Michael for trying to show off • Sandra is unimpressed that Betty asked what colour she was again. They both took very average shots. Michael says its time for some action and Sandra rolls her eyes and makes a sarcastic comment • Sandra looks annoyed when Betty asks which colour she is. Betty then hits the chip backwards. Sandra then takes a turn. Its then Michael turn, where he says about how professional he is at the game. Sandra looks embarrassed and annoyed- saying that she couldn't wait to see him shoot the shot sarcastically. • Michael is showing off and Sandra is tired of Betty asking the colour and Michael's behaviour
Partially appropriate **- (a), or (b) or (c) need to be fulfilled**	1	**(a) Sandra annoyed at Betty** **(b) Michael is bragging** **(c) Sandra is annoyed by Michael** • They were all playing a board game and Sandra got irritated at Betty's confusion and Michael not playing properly. • Sandra is annoyed that Betty is being stupid and also hates Michael. • sandra is starting to get irritated by betty • they are playing a game. sandra seems annoyed by michaels arrogance • Cliff is trying to boast saying he is the best at the game and Sandra isn't to impressed.
Inappropriate **-Neither (a), (b) or (c) can be fulfilled**	0	**(a) Sandra annoyed at Betty** **(b) Michael is bragging** **(c) Sandra is annoyed by Michael** • They are playing the board game • They have started playing the game. Both Betty and Sandra have their go but they do not have the best shot. Michael then says how he is going to show them how a professional is going to play the game • They finish eating and Sandra suggests a board game to play

MASC_Q46 (10)

Category	Score	Examples
Appropriate **- *both (a) and(b) need to be fulfilled**	2	**(a) Awareness of EITHER Betty's intention to give Sandra time alone with Cliff AND / OR Sandra's desire to spend time alone with Cliff (can be indirect by saying that Sandra "gets" to spend time with Cliff)** **(b) Mentioning of Michael's feelings such as Michael being disappointed /reluctant to leave / accepting of the fact that it didn't work out with Sandra (can be indirect - e.g., the goodbye was akward, Sandra prefers him to Michael)**

		• Michael is reluctant to leave as he is interested in Sandra. Sandra is relieved that Michael is leaving and that she gets to spend time with Cliff • Betty wanted to be alone with Michael. It was an awkward goodbye between Michael and Sandra. • Betty makes plans to leave with Michael to leave Sandra alone with Cliff since she knows that Sandra wants to get to know Cliff better and prefers him to Michael • Betty suggests her and Michael leave so that Sandra and Cliff can be alone together. Michael says affectionately goodbye to Sandra but she does not like him and is happy for him to leave. • Sandra was happy that Michael was leaving and she is thinking that hopefully she won't have to see him again, she is happy that she's staying alone with cliff because she likes him
Partially appropriate **-Either (a) or (b) need to be fulfilled**	1	**(a) Awareness of EITHER Betty's intention to give Sandra time alone with Cliff AND / OR Sandra's desire to spend time alone with Cliff (can be indirect by saying that Sandra "gets" to spend time with Cliff)** **(b) Mentioning of Michael's feelings such as Michael being disappointed /reluctant to leave / accepting of the fact that it didn't work out with Sandra (can be indirect - e.g., the goodbye was akward, Sandra prefers him to Michael)** • Michael has accepted that him and sandra will probably not date • Betty and Michael are leaving Sandra and Cliff, Michael seems slightly disappointed but accepting. • micheal is going off with betty and sandra is glad he is leaving • Betty is leaving with Micheal to give Sandra alone time with Cliff • Sandra is pleased Betty and Michael are leaving so she can be alone with Cliff • betty wants to take michael out so sandra is left alone with sliff • everyone is leaving the house but cliff is staying. betty did this so cliff and sandra have more time to spend with one another • michael realised sandra doesnt like him • Betty invited Michael out so that Sandra and Cliffe could spend time together. Michael and Betty then leave
Inappropriate **-Neither (a) not (b) are fulfilled**	0	**(a) Awareness of EITHER Betty's intention to give Sandra time alone with Cliff AND / OR Sandra's desire to spend time alone with Cliff (can be indirect by saying that Sandra "gets" to spend time with Cliff)** **(b) Mentioning of Michael's feelings such as Michael being disappointed /reluctant to leave / accepting of the fact that it didn't work out with Sandra (can be indirect - e.g., the goodbye was akward, Sandra prefers him to Michael)** • Betty finished the game and invited Michael for a night cap • saying goodbye after an evening of catching up • betty and michael are leaving • they are leaving • Betty is off for another drink with Micheal. Micheal likes Betty instead of Sandra now..

		• Michael and Betty are going out to get a drink and Cliff is staying with Sandra. The pairs say goodnight to each other and Michael and Betty leave. • Sandra says goodbye to Betty and Michael, who are leaving together. • Betty finished the game and invited Michael for a night cap • Micheal and Betty are leaving together. • Betty and mICHALE left

Appropriateness

QUALITY

General information: Decide how appropriate an answer is (regardless of the quantity of words or MST). Each answer should be scored into one of three possible ordinal categories (e.g., 0,1,2) depending on how accurately it reflects the sequence. Answers are not scored down for spelling/ grammar. If participants give several alternative responses the most appropriate alternative should be scored.

Appropriateness Scale (for the SFT and MASC):

General Guidance for Silent Film Task and Triangles Task

Category	Score	Rule
Uninterpretable	-99	• An answer that cannot be classified in terms of its appropriateness as it is not understandable (i.e., "I don't know, missing, unintelligible, containing 3 or less words)
Appropriate	2	• The answer is appropriately related to the question asked. Sufficiently detailed description that enables reconstruction of important* elements of the sequence based on the response. Includes reference to *both* interacting entities. Does not include speculation which cannot be derived from the clip.
Partially appropriate	1	• Description related to sequence, but imprecise or incomplete. Some rudimentary understanding of the reasons for characters' actions without reaching full understanding. Might only focus on one interaction partners' behaviour and include speculation which cannot be derived from the clip.
Inappropriate	0	• Misunderstanding of the question, nonsensical descriptions, wrong descriptions, focus on minor aspect of sequence, incoherent (reader unable to reconstruct clip).

Clips 1 – 5: *"What do you think happened during this clip?"*

CLIP 1: Men hiding

Category	Score	Examples
Appropriate **- *both (a) and(b) need to be fulfilled**	2	**(a) Recognition that the men try to avoid the woman in *relation* to the notice they have received** **(b) Understanding of the mens' intention behind hiding (e.g., they don't want to be found, they don't have to face her etc. so she can't ask for the rent.). This can**

		be indirect. • so I think they looked at the notice from the landlady saying the rent for two weeks was due or it was a final notice and they heard her coming and they hid under the coats so she didn't find them so they didn't have to pay the rent
Partially appropriate **-Either (a) or (b) need to be fulfilled**	1	**(a) Recognition that the men try to avoid the woman in *relation* to the notice they have received** **(b) Insight into the men's intention behind hiding (e.g., they don't want to be found, they don't have to face her etc. so she can't ask for the rent.)** • Two men were somewhere where they shouldn't be or didn't want to be seen, so hid from a maid when she came to find them.
Inappropriate **-Neither (a) not (b) are fulfilled**	0	**(a) Recognition that the men try to avoid the woman in *relation* to the notice they have received** **(b) Insight into the men's intention behind hiding (e.g., they don't want to be found, they don't have to face her etc. so she can't ask for the rent.)** • they have somebody caught and that was on the list was rent so it seems he wanted to borrow it or yeah …. I don't have any idea why the lady she hears some noise and she knocked on the door and they were hiding. No idea why they are hiding when the lady comes in. They expect that she comes in, of course, because they are hiding but I have no idea why or what happened there

