Varieties of Social Understanding

by

Emanuela Yeung
B.Sc. (Hons), University of Toronto, 2007
M.Sc., University of Victoria, 2013

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Psychology

We acknowledge with respect the Lekwungen peoples on whose traditional territory the university stands and the Songhees, Esquimalt and W̱SÁNEĆ peoples whose historical relationships with the land continue to this day.

Supervisory Committee

Varieties of Social Understanding

by

Emanuela Yeung
B.Sc. (Hons), University of Toronto, 2007
M.Sc., University of Victoria, 2013

Supervisory Committee

Dr. Ulrich Mueller, Department of Psychology
Supervisor

Dr. James Tanaka, Department of Psychology
Departmental Member

Dr. Jeremy Carpendale, Department of Psychology
Outside Member

Abstract

Supervisory Committee
Dr. Ulrich Mueller, Department of Psychology
Supervisor
Dr. James Tanaka, Department of Psychology
Departmental Member
Dr. Jeremy Carpendale, Department of Psychology
Outside Member

Philosophical and psychological theories of social understanding have largely focused on the construct of "theory of mind" (ToM) and the inferential processes that may be necessary for understanding the meaning of others' behaviour. On these traditional accounts, social understanding has often been described as a process of "mind reading" or "mentalizing", where one imputes mental states to others to make sense of their behaviour. However, recent work from social neuroscience and enactivist and phenomenological perspectives have pointed to the importance of considering non-inferential forms of social understanding that may be a more basic or foundational way in which we understand others. This dissertation investigates the relationship between these different forms of social understanding by examining the role of perceptual, motor, and conceptual processes in how we understand others. One hundred and two older adolescents and adults completed a battery of psychophysical and paper & pencil tasks. Correlations showed coherence amongst measures that assessed participants' perceptual sensitivity to social information, with minimal coherence across "theory of mind" tasks. Exploratory factor analysis conducted on 13 measures yielded a meaningful 4 factor solution that supported the distinction between conceptual or inferential

measures and more direct, perceptual forms of social understanding. Overall, the findings from this study highlight the importance of considering the variety of ways in which we can understand others and provides empirical support for a more pluralistic and comprehensive account of social understanding.

Table of Contents

List of Tables

List of Figures

Acknowledgments

I am deeply grateful to my supervisor, Ulrich Mueller, for supporting and encouraging my varied research interests over the years. Thank you, Ulrich, for your kindness, guidance, and many book recommendations – my thinking, writing, and being (or becoming?) have benefited immeasurably. Thanks also to Jeremy Carpendale and Jim Tanaka for being incredibly supportive committee members. I'm grateful for the many years of mentorship and for the generosity you've both shown in sharing your expertise and feedback on my work.

Many of the ideas in this dissertation were inspired by the time I spent as a visiting researcher at the Center for Subjectivity Research in Spring 2017. I'm grateful to Dan Zahavi and the staff at the CFS for the illuminating conversations that continue to influence my thinking today.

This work would not have been possible without the help of my dedicated and hardworking RAs. Many thanks to Chris Davies, Elisabeth Irving, Emma Malcolm, Mariah Rickman, and Amy Walsh for their assistance in recruiting and testing participants. I am also grateful to Ana Pesquita for generously sharing her task for use in this dissertation, and for offering her time to help with implementation.

My doctoral work was supported by the Social Sciences and Humanities Research Council, the University of Victoria, and the Sara Spencer Foundation, and I am grateful to have received this financial support.

Finally, I would like to express my gratitude to the wonderful community of people that have supported me all along the way. A heartfelt thank you to my friends near and far (especially Alison, Amanda, Danielle, Emily, Lesley, Pierina, and Tara) for providing encouragement, balance, and distraction, and to my family for always believing that this would get done…eventually.

Dedication

To my parents, for opening every door I’ve ever walked through.

Introduction

How do people construct an understanding of their social and psychological world and what are the processes that underlie our ability to interact with and understand others? Philosophical and psychological theories of social understanding have largely focused on the construct of "theory of mind" (ToM) and the inferential processes that may be necessary for understanding the meaning of others' behaviour. On these traditional accounts, social understanding has often been described as a process of "mind reading" or "mentalizing", where we impute mental states to others to make sense of their behaviour. However, recent work from social neuroscience and enactivist and phenomenological perspectives have pointed to the importance of considering non-inferential forms of social understanding that may be a more basic or foundational way in which we understand others.

The aim of this dissertation is to investigate the relationship between these different forms of social understanding by examining the role of perceptual, motor, and conceptual processes in how we understand others. I begin by discussing three general theoretical accounts: 1) theory theory, 2) simulation theory, and 3) interaction theory, and will then review the empirical methods that have been used to evaluate each theory. Next, I discuss theoretical frameworks that attempt to integrate these different theories by highlighting the diversity of ways in which we can understand others across different levels of analysis. Using these frameworks, I will suggest different dimensions to analyze various measures of social understanding that have been used in the literature. From there I will present findings from a study that examines this diversity in a single sample of participants. This study is proposed as an exploratory first step in empirically evaluating different forms of social understanding, the results of which will be discussed in terms of their implications for theories of social understanding and social-cognitive development. By clarifying the relationship between different ways in which people can understand others, I aim to contribute to a more pluralistic and holistic account of social understanding.

Theoretical Foundations

The question of whether and how it is possible to know another's beliefs, intentions, or mental states has been a topic of debate for hundreds of years in the philosophical literature and has typically discussed as the "problem of other minds" (e.g., Avramides, 2019). This question is rooted in a Cartesian framework that assumes a split between mind and body, "inner" and "outer", or mental and physical, and a representationalist view of knowledge and cognition (Carpendale, Atwood, & Kettner, 2013; Carpendale & Lewis, 2015). According to this perspective, social understanding is considered a "problem" because when we encounter another person all we can perceive is their physical body and behaviour (i.e., the movement of limbs that are devoid of any meaning); any mental states that 'lie behind' or cause this behaviour remain 'hidden' and must somehow be inferred. Social cognition, then, is the process by which we "solve" this problem by means of some additional inferential step, which permits us to represent the mental state of others and understand the meaning of their behaviour[1]. Traditionally, two main theories have been proposed to account for how we solve this problem – theory theory (TT) and simulation theory (ST) – and psychologists have used these as explanations for how we are able to understand and interact with others in a meaningful way.

Theory Theory

[1] The term "social cognition" has also been used in social psychology to refer to the study of how one's cognition (i.e., attitudes, preferences, etc.) or behaviour may be affected by the presence of others. This dissertation is concerned with the more fundamental issue of how we are able to understand and interact with others, and throughout this paper I will use the term "social cognition" to generally refer to cognitivist perspectives of social understanding.

According to TT, inferences regarding the mental states of others are based on the folk psychological theories that one holds about the mind. In social interaction, we use these theories in the same way that scientists do to make hypotheses about others' behaviour (e.g., if I see someone reaching out for their cup of coffee I might *infer* that they have a *desire* for coffee based on my folk theories of how intentions or desires cause certain actions or behaviour). There is considerable debate among theory-theorists concerning how these theories are used and the 'form' that they take; that is, as 'explicit' conceptual knowledge or as 'implicit' subpersonal understanding of minds (see, for example, Apperly & Butterfil, 2009; Gopnik & Wellman, 1995). From a developmental perspective, the questions have generally concerned whether theories of mind are innate (i.e., instantiated in a modular "theory of mind mechanism"; Leslie, 1994) or learned over the course of development through experience (e.g., Gopnik 1993; Wellman, 1992).

Much of the empirical work in psychology has approached social cognition from a TT perspective, and people's understanding of others has generally been assessed using tasks that evaluate their "theory of mind", or conceptual understanding of mental states and how they affect behaviour. In addition, the largest body of work has examined the development of theory of mind in children, and much less research has looked at adult social understanding. To better understand why this has been the case, I think it is helpful to briefly consider the history of this field of research.

The term “theory of mind” was first used by Premack and Woodruff (1978) in their seminal paper that asked the question, “Does a chimpanzee have a theory of mind?”. In the paper they write,

> In saying that an individual has a theory of mind we mean that the individual imputes mental states to himself and to others (either to conspecifics or to other species as well). A system of inferences of this kind is properly viewed as a theory, first because such states are not directly observable, and second because the system can be used to make predictions, specifically about the behavior of other organisms. (p. 151)

Thus, the term “theory of mind” was used deliberately to suggest that the understanding of another’s mental states (e.g., intention or purpose belief, preference, pretense, etc.) was based on the inferential use of a theory. To test whether chimpanzees attributed mental states to an agent, Premack and Woodruff (1978) used a problem comprehension paradigm in which an adult chimpanzee, Sarah, would watch a short clip of a human actor confronting a certain problem (e.g., reaching an inaccessible object that was too high). She was then presented with two photos, one of which depicted the solution to the problem (e.g., the actor stepping on a box), and asked to select a photo. Overall, Sarah was presented with 4 different problems on 6 occasions and selected the correct solution on 21 out of 24 trials. She also performed equally well on novel problem sequences for actions that she had never performed or observed before (e.g., an actor trying to play a phonograph that was not plugged in). Based on these findings, Premack & Woodruff (1978) argued that Sarah was able to succeed on these tasks because she was imputing states of mind (e.g., intention/purpose or knowledge/belief) to the human actor, and was making sense of the human’s behaviour by assuming that there were goals or beliefs that “lay behind” the actions.

Interestingly, in the commentaries to the paper, three authors independently suggested that a stronger test of the ability to impute mental states would be to demonstrate an understanding that beliefs can be *false* (Bennett, 1978; Dennett, 1978; Harman, 1978). An understanding of false belief would show that one can have a belief or representation that differs from reality, and that this belief can be used to guide behaviour. Based on these commentaries, Wimmer and Perner (1983) conducted the first assessment of false belief understanding in children using an *unexpected transfer task*. In this task, children are told a story with puppets in which a character named Maxi puts his chocolate in cupboard *x* and then leaves the scene. Unbeknownst to him, his mother moves the chocolate from cupboard *x* to cupboard *y* while he is gone, and then leaves herself. Maxi then returns and children are asked to indicate where Maxi thinks the chocolate is or where he will look for the chocolate. Hundreds of studies using various versions of this task with diverse populations of typically-developing children have shown that 4- and 5-year-olds tend to correctly answer that Maxi will look for his chocolate in the location where he left it (cupboard *x*), whereas children around 3 years of age will incorrectly answer that Maxi will look in the new location (cupboard *y*). This has been interpreted as a qualitative shift around age 4 in children's understanding of mental life, and false belief understanding has been considered an important milestone and hallmark of children's social knowledge (e.g., Wellman, 2018; Wellman & Liu, 2004).

Subsequent research has shown that ToM continues to develop into middle childhood as children begin to understand more complex aspects of mentality such as ambiguity and differences in interpretation (e.g., Chandler & Carpendale, 1996; Lalonde

& Chandler, 2002; Osterhaus, Koerber, & Sodian, 2016). Interestingly, psychological theories of adult ToM have generally been based on extrapolations from the developmental literature or in comparisons to clinical populations such as individuals with autism spectrum disorder (e.g., Happé, Cook, & Bird, 2017). As Barr and Keysar (2005) note in the title of their chapter, assessment of mindreading (or ToM) in "normal adult humans" remains "an exotic case" (p. 271). Indeed, most empirical work on adult ToM has focused on the neurobiological substrates that may underlie "mindreading" or "mentalizing" ability and less is known about the cognitive basis of ToM.