CLIP 2: Harold Van

Category	Score	Example
Appropriate **- Both (a) and(b) need to be fulfilled**	2	**(a) Description of both Harold's (sits in the van) and the driver's (drives away) behaviour.** **(b) Recognition that the driver did not know (e.g., pay attention, realise) that Harold was in the van. This can be implicit via stressing that the driver was deaf in relation to him driving away.** • The driver of the clean towels van was hard of hearing, and hence didn't realise or hear that someone had perched themselves on the back of the van to read/write comfortably, leading to the driver driving off with a man in the back of the van
Partially appropriate **- Either (a) OR (b) need to be fulfilled**	1	**(a) Description of both Harold's (sits in the van) and the driver's (drives away) behaviour** **(b) Recognition that the driver did not know (e.g., pay attention, realise) that Harold was in the van. This can be implicit via stressing that the driver was deaf in relation to him driving away** • He needed somewhere to write so he got into the truck and the truck driver drove away
Inappropriate **- Neither (a) OR (b) need to be fulfilled**	0	**(a) Description of both Harold's (sits in the van) and the driver's (drives away) behaviour** **(b) Recognition that the driver did not know (e.g., pay attention, realise) that**

		Harold was in the van. This can be implicit via stressing that the driver was deaf in relation to him driving away • The chap was delivering some clean towels to don't know, a hotel or something or other, Harold for some reason had a clipboard at the back marking up towels, counting up towels…but he's obviously not to do with the towel company or maybe actually he might be the assistant! But then the other guy wouldn't drive off without his assistant. So yeah something to do with towel delivery really

CLIP 3: Mannequin – Deliveryman

	Score	Example
Appropriate **-Both (a) and(b) need to be fulfilled**	2	**(a) Explicit reference to Harold pretending to be a mannequin in order to a accomplish goal (e.g., change time on the clock, or non-specified reasons). Can be implicit if participant mentions pretending + act of changing clock within the same sentence.** **(b) Understanding that Harold is being discovered by deliveryman plus mention of deliveryman's reaction/ emotion (e.g., scared or realises Harold was a real person).** • Harold was pretending to be a mannequin in order to gain access to that building for unknown reasons, except he sneezes leading to the man who is transporting him to realise he's not a mannequin and panic.
Partially appropriate **- Either (a) OR (b) need to be fulfilled**	1	**(a) Explicit reference to Harold pretending to be a mannequin in order to accomplish a goal (e.g., change time on the clock, or non-specified reasons). Can be implicit if participant mentions pretending + act of changing clock within the same sentence.** **(b) Understanding that Harold is being discovered by deliveryman (+ mention of deliveryman's reaction/ emotion (e.g., scared or realises Harold was a real person).** • Not quite sure. Was he putting the clock back to punch into his work or something? I couldn't quite follow it. And so he pretended to be a model to get into the building
Inappropriate **- Neither (a) OR (b) need to be fulfilled**	0	**(a) Explicit reference to Harold pretending to be a mannequin in order to accomplish a goal (e.g., change time on the clock, or non-specified reasons). Can be implicit if participant mentions pretending + act of changing clock within the same sentence.** **(b) Understanding that Harold is being discovered by deliveryman (+ mention of deliveryman's reaction/ emotion (e.g., scared or realises Harold was a real person).** • I think it was a mannequin he was coming in with and then I don't know what really happened

CLIP 4: Harold Cat

Category	Score	Example
Appropriate	2	**(a) Recognition that Harold assumes the cat/ fur is the woman's scarf (e.g., picks it up by mistake, doesn't know etc)** **(b) Description of Harold's / Woman's reaction when finding out it was a cat**

-Noth (a) and(b) need to be fulfilled		• In a crowd shop, the salesman is talking to a woman. The woman's cape suddenly drops on the flour. The salesman helps to take the cape. However, he sees the cat's tail as the cape and grab the cat to the woman and makes her a big shock. The woman put the cat down and then take the cape by herself.
Partially appropriate **-Either (a) or (b) need to be fulfilled**	1	**(a)Recognition that Harold assumes the cat/ fur is the woman's scarf (e.g., picks it up by mistake, doesn't know etc)** **(b)Description of Harold's / Woman's reaction when finding out it was a cat** • Harold accidentally knocked the woman's scarf off which fell on the floor. He attempted to pick it up but unknowingly handed the woman a similar looking cat instead. She dropped the cat and put on her scarf.
Inappropriate **-Neither (a) or (b) need to be fulfilled**	0	**(a) Recognition that Harold assumes the cat/ fur is the woman's scarf (e.g., picks it up by mistake, doesn't know etc)** **(b) Description of Harold's / Woman's reaction when finding out it was a cat** • maybe they are clothes lying there, or maybe skin - animal skin • A woman got handed a cat to wear • There was a sale and the man handed her a cat.

CLIP 5: Fan Mildred

Category	Score	Example
Appropriate **-Both (a) and (b) need to be fulfilled**	2	**(a) Harold's & / Mildred's actions (e.g., pretending to faint) in *relation* to the other people coming in** **(b) Either the reaction of the incoming people (e.g., concerned) OR insight into what the men coming in would mean (e.g., trouble).** • A man and a women were talking when another man walks in. Harold becomes scared/ shocked and tells the lady to close her eyes and sits down. He then starts fanning her and the other man looks shocked and worried about the woman. Harold then tells the man to be quiet
Partially appropriate **-Either (a) or (b) need to be fulfilled**	1	**(a) Harold's & / Mildred's actions (e.g., pretending to faint) in *relation* to the other people coming in** **(b) Either the reaction of the incoming people (e.g., concerned) OR insight into what the men coming in would mean (e.g., trouble).** • It looked like he was having some sort of rendezvous with her and she wasn't meant to be there so they're pretending that she has passed out
Inappropriate **- Neither (a), (b) or need to be fulfilled**	0	**(a) Harold's & / Mildred's actions (e.g., pretending to faint) in *relation* to the other people coming in** **(b) Either the reaction of the incoming people (e.g., concerned) OR insight into what the men coming in would mean (e.g., trouble).** • I'm not quite sure, actually! It looks like … I'm not sure if she actually fainted or pretended to faint in order to attract attention • The man was not supposed to be there with the lady so he had to pretend that he was helping her as she was not feeling well.

2. Original SFT Coding Scheme (for clip-specific questions)

Silent Film Task (Devine & Hughes, 2013)
General Guidance

Category	Score	Rule
Uninterpretable	-99	• If the candidate does not provide sufficient information in their answer, rate the answer as 'uninterpretable' or missing.
Fail	0	• Answers that miss the point or are factually incorrect. Answers that describe the actions of the clip without evidence of mentalizing. • Inappropriate mentalizing (i.e., answers that contain an explicit mental state but is wrong with respect to the clip). This includes reality errors. • Over-interpretation of the clip (e.g., attribution of intentions/mental states not evident in the content of the clip).
Partial	1	• Answers that are technically correct but do not demonstrate full understanding of mental states (e.g., answers that imply mental states but do not explicitly use mental-state terms). • Answers that refer to motivations or low-level mental states (e.g., desires rather than cognitions).
Pass	2	• An explicit mental-state attribution, appropriate to the context of the clip.

If there is more than one answer, assign the higher score. Do not penalize for poor spelling or grammar.

1. Why did the men hide?

		Examples
Fail	0	**Description/No Mentalizing:** The woman is after them. They're not meant to be in there. The woman was coming. They did something wrong. They are doing something suspicious. They owe something. They cannot pay. They need to pay. To hide from the woman (repeats question). **Inappropriate mentalizing/over-interpretation**: The men were frightened of the woman; They didn't like the lady; They might be keeping a secret; They didn't want anybody else in the room; They were stealing or spying; They've sneaked into the room. The woman will scold them.
Partial	1	**Implicit mentalizing:** They are avoiding the woman for some reason. They do not want to pay. The woman is mad/angry at them. **Under-mentalizing:** The men didn't want to get caught. The men didn't want to pay the woman. The men didn't want the woman to see them. They didn't want the woman to see them.
Pass	2	**Explicit mentalizing:** They wanted the woman to think/believe they were not in/at home. They didn't want the woman to know they were home/what they were doing. They didn't want the woman to find them (find/found are considered cognitive terms). So they do not get found by the woman.

2. What does the woman think?

		Examples
Fail	0	**Description/No Mentalizing:** She is looking for something. She is thinking (repeats question). She heard a noise. What are they doing? Terrible men! **Inappropriate mentalizing/over-interpretation**: There is someone in the house. She thinks they hid. She thinks they ran away. Who is in here? She thinks there are thieves in the room. She is worried. She is going to punish them for not giving money. The men stole something.
Partial	1	**Implicit mentalizing:** She thinks she is going crazy. She thought they were at home (facts about what she thought before coming in). She thought she heard something (facts about before she entered the room). She is confused. "What's going on?" (confusion). **Under-mentalizing:** She wants to see the men. She is angry (emotion). She wants to tell them off. She is surprised.

Pass	2	**Explicit mentalizing:** She thinks they are not at home. **Adopts voice of woman:** "Where are they?"; "They're not here". "Why is there no one here?"

3. Why did the driver lock Harold in the van?

		Examples
Fail	0	**Description/No Mentalizing:** The man is deaf/hard of hearing. He was reading on the van. He wasn't supposed to be there. The towels will fall out. He was in a rush. **Inappropriate mentalizing/over-interpretation**: The man told him to. He kidnapped him. He wanted to punish Harold.
Partial	1	**Implicit mentalizing:** He did not see/hear him. **Under-mentalizing:** He wanted to/had to continue on his rounds. He doesn't want the towels to fall out.
Pass	2	**Explicit mentalizing:** The driver didn't know Harold was in the van. He didn't mean/intend to. He did not notice Harold was in the van. It was an accident. He wasn't paying attention. He was careless

4. What is the deliveryman feeling and why?

		Examples
Fail	0	**Description/No Mentalizing:** He is dressed in women's clothes. The model moved. He is tired from carrying Harold. **Inappropriate mentalizing/over-interpretation**: He is bored as he always does the same work. He is angry because he fought with Harold. ~~He didn't know he was there~~. He has no idea. The delivery man is angry because he was tricked (suggests he knows he was tricked).
Partial	1	**Under-mentalizing:** Identifies correct emotion OR references deliveryman's lack of knowledge/being deceived. Scared because it was alive (no reference to belief/thought). Shocked because he thought it was a woman (incorrect belief). Scared because he never saw a mannequin moving (no reference to belief/thought). Shocked because it didn't look like a real man (correct emotion, no explicit mention of belief). No emotion but adopts voice of man ("Why is it moving!").
Pass	2	**Explicit mentalizing:** Correct emotion (scared/shocked/frightened/surprised/confused/anxious/worried) and reference to deliveryman's lack of knowledge/being deceived (e.g., he thought it was a mannequin/dummy but it was a man; he thought it came to life, he didn't know the man was there).

5. Why did Harold pick up the cat?

		Examples
Fail	0	**Description/No Mentalizing:** The lady dropped her cat. Cats are not allowed in the shop. He was rushing. He picked up the cat instead of the scarf. **Inappropriate mentalizing/over-interpretation**: He wanted to give her back her cat. He wanted to scare/trick/distract her. He thought it was a cat. He wanted someone to buy it. He wanted to keep the scarf for himself. He hates cats. He did it on purpose.
Partial	1	**Implicit mentalizing:** It was beside her scarf. It looked like the scarf. It was the same colour. **Under-mentalizing:** He wanted to give her back her scarf. He didn't understand.
Pass	2	**Explicit mentalizing:** He thought it was the woman's scarf. He didn't realise/know it was a cat. He was confused. He mistakenly picked up the cat. He thought it was a scarf.

6. Why does Harold fan Mildred?

		Examples
Fail	0	**Description/No Mentalizing:** Mildred was tired/hot. The room was too hot. Mildred fainted. To cool Mildred down. He worked for Mildred/It was his job. The man is coming. **Inappropriate mentalizing/over-interpretation**: To impress the man. Mildred wanted to rest. Mildred is pretending to faint/be hot. The boss/man does not want him to be with Mildred. Harold is trying to get attention. Harold wasn't working so Mildred pretended to be sick. Mildred didn't want the man to know they were talking.

Partial	1	**Implicit mentalizing:** Harold was surprised/shocked to see the man come in. To hide that he likes/fancies Mildred. He loves/likes her. To make it look like she has fainted/he is looking after her (implies attempt to alter man's perception). **Under-mentalizing:** He doesn't want his boss to catch him/get in trouble. He doesn't want the man to see them together. Harold doesn't want to get fired.
Pass	2	**Explicit mentalizing:** Harold wants the man to think Mildred fainted. Harold is pretending/acting she fainted/is hot/asleep. Harold is trying to trick the man. So that the man doesn't know what they are doing/that they are together. To trick the man coming in. They didn't want the man to notice.

Chapter 3

Dimensional Mindreading Task

Table 1.	
	Definition
MST	The total number of mental state terms used in each response (Count number)
Mentalistic explanation	If there is an explanation present, whether the explanation is related to a mental state will be coded (1 = yes; 0 = no)
Expressed Certainty	Do participants use words that indicate openness as to what could be happening? (e.g., possibly, maybe, might) (1 = yes; 0 = no)
Alternative	Do participants give at least two alternative options to explain the scenario? (1 = yes; 0 = no

Mentalistic explanation (certainty, no alternatives): There are two people walking alongside each other. They are both looking serous because they are thinking about an unpleasant interaction they had with a client earlier in the day.

Factual explanation version (certainty, no alternative):
There is an old castle next to the sea in the background and two people are walking on a pathway. The people's clothes suggest that they just had a formal meeting.

Mentalistic explanation version: There are two guys greeting each other. They are both very happy as they have a strong emotional bond but have not seen each other in a long time.

Factual explanation version:
There are two guys looking at each other and smiling. One is touching the shoulder of the other one. They are both wearing light clothes which suggests that it is a pleasant temperature outside.

Examples of possible responses according to coding scheme:

Mentalistic explanation version (certainty, no alternatives): Two colleagues are having a conversation. The way they look at each other suggests that they are enjoying the

conversation.

Mentalistic explanation version (certainty, alternatives): Two colleagues are having a conversation. The way they look at each other suggests that they are enjoying the conversation or that they are happy about some news they have just found out about earlier.

Mentalistic explanation version (uncertainty, alternatives): Two colleagues are having a conversation. The way they look at each other could mean that they are enjoying the conversation, or they could be happy about some news they may have found out about earlier.