Theory of Mind in Adults

One network of brain regions, known as the "ToM network" or "mentalizing network" has received considerable interest in the field of social neuroscience. The ToM network is composed of several cortical midline structures, including the medial prefrontal cortex (mPFC) and posterior cingulate cortex, and the bilateral temporo-parietal junction (TPJ) (e.g., Frith & Frith, 2008; Mahy, Moses, & Pfeifer, 2014; Molenberghs, Johnson, Henry, & Mattingley, 2016; Saxe, 2010). Functional magnetic resonance imaging (fMRI) studies have shown that these regions are implicated when people are asked to reason about the mental states of other people. For example, Saxe and Kanwisher (2003) had adults silently read short stories with differing types of content (i.e., false belief, false photograph, desire, non-human descriptions, and physical descriptions of people) and answer fill-in-blank questions after each one. They found greater blood-oxygen-level dependent (BOLD) response in the TPJ when participants were reading stories that implied a character's mental state (i.e., goal or desire) compared to stories about non-human objects or logically similar non-social stories (false photograph). Furthermore, the TPJ did not show an increased response to the mere presence of a person (i.e., for the physical descriptions of the person). However, a recent neuroimaging meta-analysis has shown that there is a fair amount of diversity in the operational definitions of "theory of mind" and in the tasks that have been used to assess mentalizing ability. Importantly, these differences in task parameters seem to elicit different patterns of activation in the brain (Molenberghs et al., 2016). In particular, tasks such as the Reading the Mind in Eyes Test (Castelli et al., 2010) or the Frith-Happé Animations task (Castelli, Happé, Frith, & Frith, 2000), which use more affective, visual, or implicit stimuli place additional demands on brain regions typically associated with the mirror neuron system such as premotor areas and inferior frontal gyrus (described below). By contrast, assessments of ToM that are based more on cognitive, verbal, or explicit stimuli such as verbal or written stories yield activation in brain regions specific to mentalizing, such a mPFC and TPJ.

Although neuroimaging studies in adults have been fruitful in identifying the different brain areas that may subserve mentalizing ability, much less is known about how these regions function to enable social understanding. According to Apperly (2013), this is largely due to the lack of functionalist accounts of mentalizing ability or psychological theories of the cognitive processes that may support theory of mind. As mentioned above, most empirical work in theory of mind has focused on the developmental progression during infancy and the early childhood years, and most studies tend to implicitly assume that there is not much more to mindreading than the acquisition of complex ToM concepts such as higher-order false belief (Apperly, 2013).

To date, there have only been a few studies on “mature” ToM, and these tend to focus on the cognitive processes that may affect how individuals apply mindreading concepts in a flexible, context-sensitive manner. For example, Apperly, Warren, Andrews, Grant, and Todd (2011) examined whether there is continuity in how older children and adults use theory of mind concepts to predict behaviour by using a modified false belief paradigm. In their task, participants read sentences describing a scenario that included 1) two boxes containing hidden food items, 2) which box a character believed the food to be in (either a false belief or true belief), and 3) whether the character wanted to find or avoid the food (positive or negative valence). Participants were then asked to use this information to determine which box the character would open (e.g., if the character had a false belief and a desire to avoid the food, he would open the box that had the food item because of his false belief that it was empty). The results showed that participants found false belief and negative desire (i.e., avoidance of food) to be more difficult and were slower to respond on these trials, even though they did not have to infer the belief state of the character, but rather had to use it to make a prediction about behaviour. Apperly et al. (2011) argue that this supports the idea that psychologically relevant parameters that make mental state attribution in children challenging, such as valence of belief and desire, continue to influence the effort that adults use when mentalizing. Other studies have shown that the variability in “effort” observed in adults in mentalizing tasks may be reflective of cognitive processes such as memory and executive function (e.g., German & Hehman, 2006; Lin, Keysar, & Epley, 2010).

Interestingly, some have argued that even though adults have the ability to ascribe mental states using theory of mind inferences, they do not reliably use this ability to make predictions about others' actions or in their interpretation of others' behaviour (Keysar, Lin, and Barr, 2003). For example, in a study by Keysar, Barr, Balin, and Brauner (2000), participants were instructed to play a communication game with a confederate (the "speaker") in which they had to rearrange a set of objects along a series of mutually visible shelves that were set up in a grid pattern. The speaker was given a diagram depicting how the objects should be arranged, and the participant (the "listener") was asked to follow these instructions and move objects from different shelves so they matched the diagram. Importantly, some of the objects on certain shelves were occluded such that they could only be visible to the listener and not the speaker. For example, the grid array included a large candle and a medium-sized candle that both listener and speaker could see. On one additional shelf, only visible to the listener, was a third, smaller candle (the "competitor" object). On "critical instruction" trials, the speaker asked the listener to move the "small candle". To assess whether listeners considered the competitor object, the test condition was compared to a control condition in which the competitor object was replaced with a different object, which did not match the critical instruction (a toy monkey). The results showed that even though adults are presumed to understand the perspectives of another, their behaviour was surprisingly egocentric; that is, listeners in the test condition strongly considered competitor objects as referents for the speaker's instructions, looking longer at the competitor object than in the control condition. Furthermore, in about 20% of critical instruction trials, participants would attempt to pick up the competitor object (Keysar et al., 2000). The same pattern of response was observed when the competitor object was not actually visible to the listener (i.e., placed in an opaque bag), and when the listener knew that the speaker had a false belief about what the competitor object was (Keysar et al., 2003). Thus, although adults are clearly able to understand "theory of mind" concepts such as false belief, and attribute these concepts to others, they may not do so reliably or use this information effectively in certain contexts, even in interaction with another.

Simulation Theory

In contrast to TT, simulation theorists hold the view that general psychological theories are not necessary for understanding others, and propose that we come to know another's mental states by "putting ourselves in another's mental shoes" (Goldman, 1992). Understanding the behaviour of others depends on one's ability to project oneself imaginatively into the other person's situation. Theories are not required because we have access to our own mind and the way in which our beliefs, intentions, and mental states affect our own behaviour, and we can use this as a model to understand the mental states of others.

To illustrate how simulation can allow one to understand others, Gallese and Goldman (1998) describe the example used by Tversky and Kahneman (1974) in which participants in a study were told that two travelers sharing the same limousine to the airport were caught in traffic and arrived 30 minutes after their flights were scheduled to depart. Although their planes were scheduled to leave at the same time, Mr. A was told that his flight left on time and Mr. B was told that his flight was delayed for 25 minutes and had just departed 5 minutes ago. Participants were asked who they thought would be more upset and 96% reported that it would be Mr. B. According to Gallese and Goldman (1998), in order to understand how Mr. A or B would have felt, each participant would have to put themselves in the imaginary person's 'mental shoes'. To do so, they propose that one first creates "pretend" desires, preferences, or beliefs, which are then fed into a decision-making mechanism. This mechanism then outputs a "pretend" decision, which one uses to predict the mental state of the other.

Gallese and Goldman's (1998) simulation account includes various mechanisms and processing systems (i.e., a practical reasoning system, behaviour-predicting and -explaining system, action control system, body monitoring system, and 'pretend belief and desire generator) that are thought to allow for mental state understanding. However, it should be noted that several other variants of ST have been proposed and they differ in how the simulation process is characterized. While Goldman (1993) argues that "simulation" involves the process of explicitly and deliberately adopting a simulated state of another through a "practical reasoning system" (Goldman, 2002, p. 7), others have argued that "simulation" is best understood to be an implicit, subpersonal 'mirroring', 'resonance', or 'empathy' process that may not be "inferential" at all (e.g., Gallese, 2001; Gordon, 1996; Harris, 1991).

Within the neuroscientific literature, proponents of both explicit and implicit simulation theory often point to the existence of the mirror neuron system (MNS) as a biological mechanism for the process of 'mirroring', 'simulating', or imitation (e.g., Gallese, Keysers, & Rizzolatti, 2004). Initial findings in the macaque premotor cortex showed that certain populations of neurons would fire when the monkey performed a particular goal-directed action, and when the monkey observed another monkey or human performing the same action (Gallese & Goldman, 1998). The function of mirror neurons in human adults remains controversial (e.g., Heyes, 2010; Hickok, 2013); however, the human MNS is thought to include populations of neurons in the intraparietal sulcus (IPS) and premotor areas in the inferior frontal gyrus (IFG). Most studies of the MNS have focused on action understanding. For example, in one fMRI study, participants observed videos of different goal-directed actions being performed (e.g., a precision grip or index finger pull) and were also asked to execute the same actions in the scanner. Increases in BOLD activity was observed in bilateral IFG during both action execution and observation. Furthermore, areas of the cortex that were thought to contain mirror neurons showed adaption – less activity was observed both when an action was first observed and then later executed and when the action was first executed and later observed (Kilner, Neal, Weiskopf, Friston, and Frith, 2009).

There is much less direct evidence for the role of the MNS in supporting a "simulation" process in the context of mental state attribution. However, as discussed above, some neuroimaging studies using ToM tasks such as Reading the Mind in the Eyes have observed activation in the mirror system during emotion attribution, in addition to the activation in the classic "mentalizing" network (Keysers & Gazzola, 2009; Molenbergs, Cunnington, & Mattingley, 2012).

Interaction Theory

Although TT and ST remain the dominant theoretical perspectives in the psychological literature, there has been a recent shift in the philosophical debate towards an alternative perspective for understanding social cognition. Termed the "interactive turn", researchers have begun to emphasize the embodied and embedded nature of persons, and the effect this has on how we construe social understanding (e.g., De Jaegher, Di Paolo, & Gallagher, 2010; see also Schilbach, et al., 2013). Although TT and ST have often been described as opposing accounts of social understanding because they differ in the inferential process required for imputing mental states (i.e., folk theory use or simulation), both perspectives construe social understanding as being primarily social *cognition* (i.e., as "mindreading" or "mentalizing") and both share the same meta-theoretical assumptions based on a Cartesian worldview (see Carpendale & Lewis, 2004; Gallagher, 2012). As such, both TT and ST are considered information processing or "spectorial" theories of other minds (i.e., theories of how we understand others from a third-person perspective or observational stance) and social understanding is primarily based on one's ability to *represent* the mental states of others. Whether or not this is the default way in which we understand others is a matter of debate.

According to interaction theorists, the "problem" of other minds is fundamentally misconstrued because (some) mental states can be *directly perceived* in the facial expressions and embodied behaviour of others within a given (shared) context (e.g., Gallagher 2008; 2015; Zahavi, 2011). Consider the following quote by Scheler (1954):

> *For we certainly believe ourselves to be directly acquainted with another person's joy in his laughter, with his sorrow and pain in his tears, with his shame in his blushing, with his entreaty in his outstretched hands, with his love in his look of affection, with his rage in the gnashing of his teeth, with his threats in the clenching of his fist, and with the tenor of his thoughts in the sounds of his words. If anyone tells me that this is not 'perception', for it cannot be so, in view of the fact that a perception is simply a 'complex of physical sensations', and that there is certainly no sensation of another person's mind nor any stimulus from such a source, I would beg him to turn aside from such questionable theories and address himself to the phenomenological facts.* (p. 260)

Mental states or intentions are not hidden; rather, they are expressed in, or through, our actions and behaviour, and can be directly perceived by others (Fuchs & De Jaegher, 2009). Furthermore, proponents of IT argue that when we consider the phenomenology (i.e., experience) of social interaction, very rarely do we actually have to make an inference (either through the use of a theory or simulation process) to make sense of another's actions or behaviour. In fact, it is precisely when the everyday, moment-to-moment interaction breaks down that we rely on these inferential processes (Carpendale, et al., 2013; Zahavi, 2008; 2011). Importantly, according to this direct social perception (DSP) thesis, much of our everyday social understanding results from perceptual processes rather than inference.