Factual explanation version (certainty, no alternatives):
Two colleagues are smiling and looking at each other in an office. There are other people wearing headsets working in the background, which suggests that it is a shared workspace such as a call centre.

Factual explanation version (certainty, alternatives):
Two colleagues are smiling and looking at each other in an office. There are other people wearing headsets working in the background, which suggests that it is a shared workspace such as a call centre or another type of workplace where people conduct online meetings.

Factual explanation version (uncertainty, alternatives):
Two colleagues seem to be smiling and are looking at each other in an office. There are other people wearing headsets working in the background, which suggests that it may be a shared workspace such as a call centre or another type of workplace where people conduct online meetings.

References

References

Abell, F., Happe, F., & Frith, U. (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Cognitive Development*, *15*(1), 1-16.

Abplanalp, S. J., Mote, J., Uhlman, A. C., Weizenbaum, E., Alvi, T., Tabak, B. A., & Fulford, D. (2022). Parsing social motivation: development and validation of a self-report measure of social effort. *Journal of Mental Health*, *31*(3), 366-373.

Abu-Akel, A., & Shamay-Tsoory, S. (2011). Neuroanatomical and neurochemical bases of theory of mind. *Neuropsychologia*, *49*(11), 2971-2984.

Abu-Akel, A. M., Wood, S. J., Hansen, P. C., & Apperly, I. A. (2015). Perspective-taking abilities in the balance between autism tendencies and psychosis proneness. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1808), 20150563.

Ahmed, F. S., & Stephen Miller, L. (2011). Executive function mechanisms of theory of mind. *Journal of Autism and Developmental Disorders*, *41*, 667-678.

Atran, S., & Norenzayan, A. (2004). Why minds create gods: Devotion, deception, death, and rational decision making. *Behavioral and Brain Sciences*, *27*, 754–770.

Allison, C., Auyeung, B., & Baron-Cohen, S. (2012). Toward brief "red flags" for autism screening: The short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(2), 202-212.

Ames, D. R. (2004). Strategies for social inference: a similarity contingency model of projection and stereotyping in attribute prevalence estimates. *Journal of Personality and Social Psychology*, *87*(5), 573.

Apperly, I. (2010). Mindreaders: the cognitive basis of "theory of mind." *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.

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

Apperly, I. A., & Riggs, K. J., Simpson, A., Chiavarino, C., & Samson, D. (2006). Is belief reasoning automatic?. *Psychological Science*, *17*(10), 841-844.

Apperly, I. A., & Wang, J. J. (2021). Mindreading in Adults. *The Cognitive Basis of Social Interaction Across the Lifespan*, p. 96.

Archer, D., Akert, R., & Costanzo, M. (1993). *The accurate perception of nonverbal Behavior: Questions of Theory and Research Design*.

Baron-Cohen, S. (2001). Theory of mind in normal development and autism. *Prisme*, *34*(1), 74-183.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"?. *Cognition*, *21*(1), 37-46.

Baron-Cohen, S., Campbell, R., Karmiloff-Smith, A., Grant, J., & Walker, J. (1995). Are children with autism blind to the mentalistic significance of the eyes?. *British Journal of Developmental Psychology*, *13*(4), 379-398.

Baron-Cohen, S., O'riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders*, *29*, 407-418.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(2), 241-251.

Berecz, H., Tényi, T., & Herold, R. (2016). Theory of mind in depressive disorders: A review of the literature. *Psychopathology*, *49*(3), 125-134.

Betz, N., Hoemann, K., & Barrett, L. F. (2019). Words are a context for mental inference. *Emotion*, *19*(8), 1463.

Bernhardt, B. C., & Singer, T. (2012). The neural basis of empathy. *Annual Review of Neuroscience*, *35*, 1-23.

Bera, E., Bartholomeusz, C., & Pantelis, C. (2016). Meta-analysis of Theory of Mind (ToM)

impairment in bipolar disorder. *Psychological Medicine*, *46*(2), 253-264.

Bering, J. M., & Shackelford, T. K. (2004). Supernatural agents may have provided adaptive social information. *Behavioral and Brain Sciences*, *27*, 732–733.

Boada, L., Lahera, G., Pina-Camacho, L., Merchán-Naranjo, J., Díaz-Caneja, C. M., Bellón, J. M., ... & Parellada, M. (2020). Social cognition in autism and schizophrenia spectrum disorders: the same but different?. *Journal of Autism and Developmental Disorders*, *50*, 3046-3059.

Bosacki, S., Moreira, F. P., Sitnik, V., Andrews, K., & Talwar, V. (2020). Theory of mind, self-knowledge, and perceptions of loneliness in emerging adolescents. *The Journal of Genetic Psychology*, *181*(1), 14-31.

Bradford, E. E., Hukker, V., Smith, L., & Ferguson, H. J. (2018). Assessing belief-attribution in adults with and without Autism Spectrum Disorders using a computerized false-belief task.

Brown, J. R., Donelan-McCall, N., & Dunn, J. (1996). Why talk about mental states? The significance of children's conversations with friends, siblings, and mothers. *Child Development*, *67*(3), 836-849.

Brown, T. A. (2015). Confirmatory Factor Analysis for Applied Research (2nd ed.*). London: The Guilford Press.*

Burnside, K., Ruel, A., Azar, N., & Poulin-Dubois, D. (2018). Implicit false belief across the lifespan: Non-replication of an anticipatory looking task. *Cognitive Development*, *46*, 4-11.

Cameron, C. D., Scheffer, J. A., Hadjiandreou, E., & Anderson, S. (2022). Motivated empathic choices. *Advances in Experimental Social Psychology*, *66*, 191-279.

Cameron, J., & Pierce, W. D. (2002). Rewards and intrinsic motivation: Resolving the controversy. *Bloomsbury Publishing USA.*

Cane, J. E., Ferguson, H. J., & Apperly, I. A. (2017). Using perspective to resolve reference: The impact of cognitive load and motivation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(4), 591.

Caputi, M., & Schoenborn, H. (2018). Theory of mind and internalizing symptoms during middle childhood and early adolescence: The mediating role of coping strategies. *Cogent Psychology*, *5*(1), 1487270.

Carpendale, J., & Lewis, C. (2006). How children develop social understanding. *Blackwell Publishing.*

Carstensen, L. L., Isaacowitz, D. M., & Charles, S. T. (1999). Taking time seriously: A theory of socioemotional selectivity. *American Psychologist, 54*(3), 165.

Carpenter, J. M., Green, M. C., & Vacharkulksemsuk, T. (2016). Beyond perspective-taking: Mind-reading motivation. *Motivation and Emotion, 40*, 358-374.

Cassels, T. G., & Birch, S. A. (2014). Outcomes associated with psychopathic traits in a nonclinical sample of children aged 6-13. *SOJ Psychology, 1*(1), 8.

Castelli, F., Frith, C., Happé, F., & Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain, 125*(8), 1839-1849.

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.

Chen, S. Y., Feng, Z., & Yi, X. (2017). A general introduction to adjustment for multiple comparisons. *Journal of Thoracic Disease, 9*(6), 1725.

Contrares-Huerta, L. S., Pisauro, M. A., & Apps, M. A. (2020). Effort shapes social cognition and behavior: A neuro-cognitive framework. *Neuroscience & Biobehavioral Reviews, 118*, 426-439.

Conway, J. R., Catmur, C., & Bird, G. (2019). Understanding individual differences in theory of mind via representation of minds, not mental states. Psychonomic bulletin & review, 26, 798-812.