To illustrate this point further, Zahavi (2014) argues that a more accurate example of everyday social understanding can be captured by reframing the "delayed travelers" scenario from a detached third-person perspective as described by Gallese and Goldman (1998), to one that involves second-person engagement: instead of thinking about a hypothetical scenario as a research participant in a study, you are working at the check-in desk at the airport when two individuals rush towards the counter. You soon realize that Mr. A and Mr. B have missed their flights. Whereas Mr. A seems calm and relaxed, Mr. B seems very tense and when you inform him that he had only missed his flight by five minutes, he becomes irate and begins to shout at you. At this point, if you are asked whether Mr. A or Mr. B is more upset, you would undoubtedly say Mr. B. According to Zahavi (2014), this scenario is much more akin to our everyday form of social understanding, and it seems implausible and phenomenologically unsatisfactory to say that you would have reached this conclusion via an elaborate simulation or inferential process. Instead, it seems more plausible to suggest that you directly perceived Mr. B's anger in his facial expressions or in the tenor of his voice.

Direct Social Perception

To better understand IT approaches to social understanding, and how the meaning of another's behaviour can be "directly perceived", it is important to clarify the notion of perception being employed. Interaction theorists generally hold the view that perception is enacted or enactive (e.g., Gallagher, 2008; Overgaard, 2017). This perspective challenges the cognitivist view of perception as a passive process by which an external world is "represented" as an inner state based on transduction and/or transformation of information received through sensory inputs. Instead, enactivists argue that perception is an active process based on sensorimotor action and engagement with the environment, rather than passive sensory input. “Meaning” emerges from this engagement and is enacted or “brought forth” through the sensorimotor patterns that guide action (Varela, Thompson, & Rosch, 1991).

A central feature of enactive perception is that it is “direct”. One way to understand the notion of “directness” is to consider whether any process of inference is required for meaning. As Drayson (2018) has noted, the issue of inference in perception has been a point of discussion in general theories of visual perception, and is arguably the main difference between “constructivist” accounts of vision that originated with Helmholtz (1878) and ecological theories proposed by Gibson (1967). These two approaches differ in the way in which they attempt to solve the “inverse problem” of optics – that is, how the brain solves the problem of the indeterminacy of retinal data (i.e., the same retinal image can be produced by a number of different visual percepts). Helmholtz (1878) argued that perception entails an “unconscious inference” process in which the brain uses some stored information to compute inferences about the most probable sensory stimulus that would have caused the input. By contrast, Gibson (1967) argued that the visual input to the retina includes the retinal flow or dynamic information about how certain aspects of the light array change or remain static in the environment. Importantly, these changes or constancies in the light array provide, for an active perceiver, all of the features of perception (e.g., shape, size of objects, depth, etc.) that constructivists argue need to be inferred (Drayson, 2018).

From an ecological perspective, the boundaries between perception, action, and cognition are somewhat arbitrary as these functions are thought to have evolved to prepare an organism to act and respond to its environment in increasingly complex ways (Marsh, Johnston, Richardson, & Schmidt, 2009). As such, an organism inherently perceives its environment as meaningful based on the actions that can be performed or in terms of *affordances* (Gibson, 1979; Wiltshire, Lobato, McConnell, & Fiore, 2014). These affordances are directly perceivable to the organism because the environment includes information specifying what appropriate actions may be (i.e., in terms of patterns in the optic array; Marsh et al., 2009). From this perspective, perception is construed as a relational and temporally extended process that functions primarily to enable action for an organism within its specific environment (Wiltshire et al., 2014). Enactivist accounts of perception are most similar to the ecological view in the sense that inference is not required for meaning; however, they differ in the sense that perception is thought to be non-representational or non-conceptual (i.e., perceiving is not *about* the world, rather, they are episodes of contact *with* the world; see Noë, 2004; Overgaard, 2017 for further discussion).

From a phenomenological perspective, ‘directness’ can also be understood in the sense of being “unmediated”, or as the degree of experiential acquaintance with an object. To illustrate what this sense of “directness” entails, imagine hearing about how lovely it is to sit by the water in Copenhagen with a beer on a warm spring day; contrast that with seeing a photo of that experience, and finally, contrast that with actually being there and experiencing it. On the phenomenological account, the last instance of actually being there is a more *direct* form of contact with the experience in question than seeing a photo or hearing about it. In the latter cases, the experience of being by the water is not “present” and one’s “contact” with the experience is mediated by a “re-presentation” (i.e., a photo or words). In this sense, vision or perception may be considered the paradigmatic form of direct experience (Zahavi, 2011).

With respect to social understanding, direct *social* perception describes our ability to experience or perceive (i.e., see, hear, etc.) gestures, facial expressions, and speech acts as meaningful and intentional without having to simulate or infer it; the meaning is in the action or expression itself (Gallagher, 2007; Krueger & Overgaard, 2012). From an ecological perspective, gestures or behavioural expressions can be understood as *social affordances*, specifying information regarding the potential to interact (Wiltshire et al., 2014). Consider also, the situation in which you are in the presence of another person in pain. According to the DSP thesis, you directly perceive another’s pain or sadness by seeing them writhe in pain or seeing the tears streaming down their face. By contrast, if you were to see a bottle of painkillers and an empty glass of water on the table, or notice a pile of tear-soaked tissues, you may *indirectly infer* that another person is in pain or feeling sad (Zahavi, 2011). Importantly, direct social perception should be understood at the personal level, as an activity of a perceiving subject. As such, the argument that perceptual experience is “direct” or unmediated does not negate the idea that complex neural mechanisms may be operating at the subpersonal level.

Accordingly, some interaction theorists argue that the functional activation of the MNS during action execution and observation can be interpreted from a “direct perception” perspective. As Gallagher (2007) notes, the dynamics in the MNS in response to a cue (i.e., an action performed by an other) reveal that mirror neurons fire within 30-100ms after appropriate visual stimulation. This is considered to be a short amount of time (even in a neurological sense) between activation of visual cortex and activation of premotor cortex, which raises the possibility that activation of the MNS may underpin direct perception rather than a functionally distinct "simulation" process. As Gallagher (2007) argues, given this short amount of time between activation of the visual cortex and activation of the premotor cortex, the distinction between the act of perception and the act of simulation seems to be unclear. Furthermore, even if one were to be able to draw a definitive line between neural activation, it is not clear how this distinction would translate into a step-wise progression from the act of "perception" to "simulation" at the subpersonal or personal level (Gallagher, 2007). Perhaps, then, another way to interpret mirror neuron activity is that the firing of neurons in premotor areas may support an enactive and temporally extended view of perception, rather than a multiple-step information-processing process of "vision", followed by "simulation".

DSP and Motor Theories of Social Cognition

In addition, there is evidence from the empirical literature on action perception that supports the DSP thesis. According to motor theories of social cognition, the intentions that drive our actions affect the kinematics of our movements, and these kinematic differences can be perceived by other people, giving them direct access to our intentions (Becchio, Manera, Sartori, Cavallo, and Castiello, 2012). Across a series of studies, Becchio and colleagues have shown that the kinematic patterns for certain actions vary systematically depending on whether they are based on an individual intention or social intention. For example, in the case where a participant is asked to reach and grasp an object to move it to a different location (individual intention) compared to reaching to place the object into the hand of another person (social intention), there is a significant decrease in maximal finger aperture and peak grip closing velocity in the social context (Becchio et al., 2008). Subsequent studies have shown that observers are attuned to these kinematic differences and are able to distinguish between different intentions. In one experiment, participants viewed videos in which an actor reached and grasped a wooden block with 3 different intentions: (a) to cooperate with a partner, (b) to compete against an opponent, or (c) to perform an individual action. Results showed that adults were able to accurately distinguish between these different intentions based on the movement of the actor's arm and hand alone (Sartori, Becchio, & Castiello, 2011), and were also accurate in predicting the intention when the reach-to-grasp actions were recorded as point light displays (Manera, Becchio, Cavallo, Sartori, & Castiello, 2011).

A recent study by Pesquita, Chapman, and Enns (2016) also points to our ability to pick up on very subtle kinematic cues that distinguish between different intentions in others' behaviour. In their experiment, adult participants observed videos of an actor reaching out to touch one of two buttons. Crucially, on half the trials the actor *chose* which button to touch and on the other half they were *directed* (off camera) to touch one of the buttons. Participants were asked to predict which button the actor would touch by making a key press, and also indicated whether they thought the actor had chosen or was directed to the button. Interestingly, participants were reliably quicker in predicting which button the actor would touch on trials where the actor *chose* which one to press; however, they were not able to distinguish between the two types of trials verbally (i.e., consciously).

Kinematic analysis showed that reach actions between the conditions differed on four measures. Chosen reach actions had longer mean times to peak velocity, a higher mean vertical area under the curve (corresponding to a higher vertical arch in the trajectory), a larger mean ascending angle (corresponding to a steeper ascent), and a longer mean ascending distance. According to Pesquita et al. (2016), these differences seemed to reflect decision uncertainty during the “chosen” trials. The findings suggest that participants were attuned to these cues, even though they were not able to consciously distinguish between the different reach actions (unlike in the Becchio studies). In addition, individual differences in sensitivity to kinematic cues were negatively correlated with scores on the Autism Quotient, used as a measure of global social competence. Overall, these findings support the claim that perceptual sensitivity during action prediction is an important aspect of social understanding (Pesquita et al., 2016).

Integrative Models of Social Cognition

Although TT, ST, and IT differ in the way in which they describe and explain social understanding, researchers have argued for the importance of construing social understanding as multifaceted involving various distinct process and different forms of knowing (e.g., Carpendale & Lewis, 2004; Gallagher, 2016; Musholt, 2018; Zahavi, 2011).

In the social neuroscience literature, several theoretical models have been proposed that integrate features of TT and ST with direct forms of social understanding proposed by IT. For example, Bohl and van den Bos' (2012) account uses dual-process theory to distinguish between the slower, deliberate forms of social cognition as described by ToM accounts (Type 2 processes), and fast, efficient, and stimulus-driven forms of social understanding captured by interactionist accounts (Type 1 processes). They also distinguish between personal and subpersonal levels of description and explanation, and include a third, supra-individual level, to capture the temporal dynamics, environmental, and contextual factors that may influence social understanding.

In Bohl and van den Bos' (2012) integrative model, both abstract social reasoning and direct perceptual processes are thought to be important in social interaction, and theorizing, simulation, and perception are described at both the personal and subpersonal levels. For example, personal level descriptions of Type 2 processes can be understood as theoretical and/or simulative inferences of the reasons for the other's mental state ("she must want coffee because she is tired"), while subpersonal descriptions or explanations involve the neuroscientific data on mentalization (e.g., activation of the theory of mind network) and "functional explanations of the cognitive mechanisms provided by implicit ToM" (p. 11). By contrast, personal-level descriptions of the Type 1 processes include direct perception of others' mental states ("she is sad") and subpersonal explanations include activation of the MNS or other sensorimotor areas, and "functional explanations of the cognitive mechanisms provided by interactionism" (p. 11). As the authors note, the relationship between Type 1 and Type 2 processes requires further investigation, as does the relation between personal and subpersonal levels of description and explanation. For example, on their account, it is not clear when and how these different processes are used in real time interaction, nor are the different kinds of social information provided by these processes clear.

Furthermore, although Bohl and van den Bos' (2012) model is an interesting way to think about the different forms of social understanding described in the extant literature, additional questions remain. As the authors contend, the nature of "implicit ToM" (i.e., implicit simulation or theorizing) and its characterization as subpersonal Type 1 or Type 2 processes requires further elaboration. From a phenomenological and enactivist perspective, it is difficult to understand what subpersonal theorizing or simulating (or subpersonal inference) amounts to (Gallagher, 2005, 2007; Zahavi, 2005). Although some TT and ST theorists defer to neural mechanisms as explanatory accounts, this does not really address the conceptual murkiness of how to understand "theorizing" or "simulation" as subpersonal mechanisms as these are traditionally descriptions of personal-level functions performed by a theorizing or simulating subject. In addition, there is no sufficient account for how neural processes are related to psychological mechanisms or experiences at the personal level (see Musholt, 2018 for further discussion on the personal and subpersonal distinction in the context of social cognition).

One further issue with describing mentalizing processes as being Type 2 and perceptual processes as Type 1 is that it implies that they are functionally dissociable and independent processes. In many dual system models, automatic, quick, reflexive

Type 1 processes are often described as emerging earlier both phylogenetically and ontogenetically than Type 2 processes, and some argue that these two broad ways of processing are instantiated in distinct neural and functional pathways (e.g., Adolphs, 2009; Frith & Frith, 2008). However, additional clarity regarding the emergence of, and relationship between, direct, perceptual forms of social understanding and explicit mentalizing abilities may be gleaned if we take a developmental perspective. We, as adults, are able to use language to theorize, introspect, reflect on, and imagine what it might be like to be in another's situation. However, this ability should be considered an outcome of development, and must be founded on a more basic, non-conceptual form of social understanding (e.g., Carpendale et al., 2013; Zahavi, 2011). Rather than thinking about mentalizing abilities as being functionally distinct from direct forms of social perception in adults (i.e., as separate type 1 and type 2 processes), it may be the case that direct, perceptual abilities that allow us to perceive others *as* minded provide the foundation for our ability to impute mental states in others. On this view, direct perception should not be seen as a completely separate and alternative way of understanding others. Rather, our ability to theorize or simulate may be the way in which we imagine, elaborate, or reflect on a more basic form of social understanding (see Zahavi, 2011).

The Present Study

Experimental findings across different domains of research support the proposal that social understanding comes in various forms and at different levels of awareness. This speaks to the complexity and variety of social understanding, which generally has not been considered in current experimental work on social cognition. The debate regarding TT, ST, and IT has largely remained in the philosophical and theoretical literature, and although there is some acknowledgment of the need to investigate different forms of social understanding beyond simulation or theory-use, particularly within the field of social neuroscience, few studies in the psychological domain have adopted a more comprehensive approach in studying different forms of social understanding. From an experimental perspective it is important to consider the theoretical framework within which one is working and how “social understanding” is operationalized and assessed. Theoretical perspectives are important because they may influence the way in which certain tasks are designed. For example, in many of the classic “theory of mind” tasks, such as the classic Maxi task, participants are often placed in an observational, third person stance and explicitly asked to make mental state inferences about hypothetical others. This way of assessing social understanding is quite different from predicting another person’s actions within a second person interaction, as in the Becchio and Pesquita studies. Interestingly, across these different literatures, researchers purport to be assessing “social cognition”; however, the relationship between these different forms of social understanding remains unclear.

In the present study I take a first step towards clarifying the relationship between different forms of social understanding by including a variety of different tasks used across different empirical domains in one sample of adults. I aimed to select tasks that vary in the modality of stimulus presentation and response, valence of stimuli or response (i.e., emotional or neutral), as well as the perspective engaged by the observer or participant, as these are factors that distinguish different forms of social understanding (see Table 1).

From a ToM perspective, I have included a measure of higher-order false belief in which participants hear stories with different characters and are asked to make an inference about what belief a character might have given contextual information in the scenario. This task is similar to tasks that have been used in neuroimaging studies that have found selective activation of the "mentalizing network" (e.g., Saxe & Kanwisher, 2003), and is also similar to how ToM has been assessed in children. In addition, I have included two other tasks widely used to assess ToM in both typical and atypical populations: the Reading the Mind in the Eyes and Reading the Mind in Films tasks. Both tasks are designed to assess people's ability to attribute complex emotions to others. In the case of the Films task, participants observe, from a third person perspective, scenes in which interactions are unfolding between two or more persons. In the Eyes task, participants view static images of eyes, some of which are "looking" at the observer; as such I have characterized the Eyes task as evaluating more of a first-person stance.

From an interactionist or motor perspective, I have included several tasks that rely primarily on participants' perceptual sensitivity to intention. For example, in the

Attentional Control Task used by Pesquita et al. (2016), participants are engaged in second-person interaction (albeit via a recorded video) as they are told to play a game in which they try to "beat the actor" on each trial. Furthermore, the task instructions (to predict where an actor will reach) and the requirement for a speeded motor response limits the possibility for (explicit) inference. I have also included 2 additional tasks that are based on movement cues with little additional contextual information: The Frith-Happé animations and point-light displays from the Communicative Interaction Database. Both tasks include stimuli that are non-linguistic with perceptual information based primarily on movement (i.e., between simple line drawings of triangles or between dots in the point-light displays).

Finally, as a general measure of self-reported social aptitude, I have included the Autism Quotient (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001). From this questionnaire, I will use the total score, as well as subscale scores from the social skills, communication, and imagination subscales. Based on the items that compose each scale, the social skills and communication subscales assess participants' social aptitude in the context of second-person interactions while the imagination subscale is more relevant to third person mentalizing ability (see Appendix).

While I generally take an exploratory approach, I aim to address the following questions: What is the relationship between our ability to "mentalize" or "mindread" and more direct forms of social understanding? Is our ability to attribute propositional mental states to others (including abstract, hypothetical others) fundamentally different from the way in which we engage in social interaction? To address these questions, I will examine the correlations between performance across these tasks as a measure of coherence, and use exploratory factor analysis to determine whether there is a meaningful latent structure underlying the different forms of measurement. From a TT or ST perspective, if all forms of social understanding involve a process of explicit or implicit theoretical or simulative inference, we might expect correlations amongst all of the tasks included and a unidimensional factor structure. However, if inferential and perceptual forms of social understanding are distinct but related processes, stronger correlations would be expected amongst perceptual and inferential tasks compared to between these tasks. Furthermore, these relationships would be expected to be reflected in the factor structure, such that perceptual tasks load onto one factor, while inferential tasks load onto a separate factor. Overall, the aim of this dissertation is to clarify the relationship between these various forms of social understanding in order to contribute to a more holistic and comprehensive account of how we are able to interact with and make sense of others.

Method

Participants

Participants were 102 older adolescents and adults (*M* age = 22.7 years, *SD* = 6.45 years, range = 17.1 – 62.3 years; 80 females, 22 males) recruited from the University of Victoria Psychology Research Participation System. The majority of the sample (81%) was Caucasian. All participants were fluent in English and 38% of participants reported being multilingual.

Materials

The Attentional Control, Reading the Mind in Films, Frith-Happé Animation tasks, and the Autism Quotient were administered and completed on a desktop computer with a 21.5" monitor. The Communicative Interaction Database, the Reading the Mind in the Eyes Test, and Higher-order False Belief tasks were completed on a 9.7" iPad.

Measures

Third- and Fourth-Order False Belief (FB). These tasks were adapted from Liddle and Nettle (2006) to assess complex false-belief understanding (i.e., third order – "A thinks that B thinks that C thinks…" and fourth order – "A thinks that B thinks that C thinks that D thinks…"). Participants were presented with 2 stories read aloud by the experimenter and then asked a series of forced-choice memory and ToM level 3 and 4 questions (2 questions per ToM level). Participants received a score of 1 for each correct answer to the ToM questions.

Reading the Mind in the Eyes (RME). The Eyes Test (Baron-Cohen, Wheelwright, Hill, Raste, & Pumb, 2001) was developed to assess participant's abilities to perceive complex emotions and mental states based on static images of eyes. Participants were shown 36 images of eyes, each presented individually, and asked to select 1 of 4 descriptors that best fit the emotion or mental states conveyed by the image. Each correct response received a score of 1.

Reading the Mind in Films (RMF). The RMF task (Golan, Baron-Cohen, Hill, & Golan, 2006) was designed to assess an individual's recognition of complex emotions and mental states from social scenes taken from feature films. Participants were shown 22 short scenes (5 – 30s long) that depicted an emotional interaction between 1-4 characters, and the expression of a complex emotion or mental state (e.g., "smug", "awkward", "concerned"). At the beginning of each trial, a protagonist was identified and their emotional or mental state was labelled along with 3 other "foils". Participants were asked to select their response after each video, and received a score of 1 for each correct answer.

Attentional Control Task (ACT). This task was developed by Pesquita et al. (2016) to assess sensitivity of attentional control during action prediction. On each trial, participants viewed a video clip of an actor reaching up to touch one of two buttons, either on the left or right side of the screen. The participants were told that they were playing a game in which their task was to "beat the actor" by predicting which button the actor would reach for. They were asked to respond as quickly as possible, but to also be as accurate as possible. Responses were made by pressing one of two keys, spatially mapped to the target location. On half the trials the actor chose the button they reached for ("chosen" trials) and on the other half of trials they were 'told' which button to press by a light off screen ("directed" trials). Each trial began with the participant resting their index fingers on the keyboard and their eyes on a fixation cross at the centre of the screen for 1-1.5s.

Each session began with 6 practice trials using an actor who was not used in the main test phase. For the present study, Pesquita et al.'s (2016) task was adapted by reducing the number of test trials from 400 to 200. This was done by selecting the 2 actors out of 4 that yielded the largest response time difference between chosen and directed trials and longest overall reaction times (actors 3 and 4 in the original study). The order of videos was counterbalanced across participants and presented in blocks of 50 trials (with the same actor in each block). Participants' sensitivity to the attentional state of the actors (e.g., endogenous vs. exogenous control) was computed based on differences in reaction time between the directed and chosen conditions. Participants' sensitivity to the end goal of the actor was determined based on overall accuracy across all trials.

Frith-Happé Animations (FHA). In the Frith-Happé animations task (Abell, Frith, & Happé, 2000), participants watched 12 short silent animations depicting 2 triangles moving across a white background. Each animation was approximately 1 minute in duration. In total there were 3 categories of animations, each with 4 trials: 1) random, 2) goal-directed (e.g., "fighting"), and 3) theory of mind (e.g., "teasing"). Of note are the "theory of mind" animations, which were designed to attributions of mental states or intentionality.

The task began with 3 familiarization trials in which participants were shown 1 animation for each possible category. Participants were asked to describe what they saw and feedback was given. During test trials, the presentation of each animation was followed by the same question, "what do you think was happening in the cartoon?". Verbal responses were recorded via the "speech-to-text" function on the iPad and errors in transcription were corrected in real time. For the "theory of mind" animations, participants were also asked to identify the emotion that each triangle was feeling from a list of 5 options after describing what they thought was happening in the scene. No feedback was given. Animations were presented in a pseudo-random order.

Responses were scored based on the coding scheme described in Abell et al. (2000), with responses rated 0, 1, or 2, based on their level of accuracy (i.e., the degree to which they met the intended meaning of the animation sequences). In addition, each response was assigned to 1 of 3 categories, independent of accuracy: 1) action (i.e., a simple action statement with no explicit mention of interaction between triangles), 2) interaction (i.e., a statement that mentioned interaction between triangles without psychological language or reference to a mental state), and 3) mentalizing (i.e., descriptions that included explicit psychological or mental state terms). In cases where a description could fit more than 1 category, the highest level of descriptive language was scored (i.e., mentalizing trumped interaction trumped action). For the purposes of the present study, participants' accuracy score, identification of emotion, and frequency of mentalization were used.