Corcoran, C. M., Kimhy, D., Parrilla-Escobar, M. A., Cressman, V. L., Stanford, A. D., Thompson, J., ... & Malaspina, D. (2011). The relationship of social function to depressive and negative symptoms in individuals at clinical high risk for psychosis. *Psychological Medicine, 41*(2), 251-261.

Cotter, J., Granger, K., Backx, R., Hobbs, M., Looi, C. Y., & Barnett, J. H. (2018). Social cognitive dysfunction as a clinical marker: A systematic review of meta-analyses across 30 clinical conditions. *Neuroscience & Biobehavioral Reviews, 84*, 92-99.

Cutting, A. L., & Dunn, J. (1999). Theory of mind, emotion understanding, language, and family background: Individual differences and interrelations. *Child Development, 70*(4), 853-865.

Declerck, C. H., & Bogaert, S. (2008). Social value orientation: Related to empathy and the ability to read the mind in the eyes. *The Journal of Social Psychology*, *148*(6), 711-726.

de la Osa, N., Granero, R., Domenech, J. M., Shamay-Tsoory, S., & Ezpeleta, L. (2016). Cognitive and affective components of Theory of Mind in preschoolers with oppositional defiance disorder: Clinical evidence. *Psychiatry Research*, *241*, 128–134.

Devine, R. T. (2021). Individual differences in theory of mind in middle childhood and adolescence. Theory of mind in middle childhood and adolescence: *Integrating Multiple Perspectives*, 55-76.

Devine, R. T., & Apperly, I. A. (2022). Willing and able? Theory of mind, social motivation, and social competence in middle childhood and early adolescence. *Developmental Science*, *25*(1), e13137.

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.

Devine, R. T., & Hughes, C. (2016). Measuring theory of mind across middle childhood: Reliability and validity of the silent films and strange stories tasks. *Journal of Experimental Child Psychology*, *149*, 23-40.

Devine, R. T., & Hughes, C. (2016). Theory of mind in middle childhood: Longitudinal associations with executive function and social competence. *Developmental Psychology*, *52*(5), 758.

Devine, R. T., & Hughes, C. (2019). Let's talk: Parents' mental talk (not mind-mindedness or mindreading capacity) predicts children's false belief understanding. *Child Development*, 90(4), 1236-1253.

Devine, R. T., Kovatchev, V., Grumley Traynor, I., Smith, P., & Lee, M. (2023). Machine learning and deep learning systems for automated measurement of 'advanced' theory of mind.

Devine, R. T., White, N., Ensor, R., & Hughes, C. (2016). Theory of mind in middle childhood: Longitudinal associations with executive function and social competence. *Developmental Psychology*, *52*(5), 758.

Dreisbach, G., & Goschke, T. (2004). How positive affect modulates cognitive control: reduced perseveration at the cost of increased distractibility. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(2), 343.

Dunn, J., & Brophy, M. (2005). Communication, Relationships, and Individual Differences in Children's Understanding of Mind. In *Why Language Matters for Theory of Mind, Apr, 2002, University of Toronto, Toronto, ON, Canada; This chapter originated from the aforementioned conference*. *Oxford University Press*.

Dunn, J., Brown, J., Slomkowski, C., Tesla, C., & Youngblade, L. (1991). Young children's understanding of other people's feelings and beliefs: Individual differences and their antecedents. *Child Development*, *62*(6), 1352-1366.

Dunn, J., & Cutting, A. L. (1999). Understanding others, and individual differences in friendship interactions in young children. *Social Development*, *8*(2), 201-219.

Duval, C., Piolino, P., Bejanin, A., Eustache, F., & Desgranges, B. (2011). Age effects on different components of theory of mind. *Consciousness and Cognition*, *20*(3), 627-642.

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

Deffler, S. A., Leary, M. R., & Hoyle, R. H. (2016). Knowing what you know: Intellectual humility and judgments of recognition memory. *Personality and Individual Differences*, *96*, 255-259.

Eddy, C. M., & Cavanna, A. E. (2015). Triangles, tricks and tics: Hyper-mentalizing in response to animated shapes in Tourette syndrome. *Cortex*, 71, 68-75.

Eisenberg, N. (2003). Prosocial behavior, empathy, and sympathy.

Ekerim-Akbulut, M., Selçuk, B., Slaughter, V., Hunter, J. A., & Ruffman, T. (2020). In two minds: Similarity, threat, and prejudice contribute to worse mindreading of outgroups compared with an ingroup. *Journal of Cross-Cultural Psychology*, *51*(1), 25-48.

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modelling*, *8*, 430–457.

Epley, N., Waytz, A., Akalis, S., & Cacioppo, J. T. (2008). When we need a human: Motivational determinants of anthropomorphism. *Social Cognition*, *26*(2), 143-155.

Eysenck, H. J. (1997). Personality and experimental psychology: the unification of psychology and the possibility of a paradigm. *Journal of Personality and Social Psychology*, *73*(6), 1224.

Ferguson, H. J., Apperly, I., & Cane, J. E. (2017). Eye tracking reveals the cost of switching between self and other perspectives in a visual perspective-taking task. *The Quarterly Journal of Experimental Psychology*, *70*, 1646-1660.

Fink, E., Begeer, S., Peterson, C. C., Slaughter, V., & de Rosnay, M. (2015). Friends, friendlessness, and the social consequences of gaining a theory of mind. *The British Journal of Developmental Psychology*, *33*(1), 27-30.

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1431), 459-473.

Fulmer, S. M., & Frijters, J. C. (2009). A review of self-report and alternative approaches in the measurement of student motivation. *Educational Psychology Review*, *21*(3), 219-246.

Fujita, N., Devine, R. T., & Hughes, C. (2022). Theory of mind and executive function in early childhood: A cross-cultural investigation. *Cognitive Development*, *61*, 101150.

Fussner, L. M., Mancini, K. J., & Luebbe, A. M. (2018). Depression and approach motivation: differential relations to monetary, social, and food reward. *Journal of Psychopathology and Behavioral Assessment*, *40*, 117-129.

Gable, S. L. (2006). Approach and avoidance social motives and goals. *Journal of Personality*, *74*(1), 175-222.

Gable, S. L., & Haidt, J. (2005). What (and why) is positive psychology?. *Review of General Psychology*, *9*(2), 103-110.

Gandhi, A., Mote, J., & Fulford, D. (2022). A transdiagnostic meta-analysis of physical and social anhedonia in major depressive disorder and schizophrenia spectrum disorders. *Psychiatry Research*, *309*, 114379.

Gardner, R. C., & Tremblay, P. F. (1994). On motivation, research agendas, and theoretical frameworks 1. *The Modern Language Journal*, *78*(3), 359-368.

Germine, L., Dunn, E. C., McLaughlin, K. A., & Smoller, J. W. (2015). Childhood adversity is associated with adult theory of mind and social affiliation, but not face processing. *PloS one*, *10*(6), e0129612.

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, *102*, 74–78.

Goodhew, S. C., & Edwards, M. (2019). Translating experimental paradigms into individual-differences research: Contributions, challenges, and practical recommendations. *Consciousness and Cognition*, 69, 14-25.

Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory.

Gosling, S. D., Rentfrow, P. J., & Swann Jr, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*(6), 504-528.

Guajardo, N. R., & Turley-Ames, K. J. (2004). Preschoolers' generation of different types of counterfactual statements and theory of mind understanding. *Cognitive Development*, *19*(1), 53-80.

Haidt, J., & Jussim, L. (2016). Psychological science and viewpoint diversity. APS Observer, 29.

Hamilton, A. F. D. C., Brindley, R., & Frith, U. (2009). Visual perspective taking impairment in children with autistic spectrum disorder. *Cognition*, *113*(1), 37-44.

Harvey, J. H., & Weary, G. (1984). Current issues in attribution theory and research. *Annual Review of Psychology*, *35*(1), 427-459.

Hattie, J., & Cooksey, R. W. (1984). Procedures for assessing the validities of tests using the" known-groups" method. *Applied Psychological Measurement*, *8*(3), 295-305.

Hayes, A. F. (2017). Introduction to mediation, moderation and conditional process analysis. A regression-based approach. *Guilford Publications*.

Hezel, D. M., & McNally, R. J. (2014). Theory of mind impairments in social anxiety disorder. *Behavior Therapy*, *45*(4), 530-540.

Hess, T. M. (2014). Selective engagement of cognitive resources: Motivational influences on older adults' cognitive functioning. *Perspectives on Psychological Science*, *9*(4), 388-407.

Harkness, K. L., Alavi, N., Monroe, S. M., Slavich, G. M., Gotlib, I. H., & Bagby, R. M. (2010). Gender differences in life events prior to onset of major depressive disorder: the moderating effect of age. *Journal of Abnormal Psychology*, *119*(4), 791.

Harris, M. A., Brett, C. E., Johnson, W., & Deary, I. J. (2016). Personality stability from age 14 to age 77 years. *Psychology and Aging*, *31*(8), 862.

Heimberg, R. G., & Holaway, R. M. (2007). Examination of the known-groups validity of the Liebowitz Social Anxiety Scale. *Depression and Anxiety*, *24*(7), 447-454.

Hollarek, M., & Lee, N. C. (2022). Current understanding of developmental changes in adolescent perspective taking. *Current Opinion in Psychology*, 45, 101308.

Hughes, C., Devine, R. T., Ensor, R., Koyasu, M., Mizokawa, A., & Lecce, S. (2014). Lost in translation? Comparing British, Japanese, and Italian children's theory-of-mind performance. *Child Development Research*, 2014.

Hughes, C., & Devine, R. T. (2015). Individual differences in theory of mind from preschool to adolescence: Achievements and directions. *Child Development Perspectives*, *9*(3), 149-153.

Hofmann, S. G., & Hay, A. C. (2018). Rethinking avoidance: Toward a balanced approach to avoidance in treating anxiety disorders. *Journal of Anxiety Disorders*, *55*, 14-21.

Hogan, R. T. (1991). Personality and Personality Measurement.

Henry, J. D., Phillips, L. H., Ruffman, T., & Bailey, P. E. (2013). A meta-analytic review of age differences in theory of mind. *Psychology and Aging*, *28*(3), 826.

Hughes, C., & Devine, R. T. (2015). Individual differences in theory of mind from preschool to adolescence: Achievements and directions. *Child Development Perspectives*, *9*(3), 149-153.

Hommel, B., & Colzato, L. S. (2017). The social transmission of metacontrol policies: Mechanisms underlying the interpersonal transfer of persistence and flexibility. *Neuroscience & Biobehavioral Reviews*, *81*, 43-58.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modelling: A Multidisciplinary Journal*, *6*, 1–55.

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.

Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., & Jennions, M. D. (2015). The extent and consequences of p-hacking in science. *PLoS biology*, *13*(3), e1002106.

Hruschka, D. J., Medin, D. L., Rogoff, B., & Henrich, J. (2018). Pressing questions in the study of psychological and behavioral diversity. *Proceedings of the National Academy of Sciences*, *115*(45), 11366-11368.

Harris, P. L., German, T., & Mills, P. (1996). Children's use of counterfactual thinking in causal reasoning. *Cognition*, 61(3), 233-259.

Imuta, K., Henry, J. D., Slaughter, V., Selcuk, B., & Ruffman, T. (2016). Theory of mind and prosocial behavior in childhood: A meta-analytic review. *Developmental Psychology*, *52*(8), 1192.

Inoue, Y., Yamada, K., & Kanba, S. (2006). Deficit in theory of mind is a risk for relapse of major depression. *Journal of Affective Disorders*, *95*(1-3), 125-127.

Jack, R. E., Crivelli, C., & Wheatley, T. (2018). Data-driven methods to diversify knowledge of human psychology. *Trends in Cognitive Sciences*, *22*(1), 1-5.

Kline, R. B. (2011). Convergence of structural equation modelling and multilevel modelling.

Knoll, M., & Charman, T. (2000). Teaching false belief and visual perspective taking skills in young children: Can a theory of mind be trained?. *Child Study Journal*, *30*(4), 273-273.

Koelkebeck, K., Pedersen, A., Suslow, T., Kueppers, K. A., Arolt, V., & Ohrmann, P. (2010). Theory of Mind in first-episode schizophrenia patients: correlations with cognition and personality traits. *Schizophrenia Research*, *119*(1-3), 115-123.

Koerber, S., & Osterhaus, C. (2020). Some but not all aspects of (advanced) theory of mind predict loneliness. *British Journal of Developmental Psychology*, *38*(1), 144-148.

Kouklari, E. C., Tsermentseli, S., & Auyeung, B. (2018). Executive function predicts theory of mind but not social verbal communication in school-aged children with autism spectrum disorder. *Research in Developmental Disabilities*, 76, 12–24.

Kroenke, K., & Spitzer, R. L. (2002). The PHQ-9: a new depression diagnostic and severity measure. *Psychiatric Annals*, *32*(9), 509-515.

Krupenye, C., & Call, J. (2019). Theory of mind in animals: Current and future directions. Wiley Interdisciplinary Reviews: *Cognitive Science*, *10*(6), e1503.

Lecce, S., Ceccato, I., & Cavallini, E. (2019). Theory of mind, mental state talk and social relationships in aging: The case of friendship. *Aging & Mental Health*, 23, 1105–1112.

Lei, P. W., & Shiverdecker, L. K. (2020). Performance of estimators for confirmatory factor analysis of ordinal variables with missing data. *Structural Equation Modeling: A Multidisciplinary Journal*, *27*(4), 584-601.

Lenton-Brym, A. P., Moscovitch, D. A., Vidovic, V., Nilsen, E., & Friedman, O. (2018). Theory of mind ability in high socially anxious individuals. *Anxiety, Stress, & Coping, 31*(5), 487-499.

Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in 'theory of mind'. *Trends in Cognitive Sciences, 8*(12), 528-533.

Linnenbrink-Garcia, L., & Patall, E. A. (2015). Motivation. In Handbook of educational psychology (pp. 105-117). *Routledge*.

Lin, S., Keysar, B., & Epley, N. (2010). Reflexively mindblind: Using theory of mind to interpret behavior requires effortful attention. *Journal of Experimental Social Psychology, 46*(3), 551-556.

Livingston, L. A., Shah, P., White, S. J., & Happé, F. (2021). Further developing the Frith–Happé animations: A quicker, more objective, and web-based test of theory of mind for autistic and neurotypical adults. *Autism Research, 14*(9), 1905-1912.

Lockwood, P. L., Ang, Y. S., Husain, M., & Crockett, M. J. (2017). Individual differences in empathy are associated with apathy-motivation. *Scientific Reports, 7*(1), 17293.

Long, E. L., Cuve, H. C., Conway, J. R., Catmur, C., & Bird, G. (2022). Novel theory of mind task demonstrates representation of minds in mental state inference. *Scientific Reports, 12*(1), 21133.