Communicative Interaction Database – 5AFC format (CID-5). The CID-5 is a set of point-light displays developed by Manera, von der Lühe, Schilbach, Verfaillie, and Becchio (2016) to assess an observer's ability to distinguish between the dynamics of communication and non-communicative action stimuli. The database of point-light displays consists of 21 actions performed by two agents, 14 of which are communicative and 7 of which are non-communicative individual actions.

On each trial, participants were presented with a looping video of the point-light display and asked to first decide whether the two agents were communicating or acting independently of each other (question 1, categorization), and then to select the correct action description among five response alternatives, presented in random order (question 2, action identification). A score of 1 was given for each correct response.

Autism Quotient (AQ). The Autism Quotient is a 50-item self-report questionnaire that assesses social aptitude (Baron-Cohen, Wheelwright, Skinner, et al., 2001). The questionnaire includes 5 subscales, each composed of 10 items: communication, social skills, imagination, attention to detail, and attention switching. Each item or question is scored with 1 point if the participant identifies the autistic-like behaviour either mildly or strongly (i.e., poor social skill, poor communication skills, poor imagination, exceptional attention to detail, difficulty with attention-switching). Approximately half the items are worded to produce a "disagree" response and half an "agree" response. Higher scores indicate lower self-reported social aptitude, with scores above 32 indicating potential clinically significant levels of autistic traits (though the tool is not diagnostic). For the purposes of the present study, participants' total score, and scores on the social skills, communication, and imagination subscales were used as they are most relevant to the construct of social understanding (see Appendix for subscale items).

Procedure

Participants were tested individually in a 1.5-2-hour session at the University of Victoria. Informed written consent was obtained prior to testing. Participation was voluntary and all participants received bonus course credit.

The task order was as follows: Attentional Control Task, Autism Quotient, Communicative Interaction Database, Third- and Fourth-order False Belief, Reading the Mind in Films, Frith-Happé Animations, and the Eyes Test. A fixed task order was chosen in order to separate tasks with similar cognitive demands and facilitate comparisons between tasks (see Carlson & Moses, 2001 for further justification).

Ethics Approval

The study was approved by the University of Victoria Human Research Ethics Board and was therefore performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki.

Results

Data Preparation

All variables were screened for univariate outliers, and for skewness and kurtosis. Response time (RTs) on the ACT were screened such that incorrect trials and responses > 3 SDs from the mean were excluded. No other outliers were detected in any of the other tasks. Table 2 displays the skewness and kurtosis for all variables. The FB scores and accuracy on the ACT were negatively skewed and leptokurtic. The FHA appropriateness scores, CID action score, and reaction times on the ACT were also leptokurtic. A review of individual scores and responses showed that participants seemed to be engaged in the tasks. As such, no score transformations were conducted. All other variables were reasonably normally distributed (i.e., standard scores of skewness and kurtosis were below 2).

Descriptive Statistics

Table 2 displays the descriptive statistics for all variables. There were no significant differences in performance on any tasks between females and males. The ACT had a smaller sample size due to program malfunction (n = 5) and 1 participant whose scores were excluded due to overall performance below chance level. In total this resulted in 1% of the data missing not at random. Missing data was not imputed.

Replication of the Attentional Control Task

Figure 1 shows the mean correct RT in the chosen and directed conditions overall. A paired-samples t-test showed that participants were reliably faster at correctly predicting the actor's movement in the chosen condition than in the directed condition [t(95) = 18.4, p <.001, Cohen's d = 1.87)]. Overall, participants responded correctly on 88.5% of trials (SEM = 0.88%).

To determine whether there was a relationship between participants' social aptitude and their sensitivity to the attentional state of the actors, z -scores were computed for each participant's correct RT values by subtracting each participant's RTs from the mean RTs of the participant to the corresponding actor, and dividing this difference by the SD of the participant's RTs for this actor. Each participant was then assigned a sensitivity score based on their mean difference in z-scores between the directed and chosen conditions. Contrary to Pesquita et al.'s (2016) finding, sensitivity was not correlated with social aptitude as measured by the Autism Quotient total score ($r(95) = -.14$, $p > .05$). Figure 2 shows this scatterplot.

The Relation between Different Forms of Social Understanding

To examine the relationship between different forms of social understanding, correlations amongst all tasks was examined. Table 2 displays this correlation matrix. There were significant correlations among some of tasks that assessed perceptual sensitivity, among some tasks developed the assess "theory of mind" or mental state attribution, and across these measures. Specifically, overall accuracy on the ACT was correlated with accuracy on the FHA ($r = .30$, $p < .01$) and negatively correlated with self-reported social competence as measured by the AQ total score ($r = -.20$, $p = .05$) and social skills subscale score ($r = -.20$, $p < .05$). Sensitivity on the ACT was also correlated with accuracy on the FHA ($r = .28$, $p < .01$), action identification in the CID ($r = .21$, $p < .05$), and with performance on the Eyes Test ($r = .20$, $p < .05$). Finally, action identification on the CID was also correlated with emotion identification on the FHA (r = .30, p < .01). Among the ToM tasks, performance on the RME and RMF were correlated (r = .26, p < .01), and the imagination subscale score on the AQ was negatively correlated with the RMF (r = -.30, p < .01).

To determine the latent structure underlying the data, an exploratory factor analysis (EFA) was conducted on 13 measures in jamovi (The jamovi project, 2019). The FB task was dropped due to potential ceiling effects and lack of correlations with any of the other measures. The Kaiser-Meyer-Olkin measure verified sampling adequacy for the analysis, KMO > .60 ("mediocre" but "acceptable" according to Kaiser, 1974), and all but 2 of the KMO values for the individual items were > than the acceptable limit of .50 (the KMO values for the RMF and FHA – Mentalizing were just below .40). Bartlett's test of sphericity, χ^2 (78) = 368, p < .001, indicated that correlations between items were sufficiently large for EFA. Parallel analysis with maximum likelihood extraction method with oblimin rotation yielded a 4-factor solution that accounted for 48% of the total variance. Model fit indices suggested that this model provided good fit to the data, χ^2 (32) = 43.1, p = .09, Tucker-Lewis index = 0.90, and root mean square error of approximation = .06 [.00, .08]. Table 4 shows the factor loadings after rotation. The items that cluster on the same factor suggest that factor 1 represents self-reported social aptitude, factor 2 sensitivity to communicative intentions in point-light displays, factor 3 perception of emotion and intention in others, and factor 4 attribution of mental states to others. Factors 1 and 4, and factors 2 and 3 were found to be correlated with each other (*r*s .32 and .27 respectively).

Discussion

The aim of the present study was to take a first step in clarifying the relationship between different forms of social understanding by including a variety of different tasks used across different empirical domains that varied along a number of dimensions (i.e., modality of stimulus presentation and response, valence of stimuli or response, as well as the perspective engaged by the participant). Overall, I found coherence among tasks that used dynamic visual stimuli that may assess perceptual sensitivity to social information, and a multi-factor structure that generally supports divergence in different forms of social understanding. The findings will be described in more detail below.

Replication of the Pesquita et al. (2016) study

First, the main finding from Pequsita et al.'s (2016) study on perceptual sensitivity to attentional control was replicated: observers were reliably faster in predicting an actor's direction of reach when the actor chose the direction compared to when their action was directed. In contrast to Pesquita el al.'s study, participants' sensitivity scores were not correlated with their self-reported social competence, as measured by the AQ total score. The scatterplots from both studies showed similar ranges in the sensitivity and AQ scores (i.e., there was a large degree of variability in both sensitivity and AQ scores in both studies); however, when looking at the relationship between these variables there was a wider spread across the scores and potentially more "outliers" in our sample. This could be due to the lower number of experimental trials in our study (200 trials compared to 400 trials), as it is possible that sensitivity could increase with greater exposure.

Although sensitivity was not correlated with social competence in our sample, we did find a significant negative correlation between accuracy on the ACT task (which was strongly correlated with sensitivity) and the AQ total score and social skills subscale score. Accuracy on the ACT task has been interpreted as participants' sensitivity to the end goal of the actor (Pesquita et al., 2016). Thus, in our sample, although sensitivity to the subtle kinematic cues that distinguished between endogenous and exogenous attentional control was not related to social competence, greater sensitivity to the end goal of the other person's actions was. Although the AQ includes 5 different subscales that capture different aspects of social aptitude, ACT accuracy was only correlated with the social skills subscale. One explanation for this finding could be that action prediction may be most relevant for social understanding when interacting with others or in second-person contexts. This may be reflected in the specific questions that are part of the social skills subscale of the AQ (e.g., "I find it easy to work out what someone is thinking or feeling by just looking at their face" or "I find social situations easy"; see Appendix).

Direct, Non-inferential, and Perceptual Processes

To examine the relationship between different forms of social understanding, particularly those based on the inferential and perceptual processes described by ToM and interactionist approaches respectively, I examined the pattern of correlations across all tasks and examined the latent structure underlying the different measures.

Although ACT sensitivity was not related to AQ total score, it was correlated with other measures of social understanding, namely accuracy in predicting the intended meaning of animations on the FHA task, action identification on the CID-5, and emotion recognition on the Eyes test. Accuracy on the ACT was also correlated with FHA accuracy, while FHA emotion identification was related to action identification on the CID. In addition, the EFA yielded a 4-factor solution in which the FHA and ACT measures loaded onto a single factor; this factor was also correlated with the CID factor.

From an interactionist perspective, these findings support the idea that (implicit) perceptual sensitivity to subtle kinematic cues in the actions of others may be related to more explicit forms of social understanding that are also dependent on perception of movement. In addition, I think these findings provide support for the IT claim that we can directly perceive (some) mental states or emotions in behaviour. Given that speeded responses were required in the ACT task, it is unlikely that any kind of (explicit) theorizing or simulating could be used to make predictions regarding the direction of the actor's movements. Additionally, previous findings suggest that participants are not able to consciously distinguish between these different forms of attentional control (Pesquita et al., 2016). The finding that ACT sensitivity was correlated to performance on the Eyes Test suggests that emotion recognition based on static images of eyes may, in part, be dependent on direct perceptual processes. To me this is not too surprising given that there is no additional contextual information that can be used by participants to "infer" the emotional states depicted on the Eyes Test. However, future studies that examine the phenomenology of completing this task may shed light on whether participants utilize any simulative or inferential strategies. In addition, presenting images of the eyes rapidly or requiring a speeded response may help to elucidate whether non-inferential processes are sufficient for emotion recognition on this task.

Inferential Mentalizing Processes

In the present study, 3 tasks typically described as measures of "theory of mind" were also included. Although a correlation between performance on the RME and RMF tasks was found, these tasks were minimally related to the other measures of social understanding. Interestingly, RMF scores were negatively correlated with the imagination subscale of the AQ. This subscale of the AQ mostly concerns individuals' aptitude or preference for creativity, pretence, and imagination; items on this subscale include "If I try to imagine something, I find it very easy to create a picture in my mind" and "When I'm reading a story, I find it very difficult to work out the characters' intentions" (see Appendix). This suggests that the RMF task may indeed assess participants' ability to simulate or theorize about others' intentions. However, it is unclear whether this ability is important for social interaction (for example, AQ imagination scores were not related to AQ social skills scores or any other social understanding tasks).