Luyten, P., Van Houdenhove, B., Lemma, A., Target, M., & Fonagy, P. (2012). A mentalization-based approach to the understanding and treatment of functional somatic disorders. *Psychoanalytic Psychotherapy, 26*, 121–140.

Malcolm, C., Saxton, T. K., McCarty, K., Roberts, S. G., & Pollet, T. V. (2021). Extraversion is associated with advice network size, but not network density or emotional closeness to network members. *Personality and Individual Differences, 168*, 110311.

Martinez, G., Alexandre, C., Mam-Lam-Fook, C., Bendjemaa, N., Gaillard, R., Garel, P., ... & Krebs, M. O. (2017). Phenotypic continuum between autism and schizophrenia: Evidence from the Movie for the Assessment of Social Cognition (MASC). *Schizophrenia Research, 185*, 161-166.

Mayes, L. C., Klin, A., Tercyak Jr, K. P., Cicchetti, D. V., & Cohen, D. J. (1996). Test-retest reliability for false-belief tasks. *Journal of Child Psychology and Psychiatry, 37*(3), 313-319.

McPhetres, J., & Zuckerman, M. (2018). Religiosity predicts negative attitudes towards science and lower levels of science literacy. *PloS One, 13*, e0207125.

Michael, J., & Christensen, W. (2016). Flexible goal attribution in early mindreading. *Psychological Review*, *123*(2), 219.

Meins, E., Fernyhough, C., Johnson, F., & Lidstone, J. (2006). Mind-mindedness in children: Individual differences in internal-state talk in middle childhood. *British Journal of Developmental Psychology*, *24*, 181–196.

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review*, *102*(2), 246.

Montag, C., Ehrlich, A., Neuhaus, K., Dziobek, I., Heekeren, H. R., Heinz, A., & Gallinat, J. (2010). Theory of mind impairments in euthymic bipolar patients. *Journal of Affective Disorders*, *123*(1-3), 264-269.

Mulvey, K. L., Rizzo, M. T., & Killen, M. (2016). Challenging gender stereotypes: Theory of mind and peer group dynamics. *Developmental Science*, *19*(6), 999-1010.

Naito, M. (2004). Is theory of mind a universal and unitary construct. ISSBD Newsletter, *1*(45), 9-11.

Navarro, E. (2022). What is theory of mind? A psychometric study of theory of mind and intelligence. *Cognitive Psychology*, *136*, 101495.

Nelsen, A., Gervais, W. M., & Trzesniewski, K. H. (2012). Mentalizing deficits constrain belief in a personal God. *PloS One*, *7(*5), e36880.

Nestor, B. A., Liu, Q., Sutherland, S., & Cole, D. A. (2023). The relation of depressive symptoms to theory of mind in adults. *Journal of Affective Disorders*, *339*, 194-202.

Nejati, V., Zabihzadeh, A., Maleki, G., & Tehranchi, A. (2012). Mind reading and mindfulness deficits in patients with major depression disorder. *Procedia-social and Behavioral Sciences*, *32*, 431-437.

Nettle, D., & Liddle, B. (2008). Agreeableness is related to social-cognitive, but not social-perceptual, theory of mind. European Journal of Personality: Published for the *European Association of Personality Psychology*, *22*(4), 323-335.

Nettle, D., & Saxe, R. (2021). 'If men were angels, no government would be necessary': the intuitive theory of social motivation and preference for authoritarian leaders. *Collabra: Psychology*, *7*(1), 28105.

Oakley, B. F., Brewer, R., Bird, G., & Catmur, C. (2016). Theory of mind is not theory of emotion: A cautionary note on the Reading the Mind in the Eyes Test. *Journal of Abnormal Psychology*, *125*(6), 818

O'Reilly, J., & Peterson, C. C. (2014). Theory of mind at home: Linking authoritative and authoritarian parenting styles to children's social understanding. *Early Child Development and Care*, *184*, 1934–1947.

Peterson, C., Slaughter, V., Moore, C., & Wellman, H. M. (2016). Peer social skills and theory of mind in children with autism, deafness, or typical development. *Developmental Psychology*, *52*(1), 46.

Peterson, C. C., & Wellman, H. M. (2019). Longitudinal theory of mind (ToM) development from preschool to adolescence with and without ToM delay. *Child Development*, *90*(6), 1917-1934.

Poletti, M., Enrici, I., & Adenzato, M. (2012). Cognitive and affective Theory of Mind in neurodegenerative diseases: neuropsychological, neuroanatomical and neurochemical levels. *Neuroscience & Biobehavioral Reviews*, *36*(9), 2147-2164.

Popp-Baier, U. (2010). From religion to spirituality—Megatrend in contemporary society or methodological artefact? A contribution to the secularization debate from psychology of religion. *Journal of Religion in Europe*, *3*, 34–67.

Povinelli, D. J., & Povinelli, T. J. (1996). Mindblindness. An essay on autism and theory of mind. *Trends in Neurosciences*, *19*(7), 299-300.

Preston, S. D., & De Waal, F. B. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, *25*(1), 1-20.

Preckel, F., Holling, H., & Vock, M. (2006). Academic underachievement: Relationship with cognitive motivation, achievement motivation, and conscientiousness. *Psychology in the Schools*, *43*(3), 401-411.

Preißler, S., Dziobek, I., Ritter, K., Heekeren, H. R., & Roepke, S. (2010). Social cognition in borderline personality disorder: evidence for disturbed recognition of the emotions, thoughts, and intentions of others. *Frontiers in Behavioral Neuroscience*, *4*, 182.

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

Qiu, L., Lin, H., Ramsay, J., & Yang, F. (2012). You are what you tweet: Personality expression and perception on Twitter. *Journal of Research in Personality*, *46*(6), 710-718.

Office for National Statistics, 2023. Link: https://www.ons.gov.uk/aboutus/transparencyandgovernance/freedomofinformationfoi/percentageofhomesandindividualswithtechnologicalequipment

Quesque, F., & Rossetti, Y. (2020). What do theory-of-mind tasks actually measure? Theory and practice. *Perspectives on Psychological Science*, *15*(2), 384-396.

Qureshi, A. W., Monk, R. L., Samson, D., & Apperly, I. A. (2020). Does interference between self and other perspectives in theory of mind tasks reflect a common underlying process? Evidence from individual differences in theory of mind and inhibitory control. *Psychonomic Bulletin & Review*, *27*, 178–190.

Radke, S., Volman, I., Mehta, P., van Son, V., Enter, D., Sanfey, A., ... & Roelofs, K. (2015). Testosterone biases the amygdala toward social threat approach. *Science Advances*, *1*(5), e1400074.

Ratelle, C. F., Vallerand, R. J., Chantal, Y., & Provencher, P. (2004). Cognitive adaptation and mental health: A motivational analysis. *European Journal of Social Psychology*, *34*(4), 459-476.

Repacholi, B., & Slaughter, V. (Eds.). (2004). Individual differences in theory of mind: Implications for typical and atypical development. *Psychology Press.*

Richardson, D. R., Green, L. R., & Lago, T. (1998). The relationship between perspective-taking and nonaggressive responding in the face of an attack. *Journal of Personality*, *66*(2), 235-256.

Roberts, R. P., Grady, C. L., & Addis, D. R. (2020). Creative, internally-directed cognition is associated with reduced BOLD variability.