The final ToM task, higher false belief understanding, was not related to any of the other variables. As described above, it is possible that the higher-order false belief task that I included was too simple for adults and suffered from a restricted range; with little variability, it is possible that correlations with other measures could not be detected (Bland & Altman, 2011).

Another possibility for the lack of coherence among ToM tasks is that these tasks actually measure distinct forms of social cognition and/or have high linguistic or cognitive demands that lead to the task impurity problem (e.g., Miyake et al., 2000). As such, although all of the tasks may be measuring “social understanding”, they may not show statistically significant relationships because they also assess “non-social-cognitive” processes (e.g., comprehension of complex syntax or visual acuity) that are specific to the task but unrelated to the overall construct. To rule out this possibility, future studies using larger sample sizes and a wider battery of tasks could utilize confirmatory factor analysis to compare various factor models that differ based on task characteristics or theoretically-driven models based on the forms of social understanding described by TT, ST, and IT.

Although there have been very few studies that have utilized a wide battery of measures in a single sample of adults, one recent study by Warnell and Redcay (2019) found no coherence among different theory of mind measures in a sample of adults. The tasks that were used included the RME, a higher-order ToM task that is similar to our FB task, Apperly and colleagues’ Belief Reasoning task, the Spontaneous Theory of Mind Protocol (similar to RMF but using spontaneous written descriptions instead of an alternate forced choice response), and a measure of pragmatic language comprehension. Warnell and Redcay interpreted the lack of coherence as reflecting the diversity of the ToM construct, suggesting that mental state understanding may not be a single representational ability. Instead, they argue that there are likely diverse ways in which we can understand the minds of others in real-world interactions. This suggestion is supported by the findings from this study and reiterates the importance of looking beyond conceptual forms of social understanding.

The lack of coherence among ToM measures also speaks to the importance of task selection in experimental work. As Warnell and Redcay (2019) point out, many studies that purport to measure “social cognition” either as a dependent or independent variable will often only utilize one measure, typically a ToM task that assesses mental state attribution from a third-person stance. As the findings of this study show, there are a diversity of ways in which we can “understand” others, and it is important to clarify what one intends or means when using the term “theory of mind”. As argued above, it is important to recognize that theory-use is just one limited way in which we make sense of others’ behaviour, and may not be the default mode of social understanding when interacting with others.

Limitations and Future Directions

One of the main limitations of this study is that is relies on a correlational method to examine the relationship between different forms of social understanding. As such, the lack of a statistical relationship between variables is difficult to interpret. On the one hand, a lack of correlation between ToM tasks and perceptual tasks may indicate that these are indeed distinct forms of social understanding, but on the other hand, one could argue that they are assessing completely unrelated abilities or constructs.

More generally, the lack of coherence among different measures of social understanding may reflect what Gallagher (2015; 2016) has described as the “clunky robot" problem in scientific explanation. Drawing an analogy from how engineers try to design robots, many different teams are tasked with figuring out the sub-functions of the robot based on their area of expertise – for example, computer vision scientists work on developing a sensory system, while control engineers may work on designing a motor control system, etc. When these groups come together to connect their systems, what often results is a clunky robot that doesn't function in the smooth and seamless way as predicted. Arguably, this occurs because no group was tasked with determining the dynamic relationships between the different subsystems, or how all the pieces should fit together.

According to Gallagher (2016), this problem is also applicable to scientific investigation and explanation. Although it is important for us, as empirical researchers, to decompose certain phenomena into smaller parts so they can be studied one piece at a time, we neglect to study the dynamics of how these processes are related to one another, and what results are clunky explanations. In the context of social cognition, this criticism has been clearly articulated by those working from an IT perspective. For example, De Jaegher and Di Paolo (2007) have argued that theories of social understanding cannot be reduced to explanations based on the cognitive processes within individuals. Instead, the interactive context within which social understanding occurs, and the features of the second-person interaction itself are constitutive and need to be accounted for in our theories. Thus, although the present study aimed to capture a more pluralistic and comprehensive picture of social understanding beyond "social cognition", future studies should utilize truly interactive tasks that allow for different aspects of the interaction to be measured. Perhaps one way to do so is to employ more qualitative or phenomenological methods such as those described by De Jaegher, Pieper, Clénin, and Fuchs (2017). By studying the dynamics of interactions, we may also develop more precise accounts of how different forms of social understanding are utilized and how they relate to one another.

Developmental Implications

Although the present study examined individual differences in social understanding in adults, there are important developmental implications based on this work. As described above, adults are able to use inference to theorize about or simulate mental states; however, we cannot automatically impose this adult view of the mind of pre-linguistic infants and children whose subjective life may be very different. Furthermore, if the ability to simulate or use psychological theories is an outcome of development, it cannot be the cause for development (Carpendale et al., 2013; Zahavi, 2008). Thus, although children's conceptual knowledge of others (i.e., their 'theory of mind') is an important aspect of social understanding, it is not the only way in which they can engage with others and cannot be the ontogenetic foundation for social development. Interestingly, infants' ability to "mentalize" is currently an area of much debate and active research (see, for example, Sabbagh & Paulus, 2018). However, most of this research is conducted within a ToM framework, which limits social understanding to having implicit or explicit "theories of mind" or conceptual understanding of mentality. The findings from this study support the idea that there are non-inferential ways of understanding others, which remains an unexplored area of research within the developmental literature.

Given the conceptual concerns inherent in TT and ST, I think that it would be informative to approach the empirical investigation of social-cognitive development from an IT perspective. IT also fits well constructivist approaches to development, more generally, in which the developing child constructs meaning or knowledge through interaction with others and the world (see Carpendale & Lewis, 2015; Carpendale, Lewis, & Müller, 2018). Constructivist approaches are based on a relational developmental systems perspective in which “mind” and meaning are derived from interaction and activity (Carpendale & Lewis, 2015). Whereas TT, ST, and general ToM approaches begin with the individual and presume that some sort of understanding of minds is required *for* interacting with others, this alternative perspective contends that social understanding emerges out of social interaction (Mead, 1934; Carpendale & Lewis, 2015). Rather than meaning or knowledge being an “internal representation” of the “external” world, knowledge is construed as being constructed through interacting with others and the world. As such, the emphasis is on the social process and the developmental continuity between different forms of social engagement. Thus, in following this approach, it is important to consider the different ways in which children can understand others and the different psychological processes (i.e., affective, perceptual, conceptual) they may bring to bear in their everyday social interactions Furthermore, it is also important to consider the developmental sequence in which these forms of social understanding may emerge, and whether certain later developing forms of social cognition may be dependent on earlier emerging forms of social understanding. The findings from this study suggest that it may be important to consider more direct, perceptual ways of understanding others.

Although most developmental research has focused on the conceptual understanding of others, there is a body of work that points to an affective basis for social understanding (see Hobson, 1993). For example, Reddy and colleagues have argued that a rudimentary form of social understanding can be observed in the propensity for young infants to engage in imitation ‘games’ and proto-conversations with their caregivers (e.g., Reddy & Morris, 2004; Rossmanith & Reddy, 2016). Far from being random sequences of actions, these interactions are characterized by an intimate engagement between the infant and caregiver, in which both parties mutually regulate the affective and behavioural responses of the other. In these interactions, infant and caregiver engage with one another using their whole bodies and through different modalities (e.g., gaze, touch, vocalization, movement) (Rossmanith & Reddy, 2016). Importantly, these smooth and dynamic interactions illustrate how *meaning* can emerge through affective engagement, which does not require any epistemic or cognitive relation between the infant and caregiver (Reddy & Morris, 2004). Developmentally, these action-based and affective forms of social understanding may serve as the basis for the acquisition of more conceptual or epistemic “theory of mind” knowledge. Interestingly, there is little to no research examining the social perceptual abilities of children. However, one recent study by Rice, Anderson, Velnoskey, Thompson, and Redcay (2016) found that perception of biological motion in children between 7-12 years of age was related to performance on tasks used to evaluate social-cognitive competence (the Eyes test and Strange Stories). Future studies that examine whether young children show perceptual sensitivity to the kinematics of others' actions, and the developmental course of this ability as it relates to other forms of social cognition would shed light on whether our mental state concepts are founded on a more basic perceptual sensitivity to the "mindedness" of others.

Conclusion

The aim of this study was to clarify the relationship between the different ways in which people can understand others. Although the dominant accounts of social cognition focus on our conceptual understanding of other minds and the inferential processes involved in mental state attribution, there may be a more foundational or basic form of social understanding based on perceptual processes. In the present study, correlations showed coherence amongst measures that assessed participants' perceptual sensitivity to social information, with minimal coherence across "theory of mind" tasks. Explanatory factor analysis conducted on 13 measures yielded a meaningful 4 factor solution that supported the distinction between conceptual or inferential measures and more direct, perceptual forms of social understanding. Overall, the findings from this study highlight the importance of considering the variety of ways in which we can understand others and provides empirical support for a more pluralistic and holistic account of social understanding.

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.

Adolphs, R. (2009). The social brain: neural basis of social knowledge. *Annual review of Psychology*, *60*, 693-716.

Apperly, I. A. (2013). Can theory of mind grow up? Mindreading in adults, and its implications for the neuroscience and development of mindreading. In S. Baron-Cohen, H. Tager-Flusberg, M. Lombardo (Eds.), *Understanding other minds* (3rd ed.), Oxford University Press, Oxford, UK (2013), pp. 72-92.

Apperly, I. A., Warren, F., Andrews, B. J., Grant, J., & Todd, S. (2011). Developmental continuity in theory of mind: Speed and accuracy of belief–desire reasoning in children and adults. *Child Development*, *82*, 1691-1703.

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

Avramides, A. (2019). Other minds. In E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy.* Retrieved from https://plato.stanford.edu/archives/sum2019/entries/other-minds.

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. *Journal of Child Psychology and Psychiatry*, 42, 241–251

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

Barr, D. J., & Keysar, B. (2005). Mindreading in an Exotic Case: The Normal Adult Human. In B. F. Malle & S. D. Hodges (Eds.), *Other minds: How humans bridge the divide between self and others* (pp. 271-283). New York, NY, US: Guilford Press.

Becchio, C., Manera, V., Sartori, L., Cavallo, A., & Castiello, U. (2012). Grasping intentions: from thought experiments to empirical evidence. *Frontiers in Human Neuroscience*, *6*, 117.

Becchio, C., Sartori, L., Bulgheroni, M., & Castiello, U. (2008). The case of Dr. Jekyll and Mr. Hyde: a kinematic study on social intention. *Consciousness and Cognition*, *17*(3), 557-564.

Bennett, J. (1978). Some remarks about concepts. *Behavioral and Brain Sciences*, *1,* 557-560.

Bland, J. M., & Altman, D. G. (2011). Correlation in restricted ranges of data. *BMJ*, *342*, d556.

Bohl, V., & van den Bos, W. (2012). Toward an integrative account of social cognition: marrying theory of mind and interactionism to study the interplay of Type 1 and Type 2 processes. *Frontiers in Human Neuroscience*, *6*, 274.

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72,* 1032-1053.

Carpendale, J. I., Atwood, S., & Kettner, V. (2013). Meaning and mind from the perspective of dualist versus relational worldviews: Implications for the development of pointing gestures. *Human Development*, *56*(6), 381-400.

Carpendale, J. I., & Chandler, M. J. (1996). On the distinction between false belief understanding and subscribing to an interpretive theory of mind. *Child Development*, *67*(4), 1686-1706.

Carpendale, J. I., & Lewis, C. (2004). Constructing an understanding of mind: The development of children's social understanding within social interaction. *Behavioral and Brain sciences*, *27*, 79-96.

Carpendale, J. I., & Lewis, C. (2015). The development of social understanding. In U. Müller & L. Liben (Vol. Eds.), R. M. Lerner (Series Ed.), *Handbook of Child Development and Developmental Science, Vol. 2: Cognitive Processes* (pp. 381-422). Hoboken, NJ: Wiley.

Carpendale, J. I., Lewis, C., & Müller, U. (in press, 2018). *The development of children's thinking: Its social and communicative foundations.* Los Angeles: Sage.

Carpendale, J. I., & Chandler, M. J. (1996). On the distinction between false belief understanding and subscribing to an interpretive theory of mind. *Child Development*, *67*, 1686-1706.

Castelli, I., Baglio, F., Blasi, V., Alberoni, M., Falini, A., Liverta-Sempio, O., ... & Marchetti, A. (2010). Effects of aging on mindreading ability through the eyes: an fMRI study. *Neuropsychologia*, *48*, 2586-2594.

Castelli, F., Happé, F., Frith, U., & Frith, C. (2000). Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *Neuroimage*, *12*, 314-325.

De Jaegher, H., & Di Paolo, E. (2007). Participatory sense-making. *Phenomenology and the Cognitive Sciences*, *6*(4), 485-507.

De Jaegher, H., Di Paolo, E., & Gallagher, S. (2010). Can social interaction constitute social cognition?. *Trends in Cognitive Sciences*, *14*(10), 441-447.

De Jaegher, H., Pieper, B., Clénin, D., & Fuchs, T. (2017). Grasping intersubjectivity: An invitation to embody social interaction research. *Phenomenology and the Cognitive Sciences*, *16*, 491-523.

Dennett, D. C. (1978). Beliefs about beliefs. *Behavioral and Brain Sciences*, 4, 568-570.

Drayson, Z. (2018). Direct perception and the predictive mind. *Philosophical Studies*, *175*(12), 3145-3164.

Frith, C. D., & Frith, U. (2008). Implicit and explicit processes in social cognition. *Neuron*, *60*, 503-510.

Fuchs, T., & De Jaegher, H. (2009). Enactive intersubjectivity: Participatory sense-making and mutual incorporation. *Phenomenology and the Cognitive Sciences*, *8*(4), 465-486.

Gallagher, S. (2007). Simulation trouble. *Social Neuroscience*, *2*, 353-365.

Gallagher, S. (2008). Direct perception in the intersubjective context. *Consciousness and Cognition*, *17*(2), 535-543.

Gallagher, S. (2012). In defense of phenomenological approaches to social cognition: interacting with the critics. *Review of Philosophy and Psychology*, *3*(2), 187-212.

Gallagher, S. (2016). Intercorporeity: Enaction, simulation, and the science of social cognition. In J. Reynolds & R. Sebold (Eds.), *Phenomenology and science* (pp. 161-179). Palgrave Macmillan, New York.

Gallagher, S. (2015). The new hybrids: continuing debates on social perception. *Consciousness and cognition*, *36*, 452-465.

Gallese, V. (2001). The 'shared manifold' hypothesis. From mirror neurons to empathy. *Journal of Consciousness Studies*, *8*, 33-50.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, *2*, 493-501.

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, *8*, 396-403.

German, T. P., & Hehman, J. A. (2006). Representational and executive selection resources in 'theory of mind': Evidence from compromised belief-desire reasoning in old age. *Cognition*, *101*, 129-152.

Gibson, J. J. (1967). New reasons for realism. *Synthese*, 17, 162 - 172.

Gibson, J. J. (1979). The theory of affordances. The ecological approach to visual perception. Boston, MA: Houghton Mifflin.

Golan, O., Baron-Cohen, S., Hill, J. J., & Golan, Y. (2006). The "reading the mind in films" task: complex emotion recognition in adults with and without autism spectrum conditions. *Social Neuroscience, 1*, 111-123.

Goldman, A. I. (1992). In defense of the simulation theory. *Mind & language*, *7*, 104-119.

Goldman, A. I. (1993). The psychology of folk psychology. *Behavioural and Brain Sciences*, *16,* 15–28.

Goldman, A. I. (2002). Simulation theory and mental concepts. In J. Dokic & J. Proust (Eds.), *Advances in consciousness research, Vol. 45. Simulation and knowledge of action* (pp. 1-19). Amsterdam, Netherlands: John Benjamins Publishing Company.

Gopnik, A. (1993). Theories and illusions. *Behavioural and Brain sciences*, *16*(1), 90-100.

Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind & Language*, *7*(1-2), 145-171.

Gordon, R. M. (1996). Radical simulationism. In P. Carruthers & P. K. Smith (Eds.), *Theories of theories of mind* (pp. 11–21). Cambridge, UK: Cambridge University Press.

Happé, F., Cook, J. L., & Bird, G. (2017). The structure of social cognition: In(ter)dependence of sociocognitive processes. *Annual Review of Psychology*, *68*, 243-267.

Harman, G. (1978). Studying the chimpanzee's theory of mind. *Behavioral and Brain Sciences*, *1,* 576-577.

Harris, P. L. (1991). The work of the imagination. In A. Whiten (Ed.), *Natural theories of mind: Evolution, development and simulation of everyday mindreading* (pp. 283–304). Oxford: Basil Blackwell.

Helmholtz, H. V. (1878). The origin and meaning of geometrical axioms. *Mind*, *3*, 212-225.

Heyes, C. (2010). Where do mirror neurons come from?. *Neuroscience & Biobehavioral Reviews*, *34*, 575-583.

Hickok, G. (2013). Do mirror neurons subserve action understanding?. *Neuroscience Letters*, *540*, 56-58.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*, 31-36.

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

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, *89*, 25-41.

Keysers, C., & Gazzola, V. (2009). Expanding the mirror: vicarious activity for actions, emotions, and sensations. *Current Opinion in Neurobiology*, *19*, 666-671.

Kilner, J. M., Neal, A., Weiskopf, N., Friston, K. J., & Frith, C. D. (2009). Evidence of mirror neurons in human inferior frontal gyrus. *Journal of Neuroscience*, *29*(32), 10153-10159.

Krueger, J., & Overgaard, S. (2012). Seeing subjectivity: Defending a perceptual account of other minds. *ProtoSociology, 47,* 239-262.

Lalonde, C. E., & Chandler, M. J. (2002). Children's understanding of interpretation. *New Ideas in Psychology*, *20*, 163-198.

Leslie, A. M. (1994). ToMM, ToBy, and Agency: Core architecture and domain specificity. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the Mind: Domain Specificity in Cognition and Culture* (pp.119-148). New York: Cambridge University Press.

Liddle, B., & Nettle, D. (2006). Higher-order theory of mind and social competence in school-age children. *Journal of Cultural and Evolutionary Psychology*, 4, 231–244.

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*, 551-556.

Loftus, G. R., & Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, *1*, 476-490.

Mahy, C. E., Moses, L. J., & Pfeifer, J. H. (2014). How and where: Theory-of-mind in the brain. *Developmental Cognitive Neuroscience*, *9*, 68-81.

Manera, V., Becchio, C., Cavallo, A., Sartori, L., & Castiello, U. (2011). Cooperation or competition? Discriminating between social intentions by observing prehensile movements. *Experimental Brain Research*, *211*, 547-556.

Manera, V., von der Lühe, T., Schilbach, L., Verfaillie, K., & Becchio, C. (2016). Communicative interactions in point-light displays: Choosing among multiple response alternatives. *Behavior Research Methods*, *48*, 1580-1590.

Marsh, K. L., Johnston, L., Richardson, M. J., & Schmidt, R. C. (2009). Toward a radically embodied, embedded social psychology. *European Journal of Social Psychology*, *39*, 1217-1225.

Mead, G. H. (1934). *Mind, self and society*. Chicago: University of Chicago Press.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex 'frontal lobe' tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49-100.

Molenberghs, P., Cunnington, R., & Mattingley, J. B. (2012). Brain regions with mirror properties: a meta-analysis of 125 human fMRI studies. *Neuroscience & Biobehavioral Reviews*, *36*, 341-349.

Molenberghs, P., Johnson, H., Henry, J. D., & Mattingley, J. B. (2016). Understanding the minds of others: A neuroimaging meta-analysis. *Neuroscience & Biobehavioral Reviews*, *65*, 276-291.

Musholt, K. (2018). The personal and the subpersonal in the theory of mind debate. *Phenomenology and the Cognitive Sciences*, *17*, 305-324.

Noë, A. (2004). *Action in perception*. Cambridge, MA: MIT press.

Osterhaus, C., Koerber, S., & Sodian, B. (2016). Scaling of advanced theory-of-mind tasks. *Child Development*, *87*(6), 1971-1991.

Overgaard, S. (2017). Enactivism and the perception of others' emotions. *Midwest Studies in Philosophy*, *41*, 105-129.

Pesquita, A., Chapman, C. S., & Enns, J. T. (2016). Humans are sensitive to attention control when predicting others' actions. *Proceedings of the National Academy of Sciences*, *113*(31), 8669-8674.

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

Reddy, V., & Morris, P. (2004). Participants don't need theories: knowing minds in engagement. *Theory & Psychology*, *14*(5), 647-665.

Rice, K., Anderson, L. C., Velnoskey, K., Thompson, J. C., & Redcay, E. (2016). Biological motion perception links diverse facets of theory of mind during middle childhood. *Journal of Experimental Child Psychology*, *146*, 238-246.

Rossmanith, N., & Reddy, V. (2016). Structure and Openness in the Development of Self in Infancy. *Journal of Consciousness Studies*, *23*(1-2), 237-257.

Sabbagh, M. A., & Paulus, M. (2018). Replication studies of implicit false belief with infants and toddlers. *Cognitive Development, 46,* 1-3.

Sartori, L., Becchio, C., & Castiello, U. (2011). Cues to intention: the role of movement information. *Cognition*, *119*(2), 242-252.

Saxe, R., (2010). The right temporo-parietal junction: a specific brain region for thinking about thoughts. In: Leslie, A., German, T. (Eds.), Handbook of Theory of Mind.

Saxe, R., Kanwisher, N., (2003). People thinking about thinking people. The role of the temporo-parietal junction in theory of mind. *NeuroImage, 19*, 1835–1842.

Scheler, M. (1954). *The nature of sympathy* (Peter Heath, Trans.) London: Routledge and Kegan Paul.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). A second-person neuroscience in interaction. *Behavioural and Brain Sciences*, *36*(4), 441-462.

Taylor, M. (1988). Conceptual perspective taking: Children's ability to distinguish what they know from what they see. *Child Development*, *59,* 703-718.

The jamovi project (2019). *jamovi* (Version 0.9) [Computer Software]. Retrieved from https://www.jamovi.org

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

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience.* Cambridge, MA: MIT Press.

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

Wellman, H. M. (1992). *The child's theory of mind*. Cambridge, MA, US: The MIT Press.

Wellman, H. M. (2018). Theory of mind: The state of the art. *European Journal of Developmental Psychology*, *15*, 728-755.

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, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development, 75*, 523-541.

Wiltshire, T. J., Lobato, E. J., McConnell, D. S., & Fiore, S. M. (2015). Prospects for direct social perception: a multi-theoretical integration to further the science of social cognition. *Frontiers in Human Neuroscience*, *8*, 1007.

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.

Zahavi, D. (2005). *Subjectivity and self-hood: Investigating the first-person perspective.* Cambridge: MIT Press.