Rodkin, P. C., Ryan, A. M., Jamison, R., & Wilson, T. (2013). Social goals, social behavior, and social status in middle childhood. *Developmental Psychology*, *49*(6), 1139.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modelling and more. Version 0.5–12 (BETA). *Journal of Statistical Software*, *48*, 1–36.

Rossetto, F., Baglio, F., Massaro, D., Alberoni, M., Nemni, R., Marchetti, A., & Castelli, I. (2020). Social cognition in rehabilitation context: different evolution of affective and cognitive theory of mind in mild cognitive impairment. *Behavioural Neurology*, 2020.

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.

Saxe, R. (2006). Uniquely human social cognition. Current opinion in neurobiology, *16*(2), 235-239.

Schaafsma, S. M., Pfaff, D. W., Spunt, R. P., & Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends in Cognitive Sciences*, *19*(2), 65-72.

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, *4*(8), 299-309.

Sebanz, N., & Shiffrar, M. (2009). Detecting deception in a bluffing body: The role of expertise. *Psychonomic Bulletin & Review*, *16*(1), 170-175.

Seidel, E. M., Habel, U., Kirschner, M., Gur, R. C., & Derntl, B. (2010). The impact of facial emotional expressions on behavioral tendencies in women and men. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(2), 500.

Simcharoen, S., Pinyopornpanish, M., Haoprom, P., Kuntawong, P., Wongpakaran, N., & Wongpakaran, T. (2018). Prevalence, associated factors, and impact of loneliness and interpersonal problems on internet addiction: A study in Chiang Mai medical students. *Asian Journal of Psychiatry*, *31*, 2-7.

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.

Slaughter, V., Imuta, K., Peterson, C. C., & Henry, J. D. (2015). Meta-analysis of theory of mind and peer popularity in the preschool and early school years. *Child Development*, *86*(4), 1159-1174.

Slaughter, V., & Repacholi, B. (2004). Introduction: Individual differences in theory of mind: What are we investigating?. In Individual differences in theory of mind (pp. 1-13). *Psychology Press*.

Snaith, R. P., Hamilton, M., Morley, S., Humayan, A., Hargreaves, D., & Trigwell, P. (1995). A scale for the assessment of hedonic tone the Snaith–Hamilton Pleasure Scale. *The British Journal of Psychiatry*, *167*(1), 99-103.

Sodian, B., & Kristen, S. (2016). Theory of mind. In Handbook of epistemic cognition (pp. 68-85). *Routledge.*

Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*, 166(10), 1092-1097.

Strikwerda-Brown, C., Ramanan, S., & Irish, M. (2019). Neurocognitive mechanisms of theory of mind impairment in neurodegeneration: a transdiagnostic approach. *Neuropsychiatric Disease and Treatment*, 557-573.

Swain, D., Scarpa, A., White, S., & Laugeson, E. (2015). Emotion dysregulation and anxiety in adults with ASD: Does social motivation play a role?. *Journal of Autism and Developmental Disorders*, *45*, 3971-3977.

Tager-Flusberg, H. (2007). Evaluating the theory-of-mind hypothesis of autism. *Current Directions in Psychological Science*, *16*(6), 311-315.

Tahiroglu, D., & Taylor, M. (2019). Anthropomorphism, social understanding, and imaginary companions. *British Journal of Developmental Psychology*, *37*(2), 284-299.

Tager-Flusberg, H., & Sullivan, K. (2000). A componential view of theory of mind: evidence from Williams syndrome. *Cognition*, *76*, 59–90.

Taumoepeau, M., Sadeghi, S., & Nobilo, A. (2019). Cross-cultural differences in children's theory of mind in Iran and New Zealand: The role of caregiver mental state talk. *Cognitive Development*, *51*, 32–45.

Tezci, E., Sezer, F., Gurgan, U., & Aktan, S. (2015). A study on social support and motivation. *The Anthropologist*, *22*(2), 284-292.

Trew, J. L. (2011). Exploring the roles of approach and avoidance in depression: An integrative model. *Clinical Psychology Review*, *31*(7), 1156-1168.

Urquiza-Haas, E. G., & Kotrschal, K. (2015). The mind behind anthropomorphic thinking: Attribution of mental states to other species. *Animal Behaviour*, *109*, 167-176.

Van Bavel, J. J., Mende-Siedlecki, P., Brady, W. J., & Reinero, D. A. (2016). Contextual sensitivity in scientific reproducibility. *Proceedings of the National Academy of Sciences*, *113*(23), 6454-6459.

Van Cappellen, P., Fredrickson, B. L., Saroglou, V., & Corneille, O. (2017). Religiosity and the motivation for social affiliation. *Personality and Individual Differences*, *113*, 24-31.

Villalobos, A. R., Padilla-Mora, M., & Trías, J. F. (2010). Authoritarian parenting received from mothers reveals individual differences in preschooler's false-belief, but not in advanced theory of mind. *International Journal of Psychological and Behavioral Sciences*, *4*(6), 1434-1439.

Vonk, J., & Pitzen, J. (2017). Believing in other minds: Accurate mentalizing does not predict religiosity. *Personality and Individual Differences*, *115*, 70–76.

Wagner, J., Lüdtke, O., & Robitzsch, A. (2019). Does personality become more stable with age? Disentangling state and trait effects for the big five across the life span using local structural equation modeling. *Journal of Personality and Social Psychology*, *116*(4), 666.

Warnell, K. R., & Redcay, E. (2019). Minimal coherence among varied theory of mind measures in childhood and adulthood. *Cognition*, *191*, 103997.

Wasserman, T., Wasserman, L. (2020). Motivation: state, trait, or both. Motivation, effort, and the neural network model, 93-101.

Waytz, A., Morewedge, C. K., Epley, N., Monteleone, G., Gao, J. H., & Cacioppo, J. T. (2010). Making sense by making sentient: effectance motivation increases anthropomorphism. *Journal of Personality and Social Psychology*, *99*(3), 410.

Wechsler, D. (2001). Wechsler Test of Adult Reading: WTAR. Psychological Corporation.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, 92(4), 548.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, *72*(3), 655-684.

Wellman, M. P. (1990). Fundamental concepts of qualitative probabilistic networks. *Artificial Intelligence*, *44*(3), 257-303.

Wentzel, K. R., Battle, A., Russell, S. L., & Looney, L. B. (2010). Social supports from teachers and peers as predictors of academic and social motivation. *Contemporary Educational Psychology*, *35*(3), 193-202.

Westra, E. (2019). Stereotypes, theory of mind, and the action–prediction hierarchy. Synthese, 196(7), 2821-2846.

White, S. J., Coniston, D., Rogers, & Frith, U. (2011). Developing the Frith-Happé animations: A quick and objective test of Theory of Mind for adults with autism. *Autism Research*, *4*(2), 149-154.

White, S. W., Schry, A. R., & Maddox, B. B. (2012). Brief report: The assessment of anxiety in high-functioning adolescents with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *42*, 1138-1145.

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.

Wissmann, M., & Toutenburg, H. (2007). Role of categorical variables in multicollinearity in the linear regression model.

Wongpakaran, N., Wongpakaran, T., Pinyopornpanish, M., Simcharoen, S., Suradom, C., Varnado, P., & Kuntawong, P. (2020). Development and validation of a 6-item Revised UCLA Loneliness Scale (RULS-6) using Rasch analysis. *British Journal of Health Psychology*, 25, 233–256.

Zimet, G. D., Dahlem, N. W., Zimet, S. G., & Farley, G. K. (1988). The multidimensional scale of perceived social support. *Journal of Personality Assessment*, *52*(1), 30-41.