Zahavi, D. (2008). Simulation, projection and empathy. *Consciousness and Cognition*, *17*(2), 514-522.

Zahavi, D. (2011). Empathy and direct social perception: a phenomenological proposal. *Review of Philosophy and Psychology*, *2*(3), 541-558.

Zahavi, D. (2014). *Self and other: Exploring subjectivity, empathy, and shame*. Oxford University Press, USA.

Appendix

Autism Quotient Subscale Questions

Social Skills Subscale Items

- I prefer to do things with others rather than on my own
- I find social situations easy
- I would rather go to a library than a party
- I find myself drawn more strongly to people than to things
- I find it hard to make new friends
- I find it easy to work out what someone is thinking or feeling just by looking at their face
- I enjoy social occasions
- I find it difficult to work out people's intentions
- I enjoy meeting new people
- I am a good diplomat

Communication Subscale Items

- Other people frequently tell me that what I've said is impolite, even though I think it is polite
- I enjoy social chit-chat
- When I talk, it isn't always easy for others to get in a word edgeways
- I frequently find that I don't know how to keep a conversation going
- I find it easy to "read between the lines" when someone is talking to me
- I know how to tell if someone listening to me is getting bored
- When I talk on the phone, I'm not sure when it's my turn to speak
- I am often the last to understand the point of a joke
- I am good at social chit-chat
- People often tell me that I keep going on and on about the same thing

Imagination Subscale Items

- If I try to imagine something, I find it very easy to create a picture in my mind
- When I'm reading a story, I can easily imagine what the characters might look like
- I find making up stories easy
- When I'm reading a story, I find it difficult to work out the characters' intentions
- I don't particularly enjoy reading fiction
- I would rather go to the theatre than a museum
- When I was young, I used to enjoy playing games involving pretending with other children
- I like to collect information about categories of things (e.g., types of car, types of bird, types of train, types of plant, etc.)
- I find it difficult to imagine what it would be like to be someone else
- I find it very easy to play games with children that involve pretending

Table 1. Task parameters across the different measures of social understanding

Task	Stimulus Modality	Stimulus Type	Contextual Information	Response Modality	Perspective Assessed	Form of Social Understanding
Third- and Forth Order False Belief Liddle and Nettle (2006)	Verbal	Story	Yes	Explicit Forced choice	Third	Cognitive Inferential
Reading the Mind in the Eyes Test Baron-Cohen, Wheelwright, Hill, Raste, and Pumb (2001)	Visual	Static image	No	Explicit Forced choice	First	Affective Perceptual
Reading the Mind in Films Test Golan, Baron-Cohen, Hill, and Golan, 2006	Visual	Dynamic video	Yes	Explicit Forced choice	Third	Affective Inferential
Frith-Happé Animations Abell, Frith, and Happé (2000)	Visual	Dynamic animation	No	Implicit Motor	Third	Cognitive Affective Perceptual
Attentional Control Task Pesquita, Chapman, and Enns (2016)	Visual	Dynamic video	No	Explicit Verbal	Second	Perceptual
Communicative Interaction Database Manera, von der Lühe, Schilbach, Verfaillie, & Becchio (2016)	Visual	Dynamic point light displays	No	Explicit Forced Choice	Third	Cognitive
Autism Quotient Baron-Cohen, Wheelwright, Skinner, Martin, and Clubley (2001)	-	-	-	-	-	-

Table 2. Descriptive Statistics for Measures of Social Understanding

Variable	*n*	*M*	*SD*	Range	Skew. (*SE*)	Kurt. (*SE*)
Higher-Order False Belief total correct	102	3.67	0.55	2 – 4	-1.42 (.24)	1.11 (.47)
Reading the Mind in Eyes total correct	102	28.40	3.27	21 – 36	0.11 (.24)	-0.42 (.47)
Reading the Mind in Film total correct	102	13.40	2.42	7 – 19	-0.06 (.24)	0.13 (.47)
FHA – Emotion total correct	102	6.14	1.31	2 – 8	-0.79 (.24)	0.86 (.47)
FHA – Appropriateness score	102	19.40	2.81	9 – 24	-0.99 (.24)	1.36 (.47)
FHA – Mental Attributions	102	2.69	1.32	0 – 6	-0.16 (.24)	-0.60 (.47)
CID – Category total correct	102	19.10	1.37	15 – 21	-0.67 (.24)	0.30 (.47)
CID – Action total correct	102	17.30	2.20	10 – 21	-0.98 (.24)	1.19 (.47)
ACT – Chosen RT (ms)	96	649	42.0	520 – 769	-0.25 (.25)	1.06 (.49)
ACT – Directed RT (ms)	96	675	45.70	500 – 792	-0.66 (.25)	2.43 (.49)
ACT – Overall accuracy (% correct)	96	88.5	8.63	59 – 99	-1.36 (.25)	1.43 (.49)
AQ – Social Skills score	102	2.62	2.38	0 – 9	0.87 (.24)	0.10 (.47)
AQ – Communication score	102	2.48	2.04	0 – 8	0.88 (.24)	0.34 (.47)
AQ – Imagination score	102	2.15	1.50	0 – 7	0.81 (.24)	0.82 (.47)
AQ – Total score	102	17.7	6.30	6 – 37	0.62 (.24)	0.24 (.47)

Table 3. Correlation Matrix (*N* = 102)

	FB	RME	RMF	FHA-E	FHA-A	FHA-M	CID-C	CID-A	ACT-S	ACT-A	AQ-SS	AQ-C	AQ-I
FB	-												
RME	.09	-											
RMF	.12	.26**	-										
FHA-E	.04	.03	.03	-									
FHA-A	-.05	.17	-.00	.30**	-								
FHA-M	-.13	-.06	-.04	.21*	.33**	-							
CID-C	.00	.10	-.05	.11	.12	.02	-						
CID-A	.04	.17	-.08	.30**	.18	-.07	.63***	-					
ACT-S	-.02	.20*	.00	.23	.28**	-.12	.17	.21*	-				
ACT-A	.02	.01	.13	.18	.30**	.18	.04	.08	.44***	-			
AQ-SS	-.05	-.12	.06	-.07	.10	.03	-.12	-.08	-.16	-.20*	-		
AQ-C	-.11	-.06	.02	-.01	.02	.10	-.15	-.09	-.13	-.15	.68***	-	
AQ-I	.02	-.12	-.30**	.07	-.04	-.08	-.12	-.05	-.09	-.07	.13	.22*	-
AQ-T	-.10	-.12	-.03	-.01	.07	.05	-.18	-.12	-.16	-.20†	.77***	.78***	.41***

Note. **FB** = Higher-Order False Belief; **RME** = Reading the Mind in the Eyes; **RMF** = Reading the Mind in Films; **FHA-E**= Frith-Happé Animation-Emotion Attribution; **FHA-A** = Frith-Happé Animation-Appropriateness; **FHA-M** = Frith-Happé Animation-Mental Attribution; **CID-C** = Communicative Interaction Database-Categorization; **CID-**A = Communicative Interaction Database-Action; **ACT-S** = Attentional Control Task-Sensitivity; **ACT-A** = Attentional Control Task-Accuracy; **AQ-SS** = Autism Quotient-Social Skills; **AQ-C** = Autism Quotient-Communication; **AQ-I** = Autism Quotient-Imagination; **AQ-T** = Autism Quotient Total

† p = .05, * p < .05, ** p < .01, *** p < .001

Table 4. Factor Loadings after Exploratory Factor Analysis (*N* = 96)

	Factor			
	1	**2**	**3**	**4**
RME				
RMF				-.32
FHA-Emotion			.39	
FHA-Appropriateness			.62	
FHA-Mental Attributions			.36	
CID-Category		.57		
CIA-Action		1.00		
ACT-Sensitivity			.46	
ACT-Accuracy			.60	
AQ-Social Skills	.86			
AQ-Communication	.83			
AQ-Imagination				.97
AQ-Total	.91			

Note. Maximum likelihood extraction method was used in combination with an oblimin rotation.

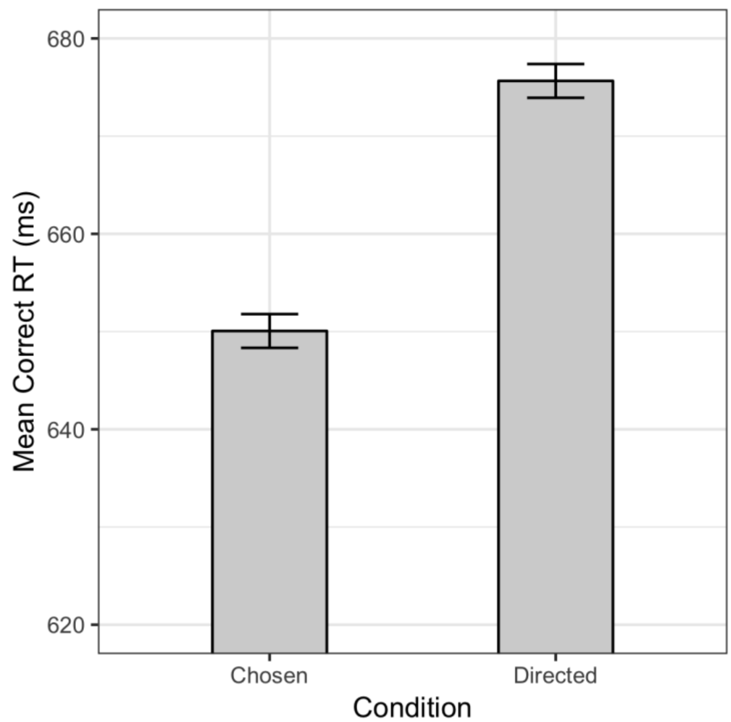

Figure 1. Mean correct response times (RT) in the chosen compared to directed conditions in the Attention Control Task. Error bars are ± 1 SEM, following the Loftus & Masson (1994) procedure for within-subjects designs.

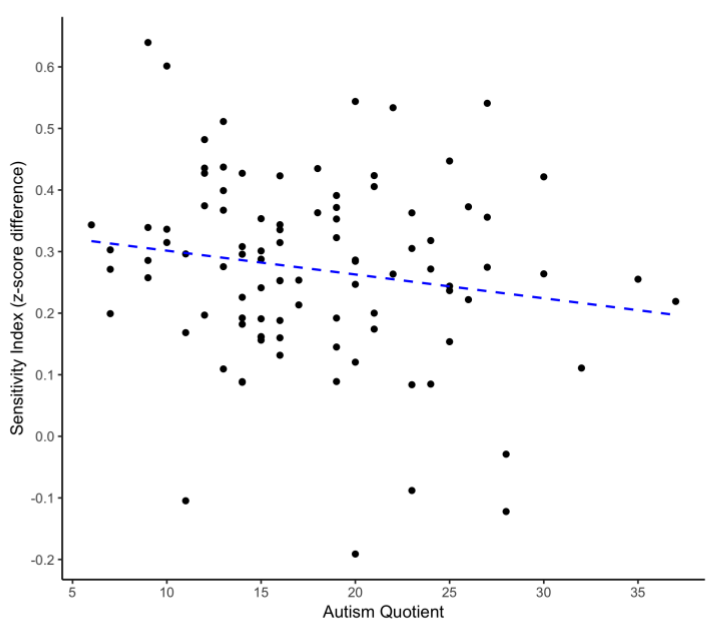

Figure 2. Scatterplot of the relationship between participants' speeded sensitivity scores and their AQ scores (N = 96).